

Molecular Basis of Amyloid Disease:  
In Silico Modelling of Beta-Amyloid  
(A $\beta$ ) Aggregation



**Beenish Khurshid**

**This dissertation is submitted for the degree of Doctor of  
Philosophy**

**December 2020**

**Department of Chemistry**



*To my mother,*

## **Declaration**

This thesis has not been submitted in support of an application for another degree at this or any other university. It is the result of my own work and includes nothing that is the outcome of work done in collaboration except where specifically indicated. Many of the ideas in this thesis were the product of discussion with my supervisor Professor Jamshed Anwar.

Beenish Khurshid

Lancaster University, UK

## List of publications

- I. **A systematic comparison of force fields for molecular simulation of hierarchical structures of  $\beta$ -amyloid peptides.**

Beenish Khurshid, Jamshed Anwar\*, David Middleton

*(Manuscript in preparation)*

- II. **Molecular Simulations Reveal a Sweet Spot for Fibril Formation of Beta Amyloid (17-42)**

Beenish Khurshid, Jamshed Anwar\*

*(Manuscript in preparation)*

- III. **Towards non-invasive diagnosis of AD molecular basis for the specificity of curcumin for fibrillar beta amyloid.**

Beenish Khurshid, Jamshed Anwar\*

*(Manuscript submitted to Angewandte Chemie)*

- IV. **Heparin-assisted amyloidogenesis uncovered through molecular dynamics simulations.**

Beenish Khurshid, Jamshed Anwar\*, Simon Boothroyd

*(Manuscript submitted to Neuroscience)*

## Abstract

Whilst the life expectancy in the world is increasing, there is an epic increase in the proportion of people with Alzheimer's disease (AD). Statistics show that over 35 million people all over the world and 0.8 million in the United Kingdom living with AD. his number is anticipated to double by 2030. AD is characterized by an immense loss of neurons leading to a loss of cognitive functions. Despite much effort and advances in scientific research and technology, we still don't know the exact mechanism of AD and as a result, there is still no cure for this detrimental disease. According to the Amyloid hypothesis, AD is caused by the aggregation of neurotoxic A $\beta$  which is cleaved from the C-terminal of a very large protein, the amyloid precursor protein (APP).

A molecular-level understanding is critical to understanding the development and progression of AD. Whilst molecular-level experimental techniques such as NMR spectroscopy, X-ray crystallography, or atomic force microscopy have advanced the field considerably, the early stages of A $\beta$  aggregation largely remain inaccessible by experiment. An effective alternative approach is the use of molecular dynamics simulation, which has been employed here.

The thesis aims to advance our molecular-level understanding of A $\beta$  aggregation and how it is modulated by selected endogenous and exogenous chemical co-factors. There are 4 aspects to the study: (i) a systematic comparison of the popular protein force fields to see whether they could reproduce various hierarchical structures of A $\beta$ ; (ii) self-assembly of A $\beta$  using both an all-atom and coarse-grained (CG) approach; (iii) interaction of the endogenous glycosaminoglycan heparan sulfate with A $\beta$ ; and (iv) interaction of curcumin (a predominant component of the spice turmeric) with A $\beta$  – curcumin is an inhibitor of A $\beta$  aggregation and also has potential for being a diagnostic for detecting amyloid.

The systematic comparison of protein force fields to see whether they could reproduce various hierarchical structures of A $\beta$  revealed that the choice depends on the size and complexity of the A $\beta$  structure and the nature of the problem being studied. For example, the GROMOS family of force fields is known and was demonstrated here to stabilize  $\beta$ -sheets. So, for studying  $\alpha$ -helix to  $\beta$ -sheet conversion during A $\beta$  oligomerization, GROMOS would be the preferred choice. Similarly, for larger more complex systems, our study shows that OPLS-aa is a better choice as it achieves a good

balance between  $\alpha$ -helix and  $\beta$ -sheet secondary structures and gives significantly less deviation (RMSD) for the larger A $\beta$  structures.

Heparan sulfate is endogenous and is known to co-exist within A $\beta$  deposits and can enhance fibril formation, suggesting a pathological connection. Here we have shown that it has a strong interaction with A $\beta$  and indeed forms a composite structure with A $\beta$ . Its flexibility is considered to be essential and we have also rationalized its molecular weight and concentration effects. Curcumin is a natural fluorescent dye and has a potential for serving as a non-invasive, highly specific diagnostic agent, along with its possible therapeutic role of inhibiting A $\beta$  aggregation. The simulations explain the nature of the high specificity of curcumin for A $\beta$  including the variation in contrast between the various A $\beta$  morphologies.

## Acknowledgments

I would like to express my sincerest gratitude to my mentor and research supervisor, Prof. Jamshed Anwar for his guidance, mentoring and support, throughout my Ph.D. I appreciate his continuous engagement with my research work and his open-door policy that gave me the confidence to discuss any hurdles in my work. I cannot thank him enough for always driving me to be my best.

My sincere gratitude to the Higher Education Commission, Pakistan, and Abdul-Wali Khan University, Mardan, Pakistan for funding my Ph.D.

I am grateful to my appraisal committee members Professor David Middleton and Dr. Michael Peach for their valuable discussions and insightful comments on my work throughout my Ph.D. With the support of the Department of Chemistry through my study period, I cannot thank enough Carol Cook for her constant support and for being always ready to help with any issue I faced during this time.

I gratefully acknowledge the support of all colleagues at the lab especially Simon Boothroyd for his help with any technical issues I faced. The use of computational facilities at computational lab and high-end computing facility (HEC), Lancaster University is also gratefully acknowledged. A special thanks to my dear friend Eleni Fitsiou for always being there whenever I needed any help with academics or personal stuff (with her lovely home-baked cakes, of course!).

My friends who made my stay at Lancaster memorable specially Arooj Mubashara Siddiqui and Zohra Mahillu, thanks to them for making it home away from home. A big thanks to my dear friend Dr. Muhammad Ali Hashmi, who was always a call away to solve technical issues related to any software or hardware. (I owe you a lot!). My mom and siblings for their unconditional love and support. My two lovely boys, Mustafa and Saad who were like a breeze of fresh air after I was stuck with running long simulations and writing the thesis. (Your smiles took my tiredness away). And finally, my dearest husband, for always believing in me and providing me with every opportunity to succeed and always giving me the freedom to use those opportunities. For his unfailing faith in me that sustained me through the toughest times. For motivating me to continue my Ph.D. after we had two kids during my Ph.D. and taking care of them while I was working long days and nights. Without you it could not have been possible!

# Contents

<b>1 INTRODUCTION.....</b>	<b>1</b>
1.1 Protein structure and function .....	1
1.2 Physiological role of protein self-assembly .....	2
1.3 Alzheimer's Disease .....	2
<i>Discovery and Definition of Alzheimer's Disease</i> .....	2
<i>Key Features of Alzheimer's disease</i> .....	4
<i>Molecular basis of Amyloidosis</i> .....	4
<i>Pathophysiology of Alzheimer's disease</i> .....	4
1.4 The mechanism of A $\beta$ aggregation .....	8
<i>Nucleation and template assisted self-assembly of amyloid peptides</i> .....	8
1.5 Structures of A $\beta$ Aggregates .....	9
1.6 Anti-AD therapies .....	10
1.7 Therapeutic issues .....	12
1.8 Role of molecular dynamics simulations in biomolecular research .....	13
1.9 Aims and objectives of the thesis.....	15
<b>2 METHODOLOGY .....</b>	<b>16</b>
2.1 Molecular simulations theory.....	16
<i>Statistical mechanics</i> .....	16
<i>Molecular simulations</i> .....	17
<i>Force Fields</i> .....	20
<i>System setting in MD simulations</i> .....	22
<i>Ewald summation</i> .....	23
<i>Periodic boundary conditions</i> .....	24
<i>Energy minimization</i> .....	25
<i>Equilibration</i> .....	27
<i>Production run</i> .....	29
2.2 MARTINI model.....	29
2.3 Analysis methods .....	30
<i>RMSD</i> .....	30
<i>RMSF</i> .....	30
<i>Inter-strand and inter-sheet distances</i> .....	31
<i>Hydrogen bond analysis</i> .....	31
<i>Clustering analysis based on rmsd</i> .....	31

<i>Twist angle</i> .....	32
<i>MM(PBSA) method</i> .....	32
2.4 Hamiltonian replica exchanges .....	33
<b>3 A SYSTEMATIC COMPARISON OF FORCE FIELDS FOR MOLECULAR SIMULATION OF HIERARCHICAL STRUCTURES OF B-AMYLOID PEPTIDES</b> .....	<b>35</b>
3.1 Abstract .....	35
3.2 Introduction .....	36
3.3 Methodology .....	41
<i>Data analysis</i> .....	42
3.4 Results .....	43
<i>A<math>\beta</math>42 monomer</i> .....	43
<i>Part-2 Fibrils</i> .....	46
<i>2BEG (U-shaped protofibril):</i> .....	47
<i>2MXU (S-shaped fibril):</i> .....	53
<i>M3Q_3-Ring fibril:</i> .....	58
3.5 Discussion .....	63
3.6 Conclusion: .....	68
3.7 References .....	69
<b>4 MOLECULAR SIMULATIONS REVEAL A SWEET SPOT FOR FIBRIL FORMATION OF BETA AMYLOID (17-42)</b> .....	<b>75</b>
4.1 Abstract .....	75
4.2 Introduction .....	76
4.3 Methodology .....	81
4.4 Results .....	85
<i>Secondary structure and topology of spontaneously aggregated A<math>\beta</math> peptides at various concentrations – all-atom MD</i> .....	85
<i>Coarse-grained molecular dynamics (MD) simulations at different concentrations</i> .....	87
4.5 Discussion: .....	91
<i>Concentration dependent A<math>\beta</math> peptides aggregation leads to the formation of various A<math>\beta</math> conformations</i> .....	91
<i>Non-specific hydrophobic interactions lead to the formation of diffused aggregates</i> .....	93
4.6 Implications .....	96
4.7 References .....	97

**5 TOWARDS NON-INVASIVE DIAGNOSIS OF AD MOLECULAR BASIS FOR THE SPECIFICITY OF CURCUMIN FOR FIBRILLAR BETA AMYLOID. 101**

5.1 Abstract ..... 101  
5.2 Introduction ..... 102  
5.3 Experimental section ..... 109  
5.4 References ..... 110

**6 HEPARIN-ASSISTED AMYLOIDOGENESIS UNCOVERED THROUGH MOLECULAR DYNAMICS SIMULATIONS..... 113**

6.1 Abstract ..... 113  
6.2 Introduction ..... 114  
6.3 Methodology ..... 117  
6.4 Results and discussion ..... 120

*KLVFFA interacts strongly with heparin and is unable to move up or down the heparin molecule..... 120*

*KLVFFA peptides aggregate to form ordered oligomers but not a fibrillar structure in 200 ns of MD simulation ..... 122*

*Aggregation of KLVFFA peptides on heparin takes place in two steps: selection and assembly..... 122*

*Heparin flexibility appears to be essential for its role in assembling peptides. 126*

*Heparin-peptide interaction is not transitory; heparin forms an integral part of the resulting protofibrils ..... 129*

*Lower heparin:peptide concentration ratio is more effective because peptide monomers do not get partitioned amongst individual heparin molecules..... 130*

*Shorter heparin chain lengths are less effective as they partition the peptides to form multiple heparin-peptide complexes that aggregate only slowly ..... 133*

*Accessing longer time scale ordering: heparin-induced assembly of pre-formed peptide dimers..... 134*

*Mapping molecular insights from heparin's assembly of KLVFFA to larger peptides and proteins ..... 135*

6.5 Summary and significance ..... 137  
6.6 References ..... 138

**7 SUMMARY AND CONCLUSIONS ..... 146**

7.1 Effects of force fields on peptides of variable size and complexity ..... 147  
7.2 Self-assembly pathways and forces driving them..... 147  
7.3 Towards non-invasive amyloid detection – interaction of curcumin with A $\beta$  species ..... 148

7.4 Endogenous species enhancing amyloid self-assembly (selection and assembly process) .....	150
7.5 Challenges and limitations of molecular simulation of A $\beta$ .....	151
<b>8 REFERENCES.....</b>	<b>155</b>
<b>9 APPENDICES .....</b>	<b>187</b>
Appendix A .....	188
9.1 A $\beta$ monomer .....	189
9.2 U-shaped protofibril.....	191
9.3 S-shaped fibril .....	193
9.4 Ring fibril.....	194
9.5 References.....	196
Appendix B .....	197
9.6 Methodology .....	197
<i>The interaction of curcumin with A<math>\beta</math>42 monomer .....</i>	<i>197</i>
<i>Effect of curcumin on self-assembly of A<math>\beta</math>42 monomers .....</i>	<i>197</i>
<i>Interaction of curcumin with A<math>\beta</math>42 fibril .....</i>	<i>197</i>
9.7 Technical details .....	198
9.8 Results.....	199
9.9 References.....	203
Appendix C .....	205

## List of Tables

Table 3.1. Average interpeptide Asp23 – Lys28 NH <sub>3</sub> <sup>+</sup> COO <sup>-</sup> distances (nm ± Std. Dev.) .....	52
Table 4.1: Simulation conditions showing Number of amyloids, water molecules, counter ions (NA), simulation box size, temperatures and simulation length.....	84
Table 4.2. The secondary structure content of various A $\beta$ systems.....	87
Table 4.3: Aspect ratio of beta amyloid aggregates; larger values show the formation of unidirectional structures.....	90
Table 6.1. Details of simulation studies carried out. This entire complement of simulations was repeated in a low ionic strength (system neutralized with counterions) and in a physiological ionic strength aqueous environment (150 mM NaCl). Unless otherwise stated the heparin chain length in terms of the degree of polymerization was dp = 24.....	119
Table 6.2. A breakdown of the free energy of binding $\Delta G$ of KLVFFA peptide with heparin in low and high ionic strength (150mM NaCl) aqueous environments calculated by the MM-PBSA method.....	122
Table 9.1 Summary of A $\beta$ structures simulated in this study.....	188
Table 9.2 Summary of simulation details .....	188
Table 9.3 The details of secondary structure content of A $\beta$ monomer.....	189
Table 9.4. Average interpeptide distances U-shaped protofibril (nm ± Std. Dev.) ...	191
Table 9.5. Average interpeptide Asp23 – Lys28 NH <sub>3</sub> <sup>+</sup> _ COO <sup>-</sup> distances (nm ± Std. Dev.).....	192
Table 9.6 Percentage of the secondary structure content S-shaped fibril calculated from DSSP program from 400ns long simulation for AmberGS, Gromos53a6 and OPLS-aa forcefields.....	193
Table 9.7 Interchain distances between HIS14, LYS17, PHE19 & 20, ASN27, ALA30 and LEU34 of two adjacent chains of S-shaped fibril. ....	194
Table 9.8. Interchain salt bridges in S-shaped fibril ( Average interpeptide Asp23 – Lys28 NH <sub>3</sub> <sup>+</sup> _ COO <sup>-</sup> distances (nm ± Std. Dev.) .....	194

Table 9.9. Average intrasheet distances ring fibril (nm $\pm$ Std. Dev.) .....	195
Table 9.10. Summary of A $\beta$ 42 models and simulations systems .....	198

## List of Figures

Figure 1.1. Cross-sections of the brain as seen from the front. Left represents a normal brain and the one on the right represents a brain with Alzheimer's disease. ( <i>Brain with Alzheimer's Disease</i>   <i>BrightFocus Foundation</i> , 2017).....	3
Figure 1.2. Mechanism of amyloid aggregation as explained by the amyloid hypothesis. Beta and gamma secretases snip off a segment of APP to yield the amyloid protein that assembles with other fragments giving rise to oligomers and then ultimately mature fibrils. These fibrils accumulate in the brain in the form of plaques. “Image courtesy of the National Institute on Aging/National Institutes of Health”(Alzheimer's Scientific Images and Video, no date).....	6
Figure 1.3. The pathway of amyloid assembly starting from A $\beta$ monomer to low-weight oligomers, protofibril, and large fibrils. (a) Solution structure of A $\beta$ 42 in the apolar microenvironment (PDB: 1iyt); (b) The structure of KLVFFA fragments of A $\beta$ 42 presenting the hydrophobic patch of A $\beta$ 42 responsible for higher-order assemblies (PDB: 3ow9) ; (c) Lower molecular weight oligomer ; (d) Structure of protofibril composed of parallel, in-register $\beta$ -sheets; (e) A $\beta$ (1-42) amyloid fibril displaying triple parallel- $\beta$ -sheet segments; (f) Key showing color codes for secondary structure content in (g); (g) The primary amino acid sequence A $\beta$ isoform (A $\beta$ 42). .....	11
Figure 2.1. (A). Lennard Jones potential; (B). Potential energy functions as a sum of its components. ....	22
Figure 2.2. Ewald summation: The sum is divided in the real space and reciprocal space by adding a Gaussian .....	24
Figure 2.3. An example of periodic boundary condition; the primary cell (grey) is surrounded by its images (white).....	25
Figure 2.4. The process of energy minimization takes place in a stepwise manner until a minimum is reached. ....	26
Figure 3.1. Free energy associated with homogenous nucleation of a sphere of radius r. .....	41

Figure 3.2. Representative snapshots of secondary structures of A $\beta$ <sub>42</sub> from a 400ns long trajectory under the influence of (upper panel): AmberGS, (middle panel): Gromos53a6 and (lower panel): OPLS-aa. ....	44
Figure 3.3. Representative data set from 400ns long simulation of A $\beta$ <sub>42</sub> : (A) Backbone RMSD of A $\beta$ <sub>42</sub> for the three force fields, (B) RMSF, (C) Radius of gyration of A $\beta$ <sub>42</sub> for the three force fields. ....	46
Figure 3.4. Structures of three fibrils of varying sizes and complexity i-e (A). U-shaped protofibril, (B). S-shaped fibril and (C) ring fibril, (D),(E) and (F) present the top views of U-shaped protofibril. S-shaped fibril and ring fibril. ....	47
Figure 3.5. Asp23 – Lys28 (COO - NH <sub>3</sub> <sup>+</sup> ) forming the salt bridge responsible for maintaining the U-shape of the U-shaped protofibril. ....	47
Figure 3.6. (A). Backbone RMSD for U-shaped protofibril from 400ns long simulations for AmberGS, Gromos53a6 and OPLS-aa forcefields. (B) Hydrophobic interactions between side chains of PHE19-GLY38 and ALA21-VAL36 of U-shaped protofibril, (C) Rg (nm) and (D) Average number of interpeptide hydrogen bonds for U-shaped protofibril in AmberGS, Gromos53a6 and OPLS-aa force fields.....	48
Figure 3.7. Representative data set from 400ns long simulation of AmberGS: (A) structure of the U-shaped protofibril after 400 ns, generated with VMD; Top, centre and lower chain's ends are denoted by red balls, (B) RMSD values of representative peptide chain, (C) RMSF values of representative peptide chains and (D) Asp23 – Lys28 (COO - NH <sub>3</sub> <sup>+</sup> ) distances for all protofibrillar chains. ...	49
Figure 3.8. Representative data set from 400ns long simulation of Gromos53a6: (A) structure of the U-shaped protofibril after 400 ns, generated with VMD; Top, centre and lower chain's ends are denoted by red balls, (B) RMSD values of representative peptide chain, (C) RMSF values of representative peptide chains and (D) Asp23 – Lys28 (COO - NH <sub>3</sub> <sup>+</sup> ) distances for all protofibrillar chains. ...	50
Figure 3.9. Representative data set from 400ns long simulation of OPLS-aa: (A) structure of the U-shaped protofibril after 400 ns, generated with VMD; Top, centre and lower chain's ends are denoted by red balls, (B) RMSD values of representative peptide chain, (C) RMSF values of representative peptide chains and (D) Asp23 – Lys28 (COO - NH <sub>3</sub> <sup>+</sup> ) distances for all protofibrillar chains. ...	51

Figure 3.10 The dihedral angle between C-alpha atoms of VAL18 and VAL24 of two adjacent $\beta$ -sheets. ....	52
Figure 3.11. Representative snapshots of S-shaped A $\beta$ fibril at 400ns. (A) Initial structure (B) AmberGS, (C) Gromos53a6, and (D) OPLS-aa force fields.....	54
Figure 3.12. Representative data set of S-shaped fibril.from 400ns long simulation of AmberGS, Gromos53a6 and OPLS-aa force fields. (A) RMSD values of representative S-shaped fibril showing the structure in AmberGS has the lowest RMSD (B) Number of residues adopting $\beta$ -sheet structure; increasing for Gromos53a6 as a function of time, (C) Radius of gyration showing OPLS-aa has a higher value of gyration showing more extended conformation (D) Total number of hydrogen bonds in the three force fields. ....	55
Figure 3.13. Representative data set from 400ns long simulation of AmberGS: (A) RMSD values of representative peptide chains; Top, centre and lower chains (B) RMSF values of representative peptide chains (C) Representative structures of Top, centre and lower chains of S-shaped fibril generated by VMD, (D) Average interpeptide distances.....	56
Figure 3.14. Representative data set from 400ns long simulation of Gromos53a6: (A) RMSD values of representative peptide chains; Top, centre and lower chains (B) RMSF values of representative peptide chains (C) Representative structures of Top, centre and lower chains of S-shaped fibril generated by VMD, (D) Average interpeptide distances.....	57
Figure 3.15. Representative data set from 400ns long simulation of OPLS-aa: (A) RMSD values of representative peptide chains; Top, centre and lower chains (B) RMSF values of representative peptide chains (C) Representative structures of Top, centre and lower chains of S-shaped fibril generated by VMD, (D) Average interpeptide distances.....	58
Figure 3.16. (A) Residues defining twist angle dihedral angle between C-alpha atoms of VAL12 and VAL18, Representative snapshots of the S-shaped fibrils, (A) AmberGS, (B) Gromos53a6 and (C) OPLS-aa. Red stick representations are to point out residue VAL18 of $\beta$ 1-sheet where twisting starts. ....	59
Figure 3.17. Representative data set of ring fibril.from 400ns long simulation of AmberGS, Gromos53a6 and OPLS-aa force fields. (A) RMSD values of	

representative ring fibril showing the structure in AmberGS have the highest RMSD, (B) Radius of gyration of ring fibril; largest for ring fibril as a function of time. ....	60
Figure 3.18. (A). Ring fibril in AmberGS, a snapshot after 400ns (B) Secondary structure evolution of ring fibril indicating an increase in $\alpha$ -helical structure (C) RMSD values for top, center, and bottom chains (D) Intra-sheet hydrogen bonds between representative chains of ring fibril.....	61
Figure 3.19. (A). Ring fibril in Gromos53a6, a snapshot after 400ns(B) Secondary structure evolution of ring fibril indicating stable $\beta$ -sheet and coil structure (C) RMSD values for top, center and bottom chains (D) Intra-sheet hydrogen bonds between representative chains of ring fibril.....	62
Figure 3.20. (A). Ring fibril in OPLS-aa, a snapshot after 400ns (B) Secondary structure evolution of ring fibril indicating stable $\beta$ -sheet, bend, and coil structure (C) RMSD values for top, center and bottom chains (D) Intra-sheet hydrogen bonds between representative chains of ring fibril. Secondary structure evolution (C) RMSD values for top, center, and bottom chains (D) Intra-sheet hydrogen bonds. ....	63
Figure 4.1. A schematic diagram of nucleation-dependent self-assembly of A $\beta$ .....	78
Figure 4.2. The initial CG structure of A $\beta$ showing backbone beads in grey and side chains in blue. ....	84
Figure 4.3. Atomistic simulations carried out at five different concentrations of A $\beta$ revealed various morphologies of A $\beta$ showing some local order but were not able to give rise to unidirectional A $\beta$ fibrils; Total number of peptides are (a) 20, (b) 40, (c) 60, (d) 80 and (e) 100 .....	86
Figure 4.4. Solvent accessible surface area (SASA in nm <sup>2</sup> ) vs time (ns) shows there is a drastic decrease in SASA in the first few nanoseconds. After which it which remains converged to an equilibrium till the end of the simulation without any further decrease. This implies a very fast aggregation of A $\beta$ peptides that gives rise to $\beta$ -sheet rich aggregates that remain trapped in their conformation throughout the simulation.....	86

Figure 4.5. Structures of aggregates formed by A $\beta$ at various peptide concentrations for 400ns .....	89
Figure 4.6. The plots show the rapid assembly of A $\beta$ into a single aggregate. (A) plot of the number of aggregates (clusters) vs time shows that higher the concentration, more quickly the aggregates are formed. (B) Zoom in to the time zone where the aggregate formation is taking place from monomers to oligomers and finally aggregates (for which number of clusters = 1); (C) Decrease in the radius of gyration; dramatic decrease at the lowest concentration because a spherical structure is formed. ....	90
Figure 4.7. Effect of temperature on the shape of the fibril. The figure shows that with an increase in temperature the unidirectional fibril like structure collapses into globular aggregate.....	91
Figure 4.8. Spherical aggregate formed by 20 peptides. Blue beads show the hydrophobic core while red beads are charged.....	95
Figure 4.9. Histograms of the potential energy distribution of the individual replicas differentiated by the colors obtained from Hamiltonian replica exchange MD for a self-assembling system containing 40 peptides in explicit water at three different exchange time-steps (a). 10; (b) 500 and (c) 1000. The curve on the extreme left is for unmodified parameters while the right one is for most strongly modified force field parameters.....	96
Figure 5.1. Structure of curcumin showing the hydrophobic linker region and the polar substituted aromatic rings. ....	103
Figure 5.2. Snapshots of the interaction of A $\beta$ 42 with curcumin 200 ns into the trajectory. (Top) Initial A $\beta$ helical structure obtained from the protein data bank (PDB:1IYT). (Bottom, Left) Without curcumin; (Bottom, Middle) A $\beta$ and curcumin in a 1:1 ratio; (Bottom, Right) A $\beta$ and curcumin in a 1:4 ratio. Curcumin is represented by a skeleton line structure in red. The peptide structure is shown in cartoon representation where yellow color represents $\beta$ -strands, red ribbon-like structure represents $\alpha$ -helix, and green and white colors show turns and coil regions respectively. It is evident from the figure that curcumin destroys the $\beta$ -sheet structure in its vicinity, and the higher the concentration of curcumin, the lower is the $\beta$ -sheet content.....	106

Figure 5.3. Binding of curcumin to A $\beta$ fibril. (a) The fibril is shown in cartoon representation (yellow- $\beta$ -sheets, green –turns, white-coil), while curcumin is represented as a surface in red to give a clear perspective of the highly-specific binding, (b) An enlarged image of the fibril end where curcumin is seen to enter into the core of the fibril. (c) A plot of the population of curcumin molecules oriented and located at various positions on the A $\beta$ fibril. ....	107
Figure 6.1. Chemical structure of the disaccharide repeating unit of heparin. The saccharide unit on the left is glucosamine (GlcN) whilst that on the right is iduronic acid (IdoA).....	116
Figure 6.2. End to end distance analysis of heparin chain at two ionic strengths in an aqueous environment. (a) Heparin chain alone. (b) Heparin in the presence of a single KLVFFA peptide.....	122
Figure 6.3. (a) Snapshot of the self-assembly of KLVFFA peptides (monomers) at 200 ns simulation time in 150 mM NaCl in water. (b) Number of peptide aggregates formed during the self-assembly of KLVFFA peptides as a function of time for both the 0 mM and 150 mM ionic strength systems. (c) Maximum aggregate size for the self-assembly of KLVFFA peptides as a function of time for both the 0 mM and 150 mM ionic strength systems. (Cluster criteria: cut-off of 0.35 nm).....	123
Figure 6.4. Evolution of the number of aggregates inclusive of heparin and peptides and peptides-only. The plot of peptides and heparin aggregates (black) reflects the initial, rapid association of the peptides with the heparin (within about 45 ns), whilst the peptides-only aggregates plot (red) reflects the relatively slower, facilitated assembly of the peptides. Note the numerous, repetitive peptide-peptide cluster making and breaking events as the heparin flexes to align and order the emergent protofilament. ....	124
Figure 6.5. (a) Snapshots of KLVFFA peptides–heparin structures in 0 mM NaCl at 200 ns for three separate simulations each started from a different random configuration. (b) Snapshot of the complex from the second simulation system (c) Snapshot of the complex from the third simulation but with a “surface” representation to show the formation of a contiguous structure. Heparin and peptides are shown in CPK and cartoon (secondary structure) representation respectively. ....	125

Figure 6.6. Snapshots of KLVFFA peptides–heparin structures in 150 mM NaCl at 200 ns for three separate simulations each started from a different random configuration, showing wide variation in the emergent morphology ..... 127

Figure 6.7. (a) Total number of aggregates (including KLVFFA peptide aggregates and heparin) as a function of time at both ionic concentrations, reflecting rapid localization of the peptide molecules onto the heparin. (b) Number of peptides-only aggregates as a function of time at both ionic concentrations. The plot shows that peptides assemble into a single structure when heparin is fully flexible within 100ns whilst there are still 3-4 loose aggregates when heparin is restrained in its extended state. (c) Snapshots of the heparin-peptides composite for the heparin molecule when fully flexible at 0 mM. (d, e) Heparin restrained in its fully extended form at 0 mM and 150 mM NaCl. For these restrained heparin systems, the lack of a contiguous peptide structure is apparent. .... 128

Figure 6.8. The KLVFFA peptide-heparin composite resulting from heparin’s assembly of KLVFFA peptides. Both the formed peptide protofilament and the heparin form helical structures that are intertwined. The heparin helix is characterized by a pitch consisting of four disaccharides with a translation of 2.0–2.1nm along the axis. (a) Heparin is represented as VDW beads (sulfur in yellow; oxygen in red; carbon in turquoise), whilst the A $\beta$  peptide protofilament is represented in surface form. (b) represents peptides’ secondary structure intertwined against heparin shown in CPK representation. .... 129

Figure 6.9. Assembly of KLVFFA peptides at various heparin: peptide molar ratios at 0mM (Left-hand side) & 150mM NaCl (Right-hand side). (a,d). 1:100; (b,e). 1:20; (c, f). 2:20 heparin: peptide molar ratio. The encircled structure in (d) shows the formation of a well-ordered fibril at 150mM..... 131

Figure 6.10. The effect of different heparin chain lengths on heparin-facilitated assembly of KLVFF peptides at 0 mM. (a) dp2, (b) dp4, (c) dp6, (d) dp8 and (e) dp24; where dp is the number of saccharide units. Longer heparin chain lengths (dp=8 and dp=24) are more effective at facilitating peptide assembly. Shorter heparin chain lengths (dp=4 and dp=6).are less effective as they form multiple heparin-peptide complexes that then aggregate only slowly. The disaccharide (dp=2) complexes show a marginal reversal in trend by aggregating together,

courtesy of their faster diffusion rate that enhances complex-complex interactions.  
 ..... 132

Figure 6.11. Number of KLVFFA peptide aggregates observed for the various heparin chain lengths (dp2-dp24) at the end of each simulation. (a) At low ionic concentration, larger chain lengths favor greater KLVFFA peptide aggregation (fewer aggregates), though shortest chain length dp2 shows distinct behavior and also promotes peptide aggregation. (b) At higher ionic concentration the aggregation increases (fewer aggregates) with an increase in chain length up to dp8, whilst dp24 doesn't enhance further aggregation ..... 134

Figure 6.12. (a) Snapshot of the self-assembly of preformed dimers of the peptide KLVFFA at 200 ns. (b) Total no. of aggregates (peptides only) as a function of simulation time for self-assembly of monomers and pre-formed peptide dimers. Dimers rapidly come together while monomer fragments take time to order and form antiparallel  $\beta$ -sheets resulting in slower aggregation. .... 135

Figure 6.13. Heparin-facilitated assembly of preformed dimers of KLVFFA in the presence of 1 molecule of heparin. The trajectory reveals two distinct stages, rapid association and localization of the peptides onto the heparin which occurs in about 5 ns, and the heparin-facilitated assembly of the associated peptides into a protofilament, which is concluded well before 200 ns. .... 136

Figure 6.14. Total no. of aggregates (inclusive of heparin and peptides) as a function of simulation time for the heparin-facilitated assembly of peptide monomers and pre-formed peptide dimers. Dimers are rapidly attracted to the heparin to form a single aggregate (only 5ns) while monomers attachment is slower. .... 137

Figure 7.1. (A) Thioflavin, (B) Congo red, (C) Curcumin. All three molecules are linear with aromatic groups separated by a linker region. .... 153

Figure 9.1. The averaged secondary-structure probability per residue in terms of regular secondary structure characteristics i-e  $\alpha$ -helices,  $\beta$ -sheets, turns and coils etc. The structure in AmberGS shows an elbow-shaped structure containing two  $\alpha$ -helices separated by a turn region at residue VAL24-ASN27, Gromos53aa6 favors the formation of  $\beta$ -sheets while OPLS-aa forms structure that is a mixture of helix and turns..... 189

Figure 9.2 The five most populated clusters for each simulated force field, the percentage of snapshots each cluster represents. (RMSD cut-off 0.25nm) and the end-to-end distance of the A $\beta$ 42 monomer showing a change in the structure of various conformations.....	190
Figure 9.3. Secondary structure of five most populated clusters of A $\beta$ 42 monomer showing change in the structure of various conformations for each simulated force field, calculated from stride* .....	190
Figure 9.4. C $\alpha$ -C $\alpha$ distance between ALA21 and VAL36 in the last two chains of U-shaped protofibril in AmberGS showing large fluctuations in the lower end of the protofibril. ....	192
Figure 9.5. C $\alpha$ -C $\alpha$ distance between PHE19 and GLY38 and ALA21 and VAL36 in the top and bottom chains of U-shaped protofibril in OPLS-aa showing fluctuations in the beginning of the simulations but later converge to an equilibrium distance. ....	193
Figure 9.6. Asp23 – Lys28 (COO - NH <sub>3</sub> <sup>+</sup> ) forming the salt bridge responsible for maintaining the U-shape of the U-shaped protofibril. ....	193
Figure 9.7. The three segments of this 3-fold symmetry fibril originally at 60° to each other but undergo a change especially for AmberGS and Gromos53a6 which give rise to distorted geometry. Out of three structures, (c) shows the least deviation from the original structure that is produced in OPLS-aa .....	194
Figure 9.8. Upper panel: (a). A typical oligomer formed at 0mM curcumin (control); (b) 16.4mM curcumin & (c) 64mM curcumin system. At higher concentration curcumin destroys $\beta$ -sheets and stabilizes off-pathway oligomers; Lower panel: The same concentrated system (c) containing 64mM curcumin and A $\beta$ aggregates. Snapshot from a 100ns long simulation shows integrated curcumin-A $\beta$ clusters via non-specific interactions. A $\beta$ are represented by Green colour while curcumin molecules are represented by red colour. (Water not shown for simplicity). ....	199
Figure 9.9. Interaction of curcumin with A $\beta$ fibril at lower concentration of curcumin. (a) Figure shows the binding of curcumin with A $\beta$ fibril in three different modes. (1 & 3) Curcumin parallel to the beta strands (BE= -30 & -25 kcal mol <sup>-1</sup> ). (2) Curcumin interacting with the loop region between the $\beta$ -strands parallel to fibril axis (BE=26 kcal mol <sup>-1</sup> ). (4) Curcumin going into the core of the fibril (BE=-27	

kcal mol<sup>-1</sup>); Curcumin is represented as line structure in red (b) Top view of curcumin entering the central cavity. (c) The G33XXXG37 motif showing the main residues involved in binding with curcumin. .... 200

Figure 9.10. Interaction of curcumin with A $\beta$  fibril at lower concentration of curcumin.

(a) Figure shows the binding of curcumin with A $\beta$  fibril in three different modes. (1 & 3) Curcumin parallel to the beta strands (BE= -30 & -25 kcal mol<sup>-1</sup>). (2) Curcumin interacting with the loop region between the  $\beta$ -strands parallel to fibril axis (BE=26 kcal mol<sup>-1</sup>). (4) Curcumin going into the core of the fibril (BE=-27 kcal mol<sup>-1</sup>); Curcumin is represented as line structure in red (b) Top view of curcumin entering the central cavity. (c) The G33XXXG37 motif showing the main residues involved in binding with curcumin. .... 201

Figure 9.11. Per-residue analysis of interaction energies of all the residues of A $\beta$  fibril with curcumins calculated by MM-PBSA method. The periodicity of peaks represents the specific residues that are involved in the binding of curcumin to A $\beta$  including methionine, glycine, valine and isoleucine. Negative values indicate preferable interaction between a residue and curcumin molecule. The highest negative peaks representing glycines imply that Curcumin prefers to interact with GLY with both of its aromatic rings with a very strong binding. .... 202

Figure 9.12. Comparison of peptide (monomer) position on heparin's saccharide units at various time intervals. (a) Fully flexible heparin (dp24) at low ionic strength. (b). Position restrained heparin (dp24) at low ionic strength. (c) Fully flexible heparin (dp24) at high ionic strength. (d) Position restrained heparin (dp24) at high ionic strength. It is clear from the plots that the movement of peptides is more restricted for the position-restrained heparin. .... 205

Figure 9.13. (a) Minimum distance between various residues of KLVFFA and negatively charged sulpho groups of heparin (cut-off = 0.6nm). The results indicate that the lysine residues of the KLVFFA peptides are involved in binding. Once the contacts are formed, they remain conserved throughout the simulation time(200ns). (b) The major contribution to the binding includes electrostatic interactions between cationic sites (lysine) on the peptides and anionic sites on heparin. (c) Peptide klvffa (dimer) binding to heparin, the interacting residues (lysines in both the strands) are shown in surface presentation. .... 206

Figure 9.14. The effect of different heparin chain lengths on heparin-facilitated assembly of KLVFFA peptides at 150mM. As the chain length increases, aggregation is also enhanced. Shorter heparin chain lengths (dp=4 and dp=6).are less effective as they form multiple heparin-peptide complexes that aggregate only slowly. (a) dp2, (b) dp4, (c) dp6, (d) dp8 and (e) dp24. Heparin and peptides are shown in CPK and cartoon (secondary structure) representation respectively.. 207

Figure 9.15. (a) The extended structure of a heparin molecule in neutral aqueous solution. The end-to-end distance of heparin is 11.5nm. (b) Snapshot of the heparin molecule in the presence of peptides (pre-dimerized) (not shown) where the end-to-end distance reduces to around 4nm. (c) Comparison of end to end distance of heparin in water and in the presence of peptides showing that it reduces significantly when heparin interacts with the peptides (dimers). ..... 208

## List of Abbreviations and Acronyms

A $\beta$	Amyloid-beta
AD	Alzheimer's disease
APP	Amyloid precursor protein
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
R <sub>g</sub>	Radius of gyration
CSF	Cerebrospinal fluid
SDS	Sodium dodecyl sulfate
FDA	Food and Drug Administration
DHA	Docosa-hexaenoic acid
ROS	Reactive oxygen species
PME	Particle mesh Ewald
SASA	Solvent accessible surface area
CNS	Central nervous system
PET	Positron emission spectroscopy
NMR	Nuclear magnetic resonance
FRET	Förster (Fluorescence) Resonance Energy Transfer
VMD	Visual molecular dynamics
GROMACS	GRoningen MACHine for Chemical Simulations
AMBER	Assisted Model Building with Energy Refinement

GROMOS GRoningen Molecular Simulation

OPLS Optimized Potentials for Liquid Simulations



# 1 Introduction

## 1.1 Protein structure and function

Proteins are undoubtedly the most diverse macro-molecules that perform a variety of functions including their role as neurotransmitters, enzymes, catalysts, signal transducers, and in cellular transport and storage. Proteins are composed of a linear chain of amino acids linked together by peptide bonds. The peptide bonds join the amino group of one amino acid to the carboxyl group of a neighboring amino acid, giving rise to the primary structure of the proteins. The chemistry of amino acids is very important because their side chains are involved in holding a specific conformation or structure of the protein. The folding of the primary structure of protein together with intra-peptide side chain bonding drives the evolution of protein's unique three-dimensional structure. Hydrogen bonding between the amide and carbonyl groups of the main chain of peptides is the basis of the secondary structure of proteins. The important secondary structures are the alpha helices, beta strands, turns, and sheets (Pauling, Corey, and Branson, 1951). There are 3.6 residues per turn in an alpha helix, which means that there is one residue every 100 degrees of rotation ( $360/3.6$ ). This regular ordering of residues allows all amino acids in the chain to form hydrogen bonds with each other giving rise to alpha helices. The hydrogen bonds are formed between the carbonyl (C=O) group of the upper coil and amide group (N-H) of the lower turn. These hydrogen-bonding interactions also give rise to parallel and anti-parallel  $\beta$ -sheets (which are often represented by flattened arrows). Two anti-parallel  $\beta$ -strands give rise to a secondary structure motif known as a  $\beta$ -hairpin. The loop between the two  $\beta$ -strands is called a  $\beta$ -turn. Turns and loops are important secondary structure components

connecting  $\beta$ -strands to each other and  $\alpha$ -helices, or helices to each other, thus playing a very important role in the folding process (Berg, Tymoczko and Stryer, 2002). The tertiary structure is formed by local folding of single linear amino acids chain into polypeptides and multiple polypeptide chains together form the quaternary structure of proteins. Proteins must be present in a folded native conformation to perform their biological functions. However, abnormal changes in amino acid sequence (point mutations) result in the misfolding of the primary structure leading to the loss of 3D structure associated with many misfolding diseases such as Alzheimer's disease (Dobson, 2004).

## **1.2 Physiological role of protein self-assembly**

A primary question concerning protein self-assembly is what is its physiological role in normal cells? The self-assembly of proteins is ubiquitous in biological systems and is crucial to the formation of highly ordered protein architectures that support cellular matrix and connective tissues (Sanders *et al.*, 2016). Some proteins undergo assembly by design e.g. virus capsid assembly (Selivanovitch and Douglas, 2019) while others do so when something goes wrong i-e. a conformational change in the protein, an amino acid mutation, or changes in the cellular environment such as solvent, temperature or pH (Carrick *et al.*, 2007). The precursor to A $\beta$  production is the protein APP (amyloid precursor protein) which itself plays a very important role in many biological activities such as regulating multiple aspects of neuronal development (Ramaker *et al.*, 2016). Other than A $\beta$  being linked with plaque formation, there is ample evidence that A $\beta$  has numerous beneficial roles in human physiology. A $\beta$  peptides may contribute to antimicrobial activity (Gosztyla, Brothers and Robinson, 2018), recovery from brain injuries (McKenzie *et al.*, 1996), maintaining neuron development (Morley and Farr, 2014), suppressing certain cancer forms (Ma *et al.*, 2014), and sealing leaks in blood-brain-barrier (Bishop and Robinson, 2002).

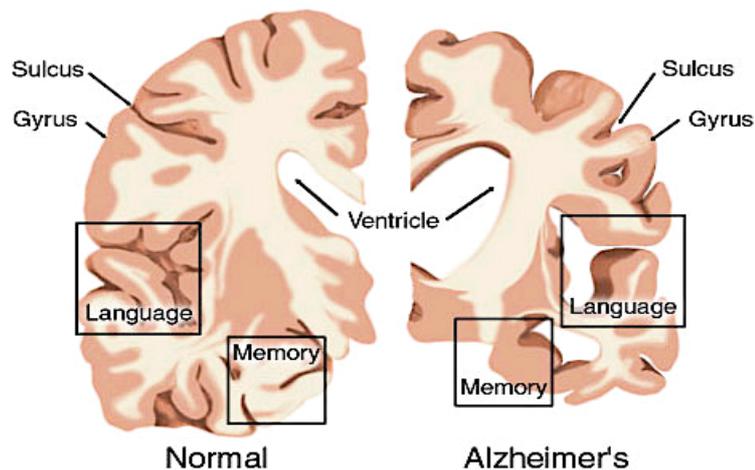
## **1.3 Alzheimer's Disease**

### Discovery and Definition of Alzheimer's Disease

Alzheimer's disease was first described in 1911 by the German psychiatrist and neuropathologist, Alois Alzheimer, who is considered as the father of neuropathology. He discovered the main characteristics of dementia i-e inexorable memory loss, aphasia,

delusions, and cognitive impairment, and linked them to physical changes in brain tissue using Bielschowsky's silver staining technique (Graeber, 1998; Zilka and Novak, 2006). Alzheimer's disease is the most common cause of dementia which is a neurodegenerative disease characterized by memory loss, cognitive impairment, and difficulties with language and problem solving. Worldwide, an estimated 54 million people are living with dementia. With respect to the UK, there are currently around 850,000 people with dementia which is expected to rise to 1.6 million and the total cost of social care of dementia patients is projected to reach £94.1 million by 2040 ('Alzheimer's disease facts and figures', 2019).

During the course of this disease, extracellular plaques and intracellular tangles are formed in the brain, which leads to the loss of connection between neurons. This eventually causes cell apoptosis and loss of brain tissue (Pant, Mukhopadhyay, and Lakshmayya, 2014). As the brain shrinks in volume because of neuronal cells death, symptoms become more severe (see Figure 1.1). During the early stages of damage, the hippocampus is affected which is involved in day-to-day memory storage. The earliest symptoms include difficulty in recalling recent events and learning new information. Often the long term memory remains unaffected (Kung, 2012). In later stages, the condition becomes worse and the patient suffers from complete cognitive impairment.



**Figure 1.1. Cross-sections of the brain as seen from the front. Left represents a normal brain and the one on the right represents a brain with Alzheimer's disease. (Brain with Alzheimer's Disease | BrightFocus Foundation, 2017)**

### Key Features of Alzheimer's disease

In AD, the normally soluble proteins accumulate in the extracellular space of various tissues as insoluble deposits of 10nm fibrils that are rich in  $\beta$ -sheet structure and have characteristic dye-binding properties (Glennner, 1980). The common structural motif of almost all amyloid fibrils is the  $\beta$  sheets in which the peptide strands are arranged perpendicular to the long axis of the fiber, irrespective of the protein sequence (Bucciantini *et al.*, 2002).

### Molecular basis of Amyloidosis

The molecular basis of amyloidosis is protein misfolding. When a polypeptide chain fails to remain in its native state or loses its 3-D structure, the potential number of misfolded or partially folded structures increases. In this aberrant state the hydrophobic residues are more exposed to the solvent, resulting in the formation of oligomers that behave as primary nucleus or seed for further aggregation. The peptide chains come closer and form a critical nucleus or a seed (nucleation phase) and then elongate to give rise to a mature fibril (elongation phase) (Stefani and Dobson, 2003). The process of protein assembly was once considered to follow a linear path that initially resulted in intermediate species that were very similar to the final native structures (Jahn and Radford, 2005). The current perspective is that the partially folded species are heterogeneous and there is no single way for them to go to the native states. On their journey towards the native state, peptides undergo many conformational changes and if suitable contacts are formed, the heterogeneity is reduced (Gilmanshin, Dyer and Callender, 1997). This is the case for smaller peptides having amino acid residues less than 100. For larger systems (amino acids greater than 100) with complex and bigger side chains compact annular species are formed as a result of the interaction of side chains with the solvents. This leads to partially folded intermediate states. Such states can either lead to native states or misfolded and aggregated states (Hartl, Bracher, and Hayer-Hartl, 2011).

### Pathophysiology of Alzheimer's disease

The three important hallmarks of AD are amyloid plaques, neurofibrillary tangles, and neuronal cell deaths. The underlying biochemical and physical pathways for AD are complex involving various chemical messengers and pathophysiological processes.

Various theories have been put forward to explain the cause of the disease. The major ones are:

- Amyloid hypothesis
- Cholinergic hypothesis
- Glutamatergic hypothesis
- Oxidative stress hypothesis
- Cholesterol
- Chronic inflammation

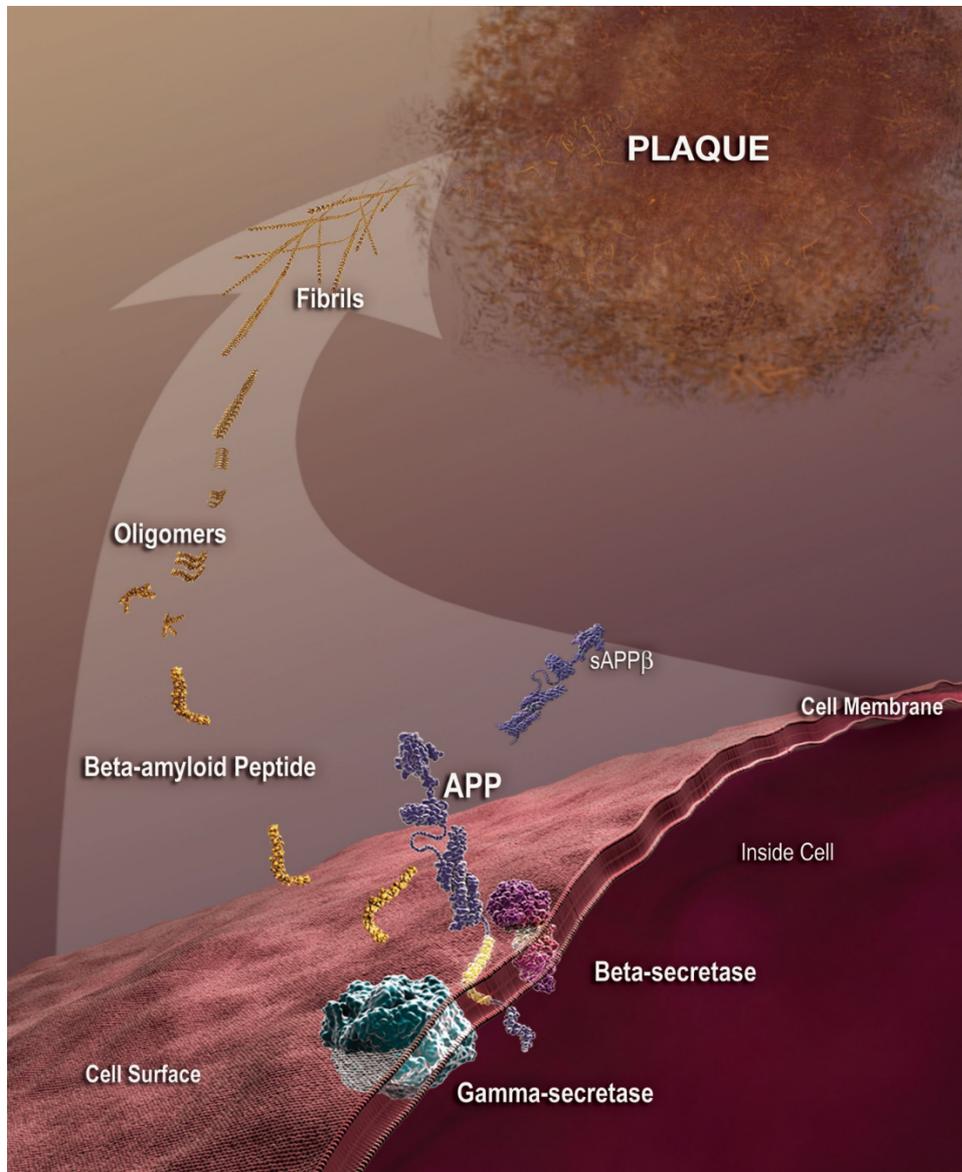
### The amyloid hypothesis

The exact molecular mechanism of AD is still not known, yet one theory, the “amyloid hypothesis” overshadows all other theories as to the factors that cause and lead to neuropathological damage. This thesis focuses on the amyloid hypothesis to investigate the process of aggregation of A $\beta$  and how it is or can be modulated by other molecules both endogenous (the glycan heparan sulfate) and exogenous (curcumin) as well their interaction with the resulting A $\beta$  assemblies. The amyloid hypothesis views extracellular A $\beta$  deposits in the hippocampus and overlying the cortical region of the brain as the fundamental cause of the disease. The major protein in neuritic plaques is A $\beta$ , a 40-42 amino acid peptide, which is derived from  $\beta$ -amyloid precursor protein (APP) after sequential cleavage by  $\beta$ -site APP cleaving enzyme 1 (BACE1). As a result, C99 fragment of APP is released giving rise to a variety of A $\beta$  peptides of varying residue lengths during subsequent cleavage by  $\beta$ -secretase (see Figure 1.2). Normally, the A $\beta$  peptides that are present in the brain in soluble form are released into the extracellular matrix and then rapidly (Abramowski *et al.*, 2008) cleared by the cerebrospinal fluid (CSF) directly into the blood (Tarasoff-Conway *et al.*, 2015). The two most common A $\beta$  isoforms present in the CSF of normal brains are A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub>. The A $\beta$  peptides form varying molecular weight oligomers that are released slowly from the brain. These oligomers aggregate to form fibrous aggregates of toxic proteins that are believed to cause Alzheimer's (Kung, 2012).

### Tau-Hypothesis

Healthy neurons are supported by microtubules, which are involved in the transport of nutrients and neurotransmitters from the body of neurons to axons. Tau protein stabilizes microtubules and it generally contains many phosphate groups. In AD, a large number of phosphate groups get attached to the tau proteins causing

“hyperphosphorylation”. As a result, Tau disconnects from microtubules and gets entangled with other tau threads from tangles. Tangles are abnormal or twisted proteins found inside the nerve cells. So the microtubules get destroyed and the internal network of transport collapses causing cell death (Chou, 2014).



**Figure 1.2. Mechanism of amyloid aggregation as explained by the amyloid hypothesis. Beta and gamma secretases snip off a segment of APP to yield the amyloid protein that assembles with other fragments giving rise to oligomers and then ultimately mature fibrils. These fibrils accumulate in the brain in the form of plaques. “Image courtesy of the National Institute on Aging/National Institutes of Health” (*Alzheimer’s Scientific Images and Video*)**

### Cholinergic Hypothesis

Acetylcholine is an important neurotransmitter in the brain, which is involved in memory and moods. In later stages of the AD number of cholinergic neurons is decreased. It is thought that by preventing degradation of cholinergic neurons and associated loss of acetylcholine neurotransmission or by increasing its level in the brain may reduce cognitive impairment seen in AD patients (Rogawski and Wenk, 2003).

### Oxidative Stress Hypothesis

Amyloid plaques are believed to cause lipid peroxidation in AD brains, which give rise to reactive oxygen and nitrogen species. These species react with other molecules and achieve a stable state releasing high-energy free electrons (free radical). These free radicals cause cell damage and react with all macromolecules including lipids, proteins, and carbohydrates. The brain is susceptible to oxidative stress because of its high oxygen consumption (Zhao and Zhao, 2013). The net effect of oxidative stress on neurons is to make them more vulnerable to damage.

### Glutamatergic Hypothesis

Postsynaptic glutamate receptors N-methyl-D-aspartate (NMDA) are associated with neuronal plasticity i-e learning and memory. Dysregulation at this receptor due to the interaction with amyloids, is thought to initiate a vicious cycle of neuronal damage. It also increases the production of APP leading to increased rates of plaque development and hyperphosphorylation of tau protein (thus NFT formation), oxidative stress followed by neuronal toxicity (Butterfield and Pocernich, 2003).

### Cholesterol

Elevated cholesterol levels appear to increase A $\beta$  production, whereas reduced cholesterol synthesis reduces A $\beta$  levels, thus the risk of AD progression (Sjögren and Blennow, 2005).

### Chronic Inflammation

As a natural response to the amyloid plaques, tangles, and damage to the brain, inflammation is stimulated. Microglia are the nerve cells, which are involved in the defense system of the brain. Thus, they release cytotoxic molecules such as cytokines and reactive oxygen species to counter the cell damage. But they do more harm than

protection. This cytotoxicity stimulates other inflammatory processes which may lead to apoptosis or cell death (Pimplikar, 2014).

## **1.4 The mechanism of A $\beta$ aggregation**

Nucleation and template-assisted self-assembly of amyloid peptides

The formation of amyloid aggregates is most likely a nucleation dependent polymerization mechanism. The rate-limiting nucleation step is the lag phase followed by rapid polymerization and fibril formation. A secondary nucleation process may also occur whereby fibril surfaces stimulate new oligomer formation, resulting in the coexistence of both fibrils and oligomers in the same solution (Cohen *et al.*, 2013). One can argue why this nucleation dependent polymerization takes place despite the extraordinary resistance of native proteins to aggregate. Evidence in the favour of aforementioned mechanism has been presented by many experimental and computational studies showing that a peptide may convert from initial conformation into disordered intermediate oligomeric states (seeds) and then their reorganization into ordered fibrils following a nucleation dependent mechanism (Auer *et al.*, 2008). To identify the forces that drive the proteins to fold or misfold is very crucial. The balance between hydrophobic forces and hydrogen bonding is different in the normal folding of peptides as compared to their amyloid aggregates. Free-energy calculations show that the minimum for the folding process is dominated by hydrophobic interaction while that for misfolding is dominated by hydrogen bonding interactions (Fitzpatrick *et al.*, 2011). During aggregation the transition from  $\alpha$ -helix to  $\beta$ -sheet depends upon hydrogen bonding that is not sequence-specific. This transition occurs for all proteins irrespective of their amino acid sequence making it a crucial step of A $\beta$  assembly (Ding *et al.*, 2003).

The partially folded or misfolded proteins are intrinsically prone to aggregation, and if they get the opportunity of favorable intermolecular contacts, their association leads to the formation of stable amyloids and ultimately causes amyloid diseases (Estácio *et al.*, 2014). A $\beta$  may not necessarily give rise to amyloid fibrils following a direct route, rather it may give rise to various oligomeric products. Many simulation studies of amyloid aggregation show that A $\beta$  dimers show a transition from parallel to antiparallel alignment during the simulation (Klimov and Thirumalai, 2003; Santini *et al.*, 2004). Zheng *et al.*'s simulations confirm this transition and proposed that parallel dimers are

metastable, and nucleation may proceed with antiparallel conformation. So it may be inferred that the antiparallel conformations are formed first, then new peptides are added to these oligomers in a parallel fashion (Zhang *et al.*, 2007). The oligomeric state, which is considered to play the main role in amyloid aggregation, is a disordered micelle-like state. Recent studies on animal models with diseased symptoms show that the amount of oligomer extracted correlates better with the severity of amyloid disease as compared to the fibrils (Sengupta, Nilson and Kaye, 2016).

FTIR studies show that the oligomers at low salt and low temperature 25°C conditions show no secondary structure characteristics. However, an increase in temperature to 37°C showed the appearance of  $\beta$ -sheets which indicates the conversion of oligomers to protofibrils (Ahmed *et al.*, 2010). MD also supports this conformational conversion at a critical temperature below which the peptides remain in  $\alpha$ -helical structures and above which they misfold into  $\beta$ -sheet rich structures with greater chances to aggregate (Ding, LaRocque and Dokholyan, 2005).

## 1.5 Structures of A $\beta$ Aggregates

Various organizational levels of A $\beta$  exist during its aggregation starting from the small, soluble monomeric form to oligomers and large insoluble protofibrils/fibrils that are visible under a light microscope (Figure 1.3). The formation of A $\beta$  fibrils is not always a straightforward process that proceeds from monomers to fibrils. In homogeneous aqueous solutions there appear to be different pathways that depend upon conditions such as temperature, pH, ionic strength, and agitation. Literature often refers to “on-pathway” or “off-pathway” structures. “On-pathway” structures are those that will eventually form fibrils while “off-pathway” are those that do not form fibrils (Verma, Vats and Taneja, 2015).

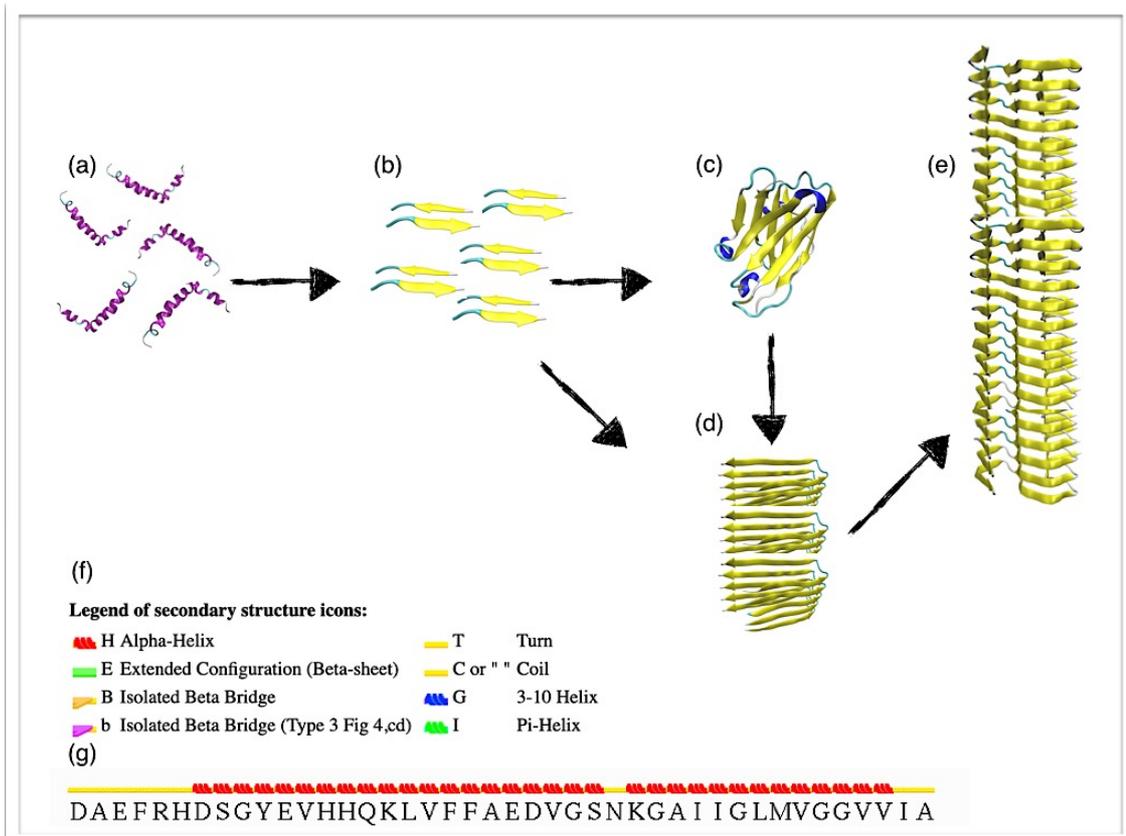
It is hard to get reliable structural information on A $\beta$  because of its conformational flexibility and its tendency to aggregate in aqueous solution. The primary amino acid sequence of A $\beta_{42}$  is shown in Figure 1.3g. The 3D NMR structure of A $\beta_{42}$  in aqueous SDS micellar solution reveals that it is predominantly  $\alpha$ -helical, comprising of two helical regions at residues 8 to 25 (helix I) and 28 to 38 (helix II) linked by a turn at residues 26-28 (Orlando Crescenzi, Tomaselli, Guerrini, Salvadori, Anna M D’Ursi, *et al.*, 2002). A general view based on many structural studies on A $\beta$  fragments and model peptides is that the A $\beta$  monomer does not possess a definite secondary structure in

aqueous solution, rather it exists in a complex conformational ensemble of collapsed coil structures (Zhang *et al.*, 2000). Molecular dynamics simulations of A $\beta$ <sub>40</sub> and A $\beta$ <sub>42</sub> suggest that the C-terminal of A $\beta$ <sub>42</sub> is more structured as compared to A $\beta$ <sub>40</sub> (which is short by two amino acids at this terminal). This hydrophobic C-terminal of the A $\beta$ <sub>42</sub> is crucial in initiating the  $\alpha$ -helical to  $\beta$ -sheet conversion that is responsible for its greater propensity to aggregate into larger amyloid fibrils (Liu *et al.*, 2010). The soluble A $\beta$  monomers aggregate to give rise to a heterogeneous size distribution ranging from low-molecular weight oligomers (100–200 kDa) to protofibrils and fibrils. Oligomers are short-lived and highly unstable, which is why little is known about their structure and conformation. These oligomers that are rich in  $\beta$ -sheets assemble to form protofibrils (approximately 5 nm in diameter). Protofibrils then linked linearly to form protofilaments, and ultimately fibrils (approximately 10 nm in diameter and may extend up to 1 $\mu$ m in length) (Johansson, 2005).

All the known A $\beta$  species in amyloid cascade are neurotoxic to different extents. An increasing number of recent studies suggest oligomers to be more toxic species than mature fibrils as they cause rapid neuronal death in vitro and lead to increased membrane conductance and rapid calcium influx that triggers apoptotic pathways. Besides, they are highly unstable, smaller in size, and have exposed hydrophobic surfaces that allow greater interaction with membranes and other cellular targets. The difference in the level of toxicity of oligomers and fibrils may reside in the variability in their structure and conformation (Verma, Vats and Taneja, 2015).

#### Anti-AD therapies

Since 1998, only five drugs have made it through rigorous clinical trials and been approved by the FDA as AD drugs (*What's New in the Alzheimer's Treatment Pipeline? | BrightFocus Foundation-2020*). These drugs however; can only reduce the disease symptoms but are unable to halt or reverse the process causing AD. Out of these five, three drugs donepezil, galantamine, and rivastigmine are cholinesterase inhibitors that control the breakdown of a chemical messenger responsible for activity and learning (Birks, 2006; Haake *et al.*, 2020). The fourth drug memantine regulates the activity of another neurotransmitter called glutamate that is considered to be important for memory, attention, reason, language, and learning processes (McShane *et al.*, 2019). The fifth drug is a combination of (donepezil) with memantine (Nikl *et al.*, 2019).



**Figure 1.3. The pathway of amyloid assembly starting from A $\beta$  monomer to low-weight oligomers, protofibril, and large fibrils. (a) Solution structure of A $\beta_{42}$  in the apolar microenvironment (PDB: 1iyt); (b) The structure of KLVFFA fragments of A $\beta_{42}$  presenting the hydrophobic patch of A $\beta_{42}$  responsible for higher-order assemblies (PDB: 3ow9); (c) Lower molecular weight oligomer; (d) Structure of protofibril composed of parallel, in-register  $\beta$ -sheets; (e) A $\beta_{42}$  amyloid fibril displaying triple parallel- $\beta$ -sheet segments; (f) Key showing color codes for secondary structure content in (g); (g) The primary amino acid sequence A $\beta$  isoform (A $\beta_{42}$ ).**

As of 2019, there were 156 trials of anti-AD therapies out of which 73% are disease-modifying drugs that are aimed at modifying the course of the disease rather than managing the symptoms (Cummings *et al.*, 2019). About 14% of the drugs are being tested for their ability to improve cognitive impairment. About 11% are aimed at improving behavioral issues related to cognitive decline such as agitation and sleep disturbances.

Disease-modifying small molecules such as Insulin-related molecules, Omega-3 fatty acid (Docosa-hexaenoic acid DHA) (Lim *et al.*, 2005), and curcumin derivatives are designed to interfere with multiple disease pathways. Epidemiological studies show that Curcumin is a good candidate for research in regards to AD because of its multiple neuroprotective properties including A $\beta$  inhibition, clearance of existing A $\beta$  plaques, reducing oxidative stress, and anti-inflammatory properties (Walker and Lue, 2007). *In*

*vitro* and *in vivo* studies points towards the beneficial role of resveratrol in reducing A $\beta$  toxicity by decreasing A $\beta$  production, enhancing the clearance of A $\beta$  peptides, and reduction A $\beta$  aggregation (Rivière *et al.*, 2007; Huang *et al.*, 2011; Granzotto and Zatta, 2014; Zhao *et al.*, 2015; Santos *et al.*, 2016). Biochemical assays show that resveratrol seems to exert its neuroprotective role through inhibiting A $\beta$  aggregation, scavenging ROS and exhibiting anti-inflammatory activities (Liu and Bitan, 2012). Similarly other herbal compounds like Chinese, Asian, or Korean ginseng which have steroid-like active compounds i.e. ginsenosides are thought to be effective for cognition and neuroprotective properties (Chen, Eckman and Eckman, 2006). Melatonin is a powerful hormone that regulates sleep, provides mitochondrial support, reduces A $\beta$  toxicity, and hyperphosphorylation (Pandi-Perumal *et al.*, 2008).

## **1.6 Therapeutic issues**

Although amyloid pathology is now decades old, (potential) drug after drug aimed at slowing down Alzheimer's disease have failed in clinical trials. There may be a few reasons for this:

Targeting either synthesis of A $\beta$  peptides, the oligomers, or only the final insoluble amyloid aggregate may fail to significantly reduce soluble load. Thus, treatments may fail unless the trial target simultaneously all the culprits in the equation, which is the "triple therapy" thinking (Mucke and Selkoe, 2012).

In a few therapeutic approaches, the material in equilibrium with the insoluble A $\beta$  in plaques is removed, hoping that it would trigger the release of soluble oligomers from the plaques a response analogous to shifting a chemical reaction to the left. If plaque removal results in the release and deposition of soluble A $\beta$  and if some soluble species are neurotoxic then this would explain why plaque-removing trials have failed (Rosenblum, 2014). The structure and dynamic characterization of toxic oligomers have been going on at a very slow pace because of their transient nature and intrinsic disorders. The nature of amyloid species is highly dynamic so to preserve them to maintain native state characteristics for the in-vitro study is a real challenge. Although computational methods have an advantage of atomistic detail, yet they lack time and length scales available in experiments.

## 1.7 Role of molecular dynamics simulations in biomolecular research

With the advent of high-performance computing, in-silico modeling has become an important tool to bridge the gap between experiment and theory. It's been called "computational microscopy" by Schulten (*HistoricaPollack, L. Klaus Schulten Reflects on his 2015 National Lecture Series*) who strongly believes that computational techniques have reached to the point where they can present a realistic molecular-level picture of cellular components that is not attainable through current NMR spectroscopy, x-ray crystallography or atomic force microscopy which are challenged by soft matter. The first MD protein simulation was of BPTI (bovine pancreatic trypsin inhibitor) in the vacuum in 1977, which were followed by many simulations based on smaller systems and shorter timescales (McCammon, Gelin and Karplus, 1977). To overcome the sampling problems, system size and timescale issues, many enhanced sampling techniques have been developed and successfully employed that include steered molecular dynamics (SMD) (Do, Lee and Le, 2018), umbrella sampling (US) (Kästner, 2011) and metadynamics (Laio and Gervasio, 2008). Besides, efficient molecular simulation packages such as Gromacs (Abraham *et al.*, 2015), Desmond (Bowers *et al.*, 2006), NAMD (Phillips *et al.*, 2005) and Amber (Case *et al.*, 2005) have led to better and faster outputs. A wide range of biological processes occur over a timescale of several orders of milliseconds that were way beyond the duration of the longest MD simulations. To overcome this problem, massively parallel machines called Anton is developed that is capable of executing millisecond-scale classical MD simulations of such biomolecular systems (Shaw *et al.*, 2008).

In this era of modern research, molecular modeling is not only making a great impact on the field of biomolecular systems but is driving many scientific endeavors. This spans a wide variety of applications ranging from understanding the structure and dynamics of small fragments of biomolecules like heptapeptide fragment GNNQQNY (Szała-Mendyk and Molski, 2020) to the modeling of large macromolecular systems such as the entire viral capsids of viruses (Fujimoto *et al.*, 2020; Jana and May, 2020; Marzinek, Huber and Bond, 2020), disease treatment such as prediction of disease associated single nucleotide polymorphisms (SNPs) (Pandey *et al.*, 2020), to biotechnological applications, e-g designing of nanocomposites and nanostructured polymer materials for energy storage (Bačová and Rissanou, 2018; Tafrihi *et al.*,

2020). A wide variety of applications include studies of structure-activity relationships of biomolecules (Rahman *et al.*, 2020), drug-design (Selivanovitch and Douglas, 2019; Kumar, Singh and Patel, 2020; Rahman *et al.*, 2020; Suárez and Díaz, 2020), binding free energy calculations for ligands/drugs to their targets (Wang, Zhang and Zheng, 2020), modeling of membrane pores (Di Scala *et al.*, 2016; Kholina *et al.*, 2020), ion-channels and transport across the membrane (Maroli and Kolandaivel, 2020) and deciphering complex kinetics of protein misfolding (Zhou and Bates, 2013; Liu *et al.*, 2020). This latter area of protein misfolding and self-assembly is of particular interest as it is related with a large number of diseases including Creutzfeldt–Jakob disease, Parkinson's disease, multiple system atrophy and Alzheimer's, disease (Walker and LeVine, 2000; Chiti and Dobson, 2009; Westermarck *et al.*, 2018).

The study of amyloid assembly is challenging as these assemblies are inherently heterogenous and complex. Molecular dynamics has played an important role in complementing the experimental techniques like XRD, NMR and cryo-EM wherever possible. All of these techniques though quite powerful in the development of structural biology still have their limitations; including the use of non-physiological conditions and lack of sufficient resolution to monitor the structural conformations and dynamics. By combining molecular dynamics with these static techniques, it has now become possible to understand the structures and dynamics biological molecules at an atomic scale. Owing to major developments in the efficiency and ease of use of MD simulations we are now able to understand to some extent the mechanism of amyloid self-assembly, nature of amyloid aggregates including small oligomers and large fibrillar assemblies, their kinetic pathways and the mechanisms of how amyloid aggregation can be inhibited. For example; a recent study shows that flavonoids directly bind to A $\beta$  both in monomeric and pentameric forms and attenuate their hydrogen bonding, salt bridges and interpeptide bonds and convert on-pathway fibrils to off-pathway fibrils without interfering with the secondary structure motifs (Gargari and Barzegar, 2020). A discrete molecular dynamics (DMD) study uncovers the molecular mechanism of how casein coated-gold nanoparticles ( $\beta$ Cas + AuNP) inhibit A $\beta$  aggregation. The simulations show that  $\beta$ Cas + AuNPs bind to A $\beta$  fibrils and stop the aggregation process by either sequestering the A $\beta$  monomers or coating the surface of the A $\beta$  fibrils and stop the secondary nucleation process (Javed *et al.*, 2019).

## 1.8 Aims and objectives of the thesis

The present thesis investigates the aggregation of A $\beta$  peptides employing molecular dynamics (MD) simulations. The principal goal of this study is to understand the aggregation behaviour of A $\beta$  peptides including full-length A $\beta$ (1-42) and shorter fragments A $\beta$ (17-22) as model systems, and how this can be modulated by endogenous and exogenous substances, using two specific examples curcumin and heparin. Aggregation inhibition is an important strategy to develop therapeutic drugs for Alzheimer's. Endogenous co-factors that promote aggregation may serve as additional intervention points.

There are four specific research objectives each of which forms a separate result chapter:

- (a). A systematic comparison of force fields used for molecular dynamics simulations of hierarchical structures of A $\beta$  peptides. (Chapter 3).
- (b). To understand the aggregation of A $\beta$  monomers into various morphologies as a function of concentration using coarse-grain molecular dynamics (Chapter 4)
- (c). To investigate the molecular basis of curcumin's specificity for A $\beta$  fibrils that makes it a promising candidate for non-invasive amyloid tracer. (Chapter 5)
- (d). To investigate the role of heparin in enhancing the process of amyloid aggregation and the influence of heparin's flexibility, molar ratio, and degree of polymerization on the formation of A $\beta$  fibrils. (Chapter 6).

# 2 Methodology

## 2.1 Molecular simulations theory

### Statistical mechanics

Molecular dynamics (MD) is a powerful tool in modern computational science that calculates the time-dependent behavior of a molecular system. For bio-molecular systems such as proteins and nucleic acids, MD simulations provide detailed information on the fluctuations and conformational changes that are not readily available otherwise.

Molecular dynamics simulations generate information at the microscopic level, including atomic positions and velocities. The conversion of this microscopic information to macroscopic observables such as thermodynamic and structural properties requires statistical mechanics. Statistical mechanics bridges the gap between microscopic motion of individual atoms and macroscopic observables such as pressure, energy, heat capacities and temperature, chemical potential, viscosity, spectra, reaction rates, etc. It provides rigorous mathematical expressions that relate macroscopic observable to atoms and molecules of the  $N$  body system. In classic mechanics, a single particle in one dimension is described by two coordinates, a position coordinate  $q$  and the momentum coordinate  $p$ . So the microstate of the  $N$ -body system in 3-D, is described by  $6N$  coordinates. As the system evolves in time the coordinates of the particles change and the system follows a trajectory in this  $6N$  dimensional phase space giving rise to fluctuations in macroscopic properties. Most often, precise microscopic details are unimportant and a concept of “ensemble” is employed which is a collection of all possible systems with different microscopic states but identical macroscopic states (BRYAN, 1902).

There exist three different equilibrium ensembles with different characteristics defined for any isolated system.

Microcanonical ensemble (NVE): The microcanonical (NVE) ensemble is the simplest of all the thermodynamic ensembles. This thermodynamic state corresponds to an

isolated system that contains a fixed number of atoms ( $N$ ), a fixed volume ( $V$ ), and precise energy ( $E$ ). The first postulate of statistical mechanics *principle of equal a priori probabilities* states that all microstates with an equal energy are equally probable (BRYAN, 1902).

Canonical Ensemble (NVT): This describes a system where the thermodynamic state is described by a fixed composition ( $N$ ), a fixed volume ( $V$ ), and a fixed temperature ( $T$ ). The system is in thermal equilibrium with a heat bath of fixed temperature.

Isobaric-Isothermal Ensemble (NPT): In this ensemble thermodynamics state has a fixed number of atoms ( $N$ ), a fixed pressure ( $P$ ), and a fixed temperature ( $T$ ).

### Molecular simulations

Molecular simulation is a powerful tool that simulates the microscopic properties of molecules in terms of structure and interaction between them that is not accessible through experiments. There are two main molecular simulation techniques are molecular dynamics (MD) and Monte Carlo (MC). In this thesis we have used MD, so we will concentrate on MD. One of the main advantages of MD is that it gives direct access to the dynamics of a system, as well as a route to measure thermodynamic and structural properties with the help of statistical mechanics.

In the classic MD, all atoms are represented as hard spheres linked by harmonic springs that represent bonds. Hard spheres present the position of the atomic nuclei. The charges are centered on these nuclei for all atoms that bear charge. Molecular dynamics simulation time-step through the trajectory of the interacting atoms. During the first step intermolecular force acting on each atom due to its neighboring atoms is calculated. The new positions and velocities are then predicted using Newton's equations of motion and the atoms are moved to their new positions. At the new positions, the intermolecular forces are re-calculated, and then another step-in time is made. This iteration continues for millions of steps in a typical simulation. The intermolecular forces may be calculated classically using molecular mechanics (force field) approximation. The integration of Newton's Laws gives rise to trajectories, linking atomic motions through time. Application of force to each atom causes an acceleration, dependent upon the atomic mass. This relationship was expressed in Newton's second law is given by:

$$F_i = m_i a_i$$

where  $F_i$  is the force exerted on particle  $i$ ,  $m_i$  is the mass of particle  $i$  and  $a_i$  is the acceleration of particle  $i$ .

Other fundamental relationships to relate force, acceleration, mass and position of the particle are as,

$$\frac{d^2 \mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} \quad (1)$$

$$\frac{d \mathbf{r}_i}{dt} = \mathbf{v}_i \quad (2)$$

$$\frac{d \mathbf{v}_i}{dt} = \frac{\mathbf{F}_i}{m_i} \quad (3)$$

The current positions ( $\mathbf{r}$ ), velocities ( $\mathbf{v}$ ) and accelerations ( $\mathbf{a}$ ) can be used to calculate the change in the position of the particle in the system. The calculation of interaction forces, in particular, the nonbonded interactions i.e. Lennard Jones and charge-charge interaction, is the most expensive part in MD simulation. The number of these pair interactions is given by  $N(N-1)/2$  where  $N$  is the number of atoms, hence increase by  $N^2$  as the size of the system increases. To compute these forces usually a cut-off is applied at some specific interatomic separation, beyond which the forces are assumed to be zero.

### Integration Algorithms

The approximate numerical method used to advance the system by one-time step is known as an integration algorithm. A good integrator should be computationally cheap, accurate, and allow a longer time step. The popular integrators are the Verlet algorithm, the Verlet leapfrog algorithm, and the velocity Verlet algorithm.

### Verlet algorithm

The equations of the Verlet algorithm are as follows.

$$r_{n+1} = 2r_n - r_{n-1} + \frac{f_n}{m} \Delta t^2 + O(\Delta t^4) \quad (4)$$

$$\mathbf{v}_n = \frac{1}{2\Delta t} [\mathbf{r}_{n+1} - \mathbf{r}_{n-1}] + O(\Delta t^2) \quad (5)$$

Suppose the system is in the middle of the simulation at time step  $n$ .  $r_n$  is the starting position for this time step,  $r_{n-1}$  is the corresponding position at a previous time step. Using Equation 5, the position of the atoms at time step  $n+1$  can be calculated. When this is done, we are ready to begin the next time step. The algorithm is straight forward, needs only modest storage, and is time-reversible.

The integrator is simple but has a few drawbacks. The positions are given for the  $n+1^{th}$  time step and velocities for the  $n^{th}$  time step. So, the velocities are one step behind the positions. In total we have to store atom positions for three consecutive time steps, which makes it demanding in terms of memory.

### Leap-frog algorithm

This is an improved and computationally cheaper version of the Verlet algorithm, as it needs to store only one set of positions and velocities. The equations defining this algorithm are as follows:

$$v_{n+\frac{1}{2}} = v_{(n)} + \frac{f_n}{m} \Delta t + O(\Delta t^3) \quad (7)$$

$$r_{n+1} = r_n + v_{n+\frac{1}{2}} \Delta t + O(\Delta t^4) \quad (8)$$

Once the forces  $f_n$  of the current time step are calculated, the algorithm calculates velocities that are half a time step ahead of the current time step  $n$ . These velocities are then used in calculating the positions of the atoms at one time step ahead using the equation. So it seems as if the velocities leap over positions, which is the origin of the algorithm's name (Van Gunsteren and Berendsen, 1988).

### The Velocity Verlet algorithm

This integrator provides both the atomic positions and velocities at the same time step, and for this reason may be regarded as the most complete form of the Verlet algorithm (Spreiter and Walter, 1999). The basic equations are as follows:

$$r_{n+1} = r_n + v_n \Delta t + \frac{f_n \Delta t^2}{2m} + O(\Delta t^3) \quad (9)$$

$$\mathbf{v}_{n+1} = \mathbf{v}_n + \frac{\Delta t}{2m} [\mathbf{f}_{n+1} + \mathbf{f}_n] + \mathcal{O}(\Delta t^3) \quad (10)$$

### Force Fields

Force fields can be classified into two families; all-atom and coarse grain force fields. All-atom force fields represent molecules in terms of each of the individual atoms while coarse grain force fields replaces atomistic detail with lower resolution, for example, the MARTINI force field represents molecules as beads mapping four heavy atoms to one interaction site (Marrink *et al.*, 2007). Though coarse grain force fields reduce the level of accuracy and information about the properties of the system, yet these are a robust approach to simulating very large systems for longer simulation time. Atomistic force fields can further be divided into two groups: all-atom and united atom. All-atom force fields represent each atom of the system explicitly for example Amber (García and Sanbonmatsu, 2002), OPLS-aa (Jorgensen and Tirado-Rives, 1988) and Charmm (Huang and Mackerell, 2013). In contrast united-atom force fields include non-polar hydrogens into their connected carbon atoms treating the CH<sub>n</sub> group as a single particle with increased mass. Polar and aromatic hydrogen are still represented explicitly. The advantage of a united-atom force field is that it allows the use of longer time-step 2fs which makes the calculations faster as compared to 1fs in case of all-atom force fields. are derived

For each force field method, the total energy consists of various energy contributions:

$$\mathbf{E}_{tot} = \mathbf{E}_{str} + \mathbf{E}_{bend} + \mathbf{E}_{tor} + \mathbf{E}_{vdw} + \mathbf{E}_{elec} + \dots \quad (11)$$

where  $E_{tot}$  is the total energy which is a combination of bonded and non-bonded energy terms.  $E_{str}$  is the energy change as a bond stretches from the ideal unstrained length and is given by

$$E_{str} = \frac{1}{2} k_b (\mathbf{b} - \mathbf{b}_0) \quad (12)$$

where  $k_b$  is the bond-stretching force constant,  $b$  and  $b_0$  are the actual and equilibrium bond lengths.

The bond-angle term is given by

$$E_{bend} = \frac{1}{2} k_q (\theta - \theta_0) \quad (13)$$

where  $\theta_0$  and  $\theta$  are the equilibrium and actual angle bending terms and  $k_q$  is the angle-bending force constant.

The dihedral potential energy term takes the form

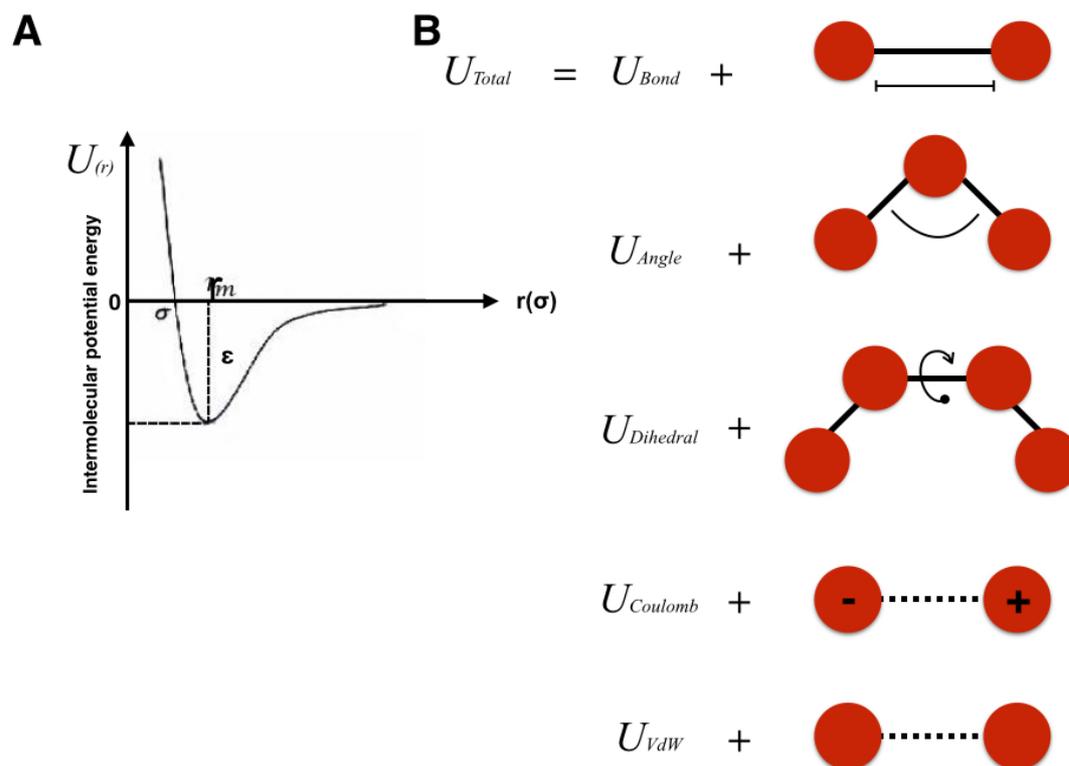
$$E_{tors} = \frac{1}{2} k_\Phi [1 + \cos(n\Phi - \gamma)] \quad (14)$$

where  $k_\Phi$  is the barrier to rotation,  $n$  is the multiplicity,  $\Phi$  is torsional angle, and  $\gamma$  is the phase factor. The multiplicity corresponds to the number of energy minima as the bond rotates through  $360^\circ$ . The phase factor  $\gamma$  is an important factor as it decides where the dihedral potential passes through its minimum value.

The van der Waals forces are calculated based on the Lennard-Jones potential, which computes the energy of interaction between two atoms as a function of the distance between them (Jones, Jones, and E., 1924).

$$V_{(r)} = 4\epsilon \left[ \left\{ \frac{\sigma}{r} \right\}^{12} - \left\{ \frac{\sigma}{r} \right\}^6 \right] \quad (15)$$

where  $r$  is the distance between two interacting atoms,  $\epsilon$  is the energy well depth and  $\sigma$  is the diameter of the atoms at zero potential.



**Figure 2.1. (A). Lennard Jones potential; (B). Potential energy functions as a sum of its components.**

The electrostatic interactions are calculated by :

$$E_{elec} = \frac{1}{4\pi\epsilon} \frac{q_1 q_2}{r_{ij}} \quad (16)$$

where  $\epsilon$  represents the dielectric constant and  $q_1$  and  $q_2$  show the charges of the interacting atoms while  $r_{ij}$  is the interatomic distance. The various parameters such as epsilon and sigma in the LJ equation, partial charges  $q_1$  and  $q_2$ , force constants, and ideal bond lengths and angles are all-atom specific and are termed as force field parameters.

#### System setting in MD simulations

Setting up an MD simulation is not a trivial task as one needs to consider many factors before setting up. This includes the starting configuration of the system under study, the

choice of force field and integration algorithms, time and length scales, boundary conditions, solvation, and then analysis methods.

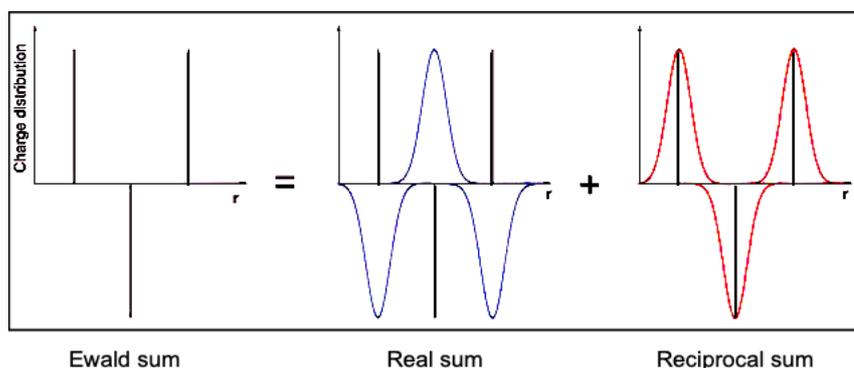
For most of the proteins, the initial structures are generally obtained from protein data bank that stores a huge number of x-ray crystallography and NMR structures. If the parameters are missing for some proteins, theoretical methods such as homology modeling is used. The systems are energy minimized to remove any net force on them. Once the initial choice on the structure of the model, force field, and MD code is made, the timescale and time step are chosen. The length of the time scale depends upon the system under study and the availability of the computational resources. The choice of time step is very important and as a general rule it should be one-tenth of the time that the shortest motion takes place. In proteins simulations this is usually the C-H bond vibrations which occurs at 10fs timescale, so generally the time step is 1fs. With the use of constraints over C-H bonds with methods like SHAKE, the time step can be doubled i.e 2fs (Swope *et al.*, 1982). MD studies are traditionally performed using NVE ensemble s but for bio-simulations, it is more practical to use NPT because it resembles experimental conditions more closely.

### Ewald summation

The laborious part of MD is to calculate long-range interactions such as the charge-charge interaction (Equation 16). It cannot be truncated using a cut-off as employed for the Lennard-Jones interaction. There are a variety of ways to address this problem. An efficient technique is to use Ewald summation that relies on the system having periodicity but is appropriate for molecular simulation whenever periodic boundaries are employed, which is the usual case. Ewald summation divides a potentially infinite sum in real space into two finite sums; one in real space and the other in reciprocal space. Generally, the model consists of a system of simple point charges interacting via the Coulombic interaction. Ewald method makes two amendments in this simple model. First, each ion is effectively neutralized at a longer range by the superimposition of a spherical Gaussian that carries an opposite charge (see Figure 2.2). The point charges and Gaussian charges form the real space part of the Ewald sum, which is now short-range. The second amendment is to superimpose a second set of Gaussian charges, but this time it has the same charges as the original point ions (so it cancels the effect of the

first set of Gaussians). The potential due to these Gaussians is calculated using Poisson's equation and is solved as a Fourier series in reciprocal Space.

This method is computationally demanding because the reciprocal sum increases with  $N^2$ , where  $N$  is the number of particles. That is the reason it is generally applied to smaller systems. For larger systems, Ewald summation is normally implemented in the simulation codes using the particle mesh Ewald (PME) technique. In this method, the charges are mapped onto the grids rather than performing direct summation of wave vectors, making the method very efficient (Ewald, 1921).



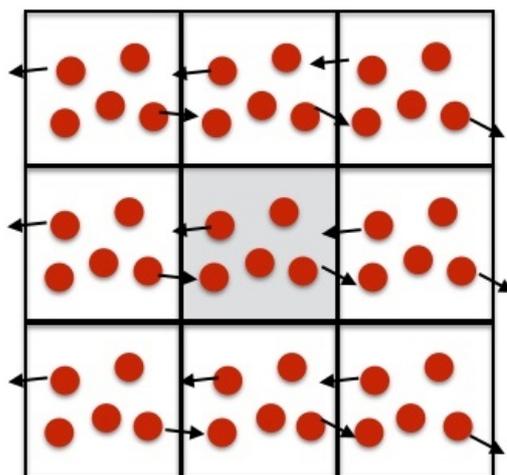
**Figure 2.2. Ewald summation: The sum is divided in the real space and reciprocal space by adding a Gaussian**

#### Periodic boundary conditions

Even with large computational resources we are still unable to simulate systems that are large enough to serve as bulk systems. Therefore, a small section of the system is taken to serve as a model/ representative of the bulk system. The boundaries of the system must be taken into consideration otherwise it would produce large surface effects. One of the solutions to avoid surface effects is to employ periodic boundary conditions (see Figure 2.3). This involves generating a box of any geometrical shape (generally cubic) and placing a small block of in a cell, and then surrounding the cell with a replica of its mirror images of the system. In doing so, now the atoms in the cell are interacting with an array of replicas and the cell is no more an isolated section in vacuum rather it is surrounded again by the bulk phase. So now if a molecule leaves the simulation box from one end, an identical molecule will enter the box from the other side. This technique works in conjunction with an idea of minimum image convention for short-range forces, according to which the molecule will interact only with the closest periodic image of its neighbors. These short-range forces are often truncated to reduce

computational consumption. This cut-off distance must be less than or equal to half the box vectors (Berendsen, van der Spoel and van Drunen, 1995). Now, the interaction energies can be calculated across cell boundaries overcoming the boundary effects.

Once the choice about MD and other parameters is made, the system is ready to set up, and start the simulations. The most important steps of MD simulations are as follows:



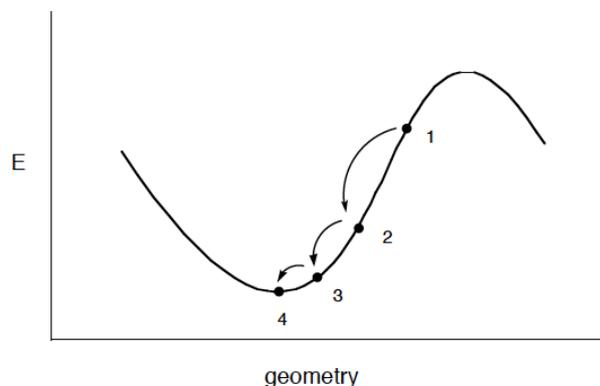
**Figure 2.3. An example of periodic boundary condition; the primary cell (grey) is surrounded by its images (white)**

### Energy minimization

Before starting any dynamics simulation, the system undergoes rigorous energy minimization to remove any net energy or high-energy overlaps between atoms. The process continues until a stable point or a minimum on the potential energy surface is reached. At a minimum, the net force on each atom vanishes. Molecular systems, in general, have a very large number of minima and in practice the system converges to a local minimum rather than the global minima.

Energy minimization involves changing the geometry of the system in a stepwise manner, so that the system can achieve a minima. In Figure 2.4 below, point 1 is the present geometry of the system, so we calculate the energy or slope of the function at this point. If this comes out to be positive, it shows that the gradient or step size is too large. So the minimization adjusts the coordinates and slope is calculated again. If the value comes out to be negative; it shows the coordinate is too small. When the slope becomes zero, a minimum has reached. There are a variety of minimization methods and I outline the three commonly employed algorithms below.

$$X_{new} = X_{old} + \textit{correction} \quad (17)$$



**Figure 2.4. The process of energy minimization takes place in a stepwise manner until a minimum is reached.**

Newton Raphson method

This is the most expensive energy minimization technique based on the Taylor series expansion of the potential energy surface.

$$x_{new} = x_{old} - \frac{E' x_{old}}{E'' x_{old}} \quad (18)$$

Here the correction term depends on the 1<sup>st</sup> derivative (slope or gradient) of the potential energy surface and also on the 2<sup>nd</sup> derivative (curvature). So one needs to calculate both of these at every step, which makes it computationally expensive, which is the biggest disadvantage of this method.

Steepest descent

This makes use of the first derivative of the energy function to locate the minima instead of relying on the 2<sup>nd</sup> derivative, which makes it faster as compared to the Newton Raphson method. It assumes 2<sup>nd</sup> derivative to be a constant hence it is based on approximation; that is why this method is less reliable as compared to the above mentioned. The energy is calculated for the initial position and the system is moved with a very small increment. The cycle is repeated many times and eventually the system reaches downhill on the energy surface. The procedure stops when it reaches

predefined threshold energy. The advantage of this method is that it is very robust and converges even when far away from the minimum.

$$\mathbf{x}_{new} = \mathbf{x}_{old} - \gamma \mathbf{E}'_{\mathbf{x}(old)} \quad (19)$$

### Conjugate gradient

The conjugate gradient algorithm accumulates information from one cycle to the next to avoid the reverse of the progress made in the previous cycle. For each step the gradient is calculated and stored. The computational requirement and storage are a lot more than steepest descent, but it is better for larger systems and gives efficient convergence. The advantage of this method is that it keeps the previously calculated gradients stored in the memory thus avoiding backtracking. It generally finds a minimum in fewer steps than steepest descent. Sometimes, it may face challenges when the initial conformation is far from a minimum. For such systems, one should first use the steepest descent algorithm and follow up with the conjugate gradient for convergence.

### Equilibration

After minimization the system is at 0K temperature. To run MD, the system needs to reach the temperature of interest. The velocities are assigned to the atoms by a random velocity generator using the constraint of the Maxwell-Boltzmann distribution. The temperature of the system is related to the average kinetic energy of the system. The most common means of temperature scaling is velocity scaling.

$$K.E = \frac{3}{2} nk_B T \quad (20)$$

where  $K.E$  is the average kinetic energy of the system under study,  $k_B$  is the Boltzmann constant. The temperature of the system is calculated using this expression. If it is less than the desired value, it is simply multiplied by the square root of  $(T_o/T)$ , where  $T_o$  is the required temperature. This step is repeated until we reach the desired temperature (Walton and Vanvliet, 2006).

The temperature of the system is controlled during a constant temperature MD as it may rise as a result of frictional forces, integration errors, or drift during equilibration. Gromacs uses weak coupling schemes such as Berendsen (Berendsen *et al.*, 1984), Andersen thermostat (Andersen, 1980), the extended ensemble Nosé-Hoover scheme (Hoover, 1985; Nosé, 2002), or a velocity-rescaling scheme (Bussi, Donadio and Parrinello, 2007) to simulate constant temperature MD.

In the Berendsen scheme, the system is weakly coupled to an external heat bath at some given temperature  $T_0$ . The thermostat will overcome any fluctuations in the kinetic energy that will result in incorrect canonical ensemble especially for smaller systems, but for very large systems the approximation yields roughly correct results for most calculated properties. The deviation of the system temperature from  $T_0$  is slowly corrected according to:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \quad (21)$$

This equation shows that temperature deviation decays exponentially with a time constant  $\tau$ . This method of temperature coupling has a gain of relaxing the system to the desired temperature efficiently but because it gives rise to incorrect sampling, it is often used in combination with the Nose-Hoover thermostat which correctly generates trajectories consistent with canonical ensembles.

For constant pressure simulations, the system is coupled to a barostat. The two most common and efficient barostat are Berendsen (Berendsen *et al.*, 1984) and Parinello Rahman barostat (M Parrinello and Rahman, 1981). The Berendsen algorithm scales coordinates of the system and box dimensions every step. These methods Berendsen can be combined with any of the temperature coupling methods above.

The Berendsen algorithm has the effect of a first-order kinetic relaxation of the pressure towards a given reference pressure  $P_0$  according to:

$$\frac{dP}{dt} = \frac{P_0 - P}{\tau_p} \quad (22)$$

For the systems where fluctuations in pressure or volume are important; for example where thermodynamic properties are being calculated, this barostat does not simulate the true NPT ensemble. GROMACS also supports constant-pressure simulations using

the Parrinello-Rahman approach which is similar to the Nosé-Hoover temperature coupling, and in theory gives the true NPT ensemble.

### Production run

Once the system properties become equilibrated at the desired temperature and pressure, one then carries out a production run in which the trajectory of the molecules is stored for future analysis. Millions of time steps are required to derive an accurate description of the phenomenon under observation.

## 2.2 MARTINI model

Atomistic MD simulation provides structural and dynamic information at a sub-nanometer length scale and femtosecond timescale. These time and length scales are computationally very expensive and cannot be extended beyond microseconds. Several approaches have sampling and size limitations. One such approach is coarse-graining which reduces the computational cost by decreasing the number of degrees of freedom for the molecules thereby reducing the complexity of the system. While using a CG model, it is essential not to over-simplify the description of the system as it will lose the important features that are required to drive a certain phenomenon. We used the popular coarse grain model “MARTINI” for studying the aggregation of A $\beta$  at various concentrations described in chapter 4. It was originally designed by Marrink and co-workers for membranes composed of lipids but later extended to proteins and other small biomolecules (Periole and Marrink, 2013). MARTINI model is based on 1:4 mapping which means that on an average four heavy atom and associated hydrogen atoms are mapped on to one interaction site which is represented as a single bead. There are four main types of interacting beads i-e polar (P), nonpolar (N), apolar (C), and charged (Q). These are further divided into subcategories based on hydrogen-bonding capabilities (donor, acceptor, both or none) and polarity ranging from 1 (low) to 5 (high). The interaction between coarse grain beads describing bond lengths, angles, and dihedrals are represented by the parameters that are typical of a classic force field. The non-bonded interactions are defined by the Lennard-Jones 12–6 potential. This force field was originally designed to be used with GROMACS, but because of easy use and implementation, the general form of the potential energy function is now efficiently

used in other MD codes such as DESMOND (Bowers *et al.*, 2006), GROMOS (Baron *et al.*, 2007) and NAMD (Shi, Izvekov and Voth, 2006).

## 2.3 Analysis methods

After the production run, the trajectories were analyzed using various analysis parameters for quality assurance using convergence parameters; root mean square deviation (RMSD) and root mean square fluctuations( RMSF) and structural analysis for example hydrogen bonds, salt bridges, inter-and intrasheet contacts and distances, clustering analysis based on RMSD and secondary structure analysis.

### RMSD

RMSD is a strong indicator of structure stability and also of any large conformational change. We used Gromacs utility gmx rms, which calculates the deviation of atoms in a structure as compared to a reference structure (either crystal or energy minimized) using the equation below:

$$RMSD(t_1, t_2) = [1/M \sum_{i=1}^N m_i ||r_i(t_1) - r_i(t_2)||^2]^{\frac{1}{2}} \dots (23)$$

where the term on the right-hand side of equation 23 describes the position of the atom  $i$  of mass  $m_i$  at certain time  $t_1$ . The second term represents the position of atom  $i$  of mass  $m_i$  at time  $t_2$ .

We used RMSD by implementing g\_rms in Gromacs to calculate the standard deviation of pre-formed structures from the reference crystal structure.

### RMSF

To get insights into the stability of the structure, the role of residue flexibility is very important. RMSF which is the square root of the variance of the fluctuation around an average position for C $\alpha$  of each participating residue is calculated as:

$$RMSF = \sqrt{1/T \sum_{t=1}^T (r_{it} - \langle r_i \rangle)^2} \dots (24)$$

where  $T$  is the total number of snapshots from the trajectory,  $r_{it}$  is the position of C $\alpha$  atom of residue  $i$  at time  $t$  and  $\langle r_i \rangle$  is the average position of C $\alpha$  atom of residue  $i$ .

### Secondary structure analysis

The secondary structure analysis program (DSSP) is a powerful tool that shows the conformational changes occurring during a simulation. This program relies on the identification of hydrogen bond patterns and broadly classifies the secondary structure motifs into three classes helix ( $\alpha$ -helix, 3-10-helix and  $\pi$ -helix),  $\beta$ -strand (isolated  $\beta$ -bridge, extended  $\beta$ -sheet) and loop (turn, bend). We used DSSP implemented in Gromacs and the web-interface of STRIDE (Dmitrij Frishman and Argos, 1995) that is also a secondary structure analysis program but also includes dihedral angle potentials along-with hydrogen bonds criteria.

### Inter-strand and inter-sheet distances

Especially for our force field study (Chapter 3), to compare the structural stability of various amyloid fibrils, we calculated inter and intrastrand and intersheet distances. The interstrand distance is calculated by averaging the distance between  $C\alpha$  atom of a residue in one strand and  $C\alpha$  atom of its corresponding residue in the adjacent strand in the same  $\beta$ -sheet which is approximately equal to  $\approx 0.48\text{nm}$ , whereas inter-sheet distance is calculated by averaging the distance between  $C\alpha$  atom of a residue in one  $\beta$ -sheet and its corresponding  $C\alpha$  atom of the other residue in the adjacent sheet ( $\approx 1.07\text{nm}$ ) (Nguyen and Hall, 2004).

### Hydrogen bond analysis

gmx hbond was employed to calculate intrapeptide hydrogen bonds. g\_hbond assigns a hydrogen bond between two atoms if the distance between a donor atom and an acceptor atom is  $\leq 0.35\text{nm}$  and the angle donor-hydrogen-acceptor  $\geq 120^\circ$  (Krishna Deepak and Sankararamkrishnan, 2016).

### Clustering analysis based on rmsd

To explore the conformation heterogeneity of  $A\beta$  monomer we used the gromos method implemented in the gmx cluster analysis program. This method counts the number of neighbors, takes the structure with the largest number of neighbors, assign it to one cluster, and eliminates it from the pool using a default cut-off (0.1nm). RMSD is used as a criteria to define the distance between structures (Daura *et al.*, 1999).

### Twist angle

The  $\beta$ -sheets generally twist in amyloid fibrils to optimize hydrogen bonds, electrostatic interactions, and side-chain interactions. Twist angles were calculated from the average dihedral angles connecting the first C $\alpha$  atom to the last C $\alpha$  atom in the same  $\beta$ -strand between adjacent  $\beta$ -sheets along the fibril axis following Jie Zheng's work. (Zheng *et al.*, 2007).

### MM(PBSA) method

The molecular mechanic Poisson-Boltzmann or the generalized born solvent accessible surface area (MM-PBSA) method is a powerful analytical technique to provide a quantitative measurement of the binding of molecules e.g. a ligand to a receptor (Kumari, Kumar and Lynn, 2014) (Baker *et al.*, 2001). We used this method as implemented in Gromacs 5.1 to calculate the binding free energy of amyloid and Curcumin. It is a single trajectory approach, which assumes no change in the conformation of the complex. The snapshots were saved from the last 30ns trajectory of MD performed, which has a stable backbone RMSD value. The binding energy is calculated using the following equation

$$\Delta G_{bind} = \langle G_C \rangle - \langle G_A \rangle - \langle G_B \rangle \dots (25)$$

where C, A and B denote the complex, molecule A and molecule B respectively.

$$\Delta G_{bind} = \langle \Delta E_{MM} \rangle + \langle \Delta G_{solv} \rangle - \langle T\Delta S \rangle \dots (26)$$

The free energy of each system is calculated as:

$$\langle G_X \rangle = \langle \Delta E_{MM} \rangle + \langle \Delta G_{solv} \rangle - T \langle S \rangle \dots (27)$$

where  $E_{MM}$  is the standard molecular mechanics' energy terms from bonded, electrostatic, and van der Waals interactions.

$$E_{MM} = E_{int} + E_{ele} + E_{vdw} \dots (28)$$

while  $\Delta G_{solv}$  represents the solvation energy comprising polar solvation energy and non-polar solvation energy.

$$\Delta G_{solv} = \Delta G_{GB} + \Delta G_{SA} \dots (29)$$

Polar solvation energy  $\Delta G_{GB}$  is calculated by using the generalized Born (GB) model in a continuum model of the solvent and non-polar solvation energy term  $\Delta G_{SA}$  linearly depends upon the solvent accessible surface area.

## 2.4 Hamiltonian replica exchanges

Classical MD suffers from insufficient sampling of conformational space. The free energy landscape for protein folding is rugged with hills, traps, and **energy** barriers. If these energy barriers are larger than the available thermal energy, they remain trapped in some local minima. To overcome this problem and increase the sampling of conformational space, several methods have been developed throughout the years. These include simulated annealing (Brünger, Adams and Rice, 1997), metadynamics (Leone *et al.*, 2010), and parallel tempering or replica exchange MD (REMD) (Qi *et al.*, 2018).

REMD has been widely used to enhance conformational sampling. In most REMD methods, temperature is used as a variable. The conformations that are trapped in a low energy minimum escape from it at a higher temperature. For an effective exchange, replicas need a sufficient overlap of potential energy at all the temperatures. This requires a very large number of replicas and in-turn increased computational time and resources.

A solution to temperature-based replica exchange is Hamiltonian exchange (H-REMD) which runs at a constant temperature and uses force field or Hamiltonian as a replica coordinate. H-REMD protocol uses the method of simultaneous switching off electrostatic and Lennard-Jones (LJ) parameters (Affentranger, Tavernelli and Di Iorio, 2006). As protein dynamics are dominated by the interparticle non-bonded interactions; modifying these parameters should in principle increase the sampling of conformational space.

The H-REMD simulations were performed with the stochastic velocity Verlet integrator. Stochastic dynamics itself control the temperature, so using the Nose-Hoover thermostat is no longer required. The Parrinello–Rahman barostat was used for pressure coupling in the NPT ensemble with a time step of 20 fs. The nonbonded peptide-peptide interactions scaled in each replica using the soft-core scaling method when van der Waals parameters were zero. The parameter  $\lambda$  allows smooth scaling of the nonbonded interactions. When  $\lambda = 0$  it represents that all the interactions are switched on and  $\lambda = 1$  where all the interactions including van der Waals and columbic are switched off and the atoms are behaving like ghosts (non-interacting). To avoid sampling of a completely unrealistic system with large overlapping of atoms, a maximum  $\lambda = 0.8$  was used with

**Molecular Basis of Amyloid Disease: In Silico Modelling of Beta-Amyloid ( $A\beta$ ) Aggregation**

$\Delta\lambda = 0.02$ . Replica exchanges were attempted every 1000 steps. The results of HREMD were analyzed by using bar and energy algorithms implemented in Gromacs.

# 3 A systematic comparison of force fields for molecular simulation of hierarchical structures of $\beta$ -amyloid peptides

Beenish Khurshid, Jamshed Anwar\*, David Middleton

Chemical Theory and Computation, Department of Chemistry, Lancaster University, Lancaster LA1 4YB, U.K.

KEYWORDS ( $\beta$ -amyloid, monomer, protofibril, fibril, force fields, MD, AmberGS, Gromos53a6, OPLS-aa).

## 3.1 Abstract

Molecular dynamics simulations are increasingly becoming widespread as an important part of biomolecular research. However, the choice of chosen parameters such as the force fields significantly affects the outcome. Making this choice is important but challenging at the same time as different force fields are calibrated and tested using different parameters, which may affect the outcome of the simulation. In this paper, we investigate the conformational behavior of four A $\beta$  peptide models varying in size and complexity using three popular force fields namely AmberGS, Gromos53a6, and OPLS-aa using brute force MD simulations. Our study rationalizes that for A $\beta$  monomer AmberGS is biased towards helical structures. Gromos53a6 favors the formation of a pair of anti-parallel  $\beta$ -sheets while OPLS-aa produces a mixture of secondary structure contents. We report that for large more complex structures such as dodeca-fibrils and ring fibrils; OPLS-aa is slightly a preferable choice because it resulted in a reasonable agreement with experiments. To our knowledge, this is the first

study to compare the effects of force fields on different A $\beta$  assemblies for longer timescale. Such a systematic comparison will offer useful guidance to others conducting molecular simulations on A $\beta$  aggregation and fibrillar polymorphism.

### **3.2 Introduction**

Statistics show that over 35 million people worldwide and 0.8 million in the UK are living with Alzheimer's disease (AD). This number is expected to double by 2030 ('Alzheimer's disease facts and figures', 2019). Unfortunately, there is no cure for AD yet and a full mechanistic understanding still eludes us. According to the popular amyloid hypothesis, it is the aggregation and deposition of neurotoxic amyloid  $\beta$ -peptide (A $\beta$ ) which gives rise to the symptoms of AD (Hayden and Teplow, 2013). The neurotoxicity in AD is often associated with the formation of  $\beta$ -sheet rich structures (Hardy and Selkoe, 2002) of A $\beta$  which means that the conversion of A $\beta$  from native  $\alpha$ -helix form to  $\beta$ -sheet could be a critical stage for intervention. Thus, it's critical to accurately predict the conformational state and structure of A $\beta$  and its various oligomeric forms. Being soft matter, the various forms and hierarchical structure of A $\beta$  exhibit considerable polymorphism and metastability, making it difficult for experimental methods such as x-ray crystallography, atomic force microscopy, or NMR spectroscopy, to characterize the structures. and provide comprehensive data on the relationship of the structure to the energetics, the effects of specific amino acid alteration, and a continuous movie of the events in real-time. Molecular simulation studies, wherein we utilize intermolecular interactions to simulate molecular dynamics, can overcome some of these limitations. However, the effectiveness of the simulation methodology critically depends upon the accuracy of the intermolecular interactions, which for biomolecules are modeled using the molecular mechanics approximation comprising a potential energy function and a set of atom-specific parameters – the so called force field (Petrov and Zagrovic, 2014).

A wide body of literature investigates the effects of force fields on amyloid peptides, (Nguyen, Li and Derreumaux, 2011)·(Cino, Choy and Karttunen, 2012)·(Watts *et al.*, 2018)·(Fluitt and De Pablo, 2015)·(Zhang *et al.*, 2000) but because each study uses a different set of parameters, simulation conditions and force fields, comparisons and interpretations of these results is quite difficult. Despite of large array of data, the area

describing the effects of force fields on intrinsically disordered peptides is still largely unexplored.

Despite growing interest in using molecular simulation to gain insights into the amyloid assembly, there is a limited and sometimes contradictory data as to which force field is best for the purpose. The four most widely used in biomolecular simulations are AMBER (García and Sanbonmatsu, 2002), (Cornell *et al.*, 1995), GROMOS (Scott *et al.*, 1999), OPLS (Jorgensen and Tirado-Rives, 1988) and CHARMM (Brooks *et al.*, 1983) families. These force fields generally differ in their parameterization, although they are regularly being reparametrized to obtain better performance. The OPLS-aa force field includes functional forms for all 20 amino acids and has the same functional form as AMBER. The parameters for OPLS were optimized to fit experimental properties of liquids such as densities and heat of vaporization. GROMOS united atoms force field was optimized with respect to the condensed phase properties of alkanes. In GROMOS87 the carbon and attached hydrogen atoms are represented as one group and van der Waals interactions were calculated from the crystal structures of hydrocarbons and amino acids using 0.8nm cutoff radii (Meacham and Green, 1996) while in GROMOS96 van der Waals interactions are reparametrized using longer nonbonded cutoff radii (1.4 nm) (Schuler, Daura and Van Gunsteren, 2001). AMBER ff94 is the most widely used force field for protein simulations in the condensed phase which is inspired by the functional form of OPLS and includes fixed partial charges calculated by fitting electrostatic potential using Hartree-Fock 6-31G\* level and no explicit treatment for hydrogen bonding. The dihedral parameters  $\phi/\psi$  are optimized to fit relative quantum-mechanical energies of alternate conformations of glycine and alanine dipeptides (Cornell *et al.*, 1995). The optimized versions of Amber ff94, particularly ff99 and ff96 were observed to over stabilize  $\alpha$ -helical (Okur *et al.*, 2003) and  $\beta$ -strand (Ono *et al.*, 2004; Kamiya, Higo and Nakamura, 2009) secondary structure content respectively. This motivated the experts to re-optimize AMBER ff94/96/99 to adjust the backbone parameters in order to fix the bias towards a particular secondary structure. One of such optimization led to the development of AMBERGS by Garcia and Sanbonmatsu where the dihedral parameters  $\phi/\psi$  are simply zeroed and as a result a greater improvement in the agreement of secondary structure content to the experimental values is observed (García and Sanbonmatsu, 2002).

The basic steps needed to develop new force fields include selection of the function i-e force field parameters i-e bonded and non-bonded interactions and adjusting these interaction potentials to reproduce selected target data such as x-ray diffraction data, force constants from vibrational spectroscopy, heats of sublimation or vaporization and densities. This is not a trivial task and is performed by specialists. End user normally selects a force field and parameters found in the literature. However, if the required parameters are missing, the process is started by borrowing parameters from a similar group and then assigning partial charges using QM calculations. This is followed by adjusting non-bonded interactions to ensure the characteristic molecular geometry. Generally, all force fields for biomolecular simulations have parameters for 20 amino acids but parameters must be available to perform simulations of systems containing other functionalities such as co-factors, prosthetic groups, inhibitors and non-naturally occurring amino acids. Certain force fields contain parameters for a range of groups that are readily transferable to other molecules (Halgren, 1996),(Lii and Allinger, 1991). Though the transferability of parameters can make simulations possible for any kind of system but they should be tested for accuracy and optimized in the same way as the original parameters set.

Previous studies on comparisons of force fields in the literature focussing on secondary structure and conformations of biomolecules, but the results are not unambiguous. For example, according to one study CHARMM, AMBER and OPLS force fields produced remarkably similar structural and dynamic properties such as SASA, Rg, RMSD and secondary structures for three proteins (bovine apo-calbindin D9K, human interleukin-4 R88Q mutant, and domain IIA of Bacillus subtilis glucose permease-) in a 2-ns long simulation(Price and Brooks, 2002). On the other hand, another study found significant differences in the ensembles of peptide conformations using CHARMM and AMBER (Yeh and Hummer, 2002).

A notable trend observed in simulations of A $\beta$  is the different structural motifs, particularly the secondary structure, being favored depending on the force field. For instance, the CHARMM all-atom force field stabilized A $\beta$  as a mixture of helices and  $\beta$ -strands with collapsed coils and loops (Siwy, Lockhart and Klimov, 2017). Lemkul and Bevan showed in one of their studies (Gerben *et al.*, 2014) that A $\beta$ <sub>40</sub> consists of helices (residues 15-23 and 28-38) and random coils in membrane mimicking environments using Gromos96 53A6 and SPC water model while AMBER produces a

helical structure. Other, later studies have shown similar results (Watts *et al.*, 2018). Barz and Urbanc carried out a series of simulations using a combination of OPLS with SPC and TIP3P water models showing that the secondary structures of both  $A\beta_{40}$  and  $A\beta_{42}$  monomers and dimers consist of  $\beta$ -strands and turns (Barz and Urbanc, 2012). The issue is further complicated by the limited timescale of such simulations, making it hard to attribute whether the differences in structure and molecular organization come from the different force fields or a lack of sampling.

Whilst the above simulation studies provided valuable insights into how a force field can affect the secondary and tertiary structures of  $A\beta$  peptides, their focus was largely on how the force fields influenced the structure of a *single*  $A\beta$  peptide or smaller peptide fragments. As simulation interest extends to the various intermediate stages of amyloid formation such as  $A\beta$  oligomers, protofibrils and fibrils, there is a need to develop a better understanding of how well current force fields can reproduce these hierarchical structures

A particular challenge in MD is to study molecular self-assembly, especially of biomolecules. The majority of events happening in the cell include self-assembly reactions, for example, protein synthesis/degradation, gene transcription, cell movement, communication, shape control, and formation of amyloids (Thomas and Schwartz, 2017). Despite its centrality, no major efforts are seen to develop such modeling tools. Though there are some recent advances in systems biology tools (Faeder, Blinov and Hlavacek, 2009), (Gruenert *et al.*, 2010) and modeling (Karr *et al.*, 2012) to incorporate parameters for complex self-assembly, major limitations remain such as sampling problem arising from extremely large conformational space and the accuracy of the force fields. Mass action differential equation (DE) models, coarse-grained (Baschek, R Klein and Schwarz, 2012) models or even methods based on Gillespie's stochastic simulation algorithm (SSA), face challenges because of very large systems containing large numbers of intermediate species, high computational cost, longer timescales and extensive simplification requirement (Frazier, Chushak and Foy, 2009). Many studies, using small peptides with well-defined structures such as ubiquitin, the villin headpiece, and the Fip35 WW domain (26–29) have been carried out to study their assemblies but they are unable to predict correct secondary and tertiary structures, folding mechanisms, and NMR chemical shifts and couplings (Fluitt and De Pablo, 2015). Amyloid aggregation is a prime example of self-assembly process causing

Alzheimer's disease but the peptides responsible there lack a unique native conformation and are inherently disordered (IDPs) (Ke *et al.*, 2017). From a computational viewpoint learning about the structure of amyloid assemblies, the nature of forces governing these interactions and amyloid aggregation pathway is very interesting but is especially challenging. Therefore, it is difficult to select a force field that is anticipated to behave well for modeling IDPs.

The polymorphism resulting from the structural inhomogeneity and instability of A $\beta$  is the basis for differences in morphology, physicochemical properties, and level of cellular toxicity. Therefore, we need to have a good understanding of the structures and conformation of various amyloid-forms i-e monomers, oligomers/protofibrils and mature fibrils involved in the aggregation process. Here we have carried out a systematic study to investigate the quality of three popular force fields, AmberGS, Gromos53a6, and OPLS-aa, for their ability to reproduce hierarchical structures of amyloid A $\beta$ . To our knowledge, this is the first study to explore the effect of force fields on the structure of large amyloid morphologies. Particular features of the study are the extended simulation times (400ns each) enabling sampling of long-timescale fluctuations and our mindfulness of the fact that peptide self-assembly is a nucleation dependent mechanism.

Classical nucleation theory predicts a critical nucleus size below which any emerging structure is liable to disintegrate and go back into to solution. Above the critical size the nucleus is stable and can proceed to develop into an appropriate solid structure. The critical nucleus size results from a free energy barrier to the formation of a stable nucleus. The barrier results from an intricate interplay between the bulk free energy and interfacial free energy of the emerging structure (Anwar and Zahn, 2011) (see Figure 3.1). When the size of the emerging structure is small (pre-critical), the overall free energy for nucleation is dominated by the interfacial free energy which in general is positive (unfavorable). The other term, the bulk free energy, is negative (favorable) for a supersaturated solution and is proportional to the number of molecules in the emerging nuclei and hence to its volume. At some stage as the particle size increases, the favorable bulk free energy begins to dominate as the  $r^3$ -dependent volume term overwhelms the  $r^2$ -dependent interfacial free energy term, where  $r$  is the radius of the emerging nuclei. Above the critical nucleus size the overall free energy begins to decline and then becomes negative, thus stabilizing the nucleus which then can proceed to the growth

phase. The critical nucleus size depends on supersaturation, being smaller for higher supersaturations. The significance of nucleation for molecular simulation of  $A\beta$  aggregates is that if the structure is below the critical nucleus size, it may disintegrate. Furthermore, we should expect the peptides at the surface to disorder in a bid to minimize the interfacial free energy. Consequently, one should be mindful that observing structural disintegration or distortions in a simulation need not be a failure of the force field, though it could be.

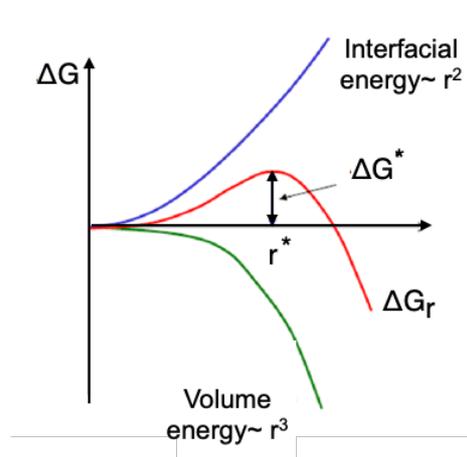


Figure 3.1. Free energy associated with homogenous nucleation of a sphere of radius  $r$ .

### 3.3 Methodology

We investigated the stability of 3 distinct peptide structures along with the  $A\beta$  monomer (1IYT(O Crescenzi *et al.*, 2002)) in MD simulations using the force fields AmberGS, Gromos53a6, and OPLS-aa. In total twelve independent simulations were performed each of a 400ns trajectory with a cumulative simulation time of 4.8 $\mu$ s. The selected structures include (i) 2BEG (Luhrs *et al.*, 2005) a protofibril of the U-shaped peptide (5-mer). This U-shaped protofibril is composed of two  $\beta$ -sheets,  $\beta_1$  and  $\beta_2$  that are tightly packed with each other. The loop region is stabilized by a salt bridge that is formed by the interaction of ASP23-LYS28. Studies show that only these three intersheet interactions can maintain the well-defined U-shaped structure of this protofibril. (ii) 2MXU (Xiao *et al.*, 2015a) a fibril of S-shaped peptides (12-mer) The ss-NMR and cryo-EM experiments show that it consists of an S-shaped triple parallel- $\beta$ -sheet segments. The structure shows three  $\beta$ -sheets at residues VAL12-VAL18( $\beta_1$ ), VAL24-ILE33 ( $\beta_2$ ) and VAL36-ILE40( $\beta_3$ ). (iii) M3Q\_3 (1-40) a ring fibril (18-mer), fiber diffraction studies show that  $A\beta$  ring fibrils consist of cross- $\beta$  structure having two

$\beta$ -sheets running parallel and  $\beta$ -strands perpendicular to the fibril axis. NMR studies show that the two  $\beta$ -sheets ( $\beta 1$  and  $\beta 2$ ) are composed of residues LYS16-ALA21 and 31ILE-VAL36, respectively, and are joined by a turn region. These  $\beta$ -sheets are held together by three specific side-chain contacts formed between residues PHE19-GLY38 and ALA21-VAL36 and a salt bridge between residues ASP23 and LYS28 (Paravastu *et al.*, 2008). The last structure was generated in Modeller V 9.13 (Webb and Sali, 2016) using 2LMQ.pdb as a starting structure (Paravastu *et al.*, 2008), whilst the other three are solution NMR structures and were taken from protein data bank (see Figure 3.4). The PDB identifiers, amino acid sequences, and initial structures are given in Table 9.1. The peptides were solvated using TIP3P, SPC, and SPCE water models for AmberGS, Gromos53a6, and OPLS-aa respectively in a cubic box with a distance of at least 1nm between the peptide and edges of the box. The simulation details are given in SI-Table 2. The system was electro-neutralized by adding suitable counter ions so that the total charge is zero. The steepest descent algorithm was used for energy minimization. The system was equilibrated stepwise in an NVT ensemble simulation followed by an NPT simulation each for 500ps.

All the simulations were carried out at 300K using the MD code Gromacs version 5.1.0 (Abraham *et al.*, 2015). The van der Waals interaction cut off was 0.8nm for AmberGS and 1.4nm for Gromso53a6 and OPLS-aa. Electrostatic interactions were handled using the particle mesh Ewald (PME) method with a grid spacing of 0.16 (Bussi, Donadio and Parrinello, 2007). The temperature and pressure were regulated using the Nose-Hoover thermostat and the Parrinello Rahman barostat respectively (M. Parrinello and Rahman, 1981). Bonds were constrained using the LINCS algorithm (Hess *et al.*, 1997).

### Data analysis

The peptide structures were characterized by a variety of analytical methods including root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), hydrogen bonds,  $C\alpha$ -  $C\alpha$  inter and intra-sheet distances, and salt bridges of peptides. The RMSD-based clustering analysis was performed to explore representative peptide conformations for the monomer  $A\beta$  from the simulated trajectories. For every conformation sampled during the simulation, residues belonging to  $\alpha$ -helices,  $\beta$ -sheets, or turns were identified using the DSSP program (D Frishman

and Argos, 1995). The secondary structure contents ( $\beta$ , helix, and coil) were analyzed using the STRIDE algorithm available online (Dmitrij Frishman and Argos, 1995). The twist angles for comparison of U-shaped and S-shaped peptides were calculated by averaging over the dihedrals between two vectors connecting the first C $\alpha$  atom to the last C $\alpha$  atom in the same  $\beta$ -strand portion between two neighboring  $\beta$ -strands within a  $\beta$ -sheet following Jie Zheng's work (Zheng *et al.*, 2007).

### 3.4 Results

#### A $\beta_{42}$ monomer

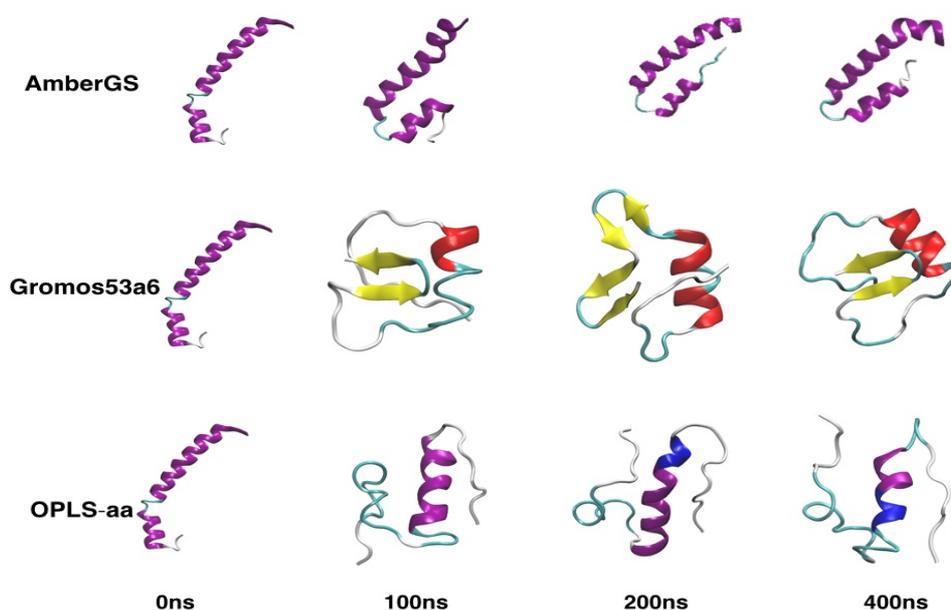
Wild type A $\beta_{42}$  peptide (-3 charge) represents the first species modeled in this study. This is the biologically relevant charged state that is present at physiological pH and under the conditions of our systems (Baumketner, 2006). The study of various conformational states of A $\beta_{42}$  during the aggregation process is a challenging task because monomer A $\beta$  is intrinsically disordered in aqueous environments (Weber and Uversky, 2017). NMR solution studies show that A $\beta(1-40)$ , or A $\beta(1-42)$  possess no  $\alpha$ -helical or  $\beta$ -sheet structure (Shao *et al.*, 1999) rather they exist predominately as collapsed coils (Zhang *et al.*, 2000). While in the apolar medium, both A $\beta(1-40)$  and A $\beta(1-42)$  are present in  $\alpha$ -helical conformations (Tomaselli *et al.*, 2006).

Overall, our simulation for the A $\beta_{42}$  monomer show similar conformation as other helical structures of full-length A $\beta$ , i.e. A $\beta_{40}$  (Vivekanandan *et al.*, 2011), however a detailed analysis of various force fields reveals significant differences. AmberGS favors helical structure while Gromos53a6 stabilizes a pair of  $\beta$ -sheets. OPLS-aa generates a variety of secondary structure motives. The averaged secondary-structure probability per residue in terms of regular secondary structure characteristics i.e.  $\alpha$ -helices,  $\beta$ -sheets, turns, and coils is presented in Figure 9.1. Our results show that AmberGS force field yields and stabilizes a well-defined elbow-shaped structure that contains two  $\alpha$ -helices separated by a turn region at residue VAL24-ASN27.

Whereas in Gromos53a6, in first 10ns, one of the  $\alpha$ -helix of A $\beta_{42}$  monomer converts into a pair of antiparallel  $\beta$ -strands and aligns itself right in front of the other  $\alpha$ -helix. During the next 100ns, another pair of anti-parallel  $\beta$ -strands appears and the  $\beta$ -sheet content increases to 14%. After 400ns simulation, the A $\beta_{42}$  monomer stabilizes a pair

of  $\beta$ -sheets at residues LEU34-VAL36 and VAL39-ILE41 and an  $\alpha$ -helix at VAL12 - VAL24.

In OPLS-aa force field peptide conformation appears mostly to be dominated by turns and coils probably due to the intrinsically disordered nature of A $\beta$ <sub>42</sub>. During the first 50ns, one of the helices is lost and converted to coil while the other helix persists at residues TYR10-VAL24. At the end of 400ns long simulation the structure consists of 22%  $\alpha$ -helix and no  $\beta$ -sheet (see Figure 3.2). The details of secondary structure content are shown in table 3 in SI.



**Figure 3.2. Representative snapshots of secondary structures of A $\beta$ <sub>42</sub> from a 400ns long trajectory under the influence of (upper panel): AmberGS, (middle panel): Gromos53a6 and (lower panel): OPLS-aa.**

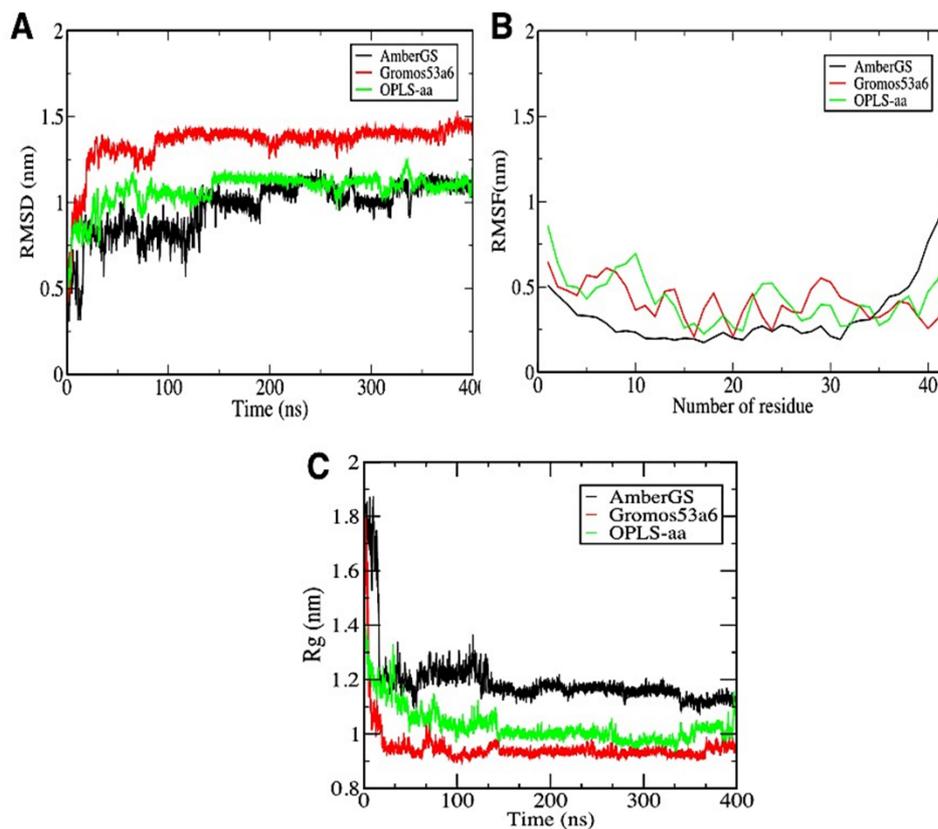
As the starting structure of the A $\beta$ <sub>42</sub> monomer used in this study consists of primarily  $\alpha$ -helices, we found from the clustering analysis (clusters indicate some population-average conformational similarity) that for AmberGS, all of the major clusters consist of two helices separated by a turn region (see Figure 9.1). Structural changes such as end-to-end distances vary in all the conformations, though their magnitude is not so prominent. There is one major cluster accounting for 48% of the conformations, whereas the other account for 20, 20, 5, and 2%. In the Gromos53a6 system, the A $\beta$ <sub>42</sub> structures became more diverse. Besides the major cluster (51%), the other clusters account for 13, 8, 7, and 3.5%, respectively. Changes in the secondary structure of the residues occur along the sequence for smaller clusters, though the most populated clusters stabilize their secondary structures. OPLS-aa structures show varying

conformational forms. The most populated cluster (27%) contains the largest  $\alpha$ -helical content while it is observed to be gradually lost in other clusters. Cluster 3 doesn't contain any  $\alpha$ -helix at all. We measured the secondary structure distribution of  $A\beta_{42}$  monomer in the three force fields by Stride and shown in Figure 9.3.

RMSD for the backbone of the  $A\beta_{42}$  for each of the force fields is plotted as a function of time in Figure 3.3. In all instances, the RMSD increases with time and then reach a plateau, confirming convergence of the simulations. The RMSD for AmberGS is the lowest being  $0.96\pm 0.1\text{nm}$  and highest for Gromos53a6 at  $1.34\pm 0.1\text{nm}$ . These values are relatively large indicating a significant change in the structure and conformation but expected as the structure samples a substantial conformational space.

RMSF reveals a high degree of variability in the residues forming the turn and the C-terminal region. During the simulation, the  $C\alpha$  atomic fluctuations for AmberGS indicated that the residues 10-30 are rigidly stable during simulation with a  $C\alpha$  RMSF of approximately 0.25 nm. The obtained results augment the results of secondary structure analysis that shows that AmberGS stabilizes the helical structure. The RMSF for residues 31–42 ranges from 0.26 to 1.0nm. This sharp increase in the RMSF value of the terminal region indicates a conformational change i-e an increase in the bend content at C-terminal is observed. Gromos53a6 and OPLS-aa show wider variation in RMSF indicating that the regions of greatest flexibility are terminal residues and the polar residues. (see Figure 3.3B).

Looking at the time evolution of secondary structures snapshots for Gromos53a6 in Figure 3.2 it is evident that, the most critical step of aggregation i-e conversion of  $\alpha$ -helix to  $\beta$ -sheet takes place in the initial 100ns of the simulation. During the next 300ns the structure looks stable. For OPLS-aa there are small fluctuations in RMSD that represent the formation of coil and bends. AmberGS has the lowest RMSD as it stabilizes the initial  $\alpha$ -helices but the fluctuation represents the change in the structure as it becomes more compact.



**Figure 3.3. Representative data set from 400ns long simulation of A $\beta$ <sub>42</sub>: (A) Backbone RMSD of A $\beta$ <sub>42</sub> for the three force fields, (B) RMSF, (C) Radius of gyration of A $\beta$ <sub>42</sub> for the three force fields.**

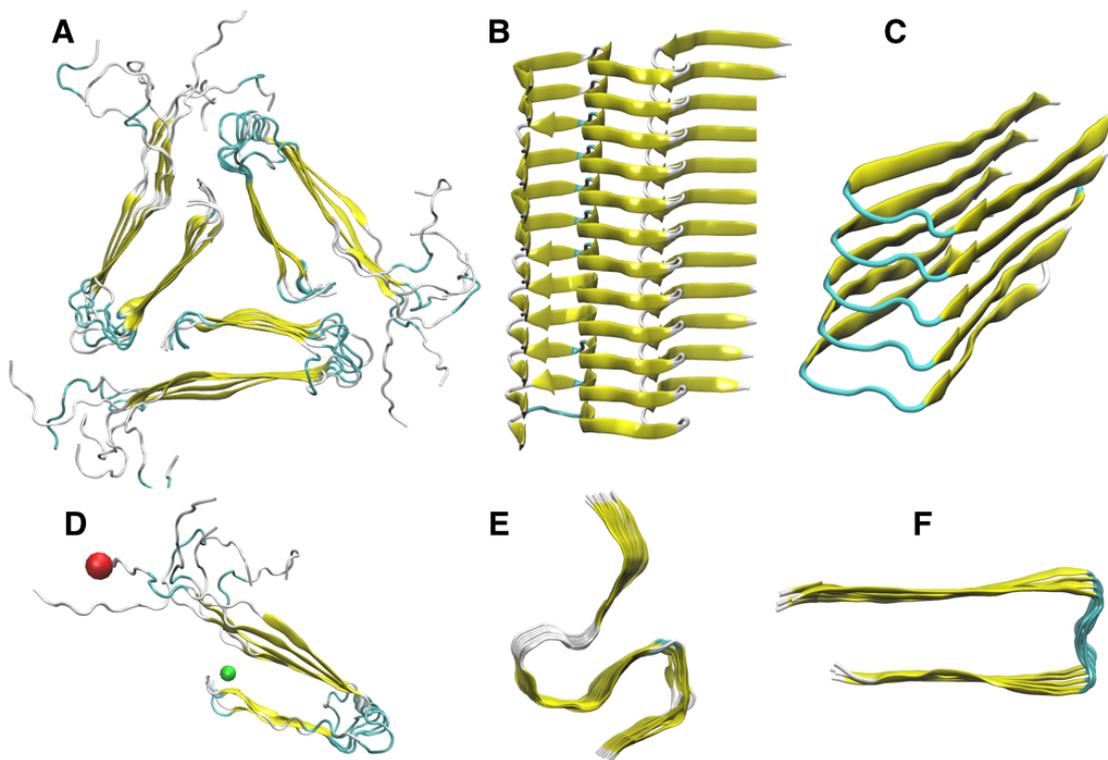
The variation in radius of gyration ( $R_g$ ) for each of the force fields is shown in Figure 3.3C. The data indicate that the A $\beta$  peptides generally converge to a more compact structure. The measured radii of gyration lies between  $0.9 \pm 0.08$  nm for Gromso53a6,  $1.0 \pm 0.09$  nm for OPLS-aa, and  $1.2 \pm 0.1$  nm for AmberGS which is in accordance with the literature(Nag *et al.*, 2011).

## Part-2 Fibrils

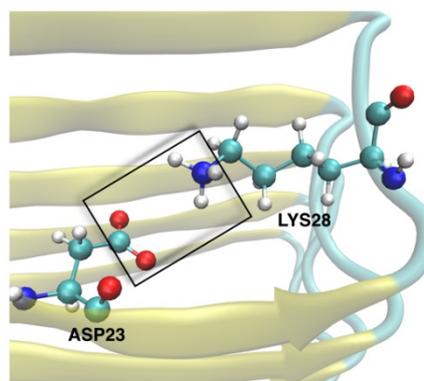
This part of the study employs MD to provide detailed structural characterization of three fibrils of varying sizes and complexity i-e U-shaped protofibril, S-shaped fibril and ring fibril, under the influence of the three force fields as mentioned earlier.

2BEG (U-shaped protofibril):

Over the course of 400ns, we observed very small deviations from the initial structure. Some twisting was seen that is evident by an initial increase in RMSD. The converged RMSD values for all three force fields range from 0.46- 0.57 nm.

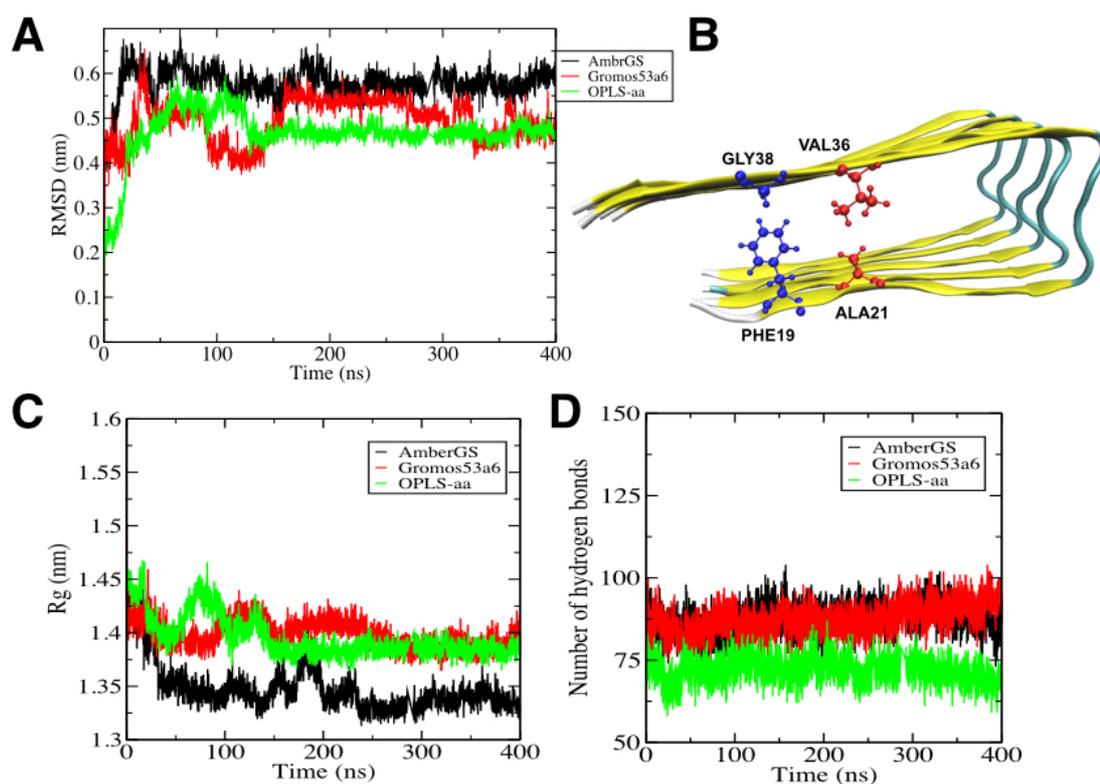


**Figure 3.4. Structures of three fibrils of varying sizes and complexity i-e (A). U-shaped protofibril, (B). S-shaped fibril and (C) ring fibril, (D),(E) and (F) present the top views of U-shaped protofibril. S-shaped fibril and ring fibril.**



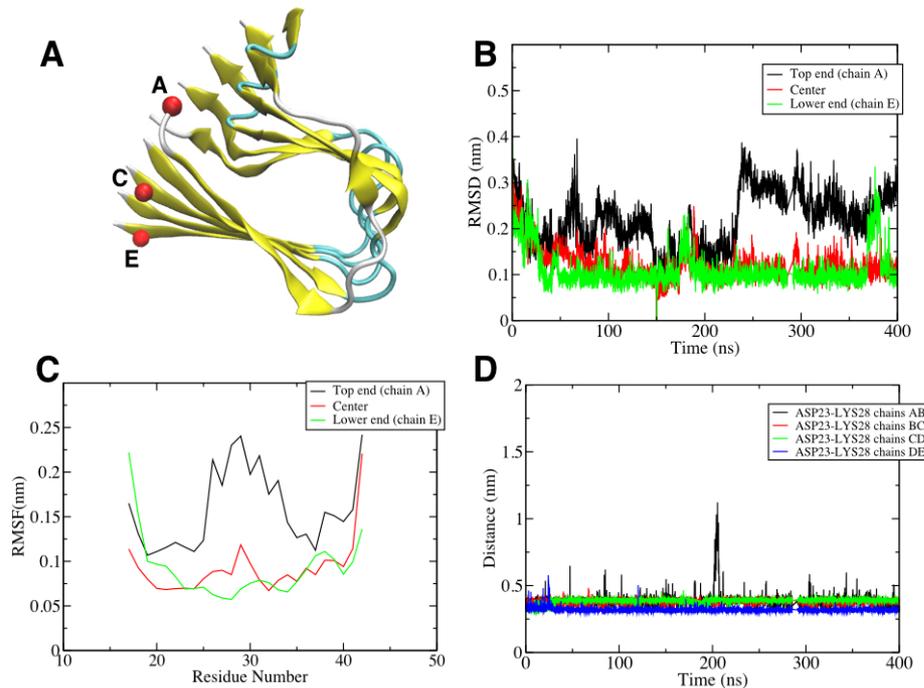
**Figure 3.5. Asp23 – Lys28 (COO<sup>-</sup> - NH<sub>3</sub><sup>+</sup>) forming the salt bridge responsible for maintaining the U-shape of the U-shaped protofibril.**

There appeared to no significant long-range fluctuations in the converged values. Having said that, the protofibril using the Gromos53a6 force field showed some fluctuations over the period 50-150ns which is due to the disruption of salt bridge ASP23-LYS28 (see Figure 3.5) in the top end (chain A) of the protofibril (will be discussed later in detail). The converged average RMSD was the highest for AmberGS  $0.57 \pm 0.03$ nm, giving rise to a compact structure ( $R_g = 1.33$ nm) (see Figure 3.6) followed by  $0.49 \pm 0.04$  for Gromos53a6, and then OPLS-aa with a value of  $0.46 \pm 0.03$ nm. The number of interchain hydrogen bonds also remain stable across all simulations, with an average of  $87.0 \pm 4$  for AmberGS and Gromos53a6 and  $72.0 \pm 4$  for OPLS-aa persisting throughout all of the 400 ns trajectories (see Figure 3.6D). The RMSD,  $R_g$ , and hydrogen bond analysis as a function of simulation time for the U-shaped protofibril for each of the force fields are shown in Figure 3.6.



**Figure 3.6. (A).** Backbone RMSD for U-shaped protofibril from 400ns long simulations for AmberGS, Gromos53a6 and OPLS-aa forcefields. **(B)** Hydrophobic interactions between side chains of PHE19-GLY38 and ALA21-VAL36 of U-shaped protofibril, **(C)**  $R_g$  (nm) and **(D)** Average number of interpeptide hydrogen bonds for U-shaped protofibril in AmberGS, Gromos53a6 and OPLS-aa force fields

Focusing on the behavior of individual chains, for all force fields, on average each protofibril chain stabilized at an RMSD value of approximately  $0.25 \pm 0.02$  nm, indicating minimal structural change except for the top chain A of the protofibril that showed large fluctuations. NMR studies show that the central chain of the pentamer has an RMSD of 0.118nm (Lührs *et al.*, 2005b). The simulations reproduced this RMSD remarkably well, the values being  $0.12 \pm 0.03$  nm,  $0.25 \pm 0.02$  nm, and  $0.12 \pm 0.06$  nm for AmberGS, Gromos53a6 and OPLS-aa respectively. Representative structures, trajectory RMSD and RMSF for selected representative peptide chains, and the variation of the Asp23 – Lys28 (COO - NH<sub>3</sub><sup>+</sup>) distances within a chain as a function of time are shown for each of the force fields in Figure 3.7 (AmberGS), Figure 3.8 (Gromos53a6) and Figure 3.9 (OPLS-aa).

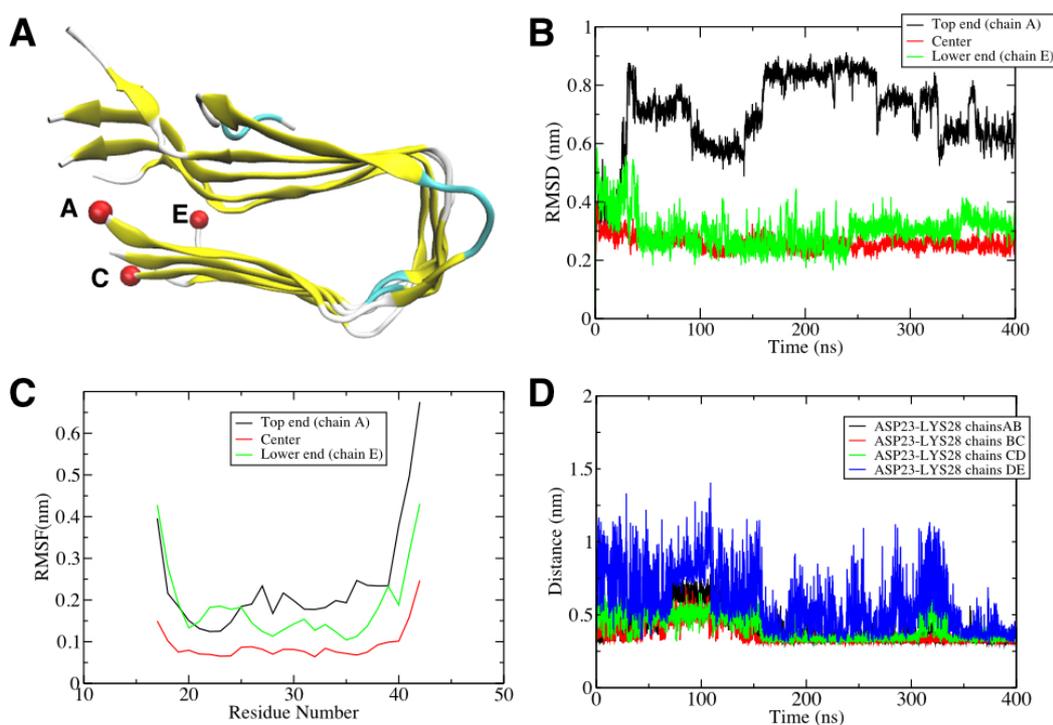


**Figure 3.7. Representative data set from 400ns long simulation of AmberGS: (A) structure of the U-shaped protofibril after 400 ns, generated with VMD; Top, centre and lower chain's ends are denoted by red balls, (B) RMSD values of representative peptide chain, (C) RMSF values of representative peptide chains and (D) Asp23 – Lys28 (COO - NH<sub>3</sub><sup>+</sup>) distances for all protofibrillar chains.**

The RMSF data shows that the most flexible regions of each chain of the U-shaped fibril are the C-terminal residues and those in close proximity of the bend region (residues 25-30) which in line with previous reports (Barale *et al.*, 2019). An exception here is the top chain in AmberGS, which shows dramatic flexibility in the whole chain

with the most prominent being the N- and C-terminal residues (see Figure 3.7C). This is due to the deformation caused at the loop region due to the over twisting of the top chain (dihedral angle  $-12^\circ$ ).

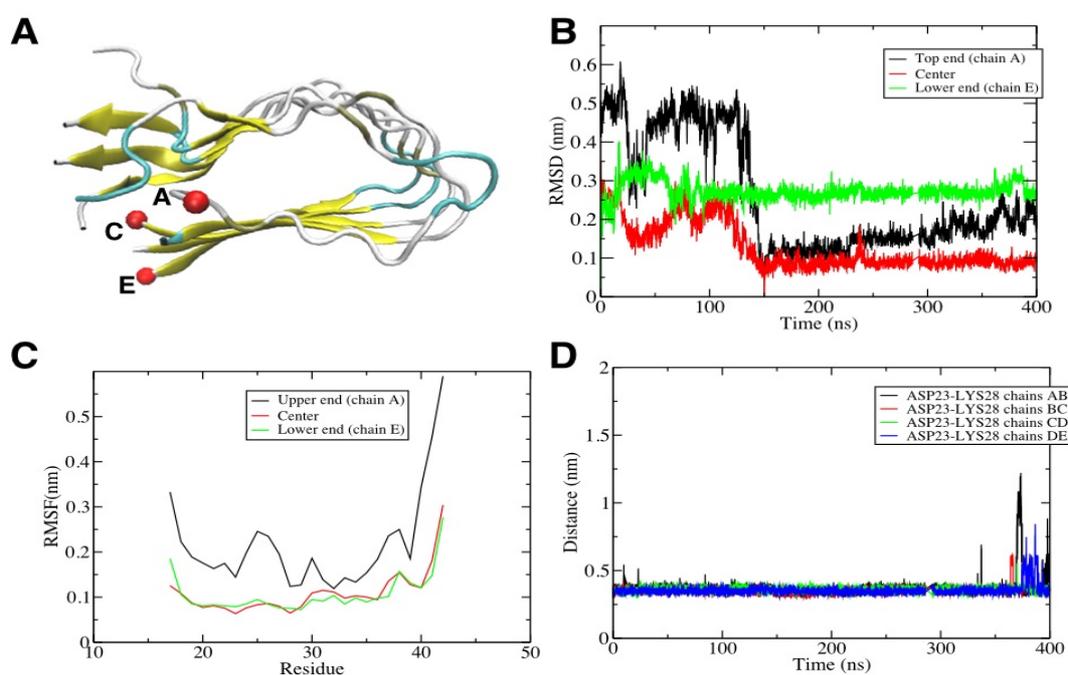
To find out what makes some of the regions of protofibril more stable than the others, the role of hydrophobic interactions between side chains of PHE19-GLY38 and ALA21-VAL36 was investigated. We calculated the interchain distances between PHE19 from  $\beta 1$  of the top chain and GLY38 of  $\beta 2$  of the next chain and so on Figure 3.6B.



**Figure 3.8. Representative data set from 400ns long simulation of Gromos53a6: (A) structure of the U-shaped protofibril after 400 ns, generated with VMD; Top, centre and lower chain's ends are denoted by red balls, (B) RMSD values of representative peptide chain, (C) RMSF values of representative peptide chains and (D) Asp23 – Lys28 (COO - NH<sub>3</sub><sup>+</sup>) distances for all protofibrillar chains.**

The U-shape of the protofibril is stabilized by these interactions, the average separation distance being  $0.75 \pm 0.1\text{nm}$ ,  $0.99 \pm 0.01\text{ nm}$ , and  $1.15 \pm 0.05\text{ nm}$  for AmberGS, Gromos53a6 and OPLS-aa respectively (distances listed in Table 4 in SI). The interactions remain stable for all the chains except for the bottom chains for the AmberGS force field where the interaction between ALA21-VAL36 show large fluctuations (see Figure 9.4). Similarly for Gromos53A6, generally the hydrophobic

interactions remain stable except for the top and lower chains that show large fluctuations as evident from the RMSD and RMSF values (see Figure 3.8C). For OPLS, all the chains show an increase in the interchain distances at the beginning of the simulation but for the rest of the simulation the distances converge to an equilibrium value and show little fluctuations (see Figure 9.5). The average distances between (F19C $\alpha$ , G38C $\alpha$ ) and (A21C $\alpha$ , V36C $\alpha$ ) is close to 1.0nm which is in good agreement with the inter-sheet inter-residue distances determined from experiments namely NMR (0.8nm) and x-ray fiber diffraction (1.0nm) (Lührs *et al.*, 2005b).



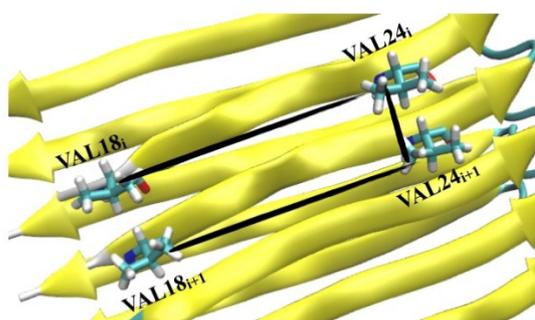
**Figure 3.9. Representative data set from 400ns long simulation of OPLS-aa: (A) structure of the U-shaped protofibril after 400 ns, generated with VMD; Top, centre and lower chain's ends are denoted by red balls, (B) RMSD values of representative peptide chain, (C) RMSF values of representative peptide chains and (D) Asp23 – Lys28 (COO - NH3<sup>+</sup>) distances for all protofibrillar chains.**

The loop region connecting  $\beta$ 1 and  $\beta$ 2 of the protofibril is stabilized by the interchain salt bridge formed by ASP23-LYS28 with an average distance of  $0.39 \pm 0.05$  nm (Barale *et al.*, 2019) (see Table 9.3 ). This salt bridge is formed by a C $_{\gamma}$  atom of residue ASP23 and N $_{\zeta}$  atom of residue LYS28 (Figure 9.6). Consistent with ssNMR data on amyloid fibrils, the structures in all of the force fields maintain interchain Asp23/Lys28 salt-bridge contacts in the loop region except for Gromso53a6 where the distance between interchain salt bridge forming residues show large fluctuations.

**Table 3.1. Average interpeptide Asp23 – Lys28 NH<sub>3</sub><sup>+</sup> COO<sup>-</sup> distances (nm  $\pm$  Std. Dev.)**

Forcefields	Chain AB	Chain BC	Chain CD	Chain DE
AmberGS	0.38 $\pm$ 0.05	0.37 $\pm$ 0.02	0.38 $\pm$ 0.02	0.31 $\pm$ 0.01
Gromos53a6	0.41 $\pm$ 0.1	0.37 $\pm$ 0.07	0.40 $\pm$ 0.08	0.58 $\pm$ 0.2
OPLS-aa	0.37 $\pm$ 0.07	0.35 $\pm$ 0.02	0.36 $\pm$ 0.01	0.35 $\pm$ 0.03

The  $\beta$ -sheet is the most prominent secondary structure feature of amyloid fibrils. As discussed above the protofibril is composed of two  $\beta$ -sheets,  $\beta$ 1 and  $\beta$ 2 containing residues 18-26 and 31-41 respectively. The interchain distances between PHE19 and GLY38 and ALA21 and VAL36 are a measure of the compactness of two  $\beta$ -sheets. The strong hydrophobic interactions between these residues provide a stable  $\beta$ -sheet structure to the protofibril. Among the three force fields, Gromos53a6 finds 68% of the residues in  $\beta$ -sheet conformation while OPLS-aa samples the lowest percentage (41%) of  $\beta$ -sheet structure but the highest percentage of coil (42%). For AmberGS the percentages of  $\beta$ -sheet structure are around 55%. The percentage of the bend population for all the force fields is in the range of 6–12%. Overall, we can rationalize that Gromos53a6 strongly favors the  $\beta$ -sheet structure while OPLS-aa provides a good balance between  $\beta$ -sheet and coil.



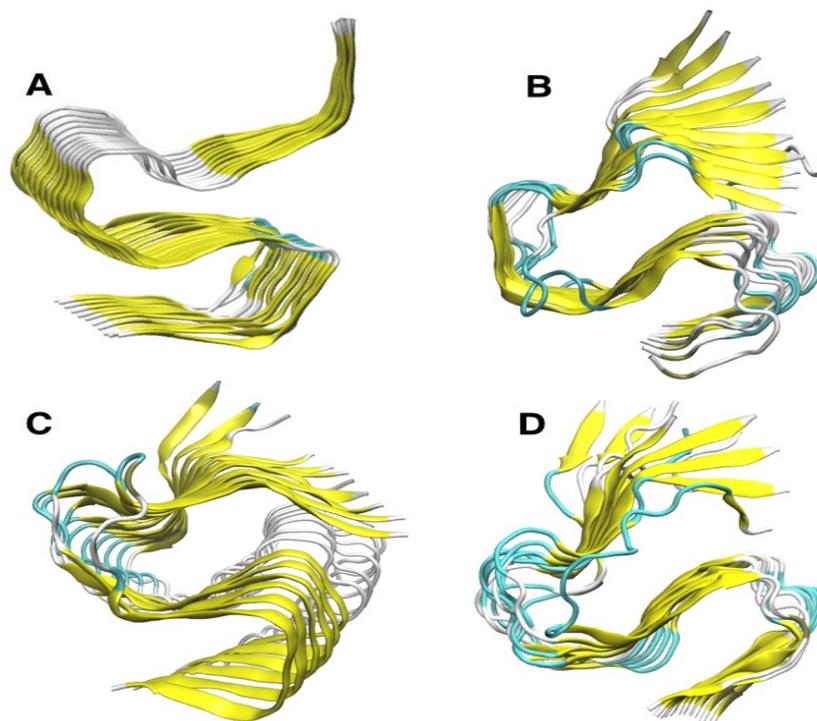
**Figure 3.10 The dihedral angle between C-alpha atoms of VAL18 and VAL24 of two adjacent  $\beta$ -sheets.**

The presence of twist along the fibrillar axis is common to amyloid fibrils. However, there appeared to be no twisting of  $\beta$ -strands in the initial structure obtained from PDB, which may be attributed to the restraints on the inter-strands backbone hydrogen bonds used during structure determination. The twist angle of the U-shaped protofibril was calculated by averaging the dihedral angle between C-alpha atoms of VAL18 and VAL24 (see Figure 3.10). We observe that the force field does affect the topology of

the protofibril. The twist angle for the initial energy minimized structure is very small but non-zero ( $-0.6^\circ$ ) between adjacent  $\beta$ -strands, after 400ns, the structure under all the force fields adopts a left-handed twist with an average value of  $-10.73^\circ$ ,  $-3.24^\circ$  and  $-9.69^\circ$  for AmberGS, Gromos53a6 and OPLS-aa respectively. Taken together with an intersheet distance of  $\approx 4.9\text{nm}$  the twist values extrapolated along the fibril axis result in a pitch of 21nm, 6.35nm, and 19nm for AmberGS, Gromos53a6 and OPLS-aa respectively. This shows that the conformational plasticity of the protofibril in AmberGS is higher but is characterized by a compact conformation as shown by a smaller value of the radius of gyration.

2MXU (S-shaped fibril):

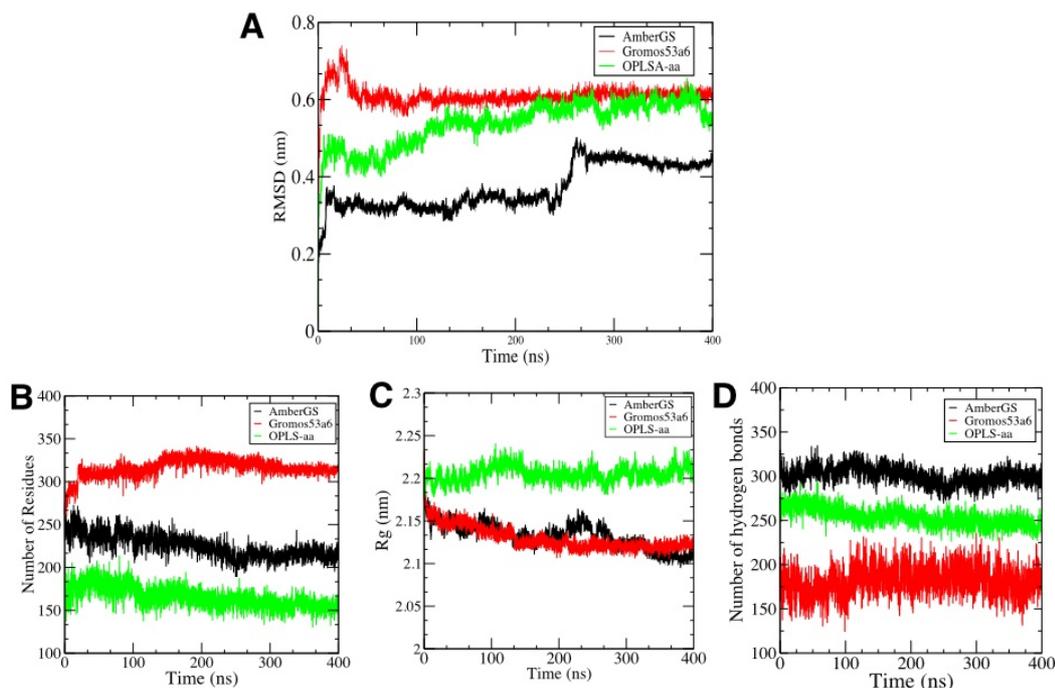
Overall, the triple- $\beta$  motif of the S-shaped protofibrils stays preserved throughout the simulation (see Figure 3.11). Compared to the structures obtained by Gromos53a6 and OPLS-aa, the AmberGS force field simulation yields a more stable structure with a lower RMSD  $0.37\pm 0.05$  nm as compared to  $0.61\pm 0.02$  nm for Gromos53a6 and  $0.53\pm 0.05$  nm for OPLS-aa. This stability is a result of the larger number of hydrogen bonds for AmberGS ( $\geq 320$ ) The RMSD for similar structures obtained by (Lee, Yoon and Shin, 2019) at 500 ns were  $\sim 0.34$  nm for CHARMM27 and  $\sim 0.5$  nm for Amber99sb-ildn, respectively. Note that at around 280ns, there is a sharp increase in the RMSD of the fibril for AmberGS due to the extraordinary flexibility of the C-terminal region of the lowest chain, while the structure in Gromos53a6 though having a higher RMSD approaches a plateau and remains stable throughout the simulation period Figure 3.12. This difference may be attributed to many factors including the larger flexibility of the end chains that induce loss of hydrogen bonds and as a result there is the loss in secondary structure as well.



**Figure 3.11. Representative snapshots of S-shaped  $A\beta$  fibril at 400ns. (A) Initial structure (B) AmberGS, (C) Gromos53a6, and (D) OPLS-aa force fields.**

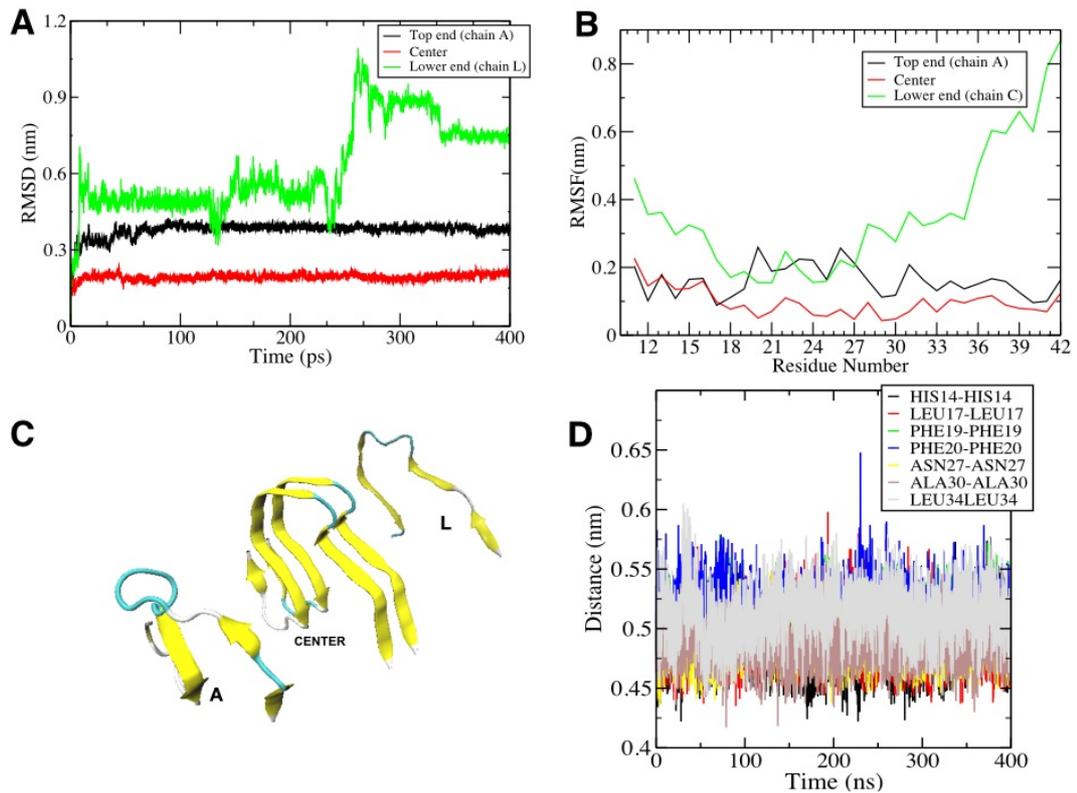
The main conformation of the structures obtained in all the force fields consists of  $\beta$ -sheets, i.e., 57% (AmberGS), 80% (Gro-mos53a6), and 42% (OPLS-aa). The content of these  $\beta$ -sheets reduces in AmberGS and OPLS-aa resulting in an increase in the other conformations, e.g., to bend structures by 11% and 16% respectively. While the fraction of bend conformation in Gromos53a6 is 4% compared to the native structure (Table 9.3).

Most of the peptides of the fibril especially in the central region forming the core are structurally stable with a deviation of less than 0.3 nm. In contrast, the ends of the fibril show plasticity. RMSD values show that both the ends, top, and bottom are relatively unstable as compared to the center. For AmberGS and OPLS-aa, the lower chains show large rmsd indicating big structural changes at the ends. The RMSF of amino acid side chains is calculated to measure local changes in the residues of the peptides forming the fibril. In both the outermost chains, upper and lower, most flexible regions of the fibril are at the N terminal and C-terminal residues.



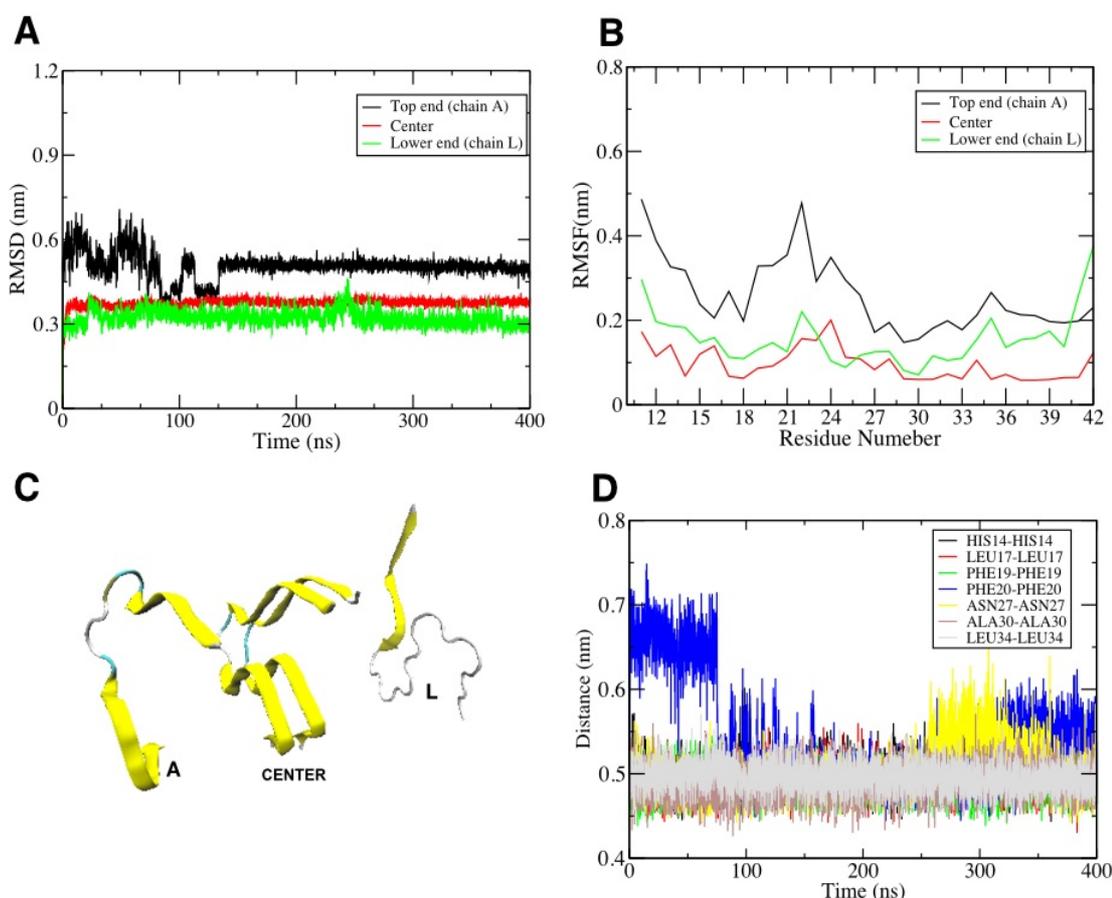
**Figure 3.12. Representative data set of S-shaped fibril from 400ns long simulation of AmberGS, Gromos53a6 and OPLS-aa force fields. (A) RMSD values of representative S-shaped fibril showing the structure in AmberGS has the lowest RMSD (B) Number of residues adopting  $\beta$ -sheet structure; increasing for Gromos53a6 as a function of time, (C) Radius of gyration showing OPLS-aa has a higher value of gyration showing more extended conformation (D) Total number of hydrogen bonds in the three force fields.**

The residues from  $\beta 1$  and  $\beta 2$  sheets are relatively stable because of the interpeptide hydrophobic contacts. In both the outermost chains, residues 21–28 show disordered conformation while residues 38–42 of the lower end are highly unstable and show disordered conformations. This region contains the C-terminal of the fibrils that are involved in the formation of the salt bridge (LYS28-ALA42) and is responsible for the structural integrity of the triple parallel- $\beta$ -sheet structure. Due to these fluctuations and other effects such as twisting, this particular interaction is no more adequate making these regions unstable. For top, centre and lower chains of the fibril, the RMSDs and RMSF values are shown in Figure 3.13, Figure 3.14, and Figure 3.15 for AmberGS, Gromos53a6, and OPLS-aa respectively.



**Figure 3.13. Representative data set from 400ns long simulation of AmberGS: (A) RMSD values of representative peptide chains; Top, centre and lower chains (B) RMSF values of representative peptide chains (C) Representative structures of Top, centre and lower chains of S-shaped fibril generated by VMD, (D) Average interpeptide distances.**

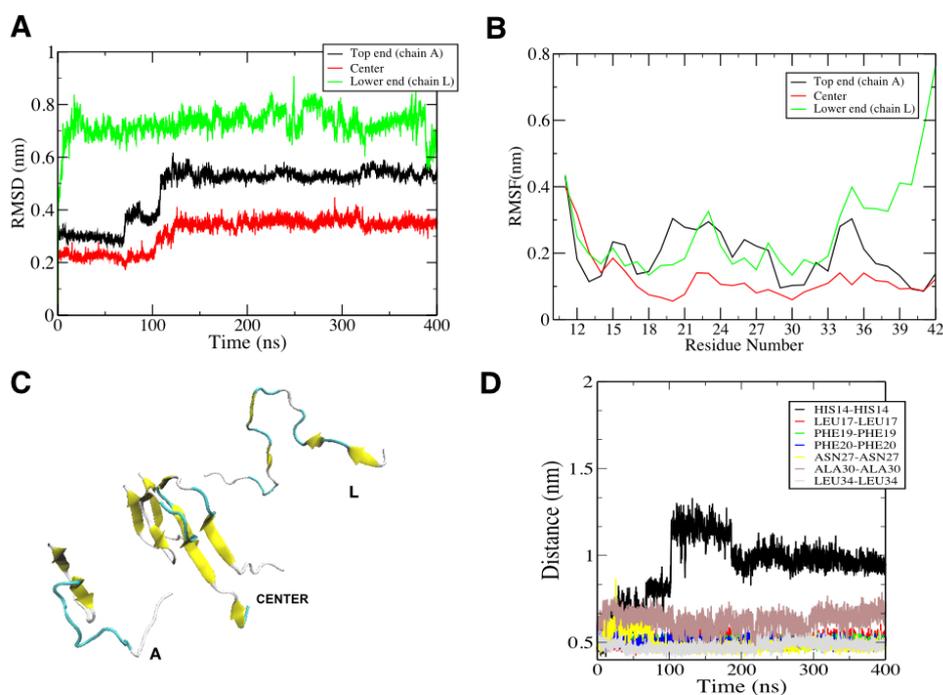
The peptides in the centre show minor deviation from the NMR structure irrespective of the instability of the ends, especially the lower end. As the data shows the central region of the fibril is mostly stable for all the force fields, we chose two chains from the core of the fibril to calculate the inter sheet residual distance to get an idea about how this interaction is affected by the effect of various force fields. The three  $\beta$ -sheet regions ( $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ ) of the fibrils form two hydrophobic cores.  $\beta_1$  and  $\beta_2$  forms the first hydrophobic core while  $\beta_2$  and  $\beta_3$  form the second hydrophobic core. These hydrophobic cores are stabilized by the interactions of residues HIS14, LYS17, PHE19 & 20, ASN27, ALA30, and LEU34 of two adjacent chains which in turn stabilizes the S-shaped conformation of the fibril (data presented in Table 9.7). These interactions generally remain stable for the central chains in all simulations of AmberGS and Gromos53A6. While the structure in OPLS-aa, show some variations in the inter-sheet residual distances. This is also confirmed by large rms deviation for the central chain and overall structure in OPLS-aa. Plots of inter-chain distances are given in Figure 3.13D, Figure 3.14D, and Figure 3.15D for AmberGS, Gromos53a6, and OPLS-aa respectively.



**Figure 3.14. Representative data set from 400ns long simulation of Gromos53a6: (A) RMSD values of representative peptide chains; Top, centre and lower chains (B) RMSF values of representative peptide chains (C) Representative structures of Top, centre and lower chains of S-shaped fibril generated by VMD, (D) Average interpeptide distances.**

A gradual left-handed twisting of  $\beta$ -strands in the  $\beta 1$  region of the fibril is observed (see Figure 3.16). To calculate the twist of the fibril during the simulations we calculated the average dihedral angles between two vectors connecting the first  $C\alpha$  atom of residue VAL12 to the last  $C\alpha$  atom of residue VAL18 in the same  $\beta$ -strand. At the end of the respective trajectories, these two strands from the core region of the fibril show a gradual twist for Gromos53a6 (total twist angle of ca.  $-78^\circ$ ) but both the ends show large twist, unlike AmerGS where the consecutive  $\beta 1$ -strands show large fluctuations in the dihedral angles and develop a total twist angle of  $-98^\circ$ . The structure in OPLS-aa is quite disordered and shows large rms deviation as discussed earlier. A similar trend is observed in the twist angle for OPLS-aa where the fibril adopts a disordered structure with an average twist angle of  $-104^\circ$ . Assuming that the distance between two consecutive peptides  $i$  and  $i+1$  in the fibril is about 0.5 nm, we obtained a pitch of 38nm,

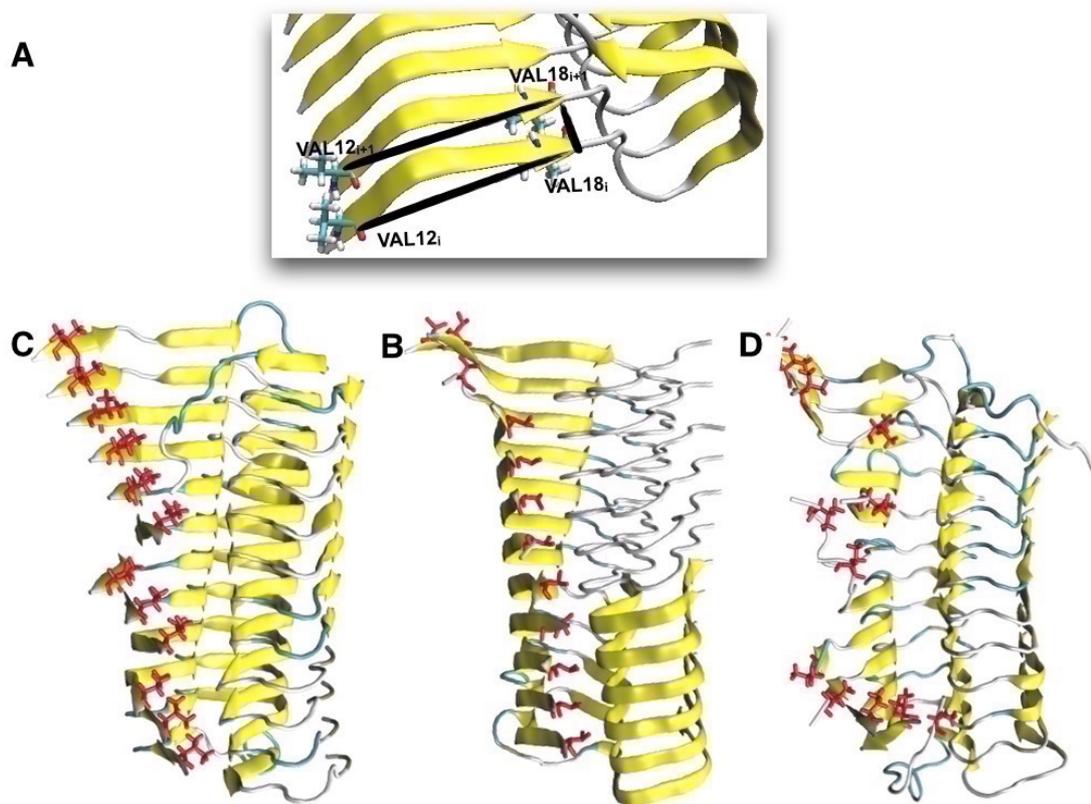
48nm, and 51nm for AmberGS, Gromos53a6 and OPLS-aa respectively which is comparable to experimental findings (Schmidt *et al.*, 2009; Zhang *et al.*, 2009).



**Figure 3.15. Representative data set from 400ns long simulation of OPLS-aa: (A) RMSD values of representative peptide chains; Top, centre and lower chains (B) RMSF values of representative peptide chains (C) Representative structures of Top, centre and lower chains of S-shaped fibril generated by VMD, (D) Average interpeptide distances.**

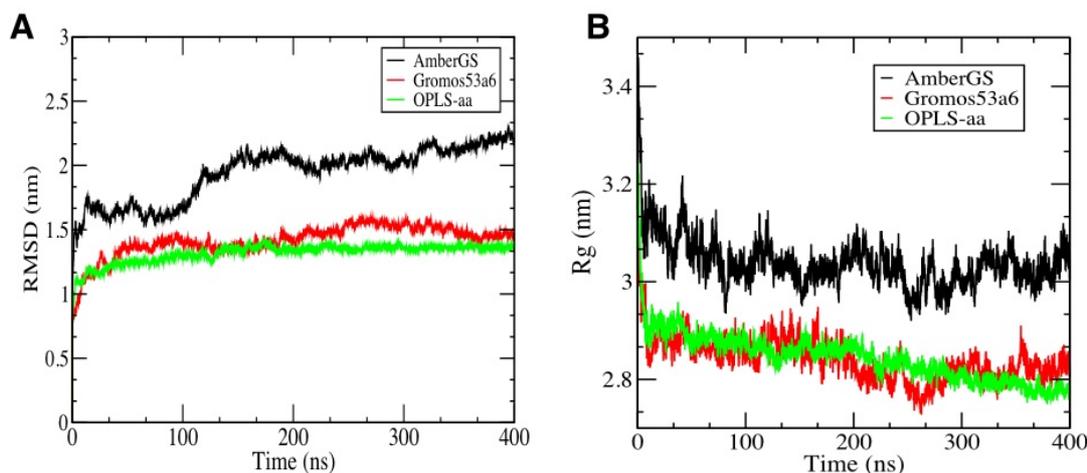
#### M3Q\_3-Ring fibril:

To investigate the relative conformational stability of ring fibril, we calculated backbone RMSD with respect to the initial energy minimized structure as a function of time. The results are plotted in Figure 3.17. For all the force fields initially RMSD increases and then fluctuates around an equilibrium value except for AmberGS (RMSD =  $1.94 \pm 0.21$  nm) that keeps on increasing. We found that RMSD for Ring fibril is lower in OPLS-aa (RMSD =  $1.31 \pm 0.07$  nm) as compared to AmberGS and Gromos53a6 ( $1.41 \pm 0.12$  nm).



**Figure 3.16. (A) Residues defining twist angle dihedral angle between C-alpha atoms of VAL12 and VAL18, Representative snapshots of the S-shaped fibrils, (A) AmberGS, (B) Gromos53a6 and (C) OPLS-aa. Red stick representations are to point out residue VAL18 of  $\beta$ 1-sheet where twisting starts.**

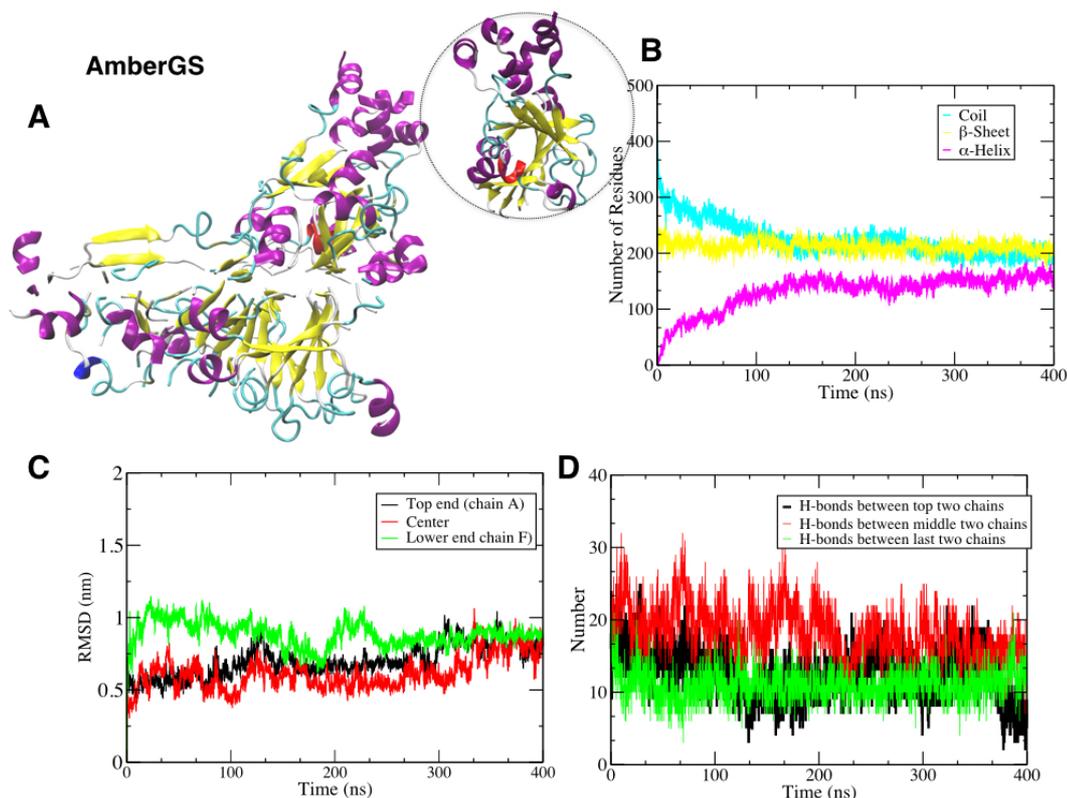
Overall large RMSD shows that the structure in all the force fields is undergoing large structural changes and systematic distortion especially for AmberGS, i.e conversion of initially present disordered regions at N-terminal to  $\alpha$ -helices. The three segments of this 3-fold symmetry fibril are originally at  $60^\circ$  to each other but undergo a change especially for AmberGS which gives rise to a distorted symmetry. Out of all the force fields, the structure in OPLS-aa can maintain the original symmetry to some extent, and here the angles between the three segments are close to  $60^\circ$  (see Figure 9.7). This is consistent with the STEM data and solid-state NMR spectra (Paravastu *et al.*, 2008). In all the simulations, the Rg initially decreases but then oscillates around the same value during 400ns. The radius of gyration fluctuates about 3.1 nm for AmberGS that has the largest RMSD values too (see Figure 3.17B)



**Figure 3.17. Representative data set of ring fibril from 400ns long simulation of AmberGS, Gromos53a6 and OPLS-aa force fields. (A) RMSD values of representative ring fibril showing the structure in AmberGS have the highest RMSD, (B) Radius of gyration of ring fibril; largest for ring fibril as a function of time.**

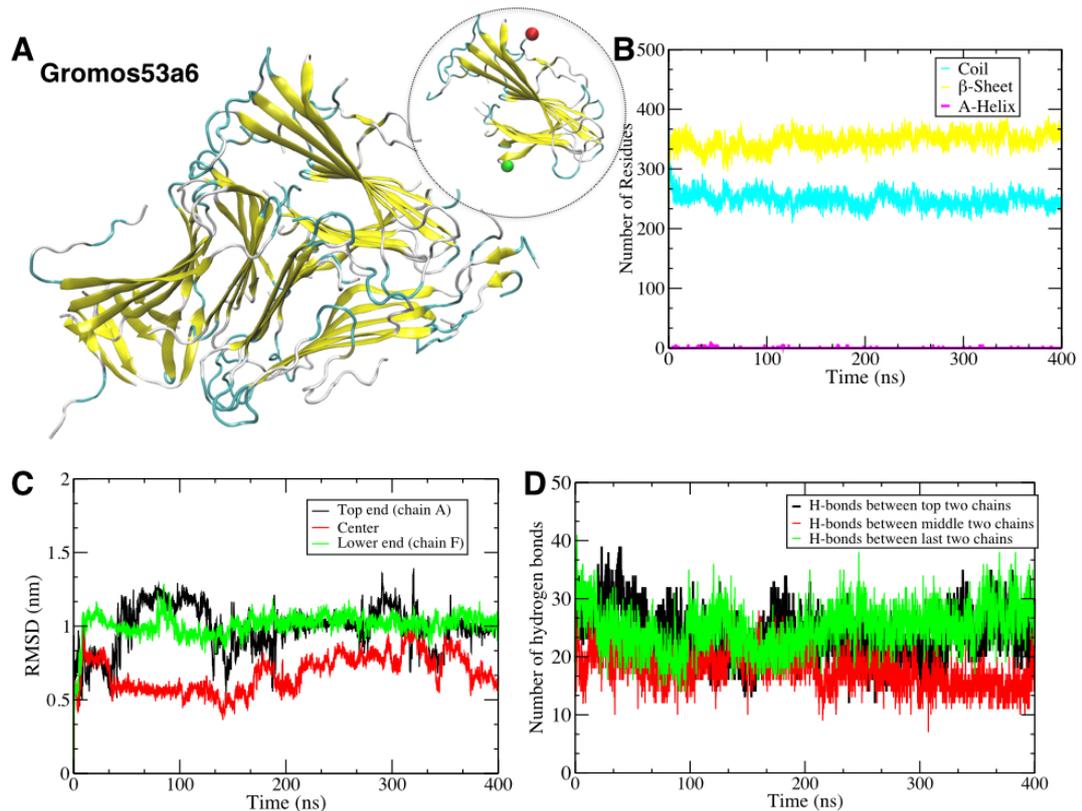
To understand the difference in the structure and conformation in detail, we performed analysis on individual chains forming the fibril segments. In the initial conformation both the  $\beta$ -sheets were closely spaced (2.0nm)(Paravastu *et al.*, 2008) but later, the ends start to open up except for OPLS-aa, increasing the distance between two adjacent  $\beta$  sheets (end-to-end distance  $\geq 3.0$ nm for AmberGS). The increase in the end-to-end distance at the lower half of the fibril and high RMSD of the lower chains together point towards a very significant phenomenon of amyloid fibrils i-e twisting.

For AmberGS, RMSD increases and shows large fluctuations for the lower end (average RMSD=0.7 $\pm$ 0.1nm) while the centre region and top-end (average RMSD 0.6 $\pm$ 0.1nm) show slight fluctuations (see Figure 3.18). To further characterize this structural fluctuation, we calculated hydrogen bonds between the top two chains, central two chains, and the last two chains (donor-acceptor distance cut-off value assigned is 0.35 nm). Time series of hydrogen bonds is given in Figure 3.18D and it indicates as expected that the increase in the RMSD for the lower chain resulting in the opening of this end which is due to a smaller number of hydrogen bonds between the last two chains. While the central region showing stable RMSD value has a maximum number of hydrogen bonds between the two chains. RMSD for the top (1.0 $\pm$ 0.1nm) and bottom end (0.9 $\pm$  0.1nm) in Gromos53a6 is even larger than the other two force fields. The structure in OPLS-aa shows the least deviation from the initial energy minimized structure for all the three regions of the fibril (see Figure 3.20).



**Figure 3.18. (A). Ring fibril in AmberGS, a snapshot after 400ns (B) Secondary structure evolution of ring fibril indicating an increase in  $\alpha$ -helical structure (C) RMSD values for top, center, and bottom chains (D) Intra-sheet hydrogen bonds between representative chains of ring fibril.**

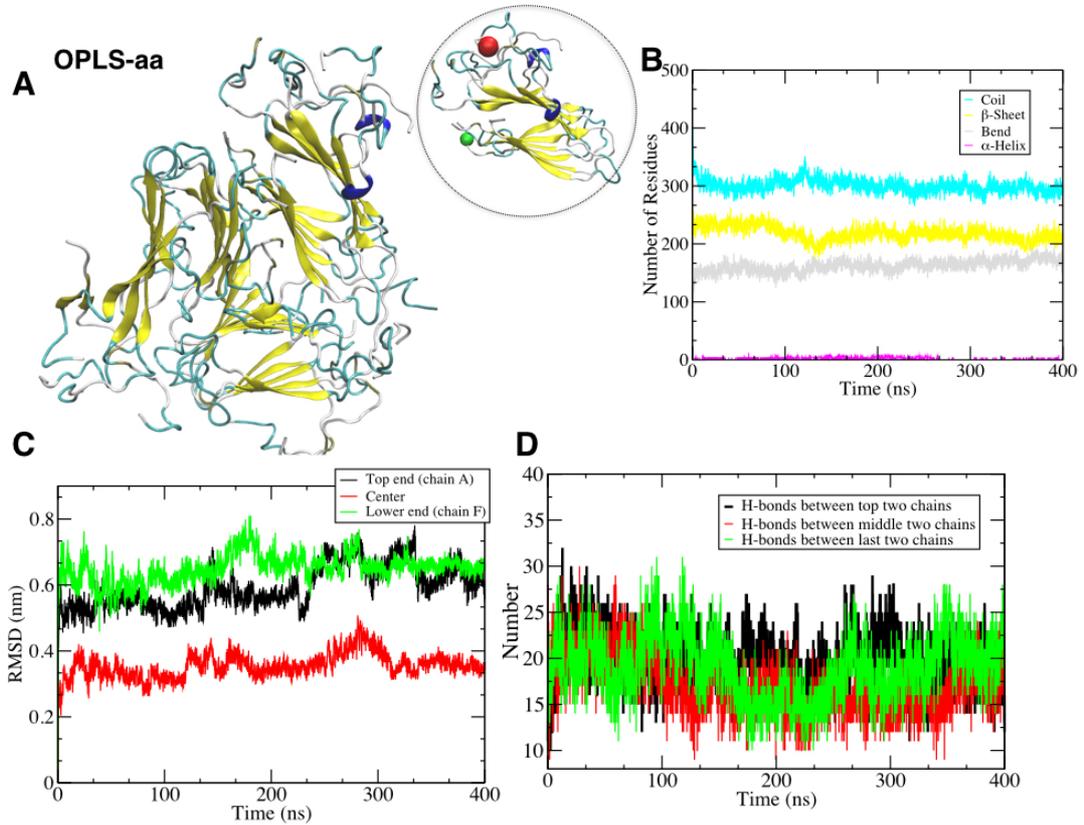
The secondary structure for each force field, is plotted in Figure 3.18B, Figure 3.19B, and Figure 3.20B for all the three force fields. For clarity, we have plotted only those secondary motives that undergo a significant change. AmberGS produces a remarkable difference in the secondary structure as compared to the other force fields in a way that there is a high percentage of helical content (17%) as compared to the structures in other force fields. Both Gromos53a6 and OPLS-aa show negligible helical content. This high propensity of helical structures is correlated to an increase in the RMSD values for ring fibril under the influence of AmberGS. The increase in the number of residues adopting helical conformation is clearly visible in Figure 3.18B. For Gromos53a6, the major secondary structure is  $\beta$ -sheet (46%) of the residues adopt it (28% for AmberGS) and it is stabilized throughout 400ns. OPLS-aa gives a mixture of coils (bend) and  $\beta$ -sheet regions (40%, 21% and 29% respectively).



**Figure 3.19. (A). Ring fibril in Gromos53a6, a snapshot after 400ns (B) Secondary structure evolution of ring fibril indicating stable  $\beta$ -sheet and coil structure (C) RMSD values for top, center and bottom chains (D) Intra-sheet hydrogen bonds between representative chains of ring fibril.**

To investigate this large RMSD fluctuation at different regions of the ring fibril, we calculated side chain intrasheet contacts i-e the  $C\alpha$ - $C\alpha$  distance between, PHE19 and GLY38, ALA21 and VAL36 and ASP23-LEU34. If the  $C\alpha$ - $C\alpha$  distance Ala21 and Val36 exceeds 2.0nm, we considered it as an open form of the fibril(Okumura and Itoh, 2016) and follow the same criteria for other residue pairs. Looking at the values of  $C\alpha$ - $C\alpha$  distances for the chains D, E, and F in table 7 in SI. It is clear that the  $C\alpha$ - $C\alpha$  distance between all the above-mentioned residue pairs noticeably increases at the lower end in all the force fields indicating that the  $\beta$ -sheets are well separated at the lower section of the fibril giving rise to an open end. While these pairs are close at the top end making the  $\beta$ -sheets at the top closer and forming a relatively closed end. We also calculated the angle  $C\alpha$ - $C\alpha$ - $C\alpha$  which indicates the stability of the loop, i-e the kink region connecting the two  $\beta$ -sheets. This angle is around 68 $^\circ$ (Okumura and Itoh, 2016) according to the NMR conformation but neither of the force field was able to reproduce this angle. This difference of the behavior at two ends of fibril is primarily due to the difference in the arrangement of the atoms of residues at two  $\beta$ -sheets (even and odd

ends). In the lower end,  $\beta 2$  forms weaker hydrogen bonds while  $\beta 1$  forms stronger hydrogen bonds that keep it closer to the other  $\beta$ -strands. This variation in the conformation at both ends can explain the unidirectional extension of the amyloid fibrils.



**Figure 3.20. (A). Ring fibril in OPLS-aa, a snapshot after 400ns (B) Secondary structure evolution of ring fibril indicating stable  $\beta$ -sheet, bend, and coil structure (C) RMSD values for top, center and bottom chains (D) Intra-sheet hydrogen bonds between representative chains of ring fibril. Secondary structure evolution (C) RMSD values for top, center, and bottom chains (D) Intra-sheet hydrogen bonds.**

### 3.5 Discussion

Molecular dynamics simulations are now a fundamental part of chemical research and increasingly essential to understanding biomolecular systems. The accuracy and hence confidence in these simulations depends on the force field parameters. Whilst the force field parameters for biomolecular assemblies such as lipid membranes are well developed, the quality and effectiveness of protein force fields for peptide and protein assemblies have largely remained investigated. This is understandable given that we do not have ample structural data that could form the basis for parameter optimization.

Here we have carried out a systematic comparison of three widely used protein force fields in MD simulations to reproduce an increasingly complex set of A $\beta$  assemblies. These simulations test the subtle and critical interplay between intra-molecular and inter-molecular interactions including hydrophobic and electrostatic interactions in stabilizing the fibrillar structures. A particularly challenging test of the parameters is the twisting observed in the larger structures. In making our inferences we also need to be mindful that nanocrystal structure instability or disordering of surface sub-structures does not necessarily imply a failure of the force field, as nucleation theory considerations need to be taken into account.

The comparison between the force fields is achieved through the use of parameters such as RMSD, RMSF, hydrogen bonding, secondary structures, inter, and intrasheet contacts. Generally, the force fields yield results that are in good agreement with experiments at a coarse level. Yet, significant differences in the behaviors of peptide systems at a finer level of detail are observed as a result of size and complexity. The convergence of MD simulations is evaluated by tracking RMSD as a function of time. These results assure the quality of the sampling and indicate if the length of the trajectories is good enough.

As shown in the results, the effects of force fields on the structures of A $\beta$  peptides are complicated, depending on the size and complexity of systems under investigation. Nevertheless, based on these results and the supporting data from previous experiments and simulations, we rationalize that for studying amyloid formation, AmberGS has strong biases toward  $\alpha$ -helical structures. Gromos53a6 that is generally employed for studying amyloid formation, favors  $\beta$ -sheet and may lead to incorrect kinetics when converting from  $\alpha$ -helix to  $\beta$ -sheet which is a crucial step of amyloid aggregation. OPLS might provide a good balance in structure as well as in kinetics and, therefore, is suitable for simulations of biomolecules. These force field effects are discussed here based on two main aspects: structure and dynamics of A $\beta$ .

The effects of force fields on the parameters associated with the structural integrity such as R<sub>g</sub> revealed little differences between the force fields, suggesting that the range of conformations sampled by each force field is, roughly, the same. Examination of C $\alpha$  - backbone stability gives some idea about the reproducibility of the experimental structure. For the monomer A $\beta$ , the peptide is undergoing a large conformational change which results in a larger RMSD value. Especially for Gromos53a6 where the initial  $\alpha$ -

helical structure is converting to a pair of antiparallel  $\beta$ -strands. For the large fibrillar structures, each of the force fields tested give comparable results in most of the cases. Generally, for all the model peptides, OPLS-aa produce stable structures with less deviations from the experimental structures.

Now talking about the fibrillar  $A\beta$ , we have observed in our simulations that no matter which force field is employed, there are structural and fluctuational differences between the ends and core of the fibrils. In many instances, even the top and bottom chains of the fibrils show significantly different behaviour that is linked with the unidirectional elongation of the fibrils (Han and Schulten, 2014). The ends of the fibril form an interface between the fibril and solution and may generally adopt a different molecular structure and character from the peptides in the bulk region. The two ends of the fibrils (even and odd based on the direction of hydrogen and oxygen atoms) expose different sides to the solvent. For example in U-shaped protofibril comprising two  $\beta$ -sheets ( $\beta 1$  and  $\beta 2$ ), the  $\beta 1$  exposed at the even end makes stronger hydrogen bonds and favors  $\beta$ -sheet formation while the  $\beta 2$  exposed at the odd end forms weak hydrogen bonds. That is the reason that one end can easily interact with the incoming monomer forming hydrogen bonds and leads to unidirectional fibril growth. Generally, it's the lower end of the  $A\beta$ , that fluctuates more adopting mostly open and sometimes a close form. While the upper end adopts a close form. The conformational flexibility of the lower chain (that is actually the odd end of the fibril) is due to the position of the  $\beta 2$  sheet (Okumura and Itoh, 2016). Our findings well agree with the results of recent high-speed AFM experiments (Watanabe-Nakayama *et al.*, 2016). From the previous studies using AMBER parm99SB force field48, it was found that there is less fluctuation at the top end (even) than at the lower (odd) end. As a result, the even end forms much stronger  $\beta$ -sheet that is favourable for any incoming peptide at this end (Kahler, Sticht and Horn, 2013). Another study predicted that the extension speed is 40 times larger for this end as a result of stable hydrophobic interactions (Han and Schulten, 2014). But there are a few studies that report much smaller difference in the behaviour of the two fibrillar ends as in reference (Schwierz *et al.*, 2016). We see in our simulations that for ring fibrillar structures in AmberGS, the lower end shows large fluctuations (average RMSD=0.7 $\pm$ 0.1nm) as a result of smaller number of hydrogen bonds between the last two chains of the fibril while the centre region and top end show slight fluctuations.  $C\alpha$ - $C\alpha$  distance between PHE19 and GLY38, ALA21 and VAL36 and ASP23-LEU34

noticeably increases (longer than 2nm) at the lower end in all the force fields indicating that the  $\beta$ -sheets are well separated at the lower section of the fibril giving rise to an open end. Similarly, for the S and U-shaped fibrils, in AmberGS and OPLS-aa, the lower chains show large rmsd indicating big structural changes at the ends while the centre shows minor deviation from the NMR structure irrespective of the instability of the ends.

Fibril twisting is fundamental to the stability of the amyloid fibrils. Generally, we observed the presence of a gradual left-handed twist of  $\beta$ -strands of the fibril that is more prominent in AmberGS and OPLS-aa (Schmidt *et al.*, 2009, 2015; Gremer *et al.*, 2017). The values of the twist angles for the S-shaped fibril in AmberGS ( $-98^\circ$ ), Gromos53a6 (ca.  $-78^\circ$ ) and OPLS-aa ( $-104^\circ$ ) show that deviations from the starting structure results from the twisting of the  $\beta$ -sheets as described in literature (Kahler, Sticht and Horn, 2013). Additionally, reduced intermolecular LYS28-ALA42 salt bridge interactions for the lower chains in these structures in all three force fields are indicative of such structural changes (Table 9.8 in SI). For the U-shaped protofibril the structure under all the force fields adopts a left-handed twist with an average value of  $-10.73^\circ$ ,  $-3.24^\circ$  and  $-9.69^\circ$  for AmberGS, Gromos53a6 and OPLS-aa respectively. Here, the magnitude of the twisting is not that high as in longer fibrils and the interchain salt bridges are well preserved except for Gromos53a6. This observation suggests that twisting of the fibrils is size dependent and the twist angles increase rapidly with length of the fibril as seen in  $A\beta$  Protofilament pair twisting in Amber force field (Kahler, Sticht and Horn, 2013). These observations also indicate that twisting of the fibrils is favoured by AmberGS and OPLS-aa force fields.

In the context of the secondary structure of the monomer, the structural data suggest that  $A\beta$  is highly  $\alpha$ -helical in the apolar environment, is largely coil form in aqueous solution (Zhang *et al.*, 2000), and forms  $\beta$ -sheets in fibrillar form (Chen *et al.*, 2017). Whatever the conformational state may be at the beginning of the assembly process, the equilibrium, conformation of  $A\beta$  monomer is predominately  $\beta$ -sheet. CD spectroscopy presents  $A\beta$ 's solution structure in water consisting of 1.1%  $\alpha$ -helix, 60.4%  $\beta$ -strand, 31.4% random coil, and 7.1%  $\beta$ -turn. Here, we found that none of the force fields could exactly produce these values and the coil structure is the most dominant conformation in both Gromos53a6 (35%) and OPLS-aa (33%) while  $\alpha$ -helix varies with the force field. The equilibrium structures are essentially compact for all the three force fields

with a radius of gyration ranging from 0.9-1.2nm (AmberGS force field that predicted the most elongated structure out of three;  $R_g=1.2\text{nm}$ ). This is consistent with the experimental hydrodynamic radius of 0.9 nm determined by single-molecule fluorescence spectroscopy (Nag *et al.*, 2011). This result points out that Gromos53a6 and OPLS-AA produce more compact structures that are in accordance with the previous study that indicates that in water  $A\beta$  is collapsed into a compact structure that is a mixture of loops, strands, and turns (Zhang *et al.*, 2000). These compact structures are believed to serve as a nucleus for further aggregation in-vivo. In addition, the  $\alpha$ -helical content is largest for AmberGS (73%) while  $\beta$ -sheet (14%) is maximum for Gromos53a6. This is comparable to experimental CD spectra of WT  $A\beta_{42}$  that contains 12%  $\beta$ -sheet and 3%  $\alpha$ -helix. Though the  $\beta$ -sheet content of  $A\beta$  in the simulations is well below the experimentally observed value but our simulation shows the formation of antiparallel  $\beta$ -sheet, which is one of the fundamental events underlying amyloid diseases. Furthermore it is also consistent with the experimental NMR chemical-shift-index (CSI) analysis in that  $\beta$ -sheet is present at the C-terminal region at residues LEU34-VAL36 and VAL39-ILE41 (Chong, Yim and Ham, 2013). The presence of a high coil, bend and turn population in the residues ILE31-LEU34 and VAL36-VAL39 in simulated  $A\beta_{42}$  is also in accord with the experimental NOE intensities, which suggest disordered structures in the hydrophilic regions. Such agreements with the experimentally observed NMR J-coupling constants, CD spectra, NMR CSI and NOE intensity analyses validate our simulation procedures for the monomer  $A\beta_{42}$ , which were also employed for simulating other amyloid models (Hou *et al.*, 2004).

For the larger more complex  $A\beta$  fibrils, a number of studies have shown that the prominent secondary structure is  $\beta$ -sheet. So, the question is, how well a certain force field is preserving the the  $\beta$ -sheet conformation of the peptides. Our simulations show that significant and predictable deviations are observed in terms of the percentage of the residues adopting  $\beta$ -sheets. For the U-shaped protofibril, in Gromos53a6, 68% of the residues adopt  $\beta$ -sheet conformation (47% for experimental solution NMR structure). Only OPLS-aa produced  $\beta$ -sheet content closer to the experimental value (41%). For the larger fibrils S-shaped fibril, again Gromos53a6 overestimates  $\beta$ -sheet content by a whopping 30% while OPLS-aa stabilizes the originally present  $\beta$ -sheets quite well. For the complex 3-fold symmetry structure, the disordered region present in the structure at the N-terminal is converted into  $\alpha$ -helix in AmberGS with the subsequent loss of coil

content while Gromos53a6 and OPLS-aa do not produce any significant  $\alpha$ -helix. This is in accordance with a number of studies showing that Amber favours helical structures over extended  $\beta$ -strand conformations, Gromos53a6 may overestimate  $\beta$ -sheet conformations, while OPLS produces a balance between  $\alpha$  and extended- $\beta$  conformations (Gerben *et al.*, 2014). These results are coherent throughout the simulations and also rationalize the other experimental and computational observations. The biases in the force fields towards a specific secondary structure motif comes from inadequate backbone dihedral term parametrization. Dihedral terms in the original force fields especially Amber are continuously being reparametrized to improve the agreement with high level QM data but only partially succeeded. Therefore, a more systematic revision is clearly needed.

### **3.6 Conclusion:**

To obtain a combined picture of the effects of force fields, we conclude that force fields for a particular system may give different answers that may not be appropriate for a given problem. Depending on the force field and the choice of parameters during the simulation, it may work very well or very poorly. In terms of biomolecular simulations in general and amyloids in specific, it is important to recognize the actual problem, expected behaviour and the properties being tested and then making a judgment in selecting a suitable force field. The next and the most important question is of-course the kind of accuracy a user is looking for? Is it qualitative, quantitative or a general behaviour or a trend? So in totality, one can argue that since A $\beta$  is known to be largely composed of  $\beta$ -strand and random coil elements, it appears that the Gromos53a6 and OPLS-aa may be a better choice for modeling A $\beta$  which is also supported by other studies on shorter A $\beta$  fragments (Smith *et al.*, 2015). Since  $\alpha$ -helix- $\beta$ -sheet conversion forms the basis of A $\beta$  oligomerization and is a crucial step of A $\beta$  aggregation, to study such a process, Gromos53a6 seems a better choice as the conversion of  $\alpha$ -helix- $\beta$ -sheet is achieved in Gromos53a6 in the available timescale. While for larger system such as A $\beta$  fibrils, for which secondary structure is already known to be stable  $\beta$ -sheets or consist of folded states, OPLS may be slightly preferred. The observed differences in the behaviours of force fields highlight the importance of continuous force field refinement. These findings may have implications for simulations of other IDPs (Intrinsically disordered proteins) of varying sizes and complexities.

### 3.7 References

- Abraham, M. J. *et al.* (2015) ‘Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers’, *SoftwareX*. doi: 10.1016/j.softx.2015.06.001.
- ‘Alzheimer’s disease facts and figures’ (2019) *Alzheimer’s & Dementia*, 3(15), pp. 321–387. doi: 10.1016/j.jalz.2019.01.010.
- Anwar, J. and Zahn, D. (2011) ‘Uncovering molecular processes in crystal nucleation and growth by using molecular simulation’, *Angewandte Chemie - International Edition*. doi: 10.1002/anie.201000463.
- Barale, S. S. *et al.* (2019) ‘Molecular Insights into Destabilization of Alzheimer’s A $\beta$  Protofibril by Arginine Containing Short Peptides: A Molecular Modeling Approach’, *ACS Omega*. doi: 10.1021/acsomega.8b02672.
- Barz, B. and Urbanc, B. (2012) ‘Dimer Formation Enhances Structural Differences between Amyloid  $\beta$ -Protein (1–40) and (1–42): An Explicit-Solvent Molecular Dynamics Study’, *PLoS ONE*, 7(4), p. e34345. doi: 10.1371/journal.pone.0034345.
- Baschek, J. E., R Klein, H. C. and Schwarz, U. S. (2012) ‘Stochastic dynamics of virus capsid formation: direct versus hierarchical self-assembly’, *BMC Biophysics*, 5(1), p. 22. doi: 10.1186/2046-1682-5-22.
- Baumketner, A. (2006) ‘Amyloid beta-protein monomer structure: A computational and experimental study’, *Protein Science*, 15(3), pp. 420–428. doi: 10.1110/ps.051762406.
- Brooks, B. R. *et al.* (1983) ‘CHARMM: A program for macromolecular energy, minimization, and dynamics calculations’, *Journal of Computational Chemistry*. doi: 10.1002/jcc.540040211.
- Bussi, G., Donadio, D. and Parrinello, M. (2007) ‘Canonical sampling through velocity rescaling’, *The Journal of Chemical Physics*. American Institute of Physics, 126(1), p. 014101. doi: 10.1063/1.2408420.
- Chen, G. F. *et al.* (2017) ‘Amyloid beta: Structure, biology and structure-based therapeutic development’, *Acta Pharmacologica Sinica*. doi: 10.1038/aps.2017.28.
- Chong, S. H., Yim, J. and Ham, S. (2013) ‘Structural heterogeneity in familial Alzheimer’s disease mutants of amyloid-beta peptides’, *Molecular BioSystems*. doi:

10.1039/c2mb25457c.

Cino, E. A., Choy, W. Y. and Karttunen, M. (2012) 'Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations', *Journal of Chemical Theory and Computation*, 8(8), pp. 2725–2740. doi: 10.1021/ct300323g.

Cornell, W. D. *et al.* (1995) 'A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules', *Journal of the American Chemical Society*. American Chemical Society, 117(19), pp. 5179–5197. doi: 10.1021/ja00124a002.

Crescenzi, O. *et al.* (2002) 'Solution structure of the Alzheimer amyloid beta-peptide (1-42) in an apolar microenvironment. Similarity with a virus fusion domain.', *EUR.J.BIOCHEM.*, 269, pp. 5642–5648. doi: 12423364.

Faeder, J. R., Blinov, M. L. and Hlavacek, W. S. (2009) 'Rule-based modeling of biochemical systems with BioNetGen', *Methods in Molecular Biology*, 500(1), pp. 113–167. doi: 10.1007/978-1-59745-525-1\_5.

Fluitt, A. M. and De Pablo, J. J. (2015) 'An Analysis of Biomolecular Force Fields for Simulations of Polyglutamine in Solution', *Biophysical Journal*. Biophysical Society, 109(5), pp. 1009–1018. doi: 10.1016/j.bpj.2015.07.018.

Frazier, J. M., Chushak, Y. and Foy, B. (2009) 'Stochastic simulation and analysis of biomolecular reaction networks', *BMC Systems Biology*, 3. doi: 10.1186/1752-0509-3-64.

Frishman, D and Argos, P. (1995) 'Knowledge-based protein secondary structure assignment.', *Proteins*, 23(4), pp. 566–79. doi: 10.1002/prot.340230412.

Frishman, Dmitrij and Argos, P. (1995) 'Knowledge-based protein secondary structure assignment', *Proteins: Structure, Function, and Bioinformatics*. doi: 10.1002/prot.340230412.

García, A. E. and Sanbonmatsu, K. Y. (2002) ' $\alpha$ -Helical stabilization by side chain shielding of backbone hydrogen bonds', *Proceedings of the National Academy of Sciences*, 99(5), pp. 2782–2787. doi: 10.1073/pnas.042496899.

Gerben, S. R. *et al.* (2014) 'Comparing atomistic molecular mechanics force fields for a difficult target: a case study on the Alzheimer's amyloid  $\beta$ -peptide.', *Journal of*

*biomolecular structure & dynamics*, 32(11), pp. 1817–32. doi: 10.1080/07391102.2013.838518.

Gremer, L. *et al.* (2017) ‘Fibril structure of amyloid- $\beta$ (1–42) by cryo–electron microscopy’, *Science*. doi: 10.1126/science.aao2825.

Gruenert, G. *et al.* (2010) ‘Rule-based spatial modeling with diffusing, geometrically constrained molecules’, *BMC Bioinformatics*, 11. doi: 10.1186/1471-2105-11-307.

Han, W. and Schulten, K. (2014) ‘Fibril elongation by A $\beta$ 17-42: Kinetic network analysis of hybrid-resolution molecular dynamics simulations’, *Journal of the American Chemical Society*. doi: 10.1021/ja507002p.

Hardy, J. and Selkoe, D. J. (2002) ‘The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics.’, *Science (New York, N.Y.)*, 297(2002), pp. 353–356. doi: 10.1126/science.1072994.

Hayden, E. Y. and Teplow, D. B. (2013) ‘Amyloid  $\beta$ -protein oligomers and Alzheimer’s disease.’, *Alzheimer’s research & therapy*, 5(6), p. 60. doi: 10.1186/alzrt226.

Hess, B. *et al.* (1997) ‘LINCS: A Linear Constraint Solver for molecular simulations’, *Journal of Computational Chemistry*. doi: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

Hou, L. *et al.* (2004) ‘Solution NMR Studies of the A $\beta$ (1-40) and A $\beta$ (1-42) Peptides Establish that the Met35 Oxidation State Affects the Mechanism of Amyloid Formation’, *Journal of the American Chemical Society*. doi: 10.1021/ja036813f.

Jorgensen, W. L. and Tirado-Rives, J. (1988) ‘The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin’, *Journal of the American Chemical Society*. doi: 10.1021/ja00214a001.

Kahler, A., Sticht, H. and Horn, A. H. C. (2013) ‘Conformational Stability of Fibrillar Amyloid-Beta Oligomers via Protofilament Pair Formation - A Systematic Computational Study’, *PLoS ONE*. doi: 10.1371/journal.pone.0070521.

Karr, J. R. *et al.* (2012) ‘A whole-cell computational model predicts phenotype from genotype’, *Cell*, 150(2), pp. 389–401. doi: 10.1016/j.cell.2012.05.044.

Ke, P. C. *et al.* (2017) ‘Implications of peptide assemblies in amyloid diseases’, *Chemical Society Reviews*. doi: 10.1039/c7cs00372b.

- Lee, M., Yoon, J. and Shin, S. (2019) 'Computational Study on Structure and Aggregation Pathway of A $\beta$ 42 Amyloid Protofibril', *Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcc.9b07195.
- Luhrs, T. *et al.* (2005) '3D structure of Alzheimer's amyloid- $\beta$ (1-42) fibrils.', *Proc.Natl.Acad.Sci.Usa*, 102, pp. 17342–17347. doi: 16293696.
- Lührs, T. *et al.* (2005) '3D structure of Alzheimer's amyloid- $\beta$ (1-42) fibrils', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0506723102.
- Nag, S. *et al.* (2011) 'Nature of the amyloid- $\beta$  monomer and the monomer-oligomer equilibrium', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M110.199885.
- Nguyen, P. H., Li, M. S. and Derreumaux, P. (2011) 'Effects of all-atom force fields on amyloid oligomerization: Replica exchange molecular dynamics simulations of the A $\beta$ 16-22 dimer and trimer', *Physical Chemistry Chemical Physics*, 13(20), pp. 9778–9788. doi: 10.1039/c1cp20323a.
- Okumura, H. and Itoh, S. G. (2016) 'Structural and fluctuational difference between two ends of A $\beta$  amyloid fibril: MD simulations predict only one end has open conformations', *Scientific Reports*. doi: 10.1038/srep38422.
- Paravastu, A. K. *et al.* (2008) 'Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0806270105.
- Parrinello, M. and Rahman, A. (1981) 'Polymorphic transitions in single crystals: A new molecular dynamics method', *Journal of Applied Physics*. doi: 10.1063/1.328693.
- Petrov, D. and Zagrovic, B. (2014) 'Are Current Atomistic Force Fields Accurate Enough to Study Proteins in Crowded Environments?', *PLoS Computational Biology*, 10(5), p. e1003638. doi: 10.1371/journal.pcbi.1003638.
- Price, D. J. and Brooks, C. L. (2002) 'Modern protein force fields behave comparably in molecular dynamics simulations', *Journal of Computational Chemistry*. doi: 10.1002/jcc.10083.
- Schmidt, M. *et al.* (2009) 'Comparison of Alzheimer A $\beta$ (1-40) and A $\beta$ (1-42) amyloid fibrils reveals similar protofilament structures', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0905007106.

- Schmidt, M. *et al.* (2015) 'Peptide dimer structure in an A $\beta$ (1–42) fibril visualized with cryo-EM', *Proceedings of the National Academy of Sciences*, 112(38), pp. 11858–11863. doi: 10.1073/pnas.1503455112.
- Schwierz, N. *et al.* (2016) 'Dynamics of Seeded A $\beta$ 40-Fibril Growth from Atomistic Molecular Dynamics Simulations: Kinetic Trapping and Reduced Water Mobility in the Locking Step', *Journal of the American Chemical Society*. doi: 10.1021/jacs.5b08717.
- Scott, W. R. P. *et al.* (1999) 'The GROMOS biomolecular simulation program package', *Journal of Physical Chemistry A*. doi: 10.1021/jp984217f.
- Shao, H. *et al.* (1999) 'Solution structures of micelle-bound amyloid  $\beta$ -(1-40) and  $\beta$ -(1-42) peptides of Alzheimer's disease', *Journal of Molecular Biology*. doi: 10.1006/jmbi.1998.2348.
- Siwy, C. M., Lockhart, C. and Klimov, D. K. (2017) 'Is the Conformational Ensemble of Alzheimer's A $\beta$ 10-40 Peptide Force Field Dependent?', *PLoS Computational Biology*. doi: 10.1371/journal.pcbi.1005314.
- Smith, M. D. *et al.* (2015) 'Force-Field Induced Bias in the Structure of A $\beta$ 21-30: A Comparison of OPLS, AMBER, CHARMM, and GROMOS Force Fields.', *Journal of chemical information and modeling*, 55(12), pp. 2587–95. doi: 10.1021/acs.jcim.5b00308.
- Thomas, M. and Schwartz, R. (2017) 'Quantitative computational models of molecular self-assembly in systems biology', *Physical Biology*, 14(3). doi: 10.1088/1478-3975/aa6cdc.
- Tomaselli, S. *et al.* (2006) 'The  $\alpha$ -to- $\beta$  conformational transition of Alzheimer's A $\beta$ -(1-42) peptide in aqueous media is reversible: A step by step conformational analysis suggests the location of  $\beta$  conformation seeding', *ChemBioChem*. doi: 10.1002/cbic.200500223.
- Vivekanandan, S. *et al.* (2011) 'A partially folded structure of amyloid-beta(1-40) in an aqueous environment', *Biochemical and Biophysical Research Communications*. doi: 10.1016/j.bbrc.2011.06.133.
- Watanabe-Nakayama, T. *et al.* (2016) 'High-speed atomic force microscopy reveals structural dynamics of amyloid  $\beta$ 1-42 aggregates', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1524807113.

Watts, C. R. *et al.* (2018) 'Effects of force fields on the conformational and dynamic properties of amyloid  $\beta$ (1-40) dimer explored by replica exchange molecular dynamics simulations', *Proteins: Structure, Function and Bioinformatics*, 86(3), pp. 279–300. doi: 10.1002/prot.25439.

Webb, B. and Sali, A. (2016) 'Comparative protein structure modeling using MODELLER', *Current Protocols in Bioinformatics*. doi: 10.1002/cpbi.3.

Weber, O. C. and Uversky, V. N. (2017) 'How accurate are your simulations? Effects of confined aqueous volume and AMBER FF99SB and CHARMM22/CMAP force field parameters on structural ensembles of intrinsically disordered proteins: Amyloid- $\beta$  42 in water', *Intrinsically Disordered Proteins*. Taylor & Francis, 5(1), p. e1377813. doi: 10.1080/21690707.2017.1377813.

Xiao, Y. *et al.* (2015) 'A $\beta$ (1-42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease.', *Nature structural & molecular biology*, 22(6), pp. 499–505. doi: 10.1038/nsmb.2991.

Yeh, I. C. and Hummer, G. (2002) 'Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent', *Journal of the American Chemical Society*. doi: 10.1021/ja025789n.

Zhang, S. *et al.* (2000) 'The Alzheimer's peptide A $\beta$  adopts a collapsed coil structure in water', *Journal of Structural Biology*. doi: 10.1006/jsbi.2000.4288.

Zheng, J. *et al.* (2007) 'Modeling the Alzheimer A $\beta$ 17-42 fibril architecture: Tight intermolecular sheet-sheet association and intramolecular hydrated cavities', *Biophysical Journal*. doi: 10.1529/biophysj.107.110700.

# 4 Molecular Simulations Reveal a Sweet Spot for Fibril Formation of Beta Amyloid (17-42)

**Copyright statement:** The contents of this chapter are currently in preparation for publication.

**Attribution:** Beenish Khurshid and Jamshed Anwar designed the research and wrote the paper. Beenish Khurshid conducted the simulations and performed the analysis.

## 4.1 Abstract

In this study, a hierarchical computational approach (all-atom to coarse grain to enhanced sampling) is introduced to study self-organizing structures of peptides as a function of concentration. Several sets of U-shaped single peptide of A $\beta$ <sub>42</sub> containing a pair of  $\beta$ -strands are considered as an example to illustrate the phenomenon of self-assembly which is believed to be involved in various amyloid diseases such as Alzheimer's disease. We observed that there is a sweet spot of concentration for fibril growth where a unidirectional fibrillar structure is formed. Above and below that diffused and spherical aggregates are formed respectively. Analysis of the gyration radius of peptides, solvent accessible surface areas, and cluster formation as a function of time provides insights into the assembly mechanism of A $\beta$ <sub>42</sub>. The aggregation propensity and aspect ratio at that particular concentration indicates that the fibril growth is unidirectional, and elongation of the structure is preferred over lateral growth. This is indicative that at a longer time scale there may be a rearrangement in the fibrillar aggregates to produce ordered fibrils.

## 4.2 Introduction

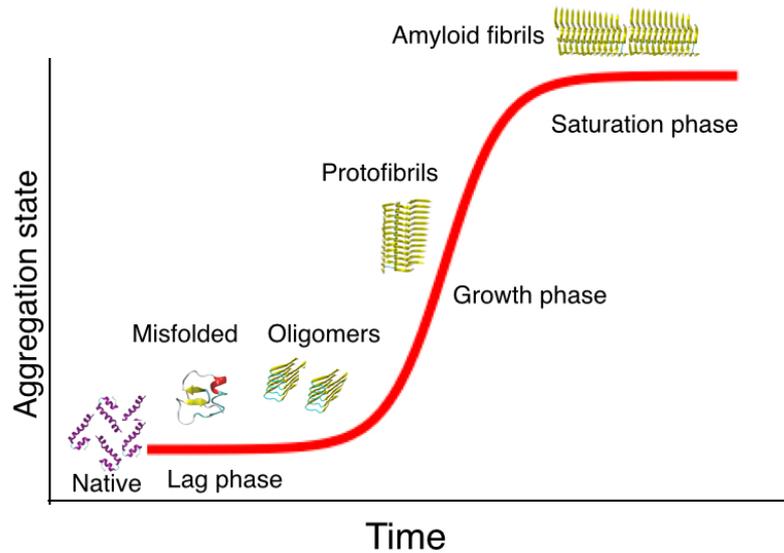
Amyloid aggregation/self-assembly is at the core of more than 25 diseases which include Alzheimer's, Parkinson's, type II diabetes, amyotrophic lateral sclerosis, and prion diseases (Chiti and Dobson, 2009). Amyloid aggregation is a spontaneous process in which small A $\beta$  peptides produced by the fragmentation of amyloid precursor protein (APP), self-assemble to form highly organized fibrillar aggregates. A $\beta$  is largely unstructured and highly flexible. It is present in the form of a collapsed coil in the aqueous solution. The lack of a structure is perhaps one of the reasons for the poorly understood function of A $\beta$  peptides. A better understanding of the structure and mechanism of its aggregation and the effect of various physiological conditions can develop a clear perspective of amyloid aggregation and its biological implications.

A large number of in vitro studies performed on amyloid aggregation clearly show that the kinetics of amyloid fibril formation is controlled by peptide concentration and the critical step of the formation of cross  $\beta$ -structures. The process involves the transition from a homogenous solution to the formation of an aggregate that co-exists with monomers in the solution phase. It follows a sigmoidal growth curve which comprises three stages: a lag phase, growth phase, and a plateau phase (Arosio, Knowles and Linse, 2015). The monomers form a  $\beta$ -sheet rich disordered oligomer during the lag phase which is thermodynamically unfavorable and thus considered a rate-limiting step. When the emerging nuclei reaches a critical size, more monomers join in and the elongation process starts that leads to the formation of ordered fibrils (Luiken and Bolhuis, 2015). According to the classical nucleation theory, small clusters (formed from monomers), are characterized by large, unfavorable interfacial energy between solute and water. The nucleation is governed by a balance between interfacial free energy and bulk-free energy. Initially when the size of the nucleus is smaller than the critical size, the nuclei is unstable because of the large interfacial energy. As the size of the cluster increases, the interaction energy between monomers in the bulk of the cluster increases decreasing the total free energy. As a result, the cluster/nuclei become stable and the probability of monomer addition increases further, leading to the growth phase. This process growth by nucleation is the key step in amyloidosis (see Figure 4.1) (Chiti and Dobson 2006). However, the nucleation in biomolecular systems may not necessarily follow classical nucleation due to the size and flexibility of proteins and the macromolecular crowding inside cells. Macromolecules including proteins, nucleic

acids, polysaccharides, and co-factors together take up around 40% of the cell volume. In-vitro experiments are conducted on solutions containing lower concentrations (less than about 1 mg/ml) of these crowders as compared to the real biochemical environment in the living systems (50–400 mg/ml) (Minton, 2001). This crowding is thought to be characterized by repulsive interactions between macromolecular species that lead to proteins adopting more compact conformations (as compared to the more expanded unfolded state) thus speeding up aggregation. In terms of intermolecular interactions, it means that crowding pushes the equilibrium towards the aggregated states that are more compact. As shown by simulations performed by (Latshaw, Cheon and Hall, 2014); that crowding can increase the aggregation rate by speeding up the formation of the fibril nucleus (the transition state) but the fibril elongation is reduced as a result of increased viscosity and excluded volume.

However, recent technological advancements especially with in-cell NMR (Li and Pielak, 2009) and fluorescence resonance energy transfer (FRET) probes (Groen *et al.*, 2015) it has become clear that the description based entirely on steric repulsion is not sufficient. Each protein is subject to multiple interactions through specific and nonspecific interactions such as hydrogen bonding, hydrophobic and electrostatic integrations, and Vander Waals. The combined effect of all these forces determines the stability of the structure formed and the process of aggregation (Sarkar, Li and Pielak, 2013). For example, repulsive electrostatic interactions between negatively charged surface patches of peptides can result in the formation of more expanded structures rather than compact states resulting in the destabilization of the structure and increase in the amount of unfolded protein in the cell. The strong dependence of protein's structure, conformation, and aggregation on the nature of the crowding agents hampers the effort to reproduce the environment that mimics the living cell. These studies are mandatory to completely understand amyloid aggregation but are highly challenging.

The physical chemistry perspective of the effects of crowding are pretty clear. Crowding implies high concentrations of the various molecules in the extracellular milieu. This in general raises the chemical potential (the affinity of a molecule for an environment) of any given molecular species. The chemical potential reflects the supersaturation and is the driving force for excluding the molecules out of the solution leading to nucleation and growth of the solid phase, in this case the A $\beta$  structures.



**Figure 4.1. A schematic diagram of nucleation-dependent self-assembly of A $\beta$**

Circular dichroism (CD), Fourier-transform infrared (FTIR) and nuclear magnetic resonance (NMR) spectroscopy are currently being used as powerful techniques to characterize structural transitions during amyloid aggregation. It has now become possible to explore the phenomenon with increase molecular understanding by the use of fluorescence methods, isotope-edited infrared Spectroscopy and quasi-elastic light scattering spectroscopy (Streets *et al.*, 2013). These methods are successful in providing useful information about the mechanisms of fibril formation, identification, and classification of peptide self-assemblies such as oligomers, proto-fibrils, amyloid fibrils and crucial parameters which are difficult to measure otherwise due to heterogeneity of the products in the system (Chiti and Dobson, 2006). The NMR diffusion experiments employing the pulse-field gradient show that A $\beta_{40}$  and A $\beta_{42}$  are present as random coils in aqueous solution, while in a lipid mimicking environment the A $\beta$  is present in helical form. The two helices present at residues 15–24 and 29–35 and the surrounding region are unstructured and flexible. Another NMR study shows that A $\beta_{40}$  adopts a  $\beta$ -hairpin structure when interacts with another binding protein which might represent the transient metastable structure needed for aggregation initiation (Hoyer *et al.*, 2008).

With the accessibility of high-end computing and improved parallelization, it has now become possible to study the aggregation of A $\beta$  peptides at a longer timescale by using explicit all-atom MD. A $\beta_{40}$  and A $\beta_{42}$  peptides prefer  $\beta$ -sheet conformation when present in the form of a dimer (the overall  $\beta$ -content is  $\approx 20\%$  while  $\alpha$ -helical content is below 10%). The conversion from  $\alpha$ -helix to  $\beta$ -sheet is a slow process and hence believed to be the rate-limiting lag phase of amyloid aggregation (Xu *et al.*, 2005). A $\beta_{42}$  has two extra amino acid residues at the C-terminus and due to higher solvent exposure at this site, it follows an aggregation pathway that is quite different from A $\beta_{40}$  (Yan and Wang, 2006). This is probably the reason behind its high cytotoxicity as compared to A $\beta_{40}$  despite it being present at a lower concentration in the CSF in AD. However, since amyloid aggregation usually happens at a much longer timescale and experiments show that the process can exceed hours, in-vitro to follow the full aggregation pathway from monomers to the fibril is beyond the current computational reach of all-atom simulations.

Atomistic models yield the required molecular resolution but because of the limited time- and length-scales that they can access, they can only probe the earliest stages of aggregation. A way forward is to utilize coarse-grained models that enable simulation of larger systems and for longer simulation times by compromise the resolution. The process of self-assembly of A $\beta$  as studied by using these coarse-grain methods show the presence of amorphous aggregates before ordered fibril formation. In the first stage, called the "burst" stage, highly mobile oligomers of varying sizes are formed which form a significant number of inter and intrachain contacts in the second step giving rise to compact but disordered structures. The presence of amorphous aggregates prior to the formation of ordered fibrils is seen in many simulation studies; for example simulations of several polyalanine peptides show their assembly into disordered oligomers and then amorphous aggregates prior to the formation of  $\beta$ -sheet rich structures (Nguyen and Hall, 2006). Coarse grain simulations rationalize the experimental observations that the end product of aggregation differs depending upon the experimental conditions such as temperature and concentration (Wu and Shea, 2011). Peptides with higher  $\beta$ -sheet propensities give rise to ordered  $\beta$ -sheet rich fibrils (on-pathway) while peptides with smaller  $\beta$ -sheet propensities can form amorphous aggregates (Bellesia and Shea, 2007, 2009). Similarly low temperatures favor amorphous aggregates as the peptide does not have sufficient energy to overcome the

kinetic barriers while at high temperatures longer and fewer  $\beta$ -sheet rich fibrils are formed (Nguyen and Hall, 2004). Auer and co-workers showed in their simulations that at higher concentrations, the formation of disordered oligomers is favored though their properties are temperature-dependent. At lower concentrations, the peptides directly convert to  $\beta$ -sheet structures without passing through the disordered oligomeric phase (Auer, Dobson and Vendruscolo, 2007).

In nature, there is a tendency for the system to be driven by thermodynamics but the organization that is formed in the biological systems is modulated by the kinetics of the process. Complex systems like proteins are characterized by large free energy landscapes that are separated by free energy barriers that are hard to escape with the available thermal energy and as a result the system is kinetically trapped. Because we wanted to explore the thermodynamics and the resultant organizational structure of the aggregate, if the system was not trapped, two approaches were considered, i.e. temperature replica exchange MD (T-REMD) and Hamiltonian replica exchange MD (H-REMD) (Sugita and Okamoto, 1999). REMD simulations have been successfully applied to many aggregation studies of  $A\beta$  dimer (Man, Nguyen, and Derreumaux, 2017). There is, however, an important limitation of temperature REMD that the range in temperature that can be covered becomes markedly reduced with an increase in system size. The configuration exchanges are based on a Monte Carlo energy criterion and as the system size increases the energy distributions decrease in their overlap, which needs to be compensated by reducing the temperature differences. This implies more replicas and a limited temperature range making the method computationally expensive and less effective. So, for larger system as ours (System-II) in explicit solvent, T-REMD is not a method of choice as the number of replicas needed to cover a given range of temperature is roughly equal to the square root of the number of degrees of freedom. And for such larger systems this can become computationally very expensive (Fukunishi, Watanabe and Takada, 2002). An alternative to this method is Hamiltonian REMD which is generally run at constant temperature and only Hamiltonians (energy functions) of the molecules of interest (the proteins) are modified but not of the rest of the system (Affentranger, Tavernelli and Di Iorio, 2006). HREMD studies carried out by Strodel and coworkers for  $A\beta_{42}$  show that the  $\beta$ -sheet content of  $A\beta$  dimer increases at lower pH, while the metal chelation can increase the exposed hydrophobic surface area, leading to an increase in aggregation and polymerization.

Current evidence shows that A $\beta$  plaques formation occurs years before the actual manifestation of the disease. Though the details about earlier events of AD are still not resolved, it is expected that in this process, at low concentrations, only monomers are present in solution. However, as the concentration exceeds a critical aggregation concentration (*cac*), larger aggregates are formed. The supersaturation drives this process and dictates the final size, shape, and morphology of the nuclei and the fibril (Yang *et al.*, 2006). Some of the deposits of A $\beta$  are diffused while others are dense core fibrillar forms based on their staining with A $\beta$  specific dyes such as Congo red and Thioflavin (Serrano-Pozo *et al.*, 2011). Diffused plaques are generally present in cognitively intact elderly people while dense core plaques are related to AD. It is still not known how long the entire process takes, but some researchers speculate that it may take years to occur in the body (C. Vickers *et al.*, 2015).

Here we explore the effects of concentration on the kinetics and the structure formed both by all-atom and coarse-grain methods. As it has been already established that it takes days in-vitro for the formation of mature fibrils, which is beyond the reach of computational resources that we have today. A way forward to access larger systems with longer timescales is to coarse grain a molecular model which is an approach we have taken here. Since, in this study MARTINI force field was used for the coarse-graining the amyloid structures, the secondary structure was conserved throughout the simulation, which does not allow the peptide structure to reorganize. After all of these outstanding issues, we can say that this study is a fair approximation of the phenomenon of self-assembly. We used H-REMD as an enhanced sampling method to investigate the structure and conformation of equilibrium structure as the timescales in the body are likely to be orders of magnitude slower.

### 4.3 Methodology

Three types of simulations were performed;

- (i) Atomistic simulations with various system sizes (peptide concentrations) to monitor A $\beta$  aggregation process in detail;
- (ii) Coarse-grained (CG)MD simulations using martini force fields with the same systems in two steps (for fast sampling). During the first step, we investigated the effect of concentration of the monomers on the shape of aggregate. This was

a screening step in distance restrained CG-MD was used. In the second step, we selected the concentration at which peptides formed 1-Dimensional fibril; A longer simulation of 2  $\mu$ s was run to look for any internal re-organization of the fibrillar structure into a more-ordered form. Another set of simulations was performed for unidirectional growing partially ordered fibril to observe the effect of temperature on the shape of the fibril;

- (iii) H-REMD simulations to see if the unidirectional structure obtained from CG is the equilibrium structure?

The monomers were initially modeled as a pair of  $\beta$ -strands and the coordinates were taken from the NMR structure of  $A\beta_{(17-42)}$  (PDB ID: 2BEG) (Lührs *et al.*, 2005a). This structure consists of parallel in register  $\beta$ -sheets that contains two  $\beta$ -strands (residues 18–26 and 31–42) connected by a turn region (residues 27–30).

The all-atom molecular dynamic (MD) simulation was performed using GROMACS package (version 5.1.1) (M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess and Team, 2015) with GROMOS96 53A6 force field and SPC/E water (Berendsen, Grigera and Straatsma, 1987). All of the systems at various peptide concentrations (number of peptides 20, 40, 60, 80, and 100) were solvated in a cubic box where the distance between solute and the edges of the box was 1nm. The systems were electro-neutralized by adding counter ions and replacing water molecules. Before starting the dynamics, the systems were energy minimized using steepest descent algorithm to remove any bad contacts. Equilibration was done in two steps using NVT and NPT. The NVT step consisted of 100 ps simulation with peptides constrained. NPT equilibration was carried out for 400 ps using Parrinello-Rahman barostat to maintain a constant pressure (1 atm) (M Parrinello and Rahman, 1981). Electrostatic interactions were calculated using the Particle Mesh Ewald summation method (PME) with 0.16 grid spacing (Darden, York and Pedersen, 1993). The cut-off radius for both electrostatic and van der Waals interactions was set to 1.2nm. LINCS method was used to constraint the bond lengths (Hess *et al.*, 1997). All the systems were simulated for 200ns with a time-step of 2fs. The structural analysis was performed using VMD (Humphrey, Dalke, and Schulten, 1996) and GROMACS analysis tools.

For the CG representation, elastic network ElnDyn, was applied to the  $A\beta$  structure, which generates harmonic bonds between the backbone beads and preserves the overall  $\beta$ -sheet structure (Periole *et al.*, 2009). The upper and lower elastic bond cut-offs were

set to 0.9nm and 0.5nm respectively and the elastic bond force constant was set to 500 kJmol<sup>-1</sup>.

All the coarse grain simulations were performed with the GROMACS package (version 5.1.1) (M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess and Team, 2015). The aggregation process was studied by randomly placing varying numbers of A $\beta$  monomers at six different concentrations of peptides (number of peptides 20, 40, 60, 80, and 100) with a distance between peptides and the box edges of at least 1.5nm. The system was solvated with pre-equilibrated CG water beads for coarse-grain systems and neutralized by adding counter-ions and replacing water molecules. Energy minimization of each system was done by using steepest descent minimization. The solvent and ions were equilibrated around the protein for 500ps in two steps i.e. NVT and NPT using Particle mesh Ewald (PME) algorithm to calculate electrostatic interactions. V-rescale algorithm was used for temperature coupling. Parrinello-Rahman algorithm was used for pressure coupling with isotropic pcouplettype. For electrostatic and van der Waals interactions, we used a straight cut off (1.2 nm). The detail of all the systems used in this study is described in **Error! Not a valid bookmark self-reference..**

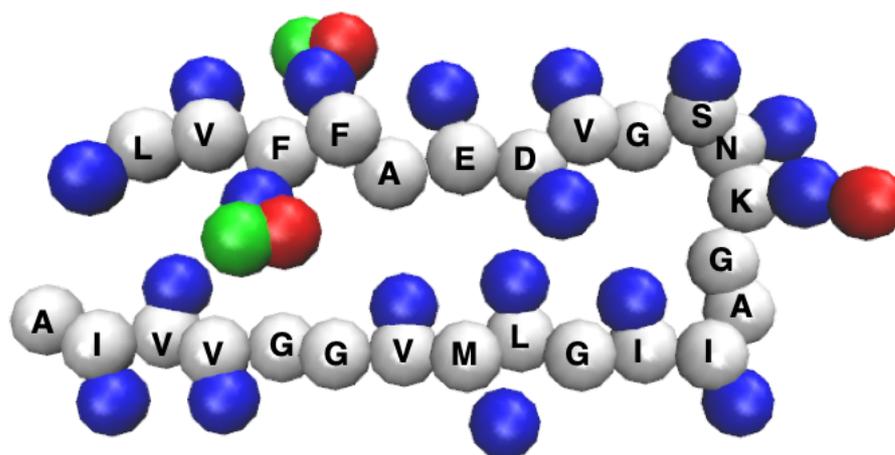
The length of coarse grain simulation was 400ns for each run except for System-II that was run for 2 $\mu$ s. The simulation temperature was set on 360K for the screening step. While for the second step of coarse grain study i-e the effect of temperature on the shape of the aggregate, three different temperatures were used i-e 300K, 400K and 500K. Analysis was performed by calculating Rg and SASA at various concentrations. Cluster analysis was done for quantitative characterize of the observed assembly process, two A $\beta$  monomers belong to the same cluster if the distance between any beads of the two monomers is less than 0.7 nm (Xu, Chen and Wang, 2014). The aspect ratios were calculated to qualify for the formation of unidirectional growth of amyloid fibril.

The replica exchange simulations were performed with the stochastic dynamics (SD) that do not need an external thermostat (Bussi, 2014). The Parrinello–Rahman barostat was used for pressure coupling in the NPT ensemble with a time step of 20 fs. The nonbonded peptide-peptide interactions scaled in each replica using the soft-core scaling method with  $\alpha = 0.5$ , softcore power  $p = 1.0$ , and interaction radius of  $\sigma = 0.3$  when van der Waals parameters were zero.  $\lambda = 0$  represents that all the interactions are switched on and  $\lambda = 1$  where all the interactions including van der Waals and columbic

are switched off. The simulations run in parallel and non-interacting with  $\Delta\lambda = 0.02$ . Hamiltonian Replica exchanges were attempted every 10, 500, and 1000 steps for three test runs. The starting coordinates of the peptides were generated by inserting them randomly in the box, running energy minimization and equilibration using the full force field. The results of HREMD were analyzed by using bar and energy algorithms implemented in Gromacs.

**Table 4.1: Simulation conditions showing Number of amyloids, water molecules, counter ions (NA), simulation box size, temperatures and simulation length**

Screening Step (1)	No. of amyloids	Counter ions	Box size (nm) <sup>3</sup>	Temp (K)	Simulation length (ns)
1	20	20	16 16 16	360	400
2	40	40	16 16 16	360	400
3	60	60	16 16 16	360	400
4	80	80	16 16 16	360	400
5	100	100	16 16 16	360	400



**Figure 4.2. The initial CG structure of A $\beta$  showing backbone beads in grey and side chains in blue.**

## 4.4 Results

Secondary structure and topology of spontaneously aggregated A $\beta$  peptides at various concentrations – all-atom MD

The self-assembly events were monitored for A $\beta$  peptides at five different concentrations i-e 20, 40, 60, 80, and 100 peptides. Representative snapshots from simulations of all the systems IA, IIA, IIIA, IVA, and VA are shown in Figure 4.3. These snapshots indicate spontaneous self-assembly of monomers initially in  $\alpha$ -helical conformation into  $\beta$ -sheet rich aggregates in 200ns long simulations minimizing their solvent accessible surface area (SASA) see Figure 4.4. SASA analysis shows a rapid initial decrease that reflects a very fast assembly process. The secondary structure content analysis shows that the self-assembly is accompanied by a conversion from  $\alpha$ -helix to  $\beta$ -sheets (see Table 4.2). For all the systems, though  $\beta$ -sheet rich aggregates were formed but not ordered and topologically quite diverse (see Figure 4.3). We observed a good amount of local order of  $\beta$ -sheets particularly at lower concentrations i-e system-I (20 peptides) and system-II (40 peptides). While at higher concentrations  $\beta$ -sheet rich conformations were formed that were completely disordered. Since time scales accessible to atomistic simulations do not allow us to follow the complete self-assembly process, we carried out CG simulations with the same system concentrations.

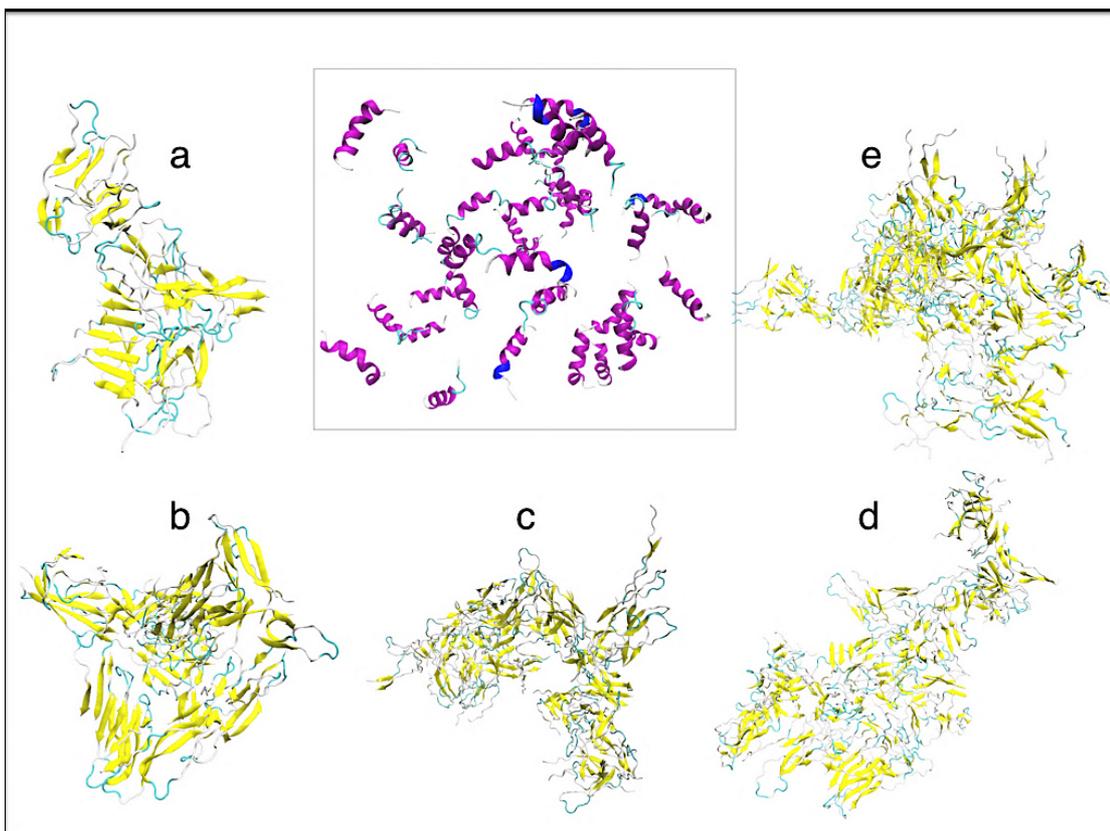


Figure 4.3. Atomistic simulations carried out at five different concentrations of A $\beta$  revealed various morphologies of A $\beta$  showing some local order but were not able to give rise to unidirectional A $\beta$  fibrils; Total number of peptides are (a) 20, (b) 40, (c) 60, (d) 80 and (e) 100

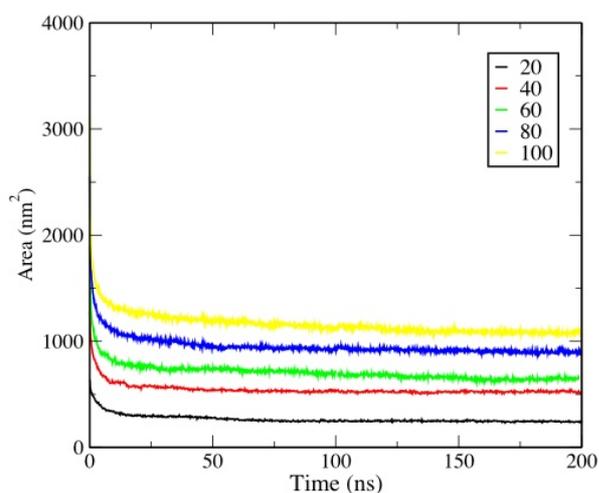


Figure 4.4. Solvent accessible surface area (SASA in nm<sup>2</sup>) vs time (ns) shows there is a drastic decrease in SASA in the first few nanoseconds. After which it which remains converged to an equilibrium till the end of the simulation without any further decrease. This implies a very fast aggregation of A $\beta$  peptides that gives rise to  $\beta$ -sheet rich aggregates that remain trapped in their conformation throughout the simulation.

**Table 4.2. The secondary structure content of various A $\beta$  systems**

Systems (No. of peptides)	Coil	$\beta$ -sheet	$\beta$ -bridge	Bend	Turn
I (20)	0.55	0.26	0.05	0.10	0.01
II (40)	0.49	0.34	0.03	0.09	0.01
III(60)	0.61	0.20	0.04	0.10	0.01
IV(80)	0.61	0.20	0.04	0.10	0.01
V(100)	0.62	0.20	0.03	0.10	0.01

#### Coarse-grained molecular dynamics (MD) simulations at different concentrations

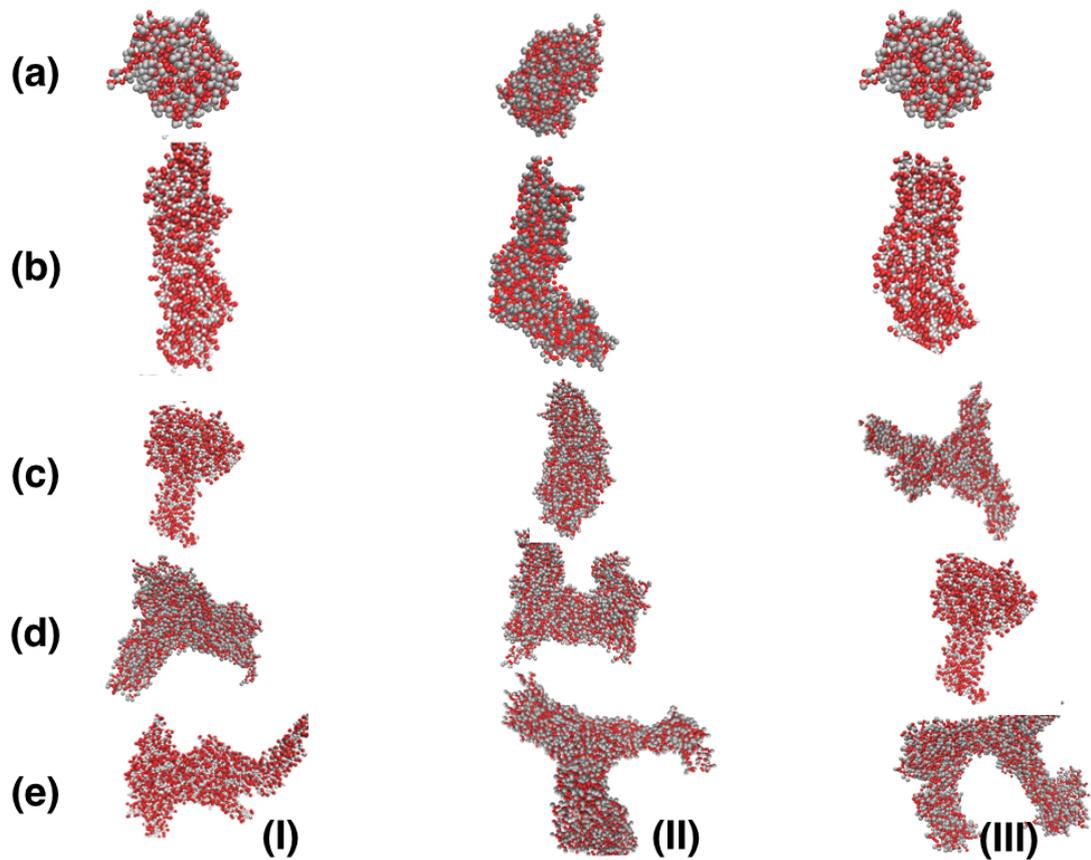
The initial coarse grain structure of the A $\beta_{17-42}$  peptide is shown in Figure 4.2. The effect of concentration on the aggregation was determined by running a series of simulations with different amyloid concentrations. A varying number of amyloids were randomly placed in the box to observe whether they aggregate to form fibrils or not. And if they do so, what is the degree of order in these aggregates. Although factors such as pH, temperature, and simulation conditions affect the outcome of such simulations, we investigated the effect of concentration and temperature only. The results show that supersaturation greatly affects the process of assembly and hence the shape of the aggregate; which is in accordance with the literature (Liang, Wang and Shen, 2016). With an increasing concentration of A $\beta$ , the hydrophobicity increases that leads to the formation of larger dense aggregates. These computational simulations of A $\beta_{17-42}$  peptide might be helpful to understand the spontaneous formation of different nanostructures via self-assembly (see Figure 4.5). It shows the snapshots of the A $\beta_{17-42}$  peptide systems at various concentrations using coarse-grained models. All the peptides aggregated in five systems, but with different aggregation morphologies ranging from spherical to longitudinal fibril like and sheet-like aggregates with hydrophobic residues residing in the interior and charged residues directed outside towards the water. Moreover, analyzing the trajectory revealed that at lower concentration (system-I), neighboring peptides aggregated into small clusters and that these small clusters merged into one large cluster in a very short time (see cluster

distribution in Figure 4.6). The resultant spherical aggregate is formed within 25ns and maintains its structure throughout 400ns. The system with 40 peptides (system-II) gave rise to laterally growing, unidirectional, disordered protofibrils immediately in the 1st few nanoseconds. This shows that at this particular concentration A $\beta_{17-42}$  peptide could self-assemble into a nano rod-like structure. The peptides at still higher concentrations could not form a unidirectional growing nanorod or fibril rather the simulations with peptide 60, 80, and 100 yielded a glass-like amorphous aggregate (here referred to as nanosheets). To investigate the stability of longitudinal fibril like structure, we extended the simulation of this system to 2 $\mu$ s. It is apparent from Figure 4.5b that the growth of elongated fibril does take place, but the structure is not ordered. The monomers 1st give rise to disordered oligomers, and then they lock into disordered fibrillar aggregate. This disordered phase is so stable that it is unable to reorganize itself into well-ordered protofibril even after 2 $\mu$ s. We did not see any dissociation in these structures once they were formed. This inference is well in line with the work done by Sorenson et.al in 2011 who carried out a similar kind of study on decapeptide hIAPP protofibrils and observed the formation of disordered assemblies (Sørensen *et al.*, 2011).

To study the process of aggregation in detail, we measured the solvent accessible surface area and radius of gyration for all the systems. The SASA greatly reduced in all the five systems. In system-I, the value of SASA reduces from 500nm to 300nm. The value of Rg for this system is very small (2.0nm) as it forms a spherical aggregate. The x, y, and z components of this aggregate are almost the same confirming the formation of a spherical structure. For system-II that contains 40 peptides gives rise to long, unidirectional unbranched but disordered fibril like structure within 6ns. The value for SASA in system-II decreases from 971nm to 610nm in the first few nanoseconds and then fluctuate around a mean value. The Rg for this system is twice the value for system-I aggregate because of its longitudinal structure. For the systems-III, IV, and V, consisting of 60, 80, and 100 peptides the value of SASA reduces from 1420nm-880nm, 1500nm-1100nm, and 1900nm-1400nm and Rg is 3.4, 4.2nm and 4.5nm respectively. Large values of Rg indicate the formation of diffuses sheet-like aggregates.

To look for 1-dimensionality of these aggregates which is generally seen in the fibril forming peptides, we calculated their moments of inertia (MOIs) along the principal axes for the largest cluster present in the system. The results showed that only system-

II is forming a 1-Dimensional fibrillar structure with an aspect ratio of 5.1. The details of aspect ratios are presented in Table 4.3.



**Figure 4.5. Structures of aggregates formed by  $A\beta$  at various peptide concentrations for 400ns in three independent replica studies.**

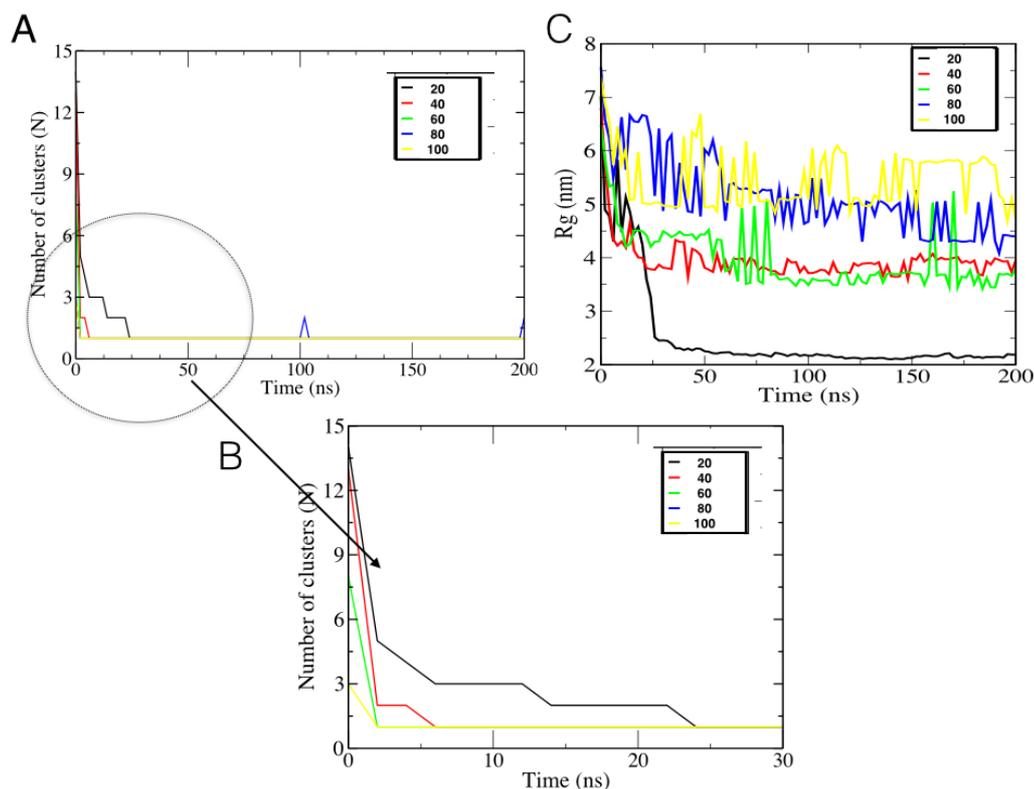


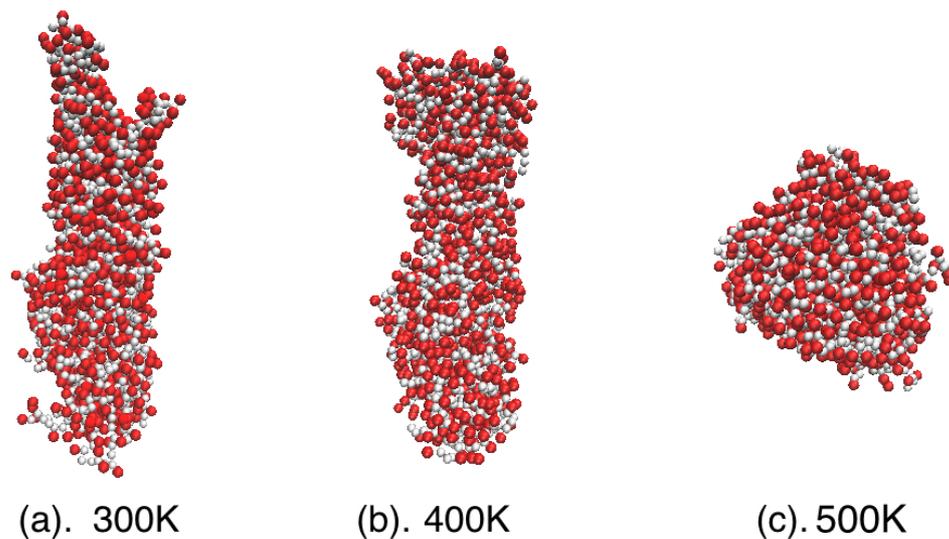
Figure 4.6. The plots show the rapid assembly of A $\beta$  into a single aggregate. (A) plot of the number of aggregates (clusters) vs time shows that higher the concentration, more quickly the aggregates are formed. (B) Zoom in to the time zone where the aggregate formation is taking place from monomers to oligomers and finally aggregates (for which number of clusters = 1); (C) Decrease in the radius of gyration; dramatic decrease at the lowest concentration because a spherical structure is formed.

Table 4.3: Aspect ratio of beta amyloid aggregates; larger values show the formation of unidirectional structures

Systems	Aspect ratio
I	2.30
II	5.1
III	2.77
IV	4.1
V	3.01

The growth of ordered fibrils was not observed in any of the simulations. For longer fibril-like structure, it was observed that the fibril once formed, gets stuck in the disordered conformation and stays as such even after 2 $\mu$ s. To study the effect of

temperature on the shape of the fibril, a set of three additional simulations was run at 300K, 400K and 500K. The structure of the fibril remained stable till 400K; however, at 500K, an amorphous aggregate was observed which indicates the denaturation of peptides. Hence it gives an idea that there is a specific concentration; temperature pair parameter, which gives rise to unidirectional pseudo fibrils (see Figure 4.7).



**Figure 4.7. Effect of temperature on the shape of the fibril. The figure shows that with an increase in temperature the unidirectional fibril like structure collapses into globular aggregate.**

## 4.5 Discussion:

Concentration-dependent A $\beta$  peptides aggregation leads to the formation of various A $\beta$  conformations

Self-assembly is an association reaction which gives rise to supramolecular structures. Formations of tertiary structures of proteins require concentrations to be above the saturation level. In fact the supersaturation should be beyond the metastable zone width to overcome the nucleation barrier. And from the protein crystallization point of view it is known that there is a sweet spot with respect to supersaturation for the crystal

nucleation and a separate one for the crystal growth. Higher supersaturation gives rise to amorphous structures whilst lower supersaturation in-principal to ordered structures, although the time scale in the latter case could be very long. Studies show that the order fibril formation could take days in-vitro. A careful inspection of the trajectory shows that there is a rapid encounter of the monomers in the first few ns followed by the formation of disordered oligomers. These oligomers form the fibrils quickly, but they do not resemble the mature ordered fibrils. This may be because of the time scale that we are using in our simulation, which restricts the formation of fully ordered mature fibrils.

Several in-vitro studies show the oligomerization of  $A\beta$  is concentration-dependent and results in the formation of aggregates of diverse sizes and lengths. At very low concentration (below 20  $\mu\text{M}$ ), no oligomers are formed even after 10 days. The concentration of  $A\beta_{42}$  (below 30  $\mu\text{M}$ ) gave rise to prefibrillar oligomers while at higher concentrations (50–75  $\mu\text{M}$ ) ordered fibrils were readily formed. The prefibrillar oligomers at lower concentrations do not form ordered fibrils but convert into another kind of oligomers of the same shape and size but different toxicity (Ladiwala *et al.*, 2012).

The unidirectional fibril like structure formed in System-II was simulated for 2 $\mu\text{s}$ . The structure remained stable throughout with no disintegration. Self-assembly of proteins is an irreversible phenomenon. But usually it means that the rate of disassembly is very slow, and the equilibrium always remains towards fibrillar form rather than monomeric form. Simulations at various temperatures show that the long fibril-like structure is thermally quite stable and retains its structure till 400K. Beyond that the structure is unable to hold its shape and collapse to a disordered globular structure. Because at this point the thermal energy becomes higher than the bonding energy, the individual molecules start moving with respect to each other, and a spherical structure is formed as it reduces the interfacial energy.

The crystallization of proteins like any other compound is driven by the minimization of free energy of the system (Weber, 1991). When the solution reaches a state of supersaturation it is thermodynamically driven to an equilibrium state that corresponds to a new energy minimum which is approximately 3-6 kcal/mole lower than the energy of the system when it was in a saturated state (Drenth and Haas, 1992). The molecules make new non-covalent interactions thus losing their degrees of freedom. This results

in the formation of aggregates that can either be crystalline or amorphous, but the amorphous aggregates are usually more favorable kinetically and hence dominate the process. The final structure depends upon the strength of interaction between the molecules. If it is above a few  $KT$ , the structure remains kinetically trapped in a particular conformation and does not lead to an ordered crystalline structure. As for  $A\beta$  peptides the binding energy between two anti-parallel  $\beta$ -strands forming a  $\beta$ -sheet is  $\approx 20$  kcal/mole, it is not possible for the peptides to undergo an internal rearrangement to form an ordered structure.

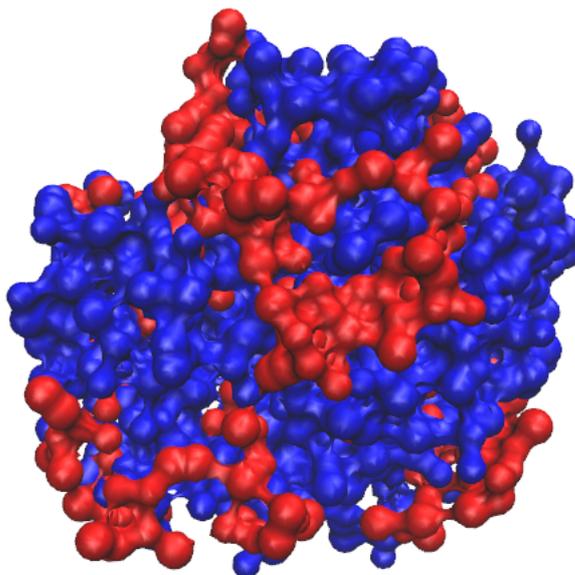
Non-specific hydrophobic interactions lead to the formation of diffused aggregates

Self-assembly of peptides is a ubiquitous phenomenon that is driven by hydrophobic interactions and hydrogen bonding. This self-assembly drives the amyloids to aggregate into first disordered oligomeric form and then fibrillar form. The peptide concentration controls the oligomerization of these peptides in the solution (Schmit, Ghosh, and Dill, 2011). The hot spots for aggregation which are charged residues on protein surfaces form complementary interactions with the same amino acid on the other protein. Several amino acids in the peptide are hydrophobic and they remain buried in the core in the native state but as the amyloid starts to unfold these hydrophobic residues reveal themselves and inter-protein associations increase. Similar effects are commonly observed in experimental studies (Shivu *et al.*, 2013).  $A\beta$  undergo aggregation in a two-step process. First, they formed globular micelle like oligomers, which then cluster to give rise to a bigger aggregate. A closer inspection shows that the main interactions, which stabilize the aggregate are hydrophobic. Residues ALA 42, ILE 41, VAL 40, VAL 39, VAL36, MET 35, LEU34, ILE 32, ILE 31, ALA30, ALA24, PHE20, PHE 19, VAL18, LEU17 are all hydrophobic. Thus, the hydrophobic interactions are the main driving force behind the self-assembly process of  $A\beta$  peptides. The non-polar residues exclude the water molecules and form a hydrophobic core (see Figure 4.8). This hydrophobic core is the hub of stability found in the cluster.

Self-assembly at one particular concentration leads to the formation of long 1-D fibrillar structures while others lead to diffused aggregates. The difference in behavior seems to lie in the way these peptides assemble. The terms “self-assembly” and “aggregation” are often used interchangeably, but in reality, both processes are different. Self-assembly is dominated by specific hydrogen bonding while aggregation is driven

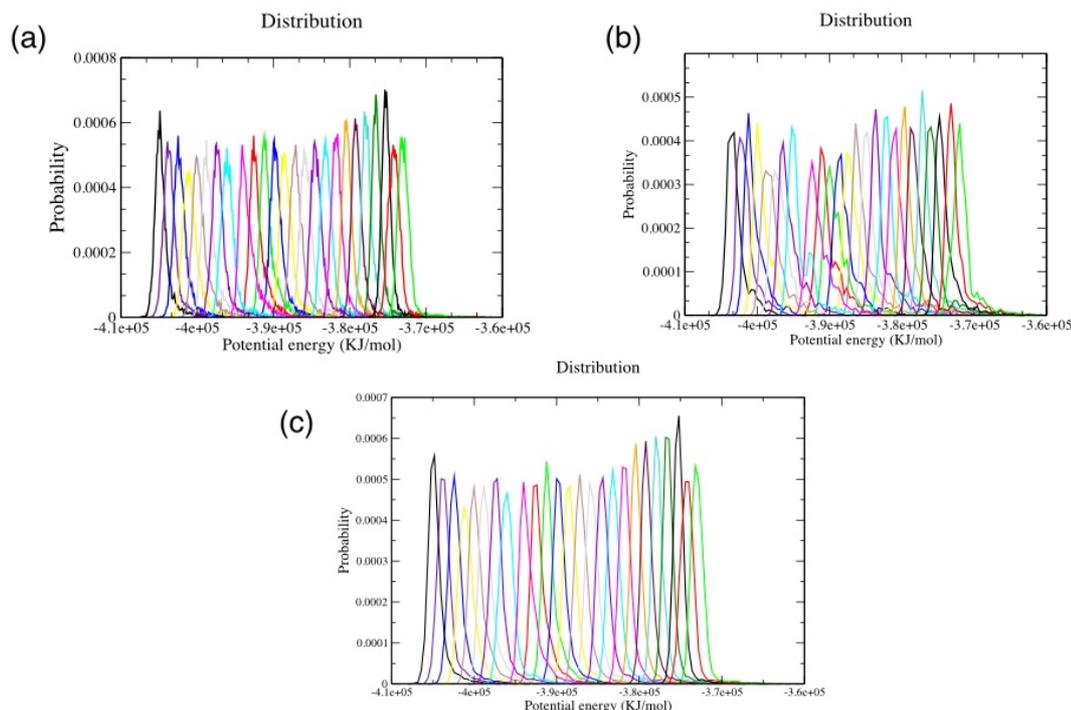
through hydrophobic interactions. That is probably the reason that hydrophobicity driven structures are amorphous or less ordered as compared to self-assembled structures. In our simulations, at high supersaturation, the increased hydrophobicity results in the rapid aggregation of  $A\beta$ . The process is very fast because non-specific hydrophobic interactions are formed rapidly while hydrogen bonding depends upon distance and angle parameters, so these are formed slowly. that the peptides don't get a chance to re-organize themselves into an ordered structure

The protein aggregation process is largely a mystery in the sense that we are still not sure if the nuclei are formed from strict monomer-monomer or monomer-dimer addition or coalescence of small clusters. It is also not clear whether the process is initiated by an ordered nucleus or there is certain rearrangement that occurs in the nuclei leading to the formation of ordered aggregates. What is generally believed is that if there are strong directional forces, the resultant morphology will reflect those directional forces. For example, in a crystal nucleus, there are interactions between the molecules that are repeated periodically. The structures having strong one-dimensional interactions will give rise to linear structures. Having strong two-dimensional forces will give rise to two-dimensional morphology. While aggregates having three-dimensional interaction forces, which are similar in magnitude (isotropic forces) will give rise to a three-dimensional structure. At the earlier stage of the nucleation, for the later one would expect a spherical nucleus because this (sphere) minimizes the interfacial free energy. The amorphous aggregates are characterized by random, non-specific interactions and they tend to form much faster than the regular crystals as seen in this study. As a result, at high supersaturation, the peptides do not get the time to re-align themselves into unidirectional structures.



**Figure 4.8. Spherical aggregate formed by 20 peptides. Blue beads show the hydrophobic core while red beads are charged.**

H-REMD may not be the method of choice for large self-assembling peptide systems: H-REMD simulations were performed to enhance conformational sampling compared to that of classic and coarse-grain MD and get an equilibrium structure. To our knowledge this is the first attempt to use H-REMD for such large systems comprising of many copies of peptides in an explicit solvent that tried to compare the results of classic MD of similar timescales with H-REMD. Previous studies, for example, have successfully used this method to study folding behaviors and self-assembly of peptides but the simulations either consist of a very small system (only containing one or two peptides) or a very short duration. For this specific problem of concentration-dependent peptide aggregation, we needed a large number of copies of the same peptide that needed to be simulated for longer times to get an organization into a fibrillar aggregate. But we could not get a good exchange rate that is a pre-requisite for a successful replica exchange. Though a very strong overlap between various replicas was found yet the exchange probability was extremely low. Several test runs were carried out to adjust the scaling factors such as  $\Delta\lambda$  and time step for exchange but none of them produced desired exchange probability. This points to the limitations of such enhanced sampling techniques for larger systems. The distribution curves obtained for various test runs are presented here in Figure 4.9.



**Figure 4.9.** Histograms of the potential energy distribution of the individual replicas differentiated by the colors obtained from Hamiltonian replica exchange MD for a self-assembling system containing 40 peptides in explicit water at three different exchange time-steps (a). 10; (b) 500 and (c) 1000. The curve on the extreme left is for unmodified parameters while the right one is for most strongly modified force field parameters.

## 4.6 Implications

Higher supersaturation, more hydrophobicity ultimately leads to higher aggregation propensity. Our results suggest that at higher supersaturation, non-specific hydrophobic interactions lead to the formation of larger aggregates. In-vivo these aggregates (amorphous) due to insolubility do not convert into dense core fibrils associated with sporadic AD which implies that these aggregates might not be toxic. To put the results in perspective, it's important to show the limitations of the methodology we used.  $\alpha$  to  $\beta$ -sheet conversion is crucial for  $A\beta$  assembly. But in CG models, secondary structure is constrained, and we cannot see the evolution of secondary structure. The absence of directional hydrogen bonding leaves no opportunity for the fibril to re-organize into an ordered structure. Also, MARTINI is thought to over-stabilize peptide-peptide interaction which may be the reason for rapid aggregate formation. Though our results still lack the time scale that is required to completely follow the aggregation process but provide a potential in the elucidation of the  $A\beta$  assembly process.

## 4.7 References

- Affentranger, R., Tavernelli, I. and Di Iorio, E. E. (2006) 'A novel hamiltonian replica exchange MD protocol to enhance protein conformational space sampling', *Journal of Chemical Theory and Computation*. doi: 10.1021/ct050250b.
- Arosio, P., Knowles, T. P. J. and Linse, S. (2015) 'On the lag phase in amyloid fibril formation.', *Physical chemistry chemical physics : PCCP*. Royal Society of Chemistry, 17(12), pp. 7606–18. doi: 10.1039/c4cp05563b.
- Auer, S., Dobson, C. M. and Vendruscolo, M. (2007) 'Characterization of the nucleation barriers for protein aggregation and amyloid formation', *HFSP Journal*. doi: 10.2976/1.2760023.
- Bellesia, G. and Shea, J. E. (2007) 'Self-assembly of B -sheet forming peptides into chiral fibrillar aggregates', *Journal of Chemical Physics*. doi: 10.1063/1.2739547.
- Bellesia, G. and Shea, J. E. (2009) 'Effect of B -sheet propensity on peptide aggregation', *Journal of Chemical Physics*. doi: 10.1063/1.3108461.
- Berendsen, H. J. C., Grigera, J. R. and Straatsma, T. P. (1987) 'The missing term in effective pair potentials', *Journal of Physical Chemistry*. doi: 10.1021/j100308a038.
- Bussi, G. (2014) 'Hamiltonian replica exchange in GROMACS: A flexible implementation', in *Molecular Physics*. doi: 10.1080/00268976.2013.824126.
- C. Vickers, J. *et al.* (2015) 'Defining the earliest pathological changes of Alzheimer's disease', *Current Alzheimer Research*. doi: 10.2174/1567205013666151218150322.
- Chiti, F. and Dobson, C. M. (2006) 'Protein Misfolding, Functional Amyloid, and Human Disease', *Annual Review of Biochemistry*. doi: 10.1146/annurev.biochem.75.101304.123901.
- Chiti, F. and Dobson, C. M. (2009) 'Amyloid formation by globular proteins under native conditions', *Nature Chemical Biology*. Nature Publishing Group, 5(1), pp. 15–22. doi: 10.1038/nchembio.131.
- Darden, T., York, D. and Pedersen, L. (1993) 'Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems', *The Journal of Chemical Physics*. doi: 10.1063/1.464397.
- Drenth, J. and Haas, C. (1992) 'Protein crystals and their stability', *Journal of Crystal*

*Growth*. doi: 10.1016/0022-0248(92)90233-9.

Fukunishi, H., Watanabe, O. and Takada, S. (2002) 'On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction', *Journal of Chemical Physics*. doi: 10.1063/1.1472510.

Groen, J. *et al.* (2015) 'Associative interactions in crowded solutions of biopolymers counteract depletion effects', *Journal of the American Chemical Society*. doi: 10.1021/jacs.5b07898.

Hess, B. *et al.* (1997) 'LINCS: A Linear Constraint Solver for molecular simulations', *Journal of Computational Chemistry*. doi: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

Hoyer, W. *et al.* (2008) 'Stabilization of a  $\beta$ -hairpin in monomeric Alzheimer's amyloid- $\beta$  peptide inhibits amyloid formation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0711731105.

Humphrey, W., Dalke, A. and Schulten, K. (1996) 'VMD: visual molecular dynamics.', *Journal of molecular graphics*, 14(1), pp. 33–8, 27–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8744570> (Accessed: 16 May 2017).

Ladiwala, A. R. A. *et al.* (2012) 'Conformational differences between two amyloid  $\beta$  oligomers of similar size and dissimilar toxicity', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.329763.

Latshaw, D. C., Cheon, M. and Hall, C. K. (2014) 'Effects of macromolecular crowding on amyloid beta (16-22) aggregation using coarse-grained simulations', *Journal of Physical Chemistry B*. doi: 10.1021/jp508970q.

Li, C. and Pielak, G. J. (2009) 'Using NMR to distinguish viscosity effects from nonspecific protein binding under crowded conditions', *Journal of the American Chemical Society*. doi: 10.1021/ja808428d.

Liang, L., Wang, L. W. and Shen, J. W. (2016) 'The self-assembly mechanism of tetrapeptides from the motif of  $\beta$ -amyloid peptides: A combined coarse-grained and all-atom molecular dynamics simulation', *RSC Advances*. doi: 10.1039/c6ra18204f.

Lührs, T. *et al.* (2005) '3D structure of Alzheimer's amyloid-beta(1-42) fibrils.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(48), pp. 17342–7. doi: 10.1073/pnas.0506723102.

- Luiken, J. A. and Bolhuis, P. G. (2015) 'Primary Nucleation Kinetics of Short Fibril-Forming Amyloidogenic Peptides', *The Journal of Physical Chemistry B*, 119(39), pp. 12568–12579. doi: 10.1021/acs.jpcc.5b05799.
- M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the G. and Team, D. (2015) 'GROMACS User Manual version 5.1'.
- Man, V. H., Nguyen, P. H. and Derreumaux, P. (2017) 'Conformational Ensembles of the Wild-Type and S8C A $\beta$ 1-42 Dimers', *Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcc.7b00267.
- Minton, A. P. (2001) 'The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media', *Journal of Biological Chemistry*. doi: 10.1074/jbc.R100005200.
- Nguyen, H. D. and Hall, C. K. (2004) 'Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides.', *Proceedings of the National Academy of Sciences of the United States of America*, 101, pp. 16180–16185. doi: 10.1073/pnas.0407273101.
- Nguyen, H. D. and Hall, C. K. (2006) 'Spontaneous fibril formation by polyalanines; discontinuous molecular dynamics simulations', *Journal of the American Chemical Society*. doi: 10.1021/ja0539140.
- Parrinello, M. and Rahman, A. (1981) 'Polymorphic transitions in single crystals : A new molecular dynamics method Polymorphic transitions in single crystals : A new molecular dynamics method', *Journal of Applied Physics*, 52(12), pp. 7182–7190. doi: 10.1063/1.328693.
- Periole, X. *et al.* (2009) 'ELNEDYN: Keep your proteins stable', *J. Chem. Th. Comp.*
- Sarkar, M., Li, C. and Pielak, G. J. (2013) 'Soft interactions and crowding', *Biophysical Reviews*. doi: 10.1007/s12551-013-0104-4.
- Schmit, J. D., Ghosh, K. and Dill, K. (2011) 'What drives amyloid molecules to assemble into oligomers and fibrils?', *Biophysical journal*, 100(2), pp. 450–8. doi: 10.1016/j.bpj.2010.11.041.
- Serrano-Pozo, A. *et al.* (2011) 'Neuropathological alterations in Alzheimer disease', *Cold Spring Harbor Perspectives in Medicine*. doi: 10.1101/cshperspect.a006189.

- Shivu, B. *et al.* (2013) 'Distinct  $\beta$ -Sheet Structure in Protein Aggregates Determined by ATR-FTIR Spectroscopy', *Biochemistry*, 52(31), pp. 5176–5183. doi: 10.1021/bi400625v.
- Sørensen, J. *et al.* (2011) 'Protofibrillar Assembly Toward the Formation of Amyloid Fibrils', *The Journal of Physical Chemistry Letters*, 2(19), pp. 2385–2390. doi: 10.1021/jz2010094.
- Streets, A. M. *et al.* (2013) 'Simultaneous Measurement of Amyloid Fibril Formation by Dynamic Light Scattering and Fluorescence Reveals Complex Aggregation Kinetics', *PLoS ONE*. doi: 10.1371/journal.pone.0054541.
- Sugita, Y. and Okamoto, Y. (1999) 'Replica-exchange molecular dynamics method for protein folding', *Chemical Physics Letters*. doi: 10.1016/S0009-2614(99)01123-9.
- Weber, P. C. (1991) 'Physical Principles of Protein Crystallization', *Advances in Protein Chemistry*. Academic Press, 41, pp. 1–36. doi: 10.1016/S0065-3233(08)60196-5.
- Wu, C. and Shea, J. E. (2011) 'Coarse-grained models for protein aggregation', *Current Opinion in Structural Biology*. doi: 10.1016/j.sbi.2011.02.002.
- Xu, L., Chen, Y. and Wang, X. (2014) 'Assembly of amyloid  $\beta$  peptides in the presence of fibril seeds: One-pot coarse-grained molecular dynamics simulations', *Journal of Physical Chemistry B*. doi: 10.1021/jp505551m.
- Xu, Y. *et al.* (2005) 'Conformational transition of amyloid  $\beta$ -peptide', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0501218102.
- Yan, Y. and Wang, C. (2006) 'A $\beta$ 42 is More Rigid than A $\beta$ 40 at the C Terminus: Implications for A $\beta$  Aggregation and Toxicity', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2006.09.046.
- Yang, G. *et al.* (2006) 'Crystal shape control by manipulating supersaturation in batch cooling crystallization', *Crystal Growth and Design*. doi: 10.1021/cg0603873.

# 5 Towards non-invasive diagnosis of AD molecular basis for the specificity of curcumin for fibrillar beta amyloid

**Copyright statement:** The contents of this chapter are currently in preparation for publication.

**Attribution:** Beenish Khurshid and Jamshed Anwar designed the research and wrote the paper. Beenish Khurshid conducted the simulations and performed the analysis.

## 5.1 Abstract

Recently, it has been shown that the amyloid burden in the retina correlates to that of the brain, and that curcumin (a natural fluorescent dye) can be used to detect the deposits *non-invasively* by retina imaging. We have investigated the molecular basis for curcumin's specificity for hierarchical A $\beta$  structures using molecular dynamics simulation, with a focus on how curcumin can detect and discriminate differing amyloid morphologies. Curcumin inhibits and breaks up  $\beta$ -sheet formation in A $\beta$  monomers. With disordered A $\beta$  structures, curcumin forms a coarse-grained composite structure. With an ordered fibril, curcumin's interaction is highly specific, the curcumin molecules depositing in the fibril groove. Curcumin has a tendency to self-aggregate, which is finely balanced with its affinity for A $\beta$  amyloid. This tendency concentrates curcumin molecules at its deposition sites potentially increasing the fluorescence signal, which is probably why curcumin is such an effective amyloid imaging agent.

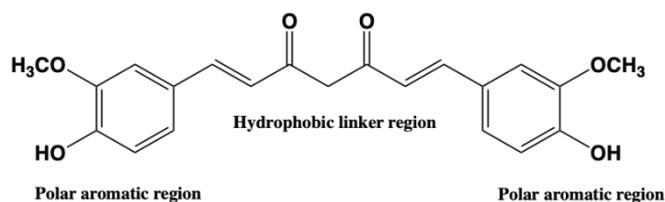
## 5.2 Introduction

Given the immense societal impact of Alzheimer's disease (AD), the identification of patients with the early form of AD (prodromal AD) is a health imperative. There is as yet no definitive diagnostic test for this pre-symptomatic phase of AD. Diagnosis of full-blown AD in the clinic involves assessment and evaluation of symptoms and cognitive skills, coupled with biochemical blood tests and brain imaging using magnetic resonance imaging (MRI), X-ray based computerized tomography (CT) or positron emission tomography (PET) which involves the use of a radiative tracer substance (Marcus, Mena, and Subramaniam, 2014). The biochemical tests and imaging are non-specific and are mostly used to rule out other conditions. Progress towards diagnosing prodromal AD is promising but still confined to research i.e. clinical trials. The current framework includes more specific PET scans using tau and beta amyloid (A $\beta$ ) binding ligands and the use of biomarkers that include A $\beta$ ,  $\beta$ -site APP-cleaving enzyme 1 gene (BACE 1), soluble A $\beta$  precursor protein (sAPP), and anti-A $\beta$  antibodies found in the cerebral spinal fluid and blood plasma (Hampel *et al.*, 2010). Beyond current technical challenges, the application of these methods (particularly amyloid-specific PET) to screen large populations in a clinical setting would be prohibitive both economically and because of safety concerns (exposure to radioactive isotopes).

A $\beta$  accumulation is considered to begin as early as 20 years before the manifestation of clinical dementia (Beason-Held *et al.*, 2013). This prodromal phase therefore represents the best opportunity-window for therapy. In recognizing the need for early therapeutic intervention, one confronts another equally significant hurdle: the need to identify at-risk patients at the earliest stages of AD development, ideally non-invasively. A recent, exciting finding is the detection of A $\beta$  deposits and p-tau in the retina, both in animal models and humans afflicted by AD (Colligris *et al.*, 2018). This has a sound basis, given that the retina shares many physiological and anatomical features with the brain and is considered to be a projection of the central nervous system (CNS). Moreover, *in vivo* studies show that the plaque burden in the retina correlates to that of the brain (Koronyo-Hamaoui *et al.*, 2011) and can be detected earlier than the detection of amyloid deposits in the brain (Zhang-Nunes *et al.*, 2006). The retina therefore offers a potentially non-invasive and accessible route to identifying at-risk patients with prodromal AD. Indeed, the concept has been demonstrated in live patients using the

pigment curcumin as the amyloid-specific fluorescence probe coupled with a modified scanning laser ophthalmoscope (Goozee *et al.*, 2016).

Curcumin, a bright yellow pigment is a constituent of the Indian spice turmeric. The structure of Curcumin is flexible with two aromatic rings at the ends with a long linker region joining both the rings (see Figure 5.1). Structural analysis shows that Curcumin exists in solution in two tautomeric forms; keto and enol (Priyadarsini, 2014). At pH below 7, Curcumin exists in keto form while above pH 7, it exists in enolate form. The Curcumin tautomers are responsible for the so much talked about hydrophobicity of Curcumin (Lee *et al.*, 2013). Due to its hydrophobicity and flexibility, Curcumin is known to bind a huge number of biomolecules. It appears that both of the aromatic rings and the rigid linker are critical to curcumin's specificity for amyloid. Removing one of the rings or altering the length/flexibility of the linker results in the loss of its activity to inhibit A $\beta$  aggregation (Reinke and Gestwicki, 2007) and by implication a loss in its binding affinity for A $\beta$ . Interestingly, other amyloid ligands e.g. congo red and chrysamine G (Maiti and Dunbar, 2018) share these features too.



**Figure 5.1. Structure of curcumin showing the hydrophobic linker region and the polar substituted aromatic rings.**

Small molecules such as thioflavin S and T and congo red are used as amyloid tracers for a very long time, but they suffer from some serious drawbacks. Thioflavin is weakly hydrophobic; that's why its binding affinity to amyloid fibrils is low (Wu, Bowers and Shea, 2011) (Wu, Bowers and Shea, 2011). While congo red is amyloid non-specific (Yakupova *et al.*, 2019) so it also stains other non-amyloid deposits such as elastin (Horobin and James, 1970), collagen, elastotic dermis, hyaline deposits (Lendrum, Slidders and Fraser, 1972) in colloid milium and lipid proteinosis (Bayer-Garner and Smoller, 2002). Curcumin, on the other hand, is not only amyloid specific as it can differentiate between AD and non-AD deposits with 80.6% specificity (Frost *et al.*, 2013) but is also able to discriminate between various A $\beta$  morphologies i-e core, neurite, diffuse, and burned-out plaques (Maiti *et al.*, 2016) Whilst its therapeutic success in clinical trials remains controversial, curcumin-based near-infrared (NIR)

fluorescence imaging probes (CRANAD-2, CRANAD-44, and CRANAD-28) have been developed and have a higher binding affinity for A $\beta$  assemblies (with a  $K_i$  of 0.07 nM for F<sup>18</sup> labeled curcumin binding for fibrillar A $\beta$ ) than the well-known molecular imaging probes, such as the Pittsburgh compound B (PiB) employed in fludeoxyglucose positron emission tomography (FDG-PET) (Ryu *et al.*, 2006; Xiong *et al.*, 2011). Unlike other A $\beta$ -specific dyes, curcumin also has an additional property of being able to inhibit amyloid aggregation (Maiti and Dunbar, 2018). It binds to A $\beta$  oligomers and fibrils and retards plaque formation (Yang *et al.*, 2005).

Here we explore the molecular level interaction of curcumin with A $\beta$  and its various morphologies employing molecular dynamics simulation (MD), to identify the molecular origin of curcumin's specificity for A $\beta$ . We have investigated the interaction of curcumin at multiple levels: (i) its interaction with a single A $\beta$  monomer; (ii) interaction with A $\beta$  molecules during their aggregation (self-assembly); and (iii) curcumin's interaction with a preformed fibril. In this way we develop a ground-up understanding of curcumin's interaction. Such an approach also enables us to rationalize how curcumin interacts with the various A $\beta$  morphologies and stages that characterize the full A $\beta$  pathway from individual A $\beta$  molecules to fully-developed fibrils.

We employed standard unbiased MD - the interaction of curcumin is strong and specific and there appears to be no ergodicity (dependence on starting configuration) issues, making the interaction trajectories wholly accessible using unbiased MD. The self-assembly systems comprised 24 monomers of A $\beta_{42}$  with a varying number of curcumin molecules namely 2, 5, 19, 77, and 308 molecules, including a control without curcumin. For the interaction of curcumin with the pre-formed fibril, we constructed A $\beta$  fibril (based on the PDB structure 2BEG) with two parallel  $\beta$ -sheets, each comprising twenty-five, antiparallel in-register  $\beta$ -strands. For the interaction of curcumin with the pre-formed fibril, we investigated two curcumin concentrations, namely 5 and 30 curcumin molecules, which were located randomly in the initial configuration. Details of the simulations are given in Supplementary Information.

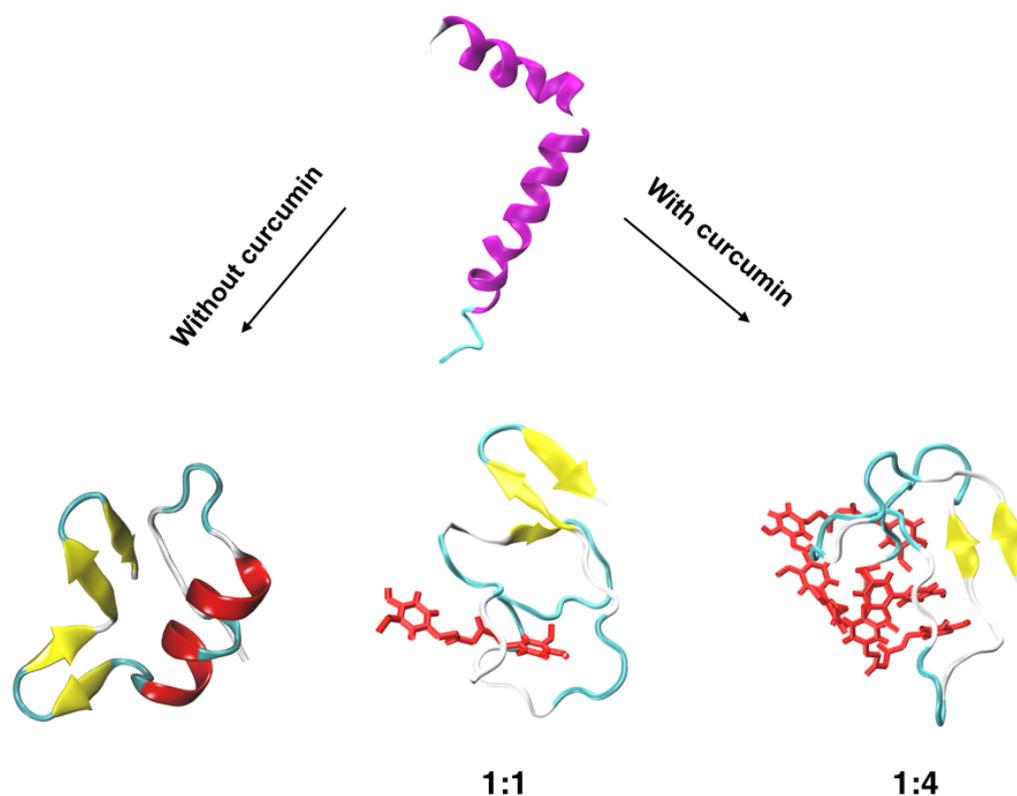
The simulations reveal that curcumin's interaction with the A $\beta_{42}$  *monomer* is non-specific, with the curcumin molecule continuously moving around and interacting with a multiple number of residues including PHE, LEU, VAL, ALA and ILE. Wherever the curcumin molecule locates itself about the A $\beta$  structure, it destroys the  $\beta$ -sheet in its vicinity. When it leaves that position, the  $\beta$ -sheet reappears. Curcumin is found to hover

over the whole structure for most of the time which is in accordance with earlier literature (Zhao *et al.*, 2012). Illustrative conformations of A $\beta$ <sub>42</sub> monomer with curcumin are shown in Figure 5.2. At the higher curcumin concentration, A $\beta$ <sub>42</sub> : curcumin ratio of 1:4, curcumin disrupts A $\beta$ <sub>42</sub> such that more than 50% of A $\beta$ <sub>42</sub> is in the coil conformation with only about 13%  $\beta$ -sheet content, compared to 26% coil and 36%  $\beta$ -sheet content in the control (without curcumin). Energetically, curcumin's interaction with A $\beta$  monomer is relatively strong, the estimated binding energy using the MM-PBSA method being  $\Delta G_{\text{binding}} = -17 \pm 5 \text{ kcal mol}^{-1}$ , which equates to about  $\approx 23 k_{\beta}T$ . The significance of expressing binding energies in terms of KT is that it is a good order-of-magnitude estimate for the energy needed for a process to occur at a particular temperature. A useful rule of thumb is that if a process needs energy ranging from 15KT to 30KT, it would occur at an appreciable rate. Above 30KT, the process would be very slow, while below 15KT the processes would be too fast to accomplish any significant phenomenon.

In the self-assembly simulations without curcumin, the A $\beta$  molecules form a disordered structure rich in  $\beta$ -sheets. Indeed, such morphology representing the early stage of A $\beta$  aggregation has been observed in earlier studies (Wei, Mousseau, and Derreumaux, 2007). In the presence of curcumin, the emergent structures are coarse composite-like and disordered (see Figure 9.8). Although, the A $\beta$  and curcumin are integrated in the aggregates, the integration is not homogeneous. Structurally, the aggregates are devoid of any  $\beta$ -sheets, in contrast to the structure formed for pure A $\beta$ . A noteworthy feature is that curcumin shows a strong affinity for itself (indeed, curcumin has a low solubility) which drives it to form large clusters of pure curcumin that are then integrated with A $\beta$  but in a coarse-grained manner. Could this concentrated curcumin density in these aggregates serve to possibly amplify the fluorescence signal?

Whilst curcumin's interaction with an A $\beta$  *monomer* is non-specific, its interaction with the pre-formed fibril is highly specific, as almost all curcumin molecules deposit within a particular groove on the fibril. In the system containing 30 curcumin molecules, whilst some individual molecules of curcumin were attracted directly to the fibril surface, others formed aggregates through their stacking (one curcumin on top of another), which then deposited on the surface of fibril in all the three replicates of the study (Figure 5.3) and figure 9.12 in the supplementary information. This is due to the

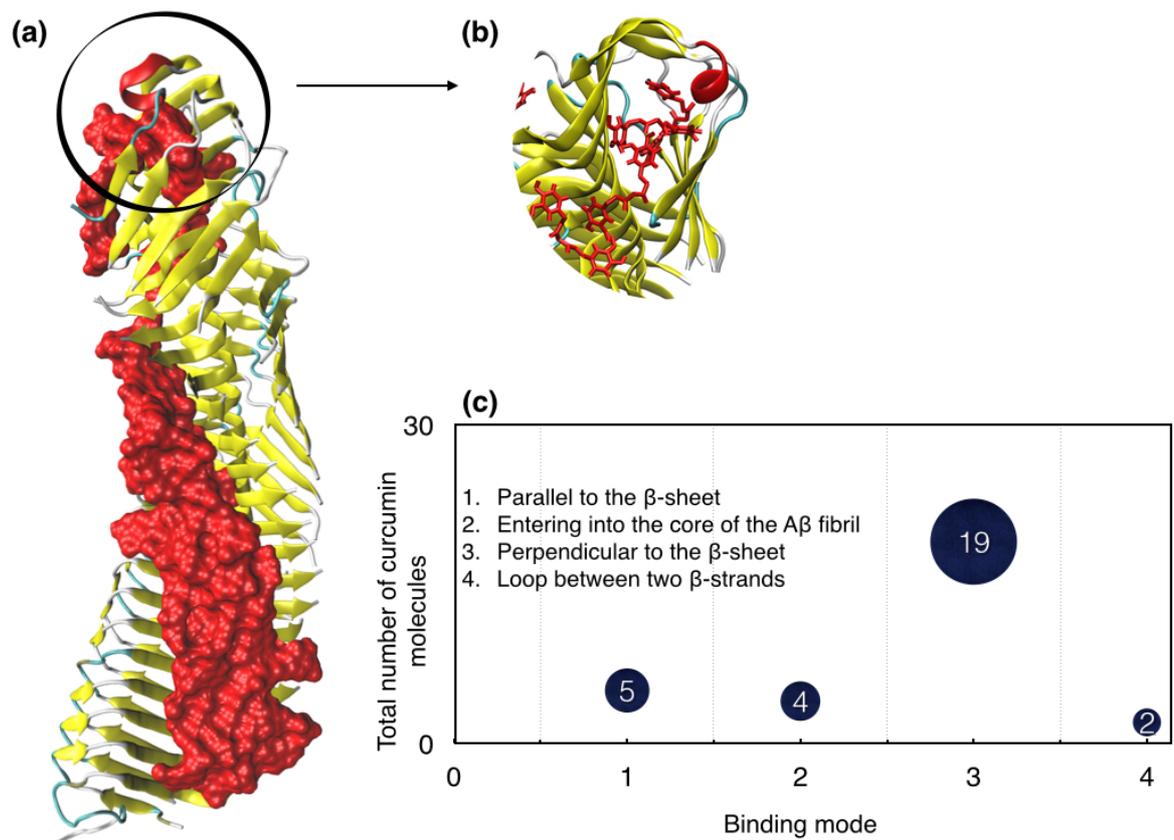
hydrophobic nature of curcumin (consistent with its low aqueous solubility), which drives its self-assembly in aqueous medium.



**Figure 5.2.** Snapshots of the interaction of A $\beta$  42 with curcumin 200 ns into the trajectory. (Top) Initial A $\beta$  helical structure obtained from the protein data bank (PDB:1IYT). (Bottom, Left) Without curcumin; (Bottom, Middle) A $\beta$  and curcumin in a 1:1 ratio; (Bottom, Right) A $\beta$  and curcumin in a 1:4 ratio. Curcumin is represented by a skeleton line structure in red. The peptide structure is shown in cartoon representation where yellow color represents  $\beta$ -strands, red ribbon-like structure represents  $\alpha$ -helix, and green and white colors show turns and coil regions respectively. It is evident from the figure that curcumin destroys the  $\beta$ -sheet structure in its vicinity, and the higher the concentration of curcumin, the lower is the  $\beta$ -sheet content.

Beyond the primary and dominant preference of the curcumin for the A $\beta$  fibril groove, the simulations reveal that a curcumin molecule can enter the hydrophobic core of the fibril via the open ends (see Figure 5.3b), and show a minor presence at the hairpin region around GLY29. Within the fibril groove, there is considerable scope for the alignment of the curcumin molecules, and through population analysis, we are able to identify 2 main modes of curcumin binding (see Figure 5.3c): (i) parallel alignment with the fibril axis; and (ii) perpendicular to the fibril axis; In the (predominant) parallel-mode, the curcumin molecules are aligned perpendicular to the  $\beta$ -sheets and intercalate

two and in some case three or even four  $\beta$ -sheet strands, preferentially binding to the two GLY33 on two different A $\beta$  units. In the perpendicular-mode, the curcumin is aligned parallel to the  $\beta$ -sheets and is localized to a low-width surface path that runs along the fibril containing hydrophobic residues specifically MET, ILE, and VAL that are present very close to the well-known G33XXXG37 motif of A $\beta$  fibril. We observed similar binding preferences at the lower concentration system where only four curcumin molecules are present. Out of four curcumin molecules, two binds to the GXXXG motif parallel to the  $\beta$ -strands, one of them goes into the core of the fibril while the other is present around the hairpin region around GLY29 (see Figure 9.9). Similar binding patterns have also been observed for other amyloid dyes— Congo Red33, BTH, and ThT (Wu *et al.*, 2008). It is due to the presence of the C-terminal residues 28-42 representing a hydrophobic domain associated with the cell membrane in APP and hydrophobicity of curcumin (Wu *et al.*, 2008).



**Figure 5.3. Binding of curcumin to A $\beta$  fibril. (a) The fibril is shown in cartoon representation (yellow-  $\beta$ -sheets, green –turns, white-coil), while curcumin is represented as a surface in red to give a clear perspective of the highly-specific binding, (b) An enlarged image of the fibril end where curcumin is seen to enter into the core of the**

**fibril. (c) A plot of the population of curcumin molecules oriented and located at various positions on the A $\beta$  fibril.**

The specific interactions were confirmed using the radial distribution function to ascertain the probability of finding a curcumin molecule at a certain distance from the individual active site residues. As anticipated, sharp peaks for residues GLY, VAL, ILE, and MET were found at 1.3 Å, 1.0 Å, 0.7 Å and 0.4 Å respectively (Figure 9.10). All three parts of curcumin i.e. two aromatic rings and the linker region bearing the diketone moiety show comparable probabilities suggesting their equal participation in binding to the A $\beta$  fibril. Earlier MD simulations and NMR experiments have proposed the importance of MET35 (Friedemann *et al.*, 2015) GLY33 (Harmer *et al.*, 2009) and the hydrophobic turn located at C-terminal GLY37–GLY38 (Lin *et al.*, 2008).

Based on sampling, the interaction (free) energy of individual curcumin molecules for the A $\beta$  fibril groove is between  $-25 \pm 5$  kcal mol<sup>-1</sup> to  $-31 \pm 7$  kcal mol<sup>-1</sup>, which is larger than the curcumin- A $\beta$  monomer interaction. This equates to about  $\approx 34 - 42k_B T$ , indicating that the interaction is strong and essentially (spontaneously) irreversible. The major contribution comes from the van der Waal's energy ( $-22$  kcal mol<sup>-1</sup>) that corresponds to the hydrophobic interaction. Free energy decomposition analysis indicates that the key residues i.e. those with the strongest interaction ( $\Delta G_{\text{binding}} > 2.0$  kcal mol<sup>-1</sup>) are GLY, VAL, ILE, and MET (Figure 9.11).

Considering the full set of simulation results, the interaction of curcumin with A $\beta$  shows distinctive features with respect to its interaction with the A $\beta$  monomer, A $\beta$  disordered structures, and with an ordered fibrillar structure. Curcumin exhibits a non-specific interaction with the A $\beta$  monomer, which is essentially a hydrophobic interaction. Being hydrophobic itself, curcumin endeavors to reduce its interface with water being attracted to the hydrophobic regions on the A $\beta$  which is in accordance with an MD study showing that curcumin in water is always present as an aggregated state (Hazra, Roy and Bagchi, 2014). The non-specific interaction implies that curcumin is likely to concentrate in any region of the brain where there is a high concentration of peptides or proteins with exposed hydrophobic stretches like A $\beta$ . Therefore, in principle, curcumin (and curcumin-based imaging probes) should be able to detect the pre-amyloid stage, though the sensitivity is likely to be low.

The self-assembly simulations yield a disordered A $\beta$  structure, with which curcumin interacts to form an integrated composite structure but at a coarse-grained level, comprising significant curcumin-only and A $\beta$ -only regions. The cause for the formation of curcumin-only regions is the tendency of curcumin to aggregate itself, due to its low solubility in aqueous media. The high concentration of curcumin within the curcumin-only regions may be responsible for the increased fluorescence signal strength, rationalizing the ability of curcumin to discriminate between the deposited (disordered) amyloid and high concentrations of A $\beta$  in solution.

The interaction of curcumin with the fibrillar structure is highly specific, with the curcumin depositing within the fibril groove. Here again, a particular feature (as a result of curcumin's strong affinity for itself) is that the curcumin does not form a monomolecular layer on the surface of the groove, but rather forms a continuous curcumin-only deposit over the whole region of the fibril groove. The surface grooves created by aligned sidechains in the fibril parallel to the growing axis presents a large surface area to the curcumin molecules for binding, thus establishing extra contacts. This likely explains the higher binding energy of this curcumin-fibril complex (-25 - -30 kcal/mol) than the curcumin-monomer complex (-17 kcal/mol). This unique form of interaction concentrating the number of curcumin molecules interacting with the fibril, in principle may serve to increase the fluorescence signal.

In summary we have explained the molecular basis for the specificity of curcumin for A $\beta$  and its amyloid structures and can rationalize how curcumin can detect and discriminate between A $\beta$  in solution and between differing amyloid morphologies. A unique feature of the curcumin molecule appears to be its tendency to self-aggregate, which is finely balanced with its affinity for A $\beta$  and its amyloid structures. The self-aggregation tendency concentrates curcumin molecules at its deposition sites serving to increase the fluorescence signal, which is probably why curcumin is such an effective amyloid imaging agent. Further, the molecular-level insights gained here would be invaluable in designing more effective and more discriminating curcumin-based imaging agents. To our knowledge our results provide new insight into how to further optimize curcumin and its derivatives for personalized treatment of AD.

### **5.3 Experimental section**

The experimental details can be found in Appendix B.

## 5.4 References

- Bayer-Garner, I. B. and Smoller, B. R. (2002) 'AL amyloidosis is not present as an incidental finding in cutaneous biopsies of patients with multiple myeloma', *Clinical and Experimental Dermatology*. doi: 10.1046/j.1365-2230.2002.01022.x.
- Beason-Held, L. L. *et al.* (2013) 'Changes in brain function occur years before the onset of cognitive impairment', *Journal of Neuroscience*. doi: 10.1523/JNEUROSCI.1402-13.2013.
- Colligris, P. *et al.* (2018) 'Ocular Manifestations of Alzheimer's and Other Neurodegenerative Diseases: The Prospect of the Eye as a Tool for the Early Diagnosis of Alzheimer's Disease', *Journal of Ophthalmology*. doi: 10.1155/2018/8538573.
- Friedemann, M. *et al.* (2015) 'Effect of methionine-35 oxidation on the aggregation of amyloid- $\beta$  peptide', *Biochemistry and Biophysics Reports*. doi: 10.1016/j.bbrep.2015.07.017.
- Frost, S. *et al.* (2013) 'Retinal vascular biomarkers for early detection and monitoring of Alzheimer's disease', *Translational Psychiatry*. doi: 10.1038/tp.2012.150.
- Goozee, K. G. *et al.* (2016) 'Examining the potential clinical value of curcumin in the prevention and diagnosis of Alzheimer's disease', *British Journal of Nutrition*, 115, pp. 449–465. doi: 10.1017/S0007114515004687.
- Hampel, H. *et al.* (2010) 'Biological markers of amyloid  $\beta$ -related mechanisms in Alzheimer's disease', *Experimental Neurology*. doi: 10.1016/j.expneurol.2009.09.024.
- Harmeier, A. *et al.* (2009) 'Role of Amyloid- $\beta$  Glycine 33 in Oligomerization, Toxicity, and Neuronal Plasticity', *Journal of Neuroscience*, 29(23).
- Hazra, M. K., Roy, S. and Bagchi, B. (2014) 'Hydrophobic hydration driven self-assembly of curcumin in water: Similarities to nucleation and growth under large metastability, and an analysis of water dynamics at heterogeneous surfaces', *Journal of Chemical Physics*. doi: 10.1063/1.4895539.
- Horobin, R. W. and James, N. T. (1970) 'The staining of elastic fibres with Direct Blue 152. A general hypothesis for the staining of elastic fibres', *Histochemie*. doi: 10.1007/BF00277460.
- Koronyo-Hamaoui, M. *et al.* (2011) 'Identification of amyloid plaques in retinas from Alzheimer's patients and noninvasive in vivo optical imaging of retinal plaques in a mouse model', *NeuroImage*. doi: 10.1016/j.neuroimage.2010.06.020.
- Lendrum, A. C., Slidders, W. and Fraser, D. S. (1972) 'Renal hyalin. A study of

amyloidosis and diabetic fibrinous vasculosis with new staining methods.’, *Journal of clinical pathology*. doi: 10.1136/jcp.25.5.373.

Lin, W. H. *et al.* (2008) ‘Amyloid- $\beta$  peptide ( $A\beta$ ) neurotoxicity is modulated by the rate of peptide aggregation:  $A\beta$  dimers and trimers correlate with neurotoxicity’, *Journal of Neuroscience*. doi: 10.1523/JNEUROSCI.3916-08.2008.

Maiti, P. *et al.* (2016) ‘A comparative study of dietary curcumin, nanocurcumin, and other classical amyloid-binding dyes for labeling and imaging of amyloid plaques in brain tissue of 5 $\times$ -familial Alzheimer’s disease mice’, *Histochemistry and Cell Biology*. doi: 10.1007/s00418-016-1464-1.

Maiti, P. and Dunbar, G. L. (2018) ‘Use of curcumin, a natural polyphenol for targeting molecular pathways in treating age-related neurodegenerative diseases’, *International Journal of Molecular Sciences*. doi: 10.3390/ijms19061637.

Marcus, C., Mena, E. and Subramaniam, R. M. (2014) ‘Brain PET in the diagnosis of Alzheimer’s disease’, *Clinical Nuclear Medicine*. doi: 10.1097/RLU.0000000000000547.

Reinke, A. A. and Gestwicki, J. E. (2007) ‘Structure-activity relationships of amyloid beta-aggregation inhibitors based on curcumin: Influence of linker length and flexibility’, *Chemical Biology and Drug Design*. doi: 10.1111/j.1747-0285.2007.00557.x.

Ryu, E. K. *et al.* (2006) ‘Curcumin and dehydrozingerone derivatives: Synthesis, radiolabeling, and evaluation for  $\beta$ -amyloid plaque imaging’, in *Journal of Medicinal Chemistry*. doi: 10.1021/jm0607193.

Wei, G., Mousseau, N. and Derreumaux, P. (2007) ‘Computational simulations of the early steps of protein aggregation.’, *Prion*. doi: 10.4161/pri.1.1.3969.

Wu, C. *et al.* (2008) ‘The Binding of Thioflavin T and Its Neutral Analog BTA-1 to Protofibrils of the Alzheimer’s Disease  $A\beta$ 16–22 Peptide Probed by Molecular Dynamics Simulations’, *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2008.09.062.

Wu, C., Bowers, M. T. and Shea, J. E. (2011) ‘On the origin of the stronger binding of PIB over Thioflavin T to protofibrils of the Alzheimer Amyloid- $\beta$  peptide: A molecular dynamics study’, *Biophysical Journal*. doi: 10.1016/j.bpj.2011.01.058.

Xiong, Z. *et al.* (2011) ‘Curcumin mediates presenilin-1 activity to reduce  $\beta$ -amyloid production in a model of Alzheimer’s disease’, *Pharmacological Reports*. doi:

10.1016/s1734-1140(11)70629-6.

Yakupova, E. I. *et al.* (2019) ‘Congo Red and amyloids: History and relationship’, *Bioscience Reports*. doi: 10.1042/BSR20181415.

Yang, F. *et al.* (2005) ‘Curcumin Inhibits Formation of Amyloid Oligomers and Fibrils, Binds Plaques, and Reduces Amyloid in Vivo’, *Journal of Biological Chemistry*, 280(7), pp. 5892–5901. doi: 10.1074/jbc.M404751200.

Zhang-Nunes, S. X. *et al.* (2006) ‘The Cerebral  $\beta$ -Amyloid Angiopathies: Hereditary and Sporadic’, *Brain Pathology*, 16(1), pp. 30–39. doi: 10.1111/j.1750-3639.2006.tb00559.x.

Zhao, L. N. *et al.* (2012) ‘The Effect of Curcumin on the Stability of A $\beta$  Dimers’, *The Journal of Physical Chemistry B*. American Chemical Society, 116(25), pp. 7428–7435. doi: 10.1021/jp3034209.

# 6 Heparin-assisted amyloidogenesis uncovered through molecular dynamics simulations

**Copyright statement:** The contents of this chapter are submitted.

**Attribution:** Beenish Khurshid and Jamshed Anwar designed the research and wrote the paper. Beenish Khurshid conducted the simulations and performed the analysis. Simon Boothroyd generated the heparin structures of various lengths using an in-house code

## 6.1 Abstract

Glycosaminoglycans (GAGs) in particular heparan sulfate and heparin are found co-localized with A $\beta$  amyloid. They have been shown to enhance fibril formation, suggesting a possible pathological connection. We have investigated heparin's assembly of the KLVFFA peptide fragment using molecular dynamics simulation, to gain a molecular-level mechanistic understanding of how GAGS enhance fibril formation. The simulations reveal an exquisite process wherein heparin accelerates peptide assembly by first 'gathering' the peptide molecules, and then assembling them. Heparin does not act as a mere template but is tightly coupled to the peptides, yielding a composite protofilament structure. The strong intermolecular interactions suggest composite formation to be a general feature of heparin's interaction with peptides. Heparin's chain flexibility is found to be essential to its fibril promotion activity, and the need for optimal heparin chain length and concentration has been rationalized. These insights yield design rules (flexibility; chain-length) and protocol guidance (heparin: peptide molar ratio) for developing effective heparin-mimetics and other functional GAGs

## 6.2 Introduction

Amyloid deposits are characterized by hierarchically structured fibrils, each comprising helices or bundles of protofilaments composed of stacked  $\beta$ -strands of the protein (Rambaran and Serpell, 2008). The formation of amyloid is a general phenomenon being exhibited by some 30 different human proteins (Stathopoulos *et al.*, 2008). Notable examples include the peptide amyloid- $\beta$  (A $\beta$ ) linked with Alzheimer's disease, and the islet amyloid polypeptide (IAPP) associated with type II diabetes. Amyloid formation is also thought to play a functional role i.e. as a part of an essential physiological response, with evidence suggesting that the process can serve to sequestrate (rogue) peptides/proteins, store hormones, or modulate mechanical properties of cells (Fowler *et al.*, 2007).

In vivo, amyloid deposits invariably contain additional 'co-factors' that include lipids, nucleic acids, proteoglycans (PG), glycosaminoglycans (GAGs), serum amyloid P component, apolipoprotein E, collagen and metal ions (Stewart and Radford, 2017), reflecting the crowded, heterogeneous biological environment. The role of GAGs in amyloidosis is now considered to be significant, with much attention focusing on heparan sulfate and heparin and how they modulate the mechanistic and kinetic pathways.

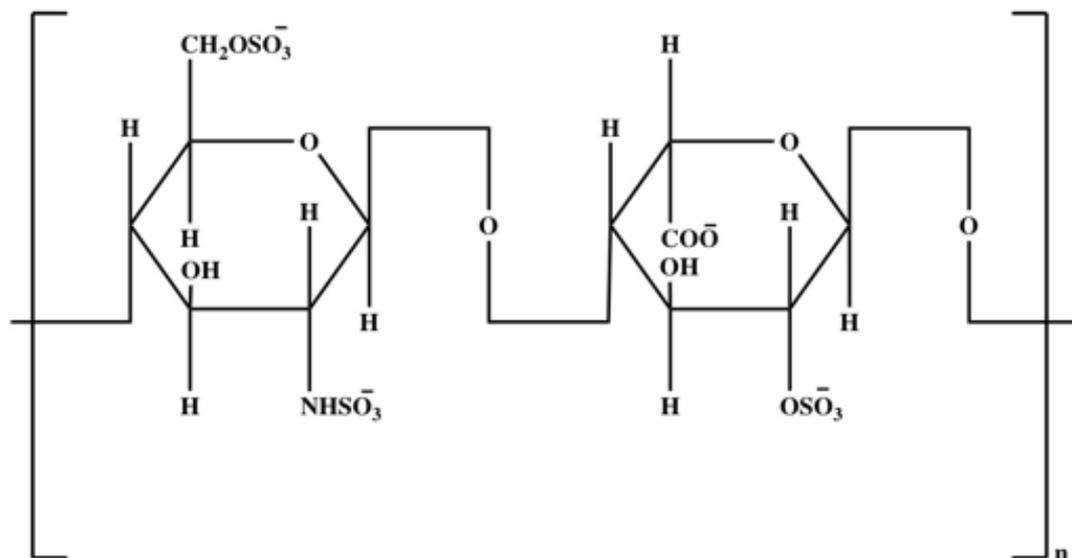
Heparin has been shown to enhance fibril formation of many amyloidogenic proteins *in-vitro* including amylin (Jha *et al.*, 2011), synuclein (Cohlberg *et al.*, 2002), transthyretin (Bourgault, James P Solomon, *et al.*, 2011), tau protein (Goedert *et al.*, 1996) and the intrinsically low-amyloidogenic peptide (PLB 1-23) acetylated cytoplasmic domain of the phospholamban transmembrane protein (Madine *et al.*, 2013). Heparan sulfate has been found to accelerate the oligomer formation of native human muscle acylphosphatase (mAcP) (Motamedi-Shad *et al.*, 2012). These enhancement effects can include a reduction in the lag phase of the nucleation step (Vilasi *et al.*, 2011), an increase in the rate and/or extent of elongation of the fibrils, and/or increase in fibril yield (Ramachandran and Udgaonkar, 2011), the exact nature of the enhancement effect being peptide-specific (Iannuzzi, Irace and Sirangelo, 2015).

The fibril enhancement role of GAGS offers an additional target for developing therapeutic agents for amyloid-based diseases. Inhibition of heparan sulfate biosynthesis in animal models can result in the complete loss of fibril formation and amyloid deposition (Elimova *et al.*, 2004; Suk *et al.*, 2006). Further, heparin has been

shown to convert toxic, soluble peptide oligomers into insoluble stable fibrils that are resistant to proteolysis, suggesting a neuroprotective role (Bergamaschini *et al.*, 2002; Bourgault, James P. Solomon, *et al.*, 2011). Indeed, GAG mimetics are being tested clinically for Alzheimer's disease, and as potential diagnostic agents for amyloid (Gervais *et al.*, 2001; Ma *et al.*, 2007).

A molecular-level mechanistic understanding of how GAGS enhances fibril formation is lacking, though essential for the further development of this field. It has been a challenge to develop a coherent mechanistic model given that GAGs can exert their action at multiple levels, from inducing  $\beta$ -sheet structure in monomers, facilitating the assembly of monomers and/or oligomers, (Suk *et al.*, 2006; Bourgault, James P Solomon, *et al.*, 2011) to enhancing extension of the protofilaments resulting in different fibril morphologies (McLaurin *et al.*, 1999). For heparan sulfate and heparin, it has been speculated that they serve as molecular scaffolds to facilitate fibril formation (Motamedi-Shad, Monsellier and Chiti, 2009). However, a recent study of heparin-facilitated aggregation of the neuropeptide  $\beta$ -endorphin has revealed that heparin does not act as a mere scaffold but is an integral component of the resulting fibril (Nespovitaya *et al.*, 2017). As to whether this could be a general phenomenon (i.e. not just restricted to  $\beta$ -endorphin) requires a better understanding of the nature and strength of the interaction of heparin with peptides.

Heparin is an unbranched, linear polymer of disaccharide units consisting of N-sulfated glucosamine (GlcNS) and iduronic acid linked by a (1–4) glycosidic bond (see Figure 6.1) (Dreyfuss *et al.*, 2009). The anionic, sulfate groups give heparin the highest negative charge density of any known biomolecule (Pita, 2015). The sulfates appear to be critical to heparin's fibril enhancement role, as their removal has been found to result in the loss of heparin's ability to promote the aggregation of  $A\beta_{40}$  and  $A\beta_{42}$ . (Castillo *et al.*, 2001).



**Figure 6.1. Chemical structure of the disaccharide repeating unit of heparin. The saccharide unit on the left is glucosamine (GlcN) whilst that on the right is iduronic acid (IdoA)**

Proteins and polyelectrolytes such as heparin form complexes primarily due to electrostatic interactions. These complexes vary in stoichiometry and architecture depending on several factors such as pH (Xu *et al.*, 2012), the charge on the protein, ionic concentration of the solution (Seyrek *et al.*, 2003), the degree of polymerization (dp) (Iannuzzi, Irace and Sirangelo, 2015), hydrophobicity (Sofronova *et al.*, 2017), stiffness and flexibility of the polyelectrolyte (Cooper *et al.*, 2005). Generally, aggregation is enhanced with increasing GAGs chain length, although for some proteins, the effect on fibril assembly plateaus when the chain length becomes sufficiently long ( $\text{dp} \gg 18$ , where dp is the number of saccharide units) (Jha *et al.*, 2011; Takase *et al.*, 2016). Short-length GAGs with  $\text{dp} \leq 4$  are essentially ineffective in accelerating fibrillogenesis. In contrast, medium-length GAGs, dp6 – dp12, can significantly reduce the lag phase and accelerate the conversion of oligomeric species into ordered fibrillar assemblies (Fraser, Darabie, and McLaurin, 2001; Quittot, Sebastiao and Bourgault, 2017). With respect to the molar heparin to peptide concentration ratio, a study on the prion-related protein fragment PrP, demonstrated that relatively low heparin to peptide ratios ranging from 0.2:1 to 0.5:1 accelerated fibril formation. At even lower molar ratios (0.1:1 and below) the effect is less pronounced. Higher molar ratios (5:1, 2:1, and 1:1) of heparin to peptide were found to inhibit peptide aggregation (Bazar and Jelinek, 2010).

Here we present a first clear perspective of how heparin interacts with KLVFFA peptide to enhance fibrillogenesis using unbiased molecular dynamics simulations. The simulations reveal that heparin does not act as a mere template but is tightly coupled to KLVFFA peptides, yielding a *composite* protofilament structure. The strong intermolecular interactions suggest composite formation to be a general feature of heparin's interaction with peptides. Also, heparin's flexibility is found to be essential to its fibril promotion activity, and we have rationalized the need for optimal chain length and heparin: KLVFFA peptide concentration ratio. These insights will underpin therapeutic approaches including the design of more effective GAG mimetics for chelating and minimizing toxicity of oligomers, and for modulating amyloidosis, both its inhibition and promotion.

### 6.3 Methodology

A significant issue with molecular simulation is the limited timescales (at best a few microseconds) that are accessible. Indeed, the self-assembly of multiple units of the full A $\beta$  into an ordered structure is outside the timescales of unbiased molecular dynamics (MD) simulation. Given this, we investigated the heparin-promoted assembly of the KLVFFA fragment peptide (lysine-leucine-valine-phenylalanine-phenylalanine-alanine), which is a stretch of hydrophobic amino acids (residues 16-21) in A $\beta$ , rather than the full A $\beta$ . The KLVFFA peptide is the shortest fragment for which experimental evidence of amyloid formation is available (Antzutkin, 2004). Solid-state NMR shows that this fragment acts as a prototype for aggregation, forming antiparallel strands leading to the formation of fibrils (Miravalle *et al.*, 2000; Ma and Nussinov, 2002; Baumketner *et al.*, 2006; Liang *et al.*, 2014).

Nine distinct sets of simulations were carried out: (i) simulation of a single heparin molecule; (ii) the interaction of a single KLVFFA monomer (PDB 2Y2A) (Colletier *et al.*, 2011) with a single heparin molecule (flexible and restrained in its extended form); (iii) the self-assembly of the KLVFFA peptides alone, which served as a control; (iv) the assembly of the KLVFFA peptides in the presence of heparin; (v) assembly of the KLVFFA peptides in the presence of a heparin molecule that was restrained in its extended conformation to investigate the role of flexibility of heparin; (vi) the effect of heparin: peptide concentration on KLVFFA assembly; (vii) the effect of heparin chain length (degree of polymerization) on heparin-promoted KLVFFA assembly. (viii) the

self-assembly of the KLVFFA *dimers* (PDB 3OW9) (Colletier *et al.*, 2011) alone; and (ix) the assembly of the KLVFFA *dimers* in the presence of heparin. The full complement of simulations carried out is listed in Table 1.

Given that the heparin molecule is significantly charged, the ionic strength of the solvent environment is likely to be critical as it would modulate the Columbic inter-molecular interactions. We, therefore, conducted the above simulation studies in a low ionic strength environment in which just sufficient  $Na^+$  counterions were added to balance the net charge of the system, and a physiological ionic strength environment that contained additional  $Na^+$  and  $Cl^-$  ions to yield a 150 mM NaCl.

Finally, on noting that pre-formed  $A\beta$  KLVFFA dimers aggregated faster than monomers (Munter *et al.*, 2007; Portillo *et al.*, 2015), we explored the assembly of pre-formed KLVFFA peptide dimers alone and in the presence of heparin (simulation studies (xiii) and (ix)). The simulations of the KLVFFA monomers showed that there are kinetic barriers to the monomer organization once the molecules become locked onto heparin. The necessary re-alignment to get anti-parallel  $\beta$ -sheet arrangements become the rate-limiting step, delaying the peptide organization process beyond the timescales of unbiased MD. In contrast, the pre-formed dimer simulations revealed rapid fibrillation of the peptides to yield partially ordered aggregates, revealing the generic features of the ordering process.

The structure of heparin dp24 (PDB 3IRJ) (Khan *et al.*, 2010) was taken from the protein data bank. The structures of shorter heparin fragments were based on heparin 18-mer (PDB 3IRI) (Khan *et al.*, 2010) with the topologies being obtained from the PRODRG server and reviewed for accuracy. The charges were assigned from literature, which had been calculated using ab-initio calculations at 6-31G\*\* (Verli and Guimarães, 2004). Unless otherwise indicated, the heparin molecule comprised 24 saccharide units (dp=24). In the self-assembly and heparin-facilitated assembly simulations, there were 20 KLVFFA monomers. Each of the assembly simulations were run with 3 replicates. The initial set of coordinates for each was randomized, with the KLVFFA peptide molecules being placed randomly around the heparin molecule to ensure that the process kinetics were not biased by the choice of starting coordinates.

To investigate whether flexibility and bending of the heparin molecule is essential for its assembling role, we simulated the facilitated peptide-assembly with the heparin

molecule position-restrained to remain in the extended form using restraints on all of the atoms with a force constant of 500 kJ mol<sup>-1</sup> nm<sup>2</sup>.

**Table 6.1. Details of simulation studies carried out. This entire complement of simulations was repeated in a low ionic strength (system neutralized with counterions) and in a physiological ionic strength aqueous environment (150 mM NaCl). Unless otherwise stated the heparin chain length in terms of the degree of polymerization was dp = 24.**

Study No.	Study details	No. of heparin molecules	No. of KLVFFA peptide(s)
(i)	Heparin alone in water	1	-
(ii)	Heparin's interaction with single KLVFFA	1	1
(iii)	Self-assembly of KLVFFA (alone) in water	-	20
(iv)	KLVFFA assembly in the presence of heparin	1	20
(v)	KLVFFA assembly in the presence of restrained heparin in extended conformation	1	1
(vi)	Effect of heparin: peptide molar concentration on KLVFFA assembly	1	20
		2	20
(vii)	Effect of heparin chain length on KLVFFA	dp2 (12)	20
		dp4 (6)	20
		dp6 (4)	20
		dp8 (3)	20
		dp24 (1)	20
(viii)	Self-assembly of pre-dimerized KLVFFA in water	-	10
(ix)	Assembly of pre-dimerized KLVFFA in the presence of heparin	1	10

Systems investigating the effects of different molar ratios of heparin: peptide on the assembly of KLVFFA peptides contained 1:20, 1:100 and 2:20 heparin: peptide molecules. For the effect of heparin chain length on heparin's propensity to promote KLVFFA peptide assembly, the heparin fragments investigated included chain lengths of 2, 4, 8, 12, and 24 saccharide units (dp2, dp4, dp8, dp12 and dp24).

All the simulations were carried out at 360K and pressure of 1 bar using Gromacs 5.0.4 (Van Der Spoel, Lindahl, Hess, Groenhof, Alan E Mark, *et al.*, 2005). The higher temperature serves to accelerate potential conformational and configurational transitions of the peptide assembly, an approach employed by the others to enhance sampling (Cino, Choy and Karttunen, 2012). The systems were equilibrated in the NVT ensemble and then in NPT ensemble using a velocity-rescaling thermostat (Bussi, Donadio and Parrinello, 2007) and the Berendsen barostat (Berendsen *et al.*, 1984) to control the temperature and pressure respectively. In the production run, the Parrinello–

Rahman barostat was used with coupling time of 2ps and an isothermal compressibility of  $4.5 \times 10^{-5} \text{ bar}^{-1}$  (M. Parrinello and Rahman, 1981). The timestep used was 2.0 fs. PME was used to calculate long-range electrostatics with a grid spacing of 0.16 nm, and a real-space cutoff of 1 nm (Di Pierro, Elber and Leimkuhler, 2015). Both Van der Waals and neighbor list cutoffs describing short-range interactions were set to 1.0 nm. The simulations were run for up to 200ns of trajectory unless stated otherwise. Heparin and the KLVFFA fragments monomer (PDB 2Y2A) and dimers (PDB 3OW9) with the ends uncapped were modeled using the Gromos96 53a6 force field as employed and validated by others (Pol-Fachin and Verli, 2008; Pol-Fachin, Fernandes and Verli, 2009; Pol-Fachin *et al.*, 2012; Wei *et al.*, 2014). The analyses of the simulation trajectories were performed using Gromacs analysis tools and VMD (Humphrey, Dalke, and Schulten, 1996). For example we used the gmx clustsize tool to follow the process of self-assembly in terms of the number of aggregates/clusters and size of the largest aggregate present in solution. The free energy of KLVFFA peptide binding to heparin was estimated using the MM-PBSA method (Kumari, Kumar, and Lynn, 2014).

## **6.4 Results and discussion**

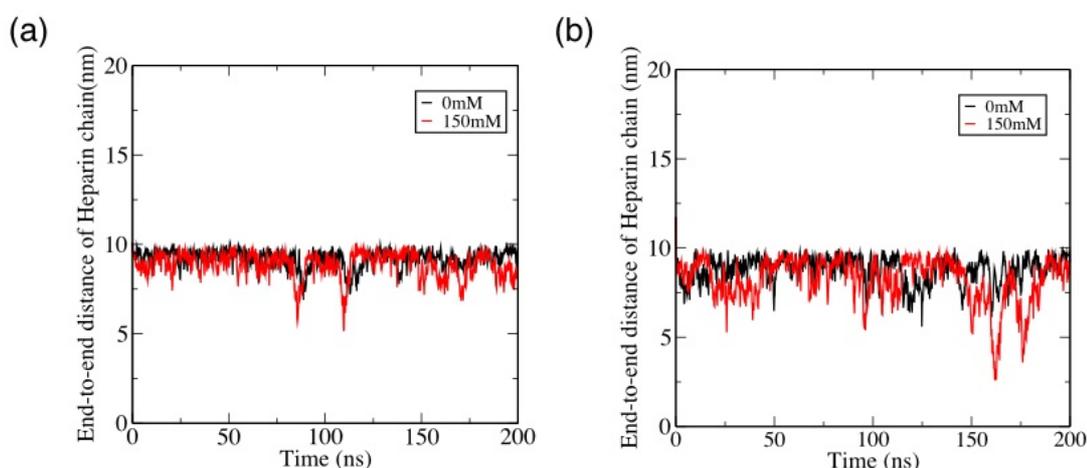
KLVFFA interacts strongly with heparin and is unable to move up or down the heparin molecule

The simulations reveal that heparin by itself in a low ionic strength environment exists in an extended chain form with an average end-to-end distance of around 9.2 nm, with a limited propensity to flex or coil. This is expected as the charged sulfate and carboxylate moieties are content in the polar aqueous environment and repulsion between the charged groups is also likely to encourage the extended form. In the higher ionic strength environment, this distance reduces slightly to an average of 8.8 nm. It appears that the repulsion between the charged heparin moieties is screened by the increased ionic strength enabling the heparin to adopt a less extended chain (Figure 6.2a).

We consider now the inclusion of a KLVFFA monomer into both systems (low and high ionic strength), whilst also investigating the role of flexibility of the heparin molecule. For both the flexible and the restrained extended-form of heparin, the KLVFFA peptide quickly attaches to one of the anionic sites on heparin through its

lysine residue. For the restrained (extended) heparin at both ionic strengths, the KLVFFA monomer, once attached, resides essentially at the same anionic site throughout the simulation (see Figure 9.11 b & d). This confirms that heparin does not act as a track on which a KLVFFA peptide can readily move up or down. For the case of heparin in its fully flexible form, KLVFFA remains localized on the heparin chain at high ionic strength but can explore a much larger displacement range at low ionic strength (Figure 9.12). This occurs as a result of the flexing of the heparin molecule, which enables the KLVFFA molecule to detach from one location to another as another strongly interacting site on the heparin molecule comes in close proximity to the KLVFFA peptide. Thus, it appears that the conformational flexibility of heparin is essential to its facilitation of the assembly of A $\beta$  peptides (see also later discussion). The attachment of the single KLFVAA monomer to heparin doesn't produce any significant change in heparin's conformation: the average end-to-end distance reduces from 9.2 nm to 8.8 nm for the low ionic strength system, and from 8.8 nm to 8.3 nm for the high ionic strength media (Figure 6.2b).

MM-PBSA calculations show that the binding of the KLVFFA monomer to heparin is strong, the binding free energy being  $-122.7 \pm 24$  kcal/mol in the low ionic strength environment and as expected (due to charge screening) lower  $-92.1 \pm 19$  kcal/mol in the high ionic strength system (see a detailed breakdown in Table 6.2). The major contribution to the total binding energy comes from the electrostatic interactions between the lysine residues of the KLVFFA and negative groups on the heparin (see Figure 9.13).



**Figure 6.2. End to end distance analysis of heparin chain at two ionic strengths in an aqueous environment. (a) Heparin chain alone. (b) Heparin in the presence of a single KLVFFA peptide.**

**Table 6.2. A breakdown of the free energy of binding  $\Delta G$  of KLVFFA peptide with heparin in low and high ionic strength (150mM NaCl) aqueous environments calculated by the MM-PBSA method**

Energy component	$\Delta G$ / kcal mol <sup>-1</sup>	$\Delta G$ / kcal mol <sup>-1</sup>
	0 mM NaCl	150 mM NaCl
van der Waals energy	-15.5	-23.0
Electrostatic energy	-254.2	-204.2
Polar solvation energy	149.42	138.5
SASA <sup>1</sup> energy	-2.5	-3.5
<b>Total binding energy</b>	<b>-122.7 ± 24</b>	<b>-92.2 ± 19</b>

<sup>1</sup> SASA (solvent accessible surface area) used to calculate  $G_{\text{nonpolar}}$  energies in mm-pbsa method

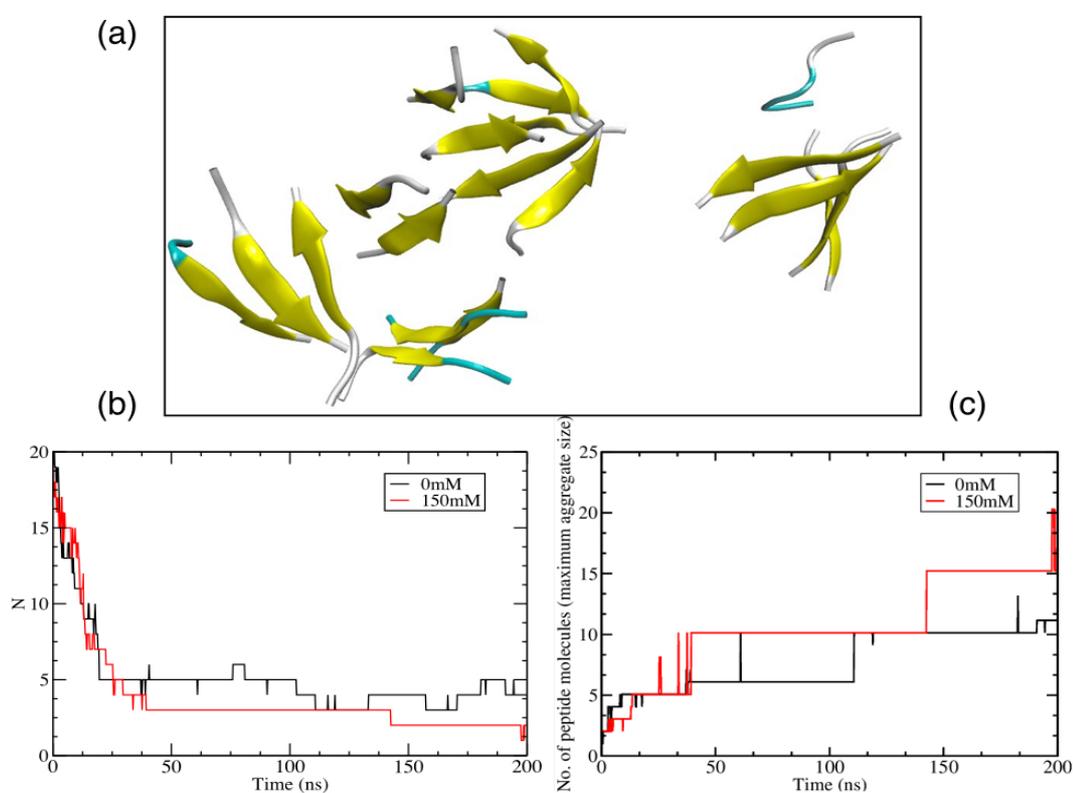
KLVFFA peptides aggregate to form ordered oligomers but not a fibrillar structure in 200 ns of MD simulation

The self-assembly of KLVFFA peptide monomers without heparin in both low and high ionic strength aqueous media results in the formation of 2-3 ordered aggregates but not a contiguous fibrillary structure (Figure 6.3). Oligomeric structures form relatively rapidly over a period of about 45ns, and then aggregate further to yield 2-3 larger structures. On formation, the oligomers begin to order internally to form antiparallel  $\beta$ -sheets. The rapid aggregation of the KLVFFA peptides is driven by strong attraction resulting from the hydrophobic effect. The aggregation is slightly faster (see Figure 6.3b) and the aggregate size is bigger (see Figure 6.3c) in the higher ionic strength system, which would be expected (and has been observed by others) (Hoyer *et al.*, 2002; Yun *et al.*, 2007) given that the chemical potential of the hydrophobic KLVFFA would be enhanced in a stronger ionic environment.

Aggregation of KLVFFA peptides on heparin takes place in two steps: selection and assembly

Considering heparin-facilitated assembly of KLVFFA peptides, in all six simulations (3 at low ionic strength and 3 at high ionic strength, each beginning from a different starting configuration), we note that there are two distinct stages: a rapid selection and localization of the peptides onto the heparin framework, followed by a cooperative

process of heparin-induced assembly of the KLVFFA peptides. The peptide localization onto heparin appears to compete with peptide-peptide interactions with a few of the peptides self-aggregating (mostly as dimers) in isolation, comprising anti-parallel  $\beta$ -sheets. In due course, these aggregates also become attached to the heparin. The isolated peptide-peptide aggregates occur mostly in the high-ionic strength simulations, presumably because the high ionic strength screens the charge-charge interaction between the KLVFFA peptides and heparin whilst also increasing the chemical potential of the hydrophobic KLVFFA peptides, thus favoring peptide-peptide interaction.

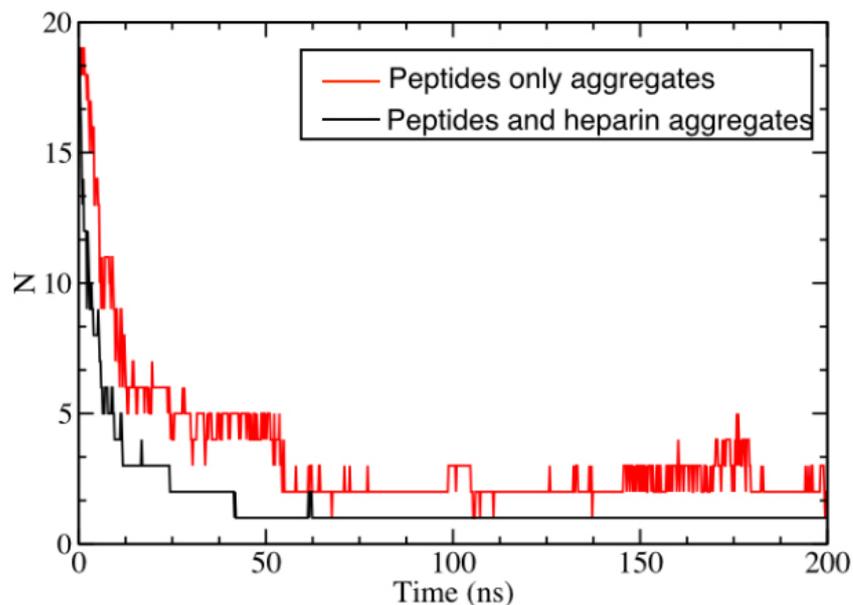


**Figure 6.3. (a) Snapshot of the self-assembly of KLVFFA peptides (monomers) at 200 ns simulation time in 150 mM NaCl in water. (b) Number of peptide aggregates formed during the self-assembly of KLVFFA peptides as a function of time for both the 0 mM and 150 mM ionic strength systems. (c) Maximum aggregate size for the self-assembly of KLVFFA peptides as a function of time for both the 0 mM and 150 mM ionic strength systems. (Cluster criteria: cut-off of 0.35 nm).**

The heparin-induced assembly process (second stage) involves bending and unbending of heparin, carrying along with it the adsorbed KLVFFA peptide units that re-align and attach to other peptide units, all the while enhancing the development of a more coherent protofilament. The emerging protofilament remains integrated with heparin throughout the process.

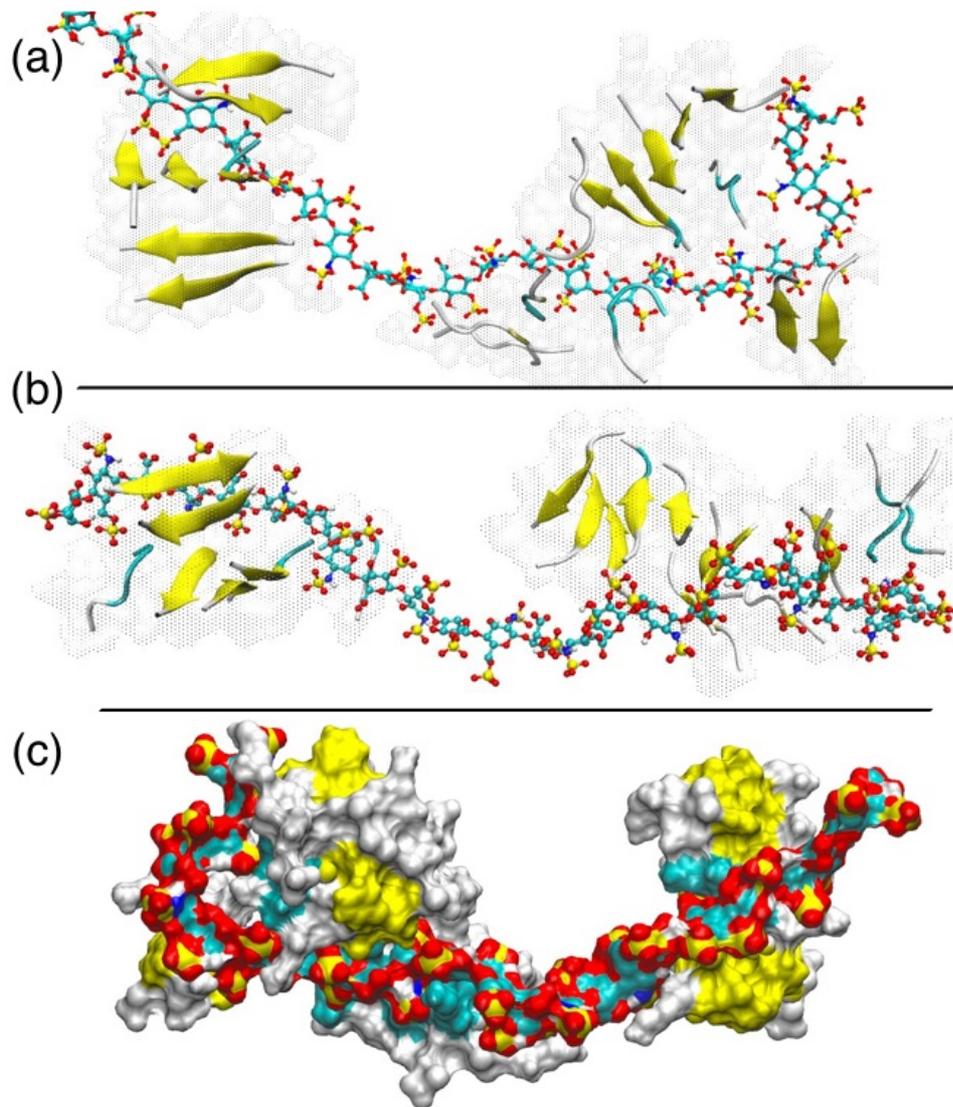
The heparin-facilitated assembly in general is faster relative to the *self*-assembly of KLVFAA peptides but shows variability that depends on the starting coordinates and is further modulated by the ionic strength of the media.

At the lower ionic strength, the peptides aggregate onto the heparin to form two intertwined helices of the stacked  $\beta$ -sheets and the heparin (Figure 6.5a, b & c), the final structure generally being extended. The aggregation of the peptides in the presence and absence of heparin is compared in Figure 4. The number of aggregates inclusive of heparin and KLVFFA peptides converges to a single structure within the time period 25-45ns in all three simulations, reflecting the rapid localization of the KLVFFA molecules onto the heparin. The peptides Only clustering (reflecting the development of a more contiguous peptide structure, albeit whilst integrated with the heparin) is a little slower and is characterized by numerous fluctuations that result from the flexing of the heparin molecule and the attachment/detachment of the peptides in a bid to form a more contiguous structure.



**Figure 6.4. Evolution of the number of aggregates inclusive of heparin and peptides and peptides-only. The plot of peptides and heparin aggregates (black) reflects the initial, rapid association of the peptides with the heparin (within about 45 ns), whilst the peptides-only aggregates plot (red) reflects the relatively slower, facilitated assembly of the peptides. Note the numerous, repetitive peptide-peptide cluster making and breaking events as the heparin flexes to align and order the emergent protofilament.**

At the higher ionic strength, the resulting peptide-heparin composite shows a greater variation in morphology (see Figure 6.6) ranging from annular to linear structures. In one instance, the heparin molecule with attached KLVFFA peptides folds itself into a ring-like structure. Additionally, in these systems, we also observe the formation of isolated KLVFFA peptide only aggregates, some of which eventually attach to the heparin molecule. As noted earlier, the greater charge screening of the higher ionic strength enables heparin to sample more retracted conformations, and also increases the chemical potential of the hydrophobic peptides thus encouraging peptide-peptide cluster formation alongside localization of the KLVFFA peptides onto the heparin molecule. Here we see both effects playing a role.

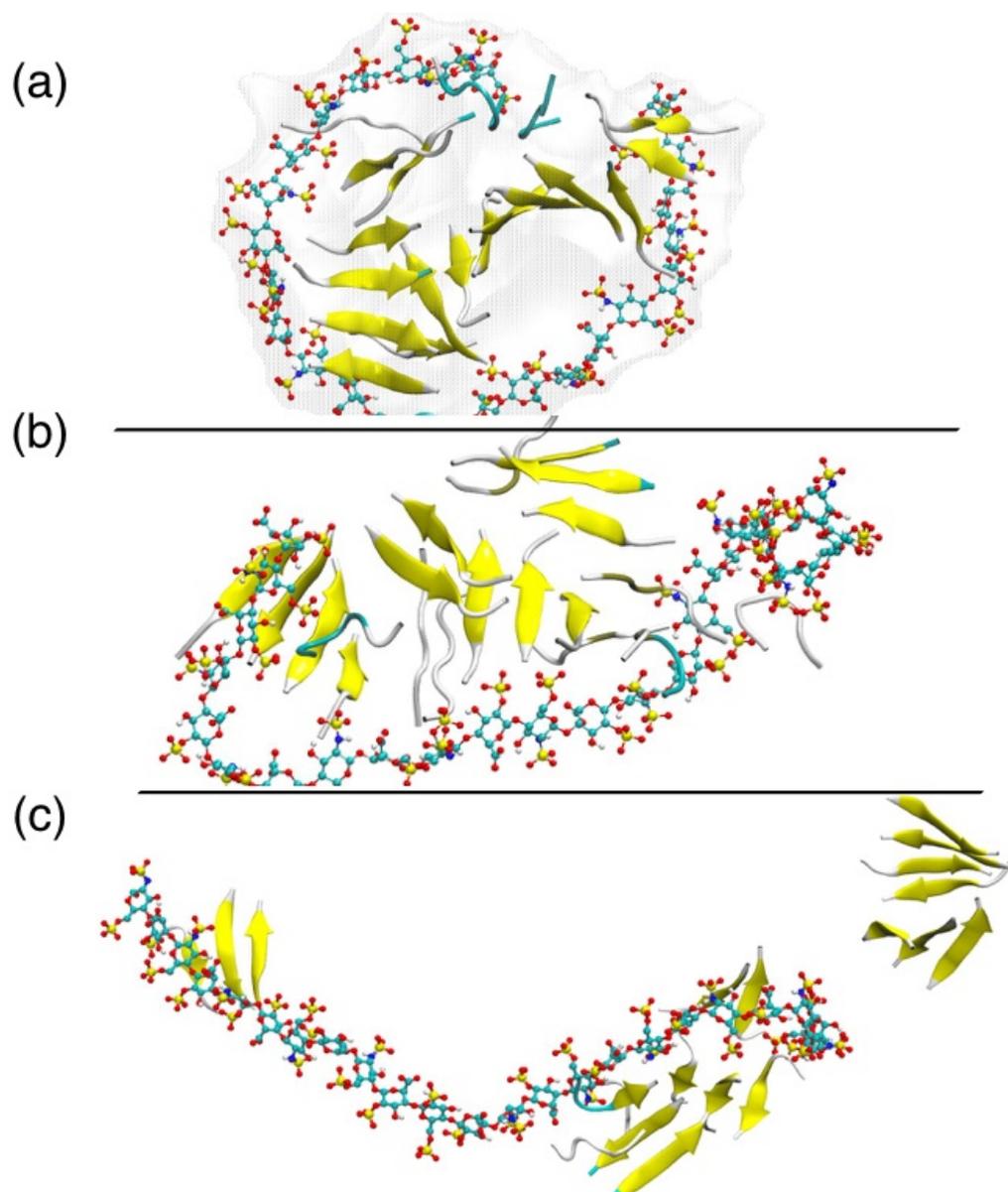


**Figure 6.5.** (a) Snapshots of KLVFFA peptides–heparin structures in 0 mM NaCl at 200 ns for three separate simulations each started from a different random configuration. (b) Snapshot of the complex from the second simulation system (c) Snapshot of the complex

**from the third simulation but with a “surface” representation to show the formation of a contiguous structure. Heparin and peptides are shown in CPK and cartoon (secondary structure) representation respectively.**

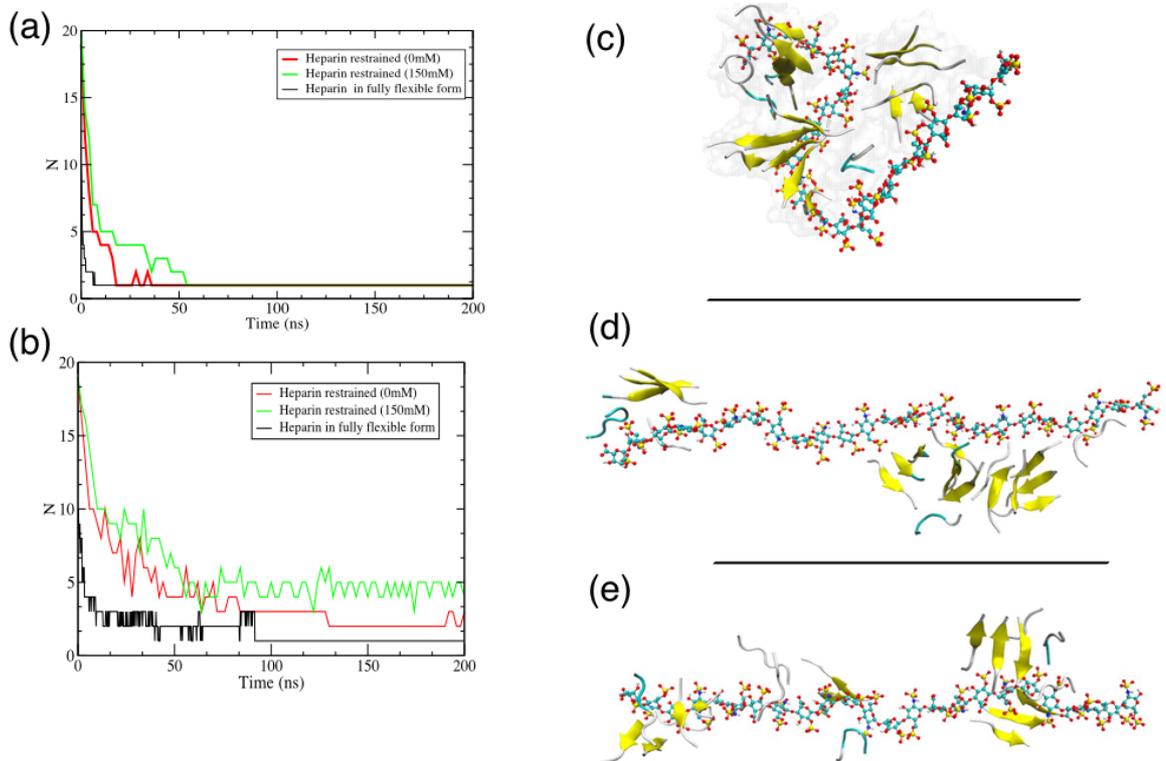
Heparin flexibility appears to be essential for its role in assembling peptides

To investigate whether the flexibility and bending of the heparin molecule are essential for its assembling role, we simulated the heparin-facilitated KLVFFA peptide assembly with the heparin molecule restrained in its fully extended form. The KLVFFA peptides quickly localized onto the heparin framework (see Figure 6.7a) but there was no subsequent organization or assembly of peptides, only some local ordering of the attached oligomers at higher ionic concentration (see Figure 6.7b; peptides only aggregates & Figure 6.7e,f). This is in contrast to the assembly process involving a fully flexible heparin molecule, which yields a more contiguous peptide protofilament albeit integrated with heparin (see Figure 6.7c).



**Figure 6.6. Snapshots of KLVFFA peptides–heparin structures in 150 mM NaCl at 200 ns for three separate simulations each started from a different random configuration, showing wide variation in the emergent morphology**

These results clearly illustrate that heparin's flexibility is essential to its role in facilitating amyloid formation. The peptide-heparin interaction is strong and limits the movement of the peptide once it is adsorbed onto the heparin. The idea of the heparin framework providing a strongly interacting but an iso-potential surface along which the peptide molecules can freely translocate (enabling it to form a contiguous structure) is appealing, but unfortunately is not borne out.

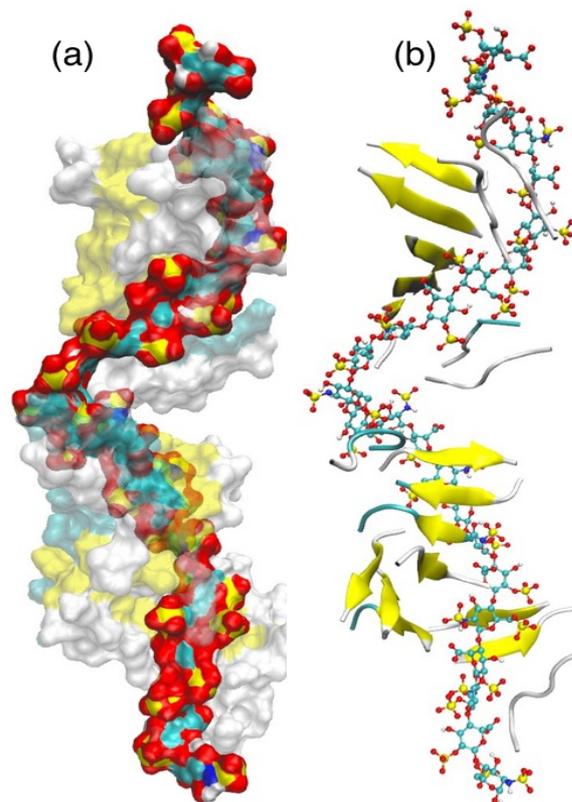


**Figure 6.7. (a) Total number of aggregates (including KLVFFA peptide aggregates and heparin) as a function of time at both ionic concentrations, reflecting rapid localization of the peptide molecules onto the heparin. (b) Number of peptides-only aggregates as a function of time at both ionic concentrations. The plot shows that peptides assemble into a single structure when heparin is fully flexible within 100ns whilst there are still 3-4 loose aggregates when heparin is restrained in its extended state. (c) Snapshots of the heparin-peptides composite for the heparin molecule when fully flexible at 0 mM. (d, e) Heparin restrained in its fully extended form at 0 mM and 150 mM NaCl. For these restrained heparin systems, the lack of a contiguous peptide structure is apparent.**

Instead, the formation of a contiguous protofilament structure comes via peptide detachment/attachment that occurs as the heparin molecule flexes, bringing peptide molecules located at one site of the heparin molecule into the interaction zone of another.

Given that heparin's preference, when alone, is to be in its extended form, what induces the heparin molecule to flex and bring the attached peptide units together? Whilst the origin of this force is difficult to disentangle, the interactions suggest that it is a combination of strong, attractive hydrophobic force between the (hydrophobic) peptide

units, and electrostatics. Unlike valence bonds, electrostatic interactions do not become saturated. Hence, the strong electrostatic interactions responsible for localizing the peptide units onto the heparin, are also able to act between an attached peptide and other regions of the heparin. It appears that the combined hydrophobic effect and electrostatic interactions are sufficient to overcome the repulsion between the anionic sulfate moieties on the heparin and any elastic energy penalty associated with its bending.



**Figure 6.8. The KLVFFA peptide-heparin composite resulting from heparin's assembly of KLVFFA peptides. Both the formed peptide protofilament and the heparin form helical structures that are intertwined. The heparin helix is characterized by a pitch consisting of four disaccharides with a translation of 2.0–2.1nm along the axis. (a) Heparin is represented as VDW beads (sulfur in yellow; oxygen in red; carbon in turquoise), whilst the A $\beta$  peptide protofilament is represented in surface form. (b) represents peptides' secondary structure intertwined against heparin shown in CPK representation.**

Heparin-peptide interaction is not transitory; heparin forms an integral part of the resulting protofibrils

It is well established that GAGs, and in particular heparan sulfate, are found co-localized with A $\beta$  amyloid (Zhang *et al.*, 2014). This begs the question of whether the

GAGs are an integral component of the fibrils or merely just co-localize within the amyloid mass. Indeed, a recent study provides evidence that heparin forms an integral component of the emerging fibrils for the endogenous opiate  $\beta$ -endorphin.

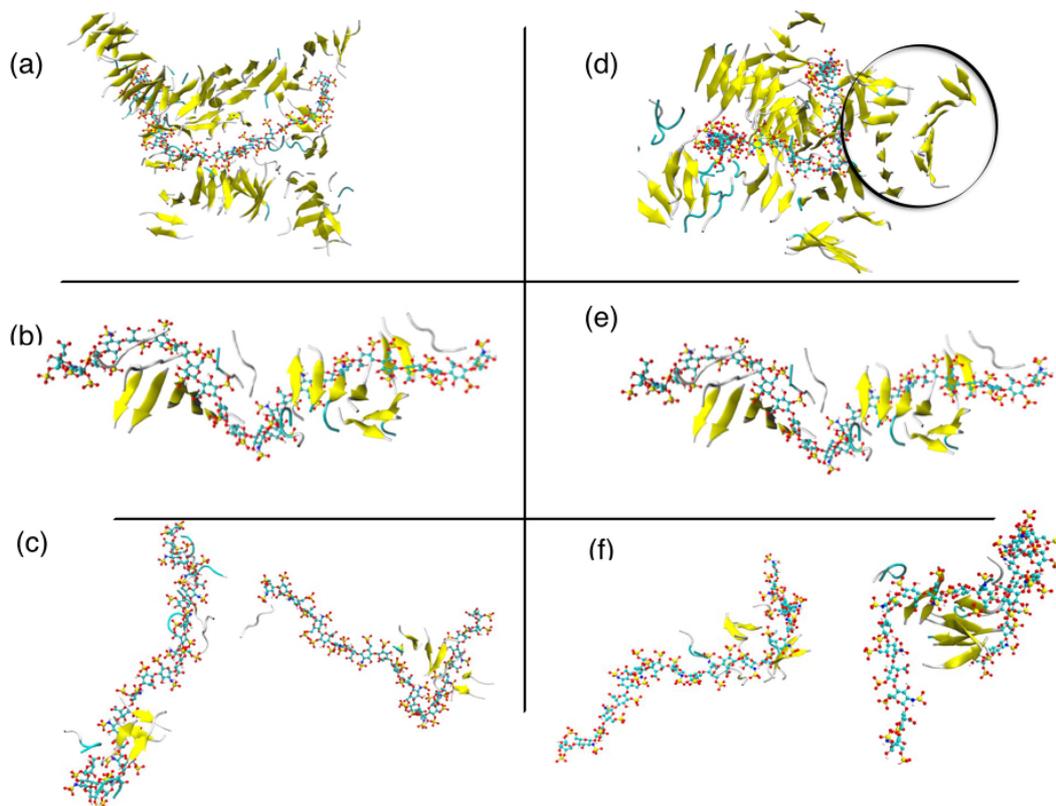
As noted above, the current simulations also reveal that heparin does not act as a mere template but is tightly coupled to the KLVFFA peptides, yielding a composite protofilament structure consisting of intertwined helices of the stacked  $\beta$ -sheets and the heparin (Figure 6.8). The binding energy between the two helices (comprising 20 KLVFFA peptides and 1 heparin molecule) was calculated to be around  $-2000\text{kcal/mol}$ , which equates to  $> 3500 k_{\beta}T$  (thermal energy at  $T=310\text{ K}$ ). Given this huge value and the intertwining nature of the two helices, there is little or no possibility of their (un-aided) dis-entanglement. However, any unwinding of the helices is likely to occur stepwise, beginning from one of the ends. The interaction energy between a single peptide and heparin ranges from  $-90\pm 19\text{ kcal/mol}$  to  $-120\pm 25\text{ kcal/mol}$  depending upon the ionic strength of the solution, whilst the inter-peptide interaction (that between the anti-parallel peptide strands) is about  $-20\text{ kcal/mol}$ . Therefore, the lowest bound for detaching the peptide strand, one residue at a time, from the composite would be  $-90\text{ kcal/mol}$  and for stripping away altogether an individual peptide molecule from the end would be  $-110\text{ kcal/mol}$ . These values equate to  $>150\text{ kT}$ , again pointing to an extremely low probability for the unwinding/detachment process. Furthermore, any stepwise disentanglement would require coordinated molecular displacements.

Given the tendency of the KLVFFA peptides and heparin to form inter-twined helices coupled with the strong, electrostatic heparin-peptide interaction, we infer that the formation of composite heparin-peptide protofilaments is a general phenomenon.

Lower heparin: peptide concentration ratio is more effective because peptide monomers do not get partitioned amongst individual heparin molecules

We investigated how the aggregation kinetics depend upon the molar concentration ratio of heparin to KLVFFA peptide. Simulations containing the low molar ratio of 1 heparin molecule:100 KLVFFA peptides (monomers) revealed very fast peptide aggregation and structuring. At low ionic strength, this results in a single dense fibrillar structure that appeared to comprise two protofilaments integrated around the heparin

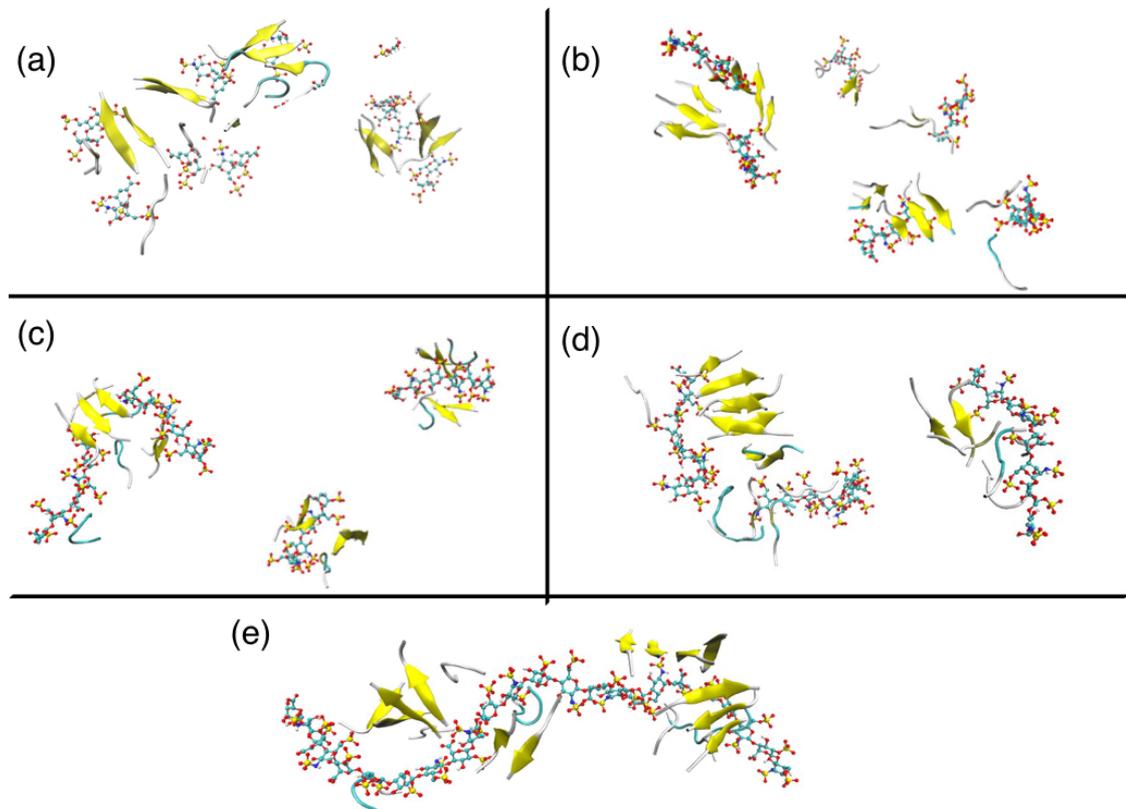
molecule (Figure 6.9a). At high ionic strength, there were two distinct pathways: formation of a heparin-peptide structure and ordered filamentous peptide-only structures (Figure 6.9d). As stated earlier, the high ionic strength enhances the hydrophobic peptide-peptide interactions whilst screening charge-charge interactions, thereby facilitating the development of peptide-only protofilaments.



**Figure 6.9. Assembly of KLVFFA peptides at various heparin: peptide molar ratios at 0mM (Left-hand side) & 150mM NaCl (Right-hand side). (a,d) 1:100; (b,e) 1:20; (c, f) 2:20 heparin: peptide molar ratio. The encircled structure in (d) shows the formation of a well-ordered fibril at 150mM.**

At a higher molar ratio of 1 heparin:20 KLVFFA peptides, for both ionic strength environments, all of the peptides assembled around the heparin giving rise to an integrated structure (see Figure 6.9b & e). At a still higher molar ratio of 2 heparin molecules:20 KLVFFA peptides (i.e. 1:10), the peptide units apportioned themselves amongst the 2 heparin molecules, from which they did not dissociate and hence were unable to form a contiguous structure (Figure 6.9c & f). Each sub-population attached

to a particular heparin molecule, being small in number can only yield at best a small local oligomeric peptide structure.



**Figure 6.10. The effect of different heparin chain lengths on heparin-facilitated assembly of KLVFF peptides at 0 mM. (a) dp2, (b) dp4, (c) dp6, (d) dp8 and (e) dp24; where dp is the number of saccharide units. Longer heparin chain lengths (dp=8 and dp=24) are more effective at facilitating peptide assembly. Shorter heparin chain lengths (dp=4 and dp=6).are less effective as they form multiple heparin-peptide complexes that then aggregate only slowly. The disaccharide (dp=2) complexes show a marginal reversal in trend by aggregating together, courtesy of their faster diffusion rate that enhances complex-complex interactions.**

These results are consistent with experimental observations and are intuitive (Bazar and Jelinek, 2010; Takase *et al.*, 2016; So *et al.*, 2017). At lower heparin: KLVFFA peptide ratio i.e. as the relative number of peptides is increased, an individual heparin molecule can attract a full complement of peptide units, the number of peptides being commensurate with the available sites on the heparin and then heparin facilitates their assembly into a protofilament structure. We surmise that at very low heparin: peptide ratios, the relatively few heparin molecules will facilitate peptide assembly, but their effect will be localized and not impact much on the bulk of the solution.

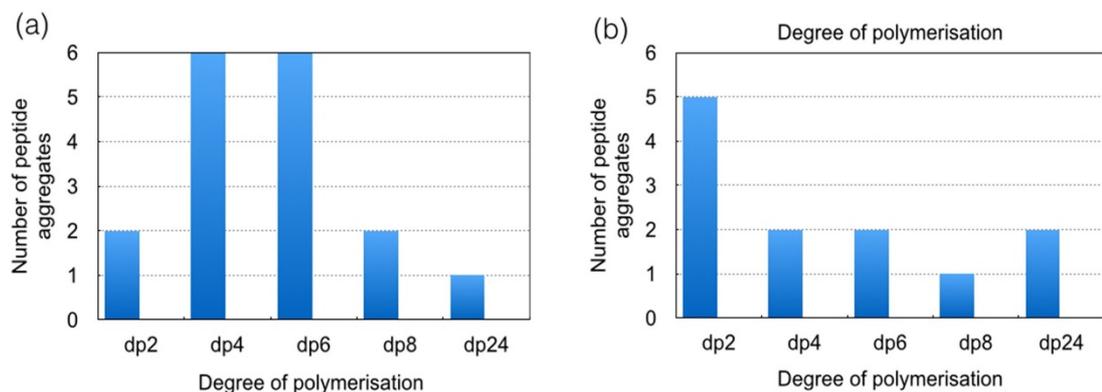
Shorter heparin chain lengths are less effective as they partition the peptides to form multiple heparin-peptide complexes that aggregate only slowly

The simulations reveal that peptide assembly (towards the development of a contiguous aggregate structure) is facilitated by the longer heparin chain length of dp24 (i.e. 24 saccharide units) at lower ionic strength. Heparin chain lengths of dp4, dp6 and dp8 did not yield a contiguous peptide structure that includes all of the peptide molecules, as the peptide molecules get apportioned amongst the individual heparin units that then do not in general show any significant tendency to come together (see Figure 6.10). For dp8, the KLVFFA peptides did yield an ordered protofilament that spanned two of the heparin dp8 units.

In contrast, the shortest chain length studied, the disaccharide heparin dp2, shows two distinct behaviors: promotion of peptide assembly for the low ionic strength environment, and little or no effect in the high ionic strength environment. The promotion at low ionic strength, whilst unexpected, does not contradict any experiments as dp2 fragments have never been studied experimentally in this context. The heparin dp2-peptide complexes formed quickly and then aggregated to form 1-2 clusters followed by some ordering of the KLVFFA peptides. It appears that being small units, further aggregation of these dp2-peptide complexes can be driven by the complex peptides courtesy of the hydrophobic effect. In principle, such a potential must also exist for the larger heparin units but will be counteracted by their greater inertia due to their size. Indeed, estimated diffusion coefficients of the various heparin chain lengths in pure water from simulation bear this out:  $D(\text{dp}2) = 4.3 \pm 1.8 \times 10^{-5} \text{ cm}^2/\text{s}$ ;  $D(\text{dp}4) = 4.8 \pm 2.0 \times 10^{-5} \text{ cm}^2/\text{s}$ ;  $D(\text{dp}6) = 4.0 \pm 3.0 \times 10^{-5} \text{ cm}^2/\text{s}$ ;  $D(\text{dp}8) = 3.1 \pm 1.7 \times 10^{-5} \text{ cm}^2/\text{s}$ ;  $D(\text{dp}24) = 0.4 \pm 0.3 \times 10^{-5} \text{ cm}^2/\text{s}$ . In the high ionic strength environment, the shorter heparin chains dp2 was ineffective in accelerating the aggregation. While larger chains dp4-dp18 significantly increase the aggregation of peptides giving rise to larger peptide aggregates. dp24 doesn't enhance further aggregation (Figure 9.14). The extent of formation of peptide aggregates for the various sized heparin units at both ionic strengths is shown in Figure 6.11.

Assessing longer time scale ordering: heparin-induced assembly of pre-formed peptide dimers

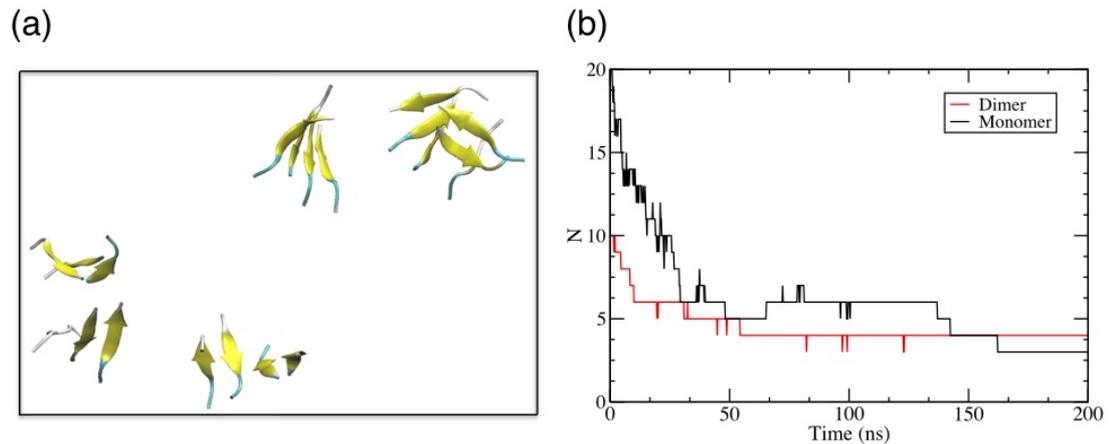
A limitation of unbiased molecular dynamics is the relatively short time scales (of the order of a microsecond) that can be accessed (Lemkul *et al.*, 2012). Experiments show that dimers aggregate remarkably faster than the monomers and there may be some subtle differences in the microstructure of monomers and dimers that is responsible for the marked increase in the aggregation propensity of dimer (Munter *et al.*, 2007; Portillo *et al.*, 2015). To observe the development of an ordered heparin- KLVFFA peptide composite protofilament in a shorter timescale, we simulated heparin-induced assembly of pre-formed KLVFFA peptide dimers in the form of anti-parallel  $\beta$ -sheets. We also looked at the interaction of a single dimer with heparin and carried out the control simulation of the self-assembly of the pre-formed peptide dimers alone i.e. without heparin.



**Figure 6.11. Number of KLVFFA peptide aggregates observed for the various heparin chain lengths (dp2-dp24) at the end of each simulation. (a) At low ionic concentration, larger chain lengths favor greater KLVFFA peptide aggregation (fewer aggregates), though shortest chain length dp2 shows distinct behavior and also promotes peptide aggregation. (b) At higher ionic concentration the aggregation increases (fewer aggregates) with an increase in chain length up to dp8, whilst dp24 doesn't enhance further aggregation**

Assembly of KLVFFA peptide dimers alone led to the formation of ordered oligomers but never a single, ordered protofilament structure over the 200 ns simulation time, just like monomers (Figure 6.12a). However, the oligomeric structures form more rapidly after about 50 ns into the trajectory (as compared to 145ns for monomeric KLVFFA), and then rearrange and order but do not associate any further (Figure 6.12b).

In the simulations investigating heparin-facilitated assembly of the preformed KLVFFA peptide dimers, the peptides rapidly localize onto the heparin which then displays a choreographic-like behavior, flexing and unflexing until the attached peptide dimers are assembled into a single stable protofilament, concluding the process just under 200 ns (Figure 6.13). This causes kinking in the heparin chain and reduces the end-to-end distance up to 4.0nm (see Figure 9.15).



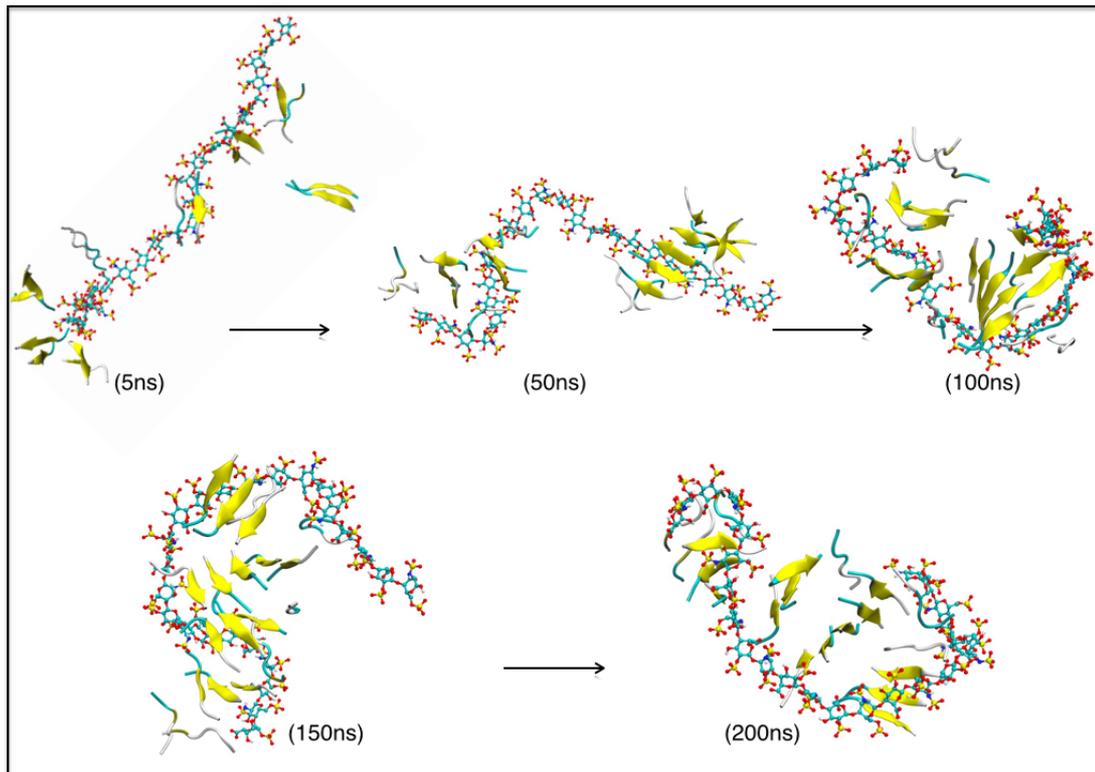
**Figure 6.12. (a) Snapshot of the self-assembly of preformed dimers of the peptide KLVFFA at 200 ns. (b) Total no. of aggregates (peptides only) as a function of simulation time for self-assembly of monomers and pre-formed peptide dimers. Dimers rapidly come together while monomer fragments take time to order and form antiparallel  $\beta$ -sheets resulting in slower aggregation.**

The protofilament structure is not ‘released’ but remains integrated with the heparin molecule. As with monomer peptides, the assembly also takes place in two distinct stages (Figure 6.14). However, in the case of dimers, the process is much faster, with the selection and localization of the peptides onto the heparin framework taking only 5 ns (Figure 6.14).

Mapping molecular insights from heparin’s assembly of KLVFFA to larger peptides and proteins

The current study looked specifically at the effect of heparin on the assembly of the peptide KLVFFA at two ionic strengths. Nevertheless, many of the simulations results fully rationalize experimental observations with respect to heparin’s interaction with other larger and complex peptides/proteins, suggesting that the inferences from the

results may be generic. Thus, the simulations not only reproduce heparin's promotion of peptide aggregation and formation of contiguous peptide aggregates (and by inference, protofibrils, and fibrils) at both low and high ionic strengths but also rationalize the effects of variations in heparin chain length as observed, for example, for the 8kDa fragment of gelsolin (65 residues), human muscle acylphosphatase (about 98 residues) (Motamedi-Shad, Monsellier and Chiti, 2009; Motamedi-Shad *et al.*, 2012) and a monomer of transthyretin (127 residues).

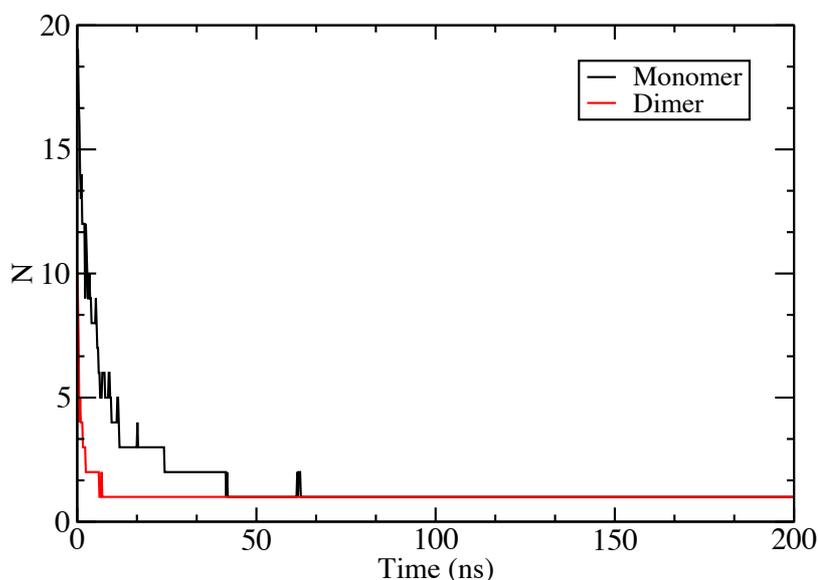


**Figure 6.13. Heparin-facilitated assembly of preformed dimers of KLVFFA in the presence of 1 molecule of heparin. The trajectory reveals two distinct stages, rapid association and localization of the peptides onto the heparin which occurs in about 5 ns, and the heparin-facilitated assembly of the associated peptides into a protofilament, which is concluded well before 200 ns.**

Further, the simulations rationalize the effects of heparin: peptide molar ratio observed for the prion-related fragment PrP (Bazar and Jelinek, 2010). Peptide/protein molecular size effects and the nature of the amino acid residues will undoubtedly modulate the mechanistic picture uncovered here for heparin-KLVFFA peptide. The optimum heparin chain length and molar ratio will depend on the commensurability of the stacked peptides with the interaction sites on the heparin (Noborn *et al.*, 2011). For large peptides, from diffusional inertia considerations, we should not expect the peptides

localizing unto the heparin molecule (as observed for KLVFFA here) but rather the heparin (being of a relatively smaller size) being attracted to the larger peptide molecules.

An important question raised in the literature concerns the first step of heparin/peptide interaction i.e. whether the peptide units localize onto the heparin framework or do they aggregate first to form oligomers before adsorbing onto heparin (Motamedi-Shad *et al.*, 2012). Whilst KLVFFA shows a tendency to adsorb on heparin's surface first and then self-assemble due to hydrophobic interactions, this does depend on the ionic strength of the aqueous environment. At 150 mM NaCl ionic strength, self-association of the peptides begins to compete with the peptide-heparin association. What happens for a particular peptide would depend on how the peptide-peptide interaction, particularly the hydrophobicity, fares relative to the peptide-heparin interaction.



**Figure 6.14. Total no. of aggregates (inclusive of heparin and peptides) as a function of simulation time for the heparin-facilitated assembly of peptide monomers and pre-formed peptide dimers. Dimers are rapidly attracted to the heparin to form a single aggregate (only 5ns) while monomers attachment is slower.**

## 6.5 Summary and significance

Whilst there is much experimental data on heparin's promotion of peptide fibrillogenesis, a molecular-level understanding of the process has been lacking. Here we have employed molecular dynamics simulation to provide a first clear perspective of how heparin promotes the initial stages of fibrillogenesis, addressing many of the outstanding questions regarding its activity and mode of action. The simulations

corroborate the experimentally observed A $\beta$  fibrillogenesis activity of heparin. Heparin achieves this by first rapidly selecting the peptides from the surrounding solution (in line with earlier inferences from experiments) (Motamedi-Shad *et al.*, 2012), and then by repeated flexing, relocates and assembles the peptide components to yield an ordered protofilaments. The heparin-peptide interactions (mostly electrostatic) are strong, and there appears to be no scope for individual peptide molecules to move up and down the heparin framework as it was a track. This explains the need for the heparin to flex as a means of relocating peptides. We were able to demonstrate that flexing is essential for heparin's assembling activity. A further consequence of the strong interaction is that the formed protofilament structure is not 'released' but remains integrated with the heparin molecule. This provides the basis for earlier experimental evidence for heparin being an integral component of the formed fibrils for the case of  $\beta$ -endorphin (Nespovitaya *et al.*, 2017), and suggests GAGs forming composite fibrils to be a more general phenomenon. The simulations rationalize the optimum heparin chain-length and the dependency of its aggregation enhancement activity on the heparin: peptide molar concentration ratio. Higher heparin: peptide ratio is less effective as the peptide molecules (sparsely) partitioned amongst the heparin molecules, reducing the probability of assembling a contiguous protofilament.

An important finding of this study is that pre-dimerized peptides aggregate remarkably faster than the monomers as supported by various experimental studies (Walsh *et al.*, 1997; Munter *et al.*, 2007; Portillo *et al.*, 2015). In our simulations, at the higher peptide concentration, the aggregation of dimers in the presence of heparin is some eight times faster than that of monomers. The monomer-dimer transition for KLVFFA is significant. These findings and insights underpin significant biological and therapeutic implications. In particular, we have identified design rules (flexibility; chain-length) and protocol guidance (heparin: peptide molar ratio) for developing effective heparin-mimetics and other functional GAGs.

## **6.6 References**

Antzutkin, O. N. (2004) 'Amyloidosis of Alzheimer's A beta peptides: solid-state nuclear magnetic resonance, electron paramagnetic resonance, transmission electron microscopy, scanning transmission electron microscopy and atomic force microscopy

studies', *Magnetic Resonance in Chemistry*, 42(2), pp. 231–246.

Baumketner, A. *et al.* (2006) 'Amyloid beta-protein monomer structure: a computational and experimental study.', *Protein science : a publication of the Protein Society*, 15(3), pp. 420–8. doi: 10.1110/ps.051762406.

Bazar, E. and Jelinek, R. (2010) 'Divergent heparin-induced fibrillation pathways of a prion amyloidogenic determinant', *ChemBioChem*, 11(14), pp. 1997–2002. doi: 10.1002/cbic.201000207.

Berendsen, H. J. C. *et al.* (1984) 'Molecular dynamics with coupling to an external bath', *The Journal of Chemical Physics*. doi: 10.1063/1.448118.

Bergamaschini, L. *et al.* (2002) 'Heparin attenuates cytotoxic and inflammatory activity of Alzheimer amyloid- $\beta$  in vitro', *Neurobiology of Aging*, 23(4), pp. 531–536. doi: 10.1016/S0197-4580(02)00003-9.

Bourgault, S., Solomon, James P, *et al.* (2011) 'Sulfated glycosaminoglycans accelerate transthyretin amyloidogenesis by quaternary structural conversion.', *Biochemistry*. NIH Public Access, 50(6), pp. 1001–15. doi: 10.1021/bi101822y.

Bourgault, S., Solomon, James P., *et al.* (2011) 'Sulfated Glycosaminoglycans Accelerate Transthyretin Amyloidogenesis by Quaternary Structural Conversion', *Biochemistry*. American Chemical Society, 50(6), pp. 1001–1015. doi: 10.1021/bi101822y.

Bussi, G., Donadio, D. and Parrinello, M. (2007) 'Canonical sampling through velocity rescaling', *The Journal of Chemical Physics*. American Institute of Physics, 126(1), p. 014101. doi: 10.1063/1.2408420.

Castillo, G. M. *et al.* (2001) 'The Sulfate Moieties of Glycosaminoglycans Are Critical for the Enhancement of  $\beta$ -Amyloid Protein Fibril Formation', *Journal of Neurochemistry*. Blackwell Science Ltd, 72(4), pp. 1681–1687. doi: 10.1046/j.1471-4159.1999.721681.x.

Cino, E. A., Choy, W. Y. and Karttunen, M. (2012) 'Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations', *Journal of Chemical Theory and Computation*, 8(8), pp. 2725–2740. doi: 10.1021/ct300323g.

Cohlberg, J. A. *et al.* (2002) 'Heparin and other glycosaminoglycans stimulate the

formation of amyloid fibrils from  $\alpha$ -synuclein in vitro', *Biochemistry*, 41(5), pp. 1502–1511. doi: 10.1021/bi011711s.

Colletier, J.-P. *et al.* (2011) 'Molecular basis for amyloid-beta polymorphism.', *Proceedings of the National Academy of Sciences of the United States of America*, 108, pp. 16938–43. doi: 10.1073/pnas.1112600108.

Cooper, C. L. *et al.* (2005) 'Polyelectrolyte-protein complexes', *Current Opinion in Colloid and Interface Science*, pp. 52–78. doi: 10.1016/j.cocis.2005.05.007.

Dreyfuss, J. L. *et al.* (2009) 'Heparan sulfate proteoglycans: structure, protein interactions and cell signaling', *Anais da Academia Brasileira de Ciências*, 81(3), pp. 409–429. doi: 10.1590/S0001-37652009000300007.

Elimova, E. *et al.* (2004) 'Amyloidogenesis recapitulated in cell culture: a peptide inhibitor provides direct evidence for the role of heparan sulfate and suggests a new treatment strategy', *The FASEB Journal*, 18(14), pp. 1749–51. doi: 10.1096/fj.03-1436fje.

Fowler, D. M. *et al.* (2007) 'Functional amyloid - from bacteria to humans', *Trends in Biochemical Sciences*, 32(5), pp. 217–224. doi: 10.1016/j.tibs.2007.03.003.

Fraser, P. E., Darabie, A. A. and McLaurin, J. (2001) 'Amyloid-beta Interactions with Chondroitin Sulfate-derived Monosaccharides and Disaccharides. IMPLICATIONS FOR DRUG DEVELOPMENT', *Journal of Biological Chemistry*, 276(9), pp. 6412–6419. doi: 10.1074/jbc.M008128200.

Gervais, F. *et al.* (2001) 'Glycosaminoglycan mimetics: a therapeutic approach to cerebral amyloid angiopathy.', *Amyloid : the international journal of experimental and clinical investigation : the official journal of the International Society of Amyloidosis*, 8 Suppl 1, pp. 28–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11676287> (Accessed: 20 July 2017).

Goedert, M. *et al.* (1996) 'Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans', *Nature*, 383(6600), pp. 550–553. doi: 10.1038/383550a0.

Hoyer, W. *et al.* (2002) 'Dependence of  $\alpha$ -synuclein aggregate morphology on solution conditions', *Journal of Molecular Biology*, 322(2), pp. 383–393. doi: 10.1016/S0022-2836(02)00775-1.

Humphrey, W., Dalke, A. and Schulten, K. (1996) 'VMD: visual molecular dynamics.', *Journal of molecular graphics*, 14(1), pp. 33–8, 27–8. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/8744570> (Accessed: 16 May 2017).

Iannuzzi, C., Irace, G. and Sirangelo, I. (2015) 'The Effect of Glycosaminoglycans (GAGs) on Amyloid Aggregation and Toxicity', *Molecules*. Multidisciplinary Digital Publishing Institute, 20(2), pp. 2510–2528. doi: 10.3390/molecules20022510.

Jha, S. *et al.* (2011) 'Mechanism of amylin fibrillization enhancement by heparin', *Journal of Biological Chemistry*, 286(26), pp. 22894–22904. doi: 10.1074/jbc.M110.215814.

Khan, S. *et al.* (2010) 'Semi-Rigid Solution Structures of Heparin by Constrained X-ray Scattering Modelling: New Insight into Heparin-Protein Complexes', *Journal of Molecular Biology*, 395(3), pp. 504–521. doi: 10.1016/j.jmb.2009.10.064.

Kumari, R., Kumar, R. and Lynn, A. (2014) 'G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations', *Journal of Chemical Information and Modeling*, 54(7), pp. 1951–1962. doi: 10.1021/ci500020m.

Lemkul, J. A. *et al.* (2012) 'Molecular Modeling of the Amyloid  $\beta$  -Peptide : Understanding the Mechanism of Alzheimer ' s Disease and the Potential for Therapeutic Intervention Molecular Modeling of the Amyloid  $\beta$  -Peptide : Understanding the Mechanism'.

Liang, C. *et al.* (2014) 'Kinetic intermediates in amyloid assembly', *Journal of the American Chemical Society*, 136(43), pp. 15146–15149. doi: 10.1021/ja508621b.

Ma, B. and Nussinov, R. (2002) 'Stabilities and conformations of Alzheimer's  $\beta$ -amyloid peptide oligomers (A $\beta$ 16-22, A $\beta$ 16-35, and A $\beta$ 10-35): Sequence effects', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.212206899.

Ma, Q. *et al.* (2007) 'Heparin Oligosaccharides as Potential Therapeutic Agents in Senile Dementia', *Curr Pharm Des.*, 13(15), pp. 1607–1616. doi: 10.1016/j.micinf.2011.07.011.Innate.

Madine, J. *et al.* (2013) 'Heparin Promotes the Rapid Fibrillization of a Peptide with Low Intrinsic Amyloidogenicity', *Biochemistry*, 52(50), pp. 8984–8992. doi: 10.1021/bi401231u.

McLaurin, J. *et al.* (1999) 'Interactions of Alzheimer amyloid-beta peptides with glycosaminoglycans effects on fibril nucleation and growth.', *European journal of biochemistry*, 266(3), pp. 1101–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10583407> (Accessed: 7 May 2017).

Miravalle, L. *et al.* (2000) 'Substitutions at codon 22 of Alzheimer's A $\beta$  peptide induce diverse conformational changes and apoptotic effects human cerebral endothelial cells', *Journal of Biological Chemistry*, 275(35), pp. 27110–27116. doi: 10.1074/jbc.M003154200.

Motamedi-Shad, N. *et al.* (2012) 'Rapid oligomer formation of human muscle acylphosphatase induced by heparan sulfate.', *Nature structural & molecular biology*, 19(5), pp. 547–54, S1-2. doi: 10.1038/nsmb.2286.

Motamedi-Shad, N., Monsellier, E. and Chiti, F. (2009) 'Amyloid formation by the model protein muscle acylphosphatase is accelerated by heparin and heparan sulphate through a scaffolding-based mechanism', *Journal of Biochemistry*, 146(6), pp. 805–814. doi: 10.1093/jb/mvp128.

Munter, L. M. *et al.* (2007) 'GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of A $\beta$ 42', *EMBO Journal*, 26(6), pp. 1702–1712. doi: 10.1038/sj.emboj.7601616.

Nespovitaya, N. *et al.* (2017) 'Heparin acts as a structural component of  $\beta$ -endorphin amyloid fibrils rather than a simple aggregation promoter', *Chem. Commun.*, 53(7), pp. 1273–1276. doi: 10.1039/C6CC09770G.

Noborn, F. *et al.* (2011) 'Heparan sulfate/heparin promotes transthyretin fibrillization through selective binding to a basic motif in the protein.', *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), pp. 5584–5589. doi: 10.1073/pnas.1101194108.

Parrinello, M. and Rahman, A. (1981) 'Polymorphic transitions in single crystals: A new molecular dynamics method', *Journal of Applied Physics*. doi: 10.1063/1.328693.

Di Pierro, M., Elber, R. and Leimkuhler, B. (2015) 'A Stochastic Algorithm for the Isobaric-Isothermal Ensemble with Ewald Summations for All Long Range Forces', *Journal of Chemical Theory and Computation*. doi: 10.1021/acs.jctc.5b00648.

Pita, R. (2015) 'The EU Regulatory Landscape of Non-Biological Complex Drugs', in,

pp. 357–380. doi: 10.1007/978-3-319-16241-6\_11.

Pol-Fachin, L. *et al.* (2012) ‘GROMOS 53A6<sub>GLYC</sub>, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates’, *Journal of Chemical Theory and Computation*. American Chemical Society, 8(11), pp. 4681–4690. doi: 10.1021/ct300479h.

Pol-Fachin, L., Fernandes, C. L. and Verli, H. (2009) ‘GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations’, *Carbohydrate Research*. doi: 10.1016/j.carres.2008.12.025.

Pol-Fachin, L. and Verli, H. (2008) ‘Depiction of the forces participating in the 2-O-sulfo- $\alpha$ -l-iduronic acid conformational preference in heparin sequences in aqueous solutions’, *Carbohydrate Research*, 343(9), pp. 1435–1445. doi: 10.1016/j.carres.2008.04.016.

Portillo, A. *et al.* (2015) ‘Role of monomer arrangement in the amyloid self-assembly’, *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1854(3), pp. 218–228. doi: 10.1016/j.bbapap.2014.12.009.

Quittot, N., Sebastiao, M. and Bourgault, S. (2017) ‘Modulation of amyloid assembly by glycosaminoglycans: from mechanism to biological significance’, *Biochemistry and Cell Biology*, 95(3), pp. 329–337. doi: 10.1139/bcb-2016-0236.

Ramachandran, G. and Udgaonkar, J. B. (2011) ‘Understanding the kinetic roles of the inducer heparin and of rod-like protofibrils during amyloid fibril formation by Tau protein.’, *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 286(45), pp. 38948–59. doi: 10.1074/jbc.M111.271874.

Rambaran, R. N. and Serpell, L. C. (2008) ‘Amyloid fibrils: abnormal protein assembly.’, *Prion*. doi: 10.4161/pri.2.3.7488.

Seyrek, E. *et al.* (2003) ‘Ionic strength dependence of protein-polyelectrolyte interactions’, *Biomacromolecules*, 4(2), pp. 273–282. doi: 10.1021/bm025664a.

So, M. *et al.* (2017) ‘Heparin-induced amyloid fibrillation of  $\beta$ 2-microglobulin explained by solubility and a supersaturation-dependent conformational phase diagram’, *Protein Science*, 26(5), pp. 1024–1036. doi: 10.1002/pro.3149.

Sofronova, A. A. *et al.* (2017) ‘Protein-polyelectrolyte complexes: Molecular dynamics simulations and experimental study’, *Polymer*, 113, pp. 39–45. doi:

10.1016/j.polymer.2017.02.047.

Van Der Spoel, D. *et al.* (2005) ‘GROMACS: fast, flexible, and free.’, *Journal of computational chemistry*, 26(16), pp. 1701–18. doi: 10.1002/jcc.20291.

Stathopoulos, P. B. *et al.* (2008) ‘Sonication of proteins causes formation of aggregates that resemble amyloid’, *Protein Science*, 13(11), pp. 3017–3027. doi: 10.1110/ps.04831804.

Stewart, K. L. and Radford, S. E. (2017) ‘Amyloid plaques beyond A $\beta$ : a survey of the diverse modulators of amyloid aggregation’, *Biophysical Reviews*, 9(4), pp. 405–419. doi: 10.1007/s12551-017-0271-9.

Suk, J. Y. *et al.* (2006) ‘Heparin accelerates gelsolin amyloidogenesis’, *Biochemistry*, 45(7), pp. 2234–2242. doi: 10.1021/bi0519295.

Takase, H. *et al.* (2016) ‘Structural requirements of glycosaminoglycans for facilitating amyloid fibril formation of human serum amyloid A’, *Amyloid*, 23(2), pp. 1350–6129. doi: 10.3109/13506129.2016.1168292.

Verli, H. and Guimarães, J. A. (2004) ‘Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution’, *Carbohydrate Research*, 339(2), pp. 281–290. doi: 10.1016/j.carres.2003.09.026.

Vilasi, S. *et al.* (2011) ‘Heparin induces harmless fibril formation in amyloidogenic W7FW14F apomyoglobin and amyloid aggregation in wild-type protein In Vitro’, *PLoS ONE*, 6(7). doi: 10.1371/journal.pone.0022076.

Walsh, D. M. *et al.* (1997) ‘Amyloid  $\beta$ -protein fibrillogenesis—Detection of a protofibrillar intermediate’, *Journal of Biological Chemistry*, 272(35), pp. 22364–22372. doi: 10.1074/jbc.272.35.22364.

Wei, Y. *et al.* (2014) ‘Design of novel cell penetrating peptides for the delivery of trehalose into mammalian cells’, *Biochimica et Biophysica Acta - Biomembranes*. doi: 10.1016/j.bbamem.2014.02.011.

Xu, Y. *et al.* (2012) ‘Effect of heparin on protein aggregation: Inhibition versus promotion’, *Biomacromolecules*, 13(5), pp. 1642–1651. doi: 10.1021/bm3003539.

Yun, S. *et al.* (2007) ‘Role of Electrostatic Interactions in Amyloid  $\beta$ -Protein (A $\beta$ ) Oligomer Formation: A Discrete Molecular Dynamics Study’, *Biophysical Journal*,

92(11), pp. 4064–4077. doi: 10.1529/biophysj.106.097766.

Zhang, G. *et al.* (2014) ‘Towards understanding the roles of heparan sulfate proteoglycans in Alzheimer’s disease.’, *BioMed research international*, 2014, p. 516028. doi: 10.1155/2014/516028.

# 7 Summary and conclusions

The cell, the smallest unit of life, together with its organelles function as factories to produce energy to sustain life. The cellular environment is very crowded and highly complex including many biomolecules such as proteins and nucleic acids. This environment is strictly supervised by a network of regulatory mechanisms to maintain stable conditions for growth, reproduction, and maintenance. Any small change in the cellular environment may affect the structure of proteins, activity of enzymes, protein binding to nucleic acids, catalytic reactions, protein assembly, and many more. One such failure leads to misfolding of proteins into harmful assemblies that are the cause of neurodegenerative diseases such as Alzheimer's disease (Jahn and Radford, 2005).

Despite billions of dollars invested in AD drug development, almost all the clinical trials have been a failure. An obvious question is why are we failing? Are we exploiting the wrong target? Is disease-intervention too late? The lack of a sufficient understanding of the process of amyloidogenic peptide assembly, the intermediates involved and the internal and external factors affecting this assembly pose is what appears to be limiting us. A molecular-level picture of the events taking place during peptide assembly is crucial to understanding this problem. Although it is often difficult to link molecular studies to in-vivo and in-vitro observations, the advances in modeling and computing power have been invaluable in developing insights into fibril formation from the molecular level (Ding, LaRocque and Dokholyan, 2005; Gsponer and Vendruscolo, 2006; Nag *et al.*, 2011; Tofoleanu and Buchete, 2012; Barale *et al.*, 2019; Martin *et al.*, 2019).

The principal aims of this thesis were to contribute to understanding how A $\beta$  molecules interact to form higher-order assemblies and the mechanism by which the peptide aggregates are modulated by certain endogenous and exogenous chemical co-factors. Below, we indicate the key contribution, and summarise and contextualize the findings and conclusions within the broader aims and the current state of the field.

## 7.1 Effects of force fields on peptides of variable size and complexity

Almost all modeling (other than first principles-based) involves the use of parameters. In molecular simulation it is the force field, comprising the potential energy function and associated parameters. The accuracy of the force field is critical to linking molecular simulations to the real world. Given this, we carried out a systematic comparison of the three popular force fields, AmberGS, Gromos53a6, and OPLS-aa, as to how well they reproduce various hierarchical structures of A $\beta$ . Force fields have been compared previously but only for the single A $\beta$  monomer. Here we have considered the structures ranging from 5-mer protofilament to 18-mer ring fibrils which test the force fields not just in terms of the monomer interactions but also interactions between protofilaments and include long-range higher-order structure features such as twisting of the structures. We have put the results in the context of nucleation theory which predicts that nanostructures, depending on their interfacial and bulk-free energies, can in principle degrade despite an accurate force field and also that structural disorder of the molecules at the structure-solvent interface are expected to be disorder to minimize the overall free energy.

The study revealed significant differences between the force fields in terms of the conformation, structural stability, and secondary structure propensity. It appears that none of these three force fields can be said to be better than the others. Each has its strengths and weaknesses, and the choice of forcefield depends on the application. Our study indicates that for the A $\beta$  monomer AmberGS is biased towards helical structures. Gromos53a6 favors the formation of a pair of anti-parallel  $\beta$ -sheets, so it seems a better choice for studying  $\alpha$ -helix- $\beta$ -sheet conversion, which is a crucial step of A $\beta$  aggregation. For large more complex structures such as dodeca-fibrils and ring fibrils, OPLS-aa is a slightly more preferable choice because it resulted in a decent agreement with experiments. We conclude that the force field bias during MD simulations of biomolecules cannot be ignored and hence, there is a need for further reparameterization and force field development.

## 7.2 Self-assembly pathways and forces driving them

This study was designed with a viewpoint to understand the nucleation-based polymerization of A $\beta$  peptides. A $\beta$  fibrils are highly ordered pseudocrystalline

structures that are the end-product of A $\beta$  aggregation. This whole process of nucleation and growth is not straightforward; it generates a heterogeneous mixture of A $\beta$  peptides ranging from misfolded peptides to small disordered oligomers to large fibrils. The process is driven by hydrophobic interactions and hydrogen bonding. And there is an intricate interplay between them. If the process is dominated by hydrogen bonding, classic nucleation takes place that results in the formation of ordered fibrils. But if it is driven predominantly by hydrophobicity, there is a rapid coalesce of peptide into disordered aggregates that are believed to be formed without passing through a lag phase.

In our study, we used five different concentrations of A $\beta$  peptides ranging from 20 -100 peptides in explicit solvent. All the systems aggregated very fast and resulted in structures that vary in morphology depending on the concentration. At lower concentrations, small spherical aggregates were formed while at very high supersaturation, amorphous aggregates were formed. There was a sweet spot for the evolution of a single, unidirectional growing fibril like aggregate which lies somewhere in the middle range concentration. The structure was highly stable and maintained its conformation even at 400K. But at 500K, it collapsed into a spherical structure. Our study clearly shows that at very high supersaturation the self-assembly of A $\beta$  is governed by hydrophobicity resulting in amorphous sheet-like structures.

An implication of this phenomenon in the body may be that at very high supersaturation, the formation of dense core fibrils is limited, and only diffused aggregates are formed. That may be one of the reasons that diffused plaques are observed in healthy elderly people while dense-core plaques are related to AD.

### **7.3 Towards non-invasive amyloid detection – interaction of curcumin with A $\beta$ species**

An issue often raised with current clinical trials for Alzheimer's therapy is that in almost all studies the patients being treated have pathologies which so severe and advanced that it is simply impossible to regress. Having acknowledged that it is then essential we can identify patients at the earliest stages of the disease. It is now thought that Alzheimer's onset can occur some 20 years before symptoms begin to show up, the so-called prodromal stage. So how does one identify at-risk patients? In 2017, a proof of concept trial successfully used a curcumin based probe for amyloid imaging in retina of



associated with reduced fibril formation. Indeed, this study supports the rationale that curcumin has both protective and imaging role in amyloidosis.

## **7.4 Endogenous species enhancing amyloid self-assembly (selection and assembly process)**

Glycosaminoglycans (GAGs) are found to co-localize with amyloid plaques in the brain suggesting that they may enhance amyloid formation (McLaurin *et al.*, 1999; Bourgault, James P Solomon, *et al.*, 2011; Noborn *et al.*, 2011; Takase *et al.*, 2016). A key endogenous GAG is heparan sulfate. The anticoagulant heparin is similar to heparan sulfate and has been used in in-vitro studies as a proxy for heparan sulfate. Whilst there is plenty of experimental data on the interaction heparin with A $\beta$ , the molecular perspective and understanding has been lacking. The study in Chapter 6 attempted to develop a molecular-level picture of how heparin promotes amyloid assembly.

Our MD simulation study rationalizes the experimentally observed A $\beta$  fibrillogenesis enhancement activity of heparin. Without heparin, A $\beta$  peptides could only form small oligomers in the timescale of the simulation but in the presence of heparin an intertwined composite of peptide-heparin is observed. Heparin interacts with A $\beta$  via electrostatic interactions to form a contiguous structure. This structure is not released but remains integrated with A $\beta$  with strong binding energy -92.1 kcal/mol to -122.7 kcal/mol. This process of integration occurs in two steps, selections and assembly. In the first step there is a rapid localization of A $\beta$  on heparin while during the second step, the continuous flexing and bending of heparin chain assembles the peptides into an ordered aggregate. This observation also led us to our second question, i-e if flexibility of heparin plays any role in facilitating amyloid assembly. We were able to demonstrate that chain flexibility of heparin is essential for its activity and if it is restrained into extended conformation, larger ordered aggregates are not formed rather small oligomers are observed. The chain length of heparin is also critical to modulate its aggregation enhancement activity. Shorter chain lengths (dp4, dp6 and dp8) are less effective as they partition the peptides to form multiple heparin-peptide complexes that aggregate only slowly while longer dp24 chain facilitate assembly. Similarly, higher heparin: peptide ratio is less effective as the peptide molecules (sparsely) partitioned amongst the heparin molecules, reducing the probability of assembling a contiguous protofilament. An important finding of this study is that rate of fibrilization of pre-dimerized peptides is

some eight times faster than that of monomers. The monomer-dimer transition for KLVFFA is significant. These findings and insights underpin significant design rules for developing effective heparin-mimetics and other functional GAGs.

## 7.5 Challenges and limitations of molecular simulation of A $\beta$

Molecular dynamics simulations provide a great deal of information about the protein folding mechanism, self-assembly, and interaction with different endogenous and exogenous species at a molecular level. But a considerable amount of time and computational resources are required to produce this kind of information even with highly parallelized and efficient MD codes (Teodoro *et al.*, 2009). With the recent advances in CPU and GPU hardware, there is some progress towards sufficient data collection yet there is still a great need of improvement with respect to MD codes, GPU and CPUs, so that one could observe those processes quickly and efficiently that, at the moment may require weeks or even months of real-time.

The present force fields generally have parameters available for major biomolecules such as protein, nucleic acids, and lipids but lack those for smaller molecules and ligands. Topology generation for biomolecules is easier as the available parameters can be applied efficiently but the problem arises where any foreign molecule even the common organic functional groups are involved. It is not a trivial task to generate topologies that are compatible with the parent force field. There are a few webservers available that generate topologies compatible with the Gromos force field but those parameters are advised to validate thoroughly (Schüttelkopf and Van Aalten, 2004; Malde *et al.*, 2011). Parameterization of these topologies with other force fields requires quantum mechanical calculations and charge calculations using empirical fittings to reproduce known properties which is very time-consuming.

There are some problems with the simulations conducted for A $\beta$  peptide and its interaction with anti-aggregation agents that have not been properly addressed. For example, the choice of structure of A $\beta$  is not unambiguous; the starting structure used in MD simulations in aqueous medium is obtained from NMR that was generated in non-aqueous solvents. Both the structures are different i-e a collapsed coil structure in aqueous solution (Zhang *et al.*, 2000) and an  $\alpha$ -helical structure in polar solvent (O Crescenzi *et al.*, 2002). A $\beta$  peptides are intrinsically disordered and can acquire has

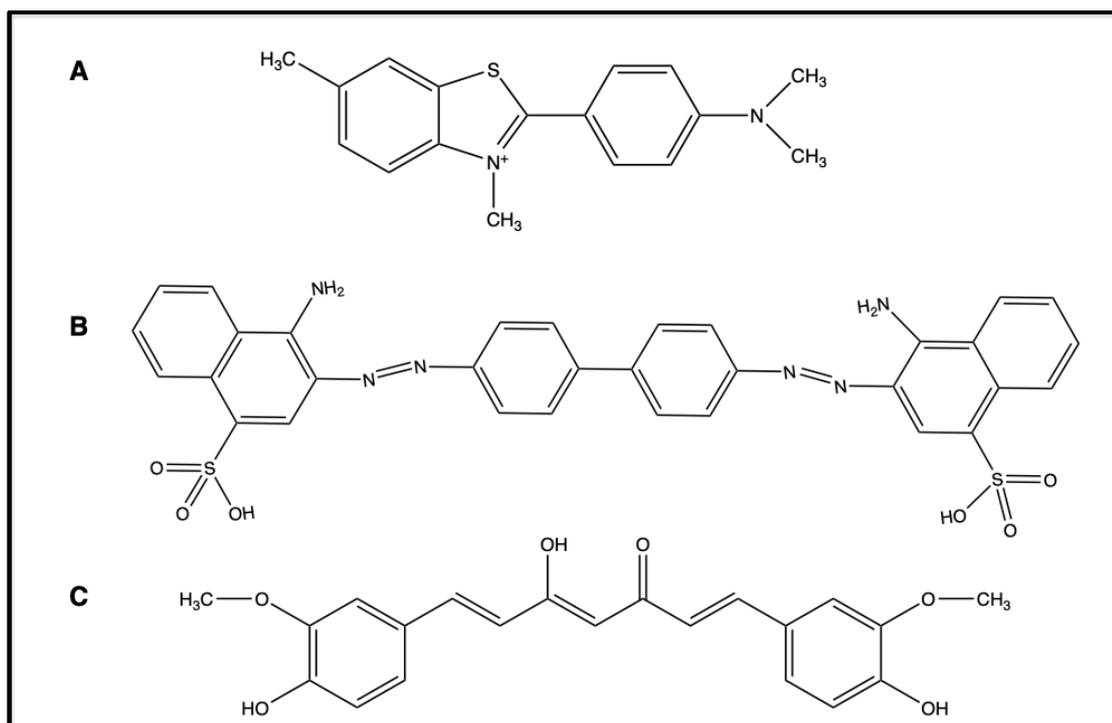
many different conformations in water, so it is not clear if the starting structure that is being used in the MD simulations represents the real structure present in-vivo or not.

Another problem that is often not addressed is the lack of sufficient MD sampling. Most of the simulations involve simplified models and shorter durations that are not sufficient to observe any relevant phenomenon. To observe  $A\beta$  dynamics fully in aqueous solution, a good starting structure, a sufficiently long simulation for exhaustive sampling, and a number of replicates for reliability must be employed.

One of the main reasons for the failure of AD drugs is the choice of target for anti-aggregation therapy. While exploiting the monomer as the primary causative agent in AD might be an attractive rational, but there is a plethora of experimental evidences that suggests  $A\beta$  oligomers as the main neuro-toxic species involved in AD (Benilova, Karran and De Strooper, 2012). Some studies also emphasize the importance of mature  $A\beta$  fibrils as a potential target because they serve as reservoirs of toxic oligomers (Cohen *et al.*, 2013). So for a successful anti-aggregation/anti-amyloid one needs to target all the culprits that are thought to be involved in this vicious cycle i-e monomers, toxic oligomers, and mature fibrils. A multi-targeted therapy against amyloid aggregation including inhibition of aggregation of  $A\beta$  monomers; disruption of toxic  $\beta$ -sheets in amyloid oligomers and the compact structure of  $A\beta$  fibrils to prevent their surface to act as secondary nucleation sites should be considered.

Molecular simulations have three major limitations: the accuracy of force fields, limited system sizes, and limited accessible timescale. We can model large (biological) systems but have to compromise molecular detail, an approach commonly employed to address this issue being coarse-grained simulation. Timescale is a big issue, particularly when we are dealing with phase transitions such as nucleation and growth, as in this thesis. Nucleation, the formation of an embryo of an ordered structure, is a statistically rare event process. This makes it challenging to simulate using standard simulation methods. In standard simulations we overcome this (badly) by increasing the supersaturation markedly to cause the event to occur in simulation timescales. Study this phenomenon requires the use of advanced simulation methods such as umbrella sampling (Kästner, 2011) and metadynamics (Laio and Parrinello, 2002), which have generally been limited in their application to smaller molecule systems. These methods need to be extended in application to the biomolecules like  $A\beta$ .

In our coarse-grain simulations using MARTINI, we employed elastic networks to preserve the secondary structure of the peptides used because MARTINI does not have directional hydrogen bonds. As MARTINI does not have any special interactions to model directional hydrogen bonds, the secondary structure cannot change or evolve. Indeed, it has to be fixed at the outset. As a result the MARTINI force field cannot be used to study A $\beta$  aggregation starting from a monomer which needs to be converted into a  $\beta$ -sheet for aggregation. However, to study the late-stage aggregation where  $\beta$ -sheets are already formed, MARTINI can be used to get insights into the assembly process. In this perspective, it would be good to specifically tune the interactions between hydrogen bonding beads. One approach could be to introduce additional beads representing hydrogen bond donors and acceptors and increase their interaction strengths. Another possibility might be to make patchy particles/beads with opposite charges acting as hydrogen bond donors and acceptors based on their underlying chemical group.



**Figure 7.1. (A) Thioflavin, (B) Congo red, (C) Curcumin. All three molecules are linear with aromatic groups separated by a linker region.**

We have shown in one of our studies (Chapter 5) that curcumin can not only provide easy and inexpensive means of detecting A $\beta$  in the retina but also acts as an anti-aggregation agent at different levels of A $\beta$  organization i-e monomers, oligomers, and fibrils. Experiments show that curcumin and curcumin-based derivatives are more

sensitive for labeling and imaging A $\beta$  plaques in the brain tissues of AD mice than classic A $\beta$  binding dyes such as thioflavin and congo-red. Moreover, curcumin derivatives can discriminate between various morphologies of A $\beta$  such as core, neuritic, diffuse, and burned-out, to a greater degree than other amyloid-binding dyes. In this perspective it is worth designing a study based on the comparison of curcumin and other classic dyes (thioflavin and congo red) to investigate what makes curcumin more sensitive than the others despite having the same chemical features.

# 8 References

Abraham, M. J. *et al.* (2015) ‘Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers’, *SoftwareX*. doi: 10.1016/j.softx.2015.06.001.

Abramowski, D. *et al.* (2008) ‘Dynamics of A $\beta$  turnover and deposition in different  $\beta$ -amyloid precursor protein transgenic mouse models following  $\gamma$ -secretase inhibition’, *Journal of Pharmacology and Experimental Therapeutics*. doi: 10.1124/jpet.108.140327.

Affentranger, R., Tavernelli, I. and Di Iorio, E. E. (2006) ‘A novel hamiltonian replica exchange MD protocol to enhance protein conformational space sampling’, *Journal of Chemical Theory and Computation*. doi: 10.1021/ct050250b.

Ahmed, M. *et al.* (2010) ‘Structural conversion of neurotoxic amyloid-beta(1-42) oligomers to fibrils.’, *Nature structural & molecular biology*. Nature Publishing Group, 17(5), pp. 561–7. doi: 10.1038/nsmb.1799.

‘Alzheimer’s disease facts and figures’ (2019) *Alzheimer’s & Dementia*, 3(15), pp. 321–387. doi: 10.1016/j.jalz.2019.01.010.

*Alzheimer’s Scientific Images and Video* (no date) *Alzheimer’s Scientific Images and Video*. Available at: <https://www.nia.nih.gov/alzheimers/alzheimers-scientific-images-and-video> (Accessed: 26 March 2020).

Andersen, H. C. (1980) ‘Molecular dynamics simulations at constant pressure and/or temperature’, *The Journal of Chemical Physics*. American Institute of Physics, 72(4), pp. 2384–2393. doi: 10.1063/1.439486.

Antzutkin, O. N. (2004) ‘Amyloidosis of Alzheimer’s A beta peptides: solid-state nuclear magnetic resonance, electron paramagnetic resonance, transmission electron microscopy, scanning transmission electron microscopy and atomic force microscopy studies’, *Magnetic Resonance in Chemistry*, 42(2), pp. 231–246.

- Anwar, J. and Zahn, D. (2011) 'Uncovering molecular processes in crystal nucleation and growth by using molecular simulation', *Angewandte Chemie - International Edition*. doi: 10.1002/anie.201000463.
- Arosio, P., Knowles, T. P. J. and Linse, S. (2015) 'On the lag phase in amyloid fibril formation.', *Physical chemistry chemical physics : PCCP*. Royal Society of Chemistry, 17(12), pp. 7606–18. doi: 10.1039/c4cp05563b.
- Auer, S. *et al.* (2008) 'Self-Templated Nucleation in Peptide and Protein Aggregation', *Physical Review Letters*, 101(25), p. 258101. doi: 10.1103/PhysRevLett.101.258101.
- Auer, S., Dobson, C. M. and Vendruscolo, M. (2007) 'Characterization of the nucleation barriers for protein aggregation and amyloid formation', *HFSP Journal*. doi: 10.2976/1.2760023.
- Báčová, P. and Rissanou, A. N. (2018) 'Modelling of novel polymer materials through atomistic molecular dynamics simulations', *Procedia Computer Science*. Elsevier, 136, pp. 341–350. doi: 10.1016/J.PROCS.2018.08.280.
- Baker, N. A. *et al.* (2001) 'Electrostatics of nanosystems: application to microtubules and the ribosome.', *Proceedings of the National Academy of Sciences of the United States of America*. National Academy of Sciences, 98(18), pp. 10037–41. doi: 10.1073/pnas.181342398.
- Barale, S. S. *et al.* (2019) 'Molecular Insights into Destabilization of Alzheimer's  $\text{A}\beta$  Protofibril by Arginine Containing Short Peptides: A Molecular Modeling Approach', *ACS Omega*. doi: 10.1021/acsomega.8b02672.
- Baron, R. *et al.* (2007) 'Comparison of thermodynamic properties of coarse-grained and atomic-level simulation models', *ChemPhysChem*. doi: 10.1002/cphc.200600658.
- Barz, B. and Urbanc, B. (2012) 'Dimer Formation Enhances Structural Differences between Amyloid  $\beta$ -Protein (1–40) and (1–42): An Explicit-Solvent Molecular Dynamics Study', *PLoS ONE*, 7(4), p. e34345. doi: 10.1371/journal.pone.0034345.
- Baschek, J. E., Klein, H. C. and Schwarz, U. S. (2012) 'Stochastic dynamics of virus capsid formation: direct versus hierarchical self-assembly', *BMC Biophysics*, 5(1), p. 22. doi: 10.1186/2046-1682-5-22.
- Baumketner, A. *et al.* (2006) 'Amyloid beta-protein monomer structure: a computational and experimental study.', *Protein science : a publication of the Protein*

- Society*, 15(3), pp. 420–8. doi: 10.1110/ps.051762406.
- Baumketner, A. (2006) ‘Amyloid beta-protein monomer structure: A computational and experimental study’, *Protein Science*, 15(3), pp. 420–428. doi: 10.1110/ps.051762406.
- Bayer-Garner, I. B. and Smoller, B. R. (2002) ‘AL amyloidosis is not present as an incidental finding in cutaneous biopsies of patients with multiple myeloma’, *Clinical and Experimental Dermatology*. doi: 10.1046/j.1365-2230.2002.01022.x.
- Bazar, E. and Jelinek, R. (2010) ‘Divergent heparin-induced fibrillation pathways of a prion amyloidogenic determinant’, *ChemBioChem*, 11(14), pp. 1997–2002. doi: 10.1002/cbic.201000207.
- Beason-Held, L. L. *et al.* (2013) ‘Changes in brain function occur years before the onset of cognitive impairment’, *Journal of Neuroscience*. doi: 10.1523/JNEUROSCI.1402-13.2013.
- Bellesia, G. and Shea, J. E. (2007) ‘Self-assembly of  $\beta$ -sheet forming peptides into chiral fibrillar aggregates’, *Journal of Chemical Physics*. doi: 10.1063/1.2739547.
- Bellesia, G. and Shea, J. E. (2009) ‘Effect of  $\beta$ -sheet propensity on peptide aggregation’, *Journal of Chemical Physics*. doi: 10.1063/1.3108461.
- Benilova, I., Karran, E. and De Strooper, B. (2012) ‘The toxic A $\beta$  oligomer and Alzheimer’s disease: an emperor in need of clothes.’, *Nature neuroscience*. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved., 15(3), pp. 349–57. doi: 10.1038/nn.3028.
- Berendsen, H. J. C. *et al.* (1984) ‘Molecular dynamics with coupling to an external bath’, *The Journal of Chemical Physics*. doi: 10.1063/1.448118.
- Berendsen, H. J. C., Grigera, J. R. and Straatsma, T. P. (1987) ‘The missing term in effective pair potentials’, *Journal of Physical Chemistry*. doi: 10.1021/j100308a038.
- Berendsen, H. J. C., van der Spoel, D. and van Drunen, R. (1995) ‘GROMACS: A message-passing parallel molecular dynamics implementation’, *Computer Physics Communications*, 91(1–3), pp. 43–56. doi: 10.1016/0010-4655(95)00042-E.
- Berg, J. M., Tymoczko, J. L. and Stryer, L. (2002) ‘Secondary Structure: Polypeptide Chains Can Fold Into Regular Structures Such as the Alpha Helix, the Beta Sheet, and Turns and Loops’. W H Freeman. Available at:

<https://www.ncbi.nlm.nih.gov/books/NBK22580/> (Accessed: 26 March 2020).

Bergamaschini, L. *et al.* (2002) ‘Heparin attenuates cytotoxic and inflammatory activity of Alzheimer amyloid- $\beta$  in vitro’, *Neurobiology of Aging*, 23(4), pp. 531–536. doi: 10.1016/S0197-4580(02)00003-9.

Biarnés, X. *et al.* (2012) ‘METAGUI. A VMD interface for analyzing metadynamics and molecular dynamics simulations’, *Computer Physics Communications*. Elsevier B.V., 183(1), pp. 203–211. doi: 10.1016/j.cpc.2011.08.020.

Birks, J. S. (2006) ‘Cholinesterase inhibitors for Alzheimer’s disease’, in Birks, J. S. (ed.) *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. doi: 10.1002/14651858.CD005593.

Bishop, G. M. and Robinson, S. R. (2002) ‘The amyloid hypothesis: Let sleeping dogmas lie?’, *Neurobiology of aging*. doi: 10.1016/S0197-4580(02)00050-7.

Bourgault, S., Solomon, James P, *et al.* (2011) ‘Sulfated glycosaminoglycans accelerate transthyretin amyloidogenesis by quaternary structural conversion.’, *Biochemistry*. NIH Public Access, 50(6), pp. 1001–15. doi: 10.1021/bi101822y.

Bourgault, S., Solomon, James P., *et al.* (2011) ‘Sulfated Glycosaminoglycans Accelerate Transthyretin Amyloidogenesis by Quaternary Structural Conversion’, *Biochemistry*. American Chemical Society, 50(6), pp. 1001–1015. doi: 10.1021/bi101822y.

Bowers, K. J. *et al.* (2006) ‘Scalable algorithms for molecular dynamics simulations on commodity clusters’, in *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing, SC’06*. doi: 10.1145/1188455.1188544.

*Brain with Alzheimer’s Disease | BrightFocus Foundation* (2017). Available at: <https://www.brightfocus.org/alzheimers/infographic/brain-alzheimers-disease> (Accessed: 4 January 2019).

Brooks, B. R. *et al.* (1983) ‘CHARMM: A program for macromolecular energy, minimization, and dynamics calculations’, *Journal of Computational Chemistry*. doi: 10.1002/jcc.540040211.

Brünger, A. T., Adams, P. D. and Rice, L. M. (1997) ‘New applications of simulated annealing in X-ray crystallography and solution NMR’, *Structure*. doi: 10.1016/S0969-2126(97)00190-1.

- BRYAN, G. H. (1902) 'Elementary Principles in Statistical Mechanics', *Nature*. doi: 10.1038/066291a0.
- Bucciantini, M. *et al.* (2002) 'Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases.', *Nature*, 416(6880), pp. 507–11. doi: 10.1038/416507a.
- Bussi, G. (2014) 'Hamiltonian replica exchange in GROMACS: A flexible implementation', in *Molecular Physics*. doi: 10.1080/00268976.2013.824126.
- Bussi, G., Donadio, D. and Parrinello, M. (2007) 'Canonical sampling through velocity rescaling', *The Journal of Chemical Physics*. American Institute of Physics, 126(1), p. 014101. doi: 10.1063/1.2408420.
- Butterfield, D. A. and Pocernich, C. B. (2003) 'The glutamatergic system and Alzheimer's disease: therapeutic implications.', *CNS drugs*, 17(9), pp. 641–52.
- C. Vickers, J. *et al.* (2015) 'Defining the earliest pathological changes of Alzheimer's disease', *Current Alzheimer Research*. doi: 10.2174/1567205013666151218150322.
- Carrick, L. M. *et al.* (2007) 'Effect of ionic strength on the self-assembly, morphology and gelation of pH responsive  $\beta$ -sheet tape-forming peptides', *Tetrahedron*. doi: 10.1016/j.tet.2007.05.036.
- Case, D. A. *et al.* (2005) 'The Amber biomolecular simulation programs', *Journal of Computational Chemistry*. doi: 10.1002/jcc.20290.
- Castillo, G. M. *et al.* (2001) 'The Sulfate Moieties of Glycosaminoglycans Are Critical for the Enhancement of  $\beta$ -Amyloid Protein Fibril Formation', *Journal of Neurochemistry*. Blackwell Science Ltd, 72(4), pp. 1681–1687. doi: 10.1046/j.1471-4159.1999.721681.x.
- Chen, F., Eckman, E. A. and Eckman, C. B. (2006) 'Reductions in levels of the Alzheimer's amyloid beta peptide after oral administration of ginsenosides.', *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 20(8), pp. 1269–71. doi: 10.1096/fj.05-5530fje.
- Chen, G. F. *et al.* (2017) 'Amyloid beta: Structure, biology and structure-based therapeutic development', *Acta Pharmacologica Sinica*. doi: 10.1038/aps.2017.28.
- Chiti, F. and Dobson, C. M. (2006) 'Protein Misfolding, Functional Amyloid, and

Human Disease', *Annual Review of Biochemistry*. doi: 10.1146/annurev.biochem.75.101304.123901.

Chiti, F. and Dobson, C. M. (2009) 'Amyloid formation by globular proteins under native conditions', *Nature Chemical Biology*. Nature Publishing Group, 5(1), pp. 15–22. doi: 10.1038/nchembio.131.

Chong, S. H., Yim, J. and Ham, S. (2013) 'Structural heterogeneity in familial Alzheimer's disease mutants of amyloid-beta peptides', *Molecular BioSystems*. doi: 10.1039/c2mb25457c.

Chou, E. (2014) 'Alzheimer's Disease: Background, Current and Future Treatments', *International Journal of Medical Students*, pp. 56–63.

Cino, E. A., Choy, W. Y. and Karttunen, M. (2012) 'Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations', *Journal of Chemical Theory and Computation*, 8(8), pp. 2725–2740. doi: 10.1021/ct300323g.

Cohen, S. I. A. *et al.* (2013) 'Proliferation of amyloid-42 aggregates occurs through a secondary nucleation mechanism', *Proceedings of the National Academy of Sciences*, 110(24), pp. 9758–9763. doi: 10.1073/pnas.1218402110.

Cohlberg, J. A. *et al.* (2002) 'Heparin and other glycosaminoglycans stimulate the formation of amyloid fibrils from A $\beta$ -synuclein in vitro', *Biochemistry*, 41(5), pp. 1502–1511. doi: 10.1021/bi011711s.

Colletier, J.-P. *et al.* (2011) 'Molecular basis for amyloid-beta polymorphism.', *Proceedings of the National Academy of Sciences of the United States of America*, 108, pp. 16938–43. doi: 10.1073/pnas.1112600108.

Colligris, P. *et al.* (2018) 'Ocular Manifestations of Alzheimer's and Other Neurodegenerative Diseases: The Prospect of the Eye as a Tool for the Early Diagnosis of Alzheimer's Disease', *Journal of Ophthalmology*. doi: 10.1155/2018/8538573.

Cooper, C. L. *et al.* (2005) 'Polyelectrolyte-protein complexes', *Current Opinion in Colloid and Interface Science*, pp. 52–78. doi: 10.1016/j.cocis.2005.05.007.

Cornell, W. D. *et al.* (1995) 'A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules', *Journal of the American Chemical Society*. American Chemical Society, 117(19), pp. 5179–5197. doi:

10.1021/ja00124a002.

Crescenzi, Orlando, Tomaselli, S., Guerrini, R., Salvadori, S., D'Ursi, Anna M., *et al.* (2002) 'Solution structure of the Alzheimer amyloid beta-peptide (1-42) in an apolar microenvironment. Similarity with a virus fusion domain.', *European journal of biochemistry / FEBS*, 269(22), pp. 5642–8.

Crescenzi, O *et al.* (2002) 'Solution structure of the Alzheimer amyloid beta-peptide (1-42) in an apolar microenvironment. Similarity with a virus fusion domain.', *EUR.J.BIOCHEM.*, 269, pp. 5642–5648. doi: 12423364.

Crescenzi, Orlando, Tomaselli, S., Guerrini, R., Salvadori, S., D'Ursi, Anna M., *et al.* (2002) 'Solution structure of the Alzheimer amyloid  $\beta$ -peptide (1-42) in an apolar microenvironment: Similarity with a virus fusion domain', *European Journal of Biochemistry*. doi: 10.1046/j.1432-1033.2002.03271.x.

Cummings, J. *et al.* (2019) 'Alzheimer's disease drug development pipeline: 2019', *Alzheimer's and Dementia: Translational Research and Clinical Interventions*. doi: 10.1016/j.trci.2019.05.008.

Darden, T., York, D. and Pedersen, L. (1993) 'Particle mesh Ewald: An  $N \cdot \log(N)$  method for Ewald sums in large systems', *The Journal of Chemical Physics*. doi: 10.1063/1.464397.

Daura, X. *et al.* (1999) 'Peptide Folding: When Simulation Meets Experiment', *Angewandte Chemie International Edition*. doi: 10.1002/(sici)1521-3773(19990115)38:1/2<236::aid-anie236>3.0.co;2-m.

Ding, F. *et al.* (2003) 'Mechanism for the alpha-helix to beta-hairpin transition.', *Proteins*, 53(2), pp. 220–228. doi: 10.1002/prot.10468.

Ding, F., LaRocque, J. J. and Dokholyan, N. V (2005) 'Direct observation of protein folding, aggregation, and a prion-like conformational conversion.', *The Journal of biological chemistry*, 280(48), pp. 40235–40. doi: 10.1074/jbc.M506372200.

Do, P. C., Lee, E. H. and Le, L. (2018) 'Steered Molecular Dynamics Simulation in Rational Drug Design', *Journal of Chemical Information and Modeling*. doi: 10.1021/acs.jcim.8b00261.

Dobson, C. M. (2004) 'Principles of protein folding, misfolding and aggregation', in *Seminars in Cell and Developmental Biology*. doi: 10.1016/j.semcd.2003.12.008.

Drenth, J. and Haas, C. (1992) 'Protein crystals and their stability', *Journal of Crystal Growth*. doi: 10.1016/0022-0248(92)90233-9.

Dreyfuss, J. L. *et al.* (2009) 'Heparan sulfate proteoglycans: structure, protein interactions and cell signaling', *Anais da Academia Brasileira de Ciências*, 81(3), pp. 409–429. doi: 10.1590/S0001-37652009000300007.

Elimova, E. *et al.* (2004) 'Amyloidogenesis recapitulated in cell culture: a peptide inhibitor provides direct evidence for the role of heparan sulfate and suggests a new treatment strategy', *The FASEB Journal*, 18(14), pp. 1749–51. doi: 10.1096/fj.03-1436fje.

Estácio, S. G. *et al.* (2014) 'A simulated intermediate state for folding and aggregation provides insights into  $\Delta$ N6  $\beta$ 2-microglobulin amyloidogenic behavior.', *PLoS computational biology*. Public Library of Science, 10(5), p. e1003606. doi: 10.1371/journal.pcbi.1003606.

Ewald, P. P. (1921) 'Die Berechnung optischer und elektrostatischer Gitterpotentiale', *Annalen der Physik*, 369(3), pp. 253–287. doi: 10.1002/andp.19213690304.

Faeder, J. R., Blinov, M. L. and Hlavacek, W. S. (2009) 'Rule-based modeling of biochemical systems with BioNetGen', *Methods in Molecular Biology*, 500(1), pp. 113–167. doi: 10.1007/978-1-59745-525-1\_5.

Fitzpatrick, A. W. *et al.* (2011) 'Inversion of the Balance between Hydrophobic and Hydrogen Bonding Interactions in Protein Folding and Aggregation', *PLoS Computational Biology*. Edited by V. S. Pande, 7(10), p. e1002169. doi: 10.1371/journal.pcbi.1002169.

Fluitt, A. M. and De Pablo, J. J. (2015) 'An Analysis of Biomolecular Force Fields for Simulations of Polyglutamine in Solution', *Biophysical Journal*. Biophysical Society, 109(5), pp. 1009–1018. doi: 10.1016/j.bpj.2015.07.018.

Fowler, D. M. *et al.* (2007) 'Functional amyloid - from bacteria to humans', *Trends in Biochemical Sciences*, 32(5), pp. 217–224. doi: 10.1016/j.tibs.2007.03.003.

Fraser, P. E., Darabie, A. A. and McLaurin, J. (2001) 'Amyloid-beta Interactions with Chondroitin Sulfate-derived Monosaccharides and Disaccharides. IMPLICATIONS FOR DRUG DEVELOPMENT', *Journal of Biological Chemistry*, 276(9), pp. 6412–6419. doi: 10.1074/jbc.M008128200.

- Frazier, J. M., Chushak, Y. and Foy, B. (2009) 'Stochastic simulation and analysis of biomolecular reaction networks', *BMC Systems Biology*, 3. doi: 10.1186/1752-0509-3-64.
- Friedemann, M. *et al.* (2015) 'Effect of methionine-35 oxidation on the aggregation of amyloid- $\beta$  peptide', *Biochemistry and Biophysics Reports*. doi: 10.1016/j.bbrep.2015.07.017.
- Frishman, D and Argos, P. (1995) 'Knowledge-based protein secondary structure assignment.', *Proteins*, 23(4), pp. 566–79. doi: 10.1002/prot.340230412.
- Frishman, Dmitriy and Argos, P. (1995) 'Knowledge-based protein secondary structure assignment', *Proteins: Structure, Function, and Bioinformatics*. doi: 10.1002/prot.340230412.
- Frost, S. *et al.* (2013) 'Retinal vascular biomarkers for early detection and monitoring of Alzheimer's disease', *Translational Psychiatry*. doi: 10.1038/tp.2012.150.
- Fujimoto, K. *et al.* (2020) 'Free energy profile of permeation of Entecavir through Hepatitis B virus capsid studied by molecular dynamics calculation', *Pure and Applied Chemistry*. doi: 10.1515/pac-2020-0109.
- Fukunishi, H., Watanabe, O. and Takada, S. (2002) 'On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction', *Journal of Chemical Physics*. doi: 10.1063/1.1472510.
- García, A. E. and Sanbonmatsu, K. Y. (2002) ' $\alpha$ -Helical stabilization by side chain shielding of backbone hydrogen bonds', *Proceedings of the National Academy of Sciences*, 99(5), pp. 2782–2787. doi: 10.1073/pnas.042496899.
- Gargari, S. A. and Barzegar, A. (2020) 'Simulations on the dual effects of flavonoids as suppressors of A $\beta$ 42 fibrillogenesis and destabilizers of mature fibrils', *Scientific Reports*. doi: 10.1038/s41598-020-72734-9.
- Gerben, S. R. *et al.* (2014) 'Comparing atomistic molecular mechanics force fields for a difficult target: a case study on the Alzheimer's amyloid  $\beta$ -peptide.', *Journal of biomolecular structure & dynamics*, 32(11), pp. 1817–32. doi: 10.1080/07391102.2013.838518.
- Gervais, F. *et al.* (2001) 'Glycosaminoglycan mimetics: a therapeutic approach to cerebral amyloid angiopathy.', *Amyloid: the international journal of experimental and*

*clinical investigation : the official journal of the International Society of Amyloidosis*, 8 Suppl 1, pp. 28–35. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11676287> (Accessed: 20 July 2017).

Gilmanshin, R., Dyer, R. B. and Callender, R. H. (1997) ‘Structural heterogeneity of the various forms of apomyoglobin: implications for protein folding.’, *Protein science : a publication of the Protein Society*, 6(10), pp. 2134–42. doi: 10.1002/pro.5560061008.

Glenner, G. G. (1980) ‘Amyloid deposits and amyloidosis. The beta-fibrilloses (first of two parts).’, *The New England journal of medicine*, 302(23), pp. 1283–92. doi: 10.1056/NEJM198006053022305.

Goedert, M. *et al.* (1996) ‘Assembly of microtubule-associated protein tau into Alzheimer-like filaments induced by sulphated glycosaminoglycans’, *Nature*, 383(6600), pp. 550–553. doi: 10.1038/383550a0.

Goozee, K. G. *et al.* (2016) ‘Examining the potential clinical value of curcumin in the prevention and diagnosis of Alzheimer’s disease’, *British Journal of Nutrition*, 115, pp. 449–465. doi: 10.1017/S0007114515004687.

Gosztyla, M. L., Brothers, H. M. and Robinson, S. R. (2018) ‘Alzheimer’s Amyloid- $\beta$  is an Antimicrobial Peptide: A Review of the Evidence’, *Journal of Alzheimer’s Disease*. doi: 10.3233/JAD-171133.

Graeber, M. B. (1998) ‘The case described by Alois Alzheimer in 1911’, *Psychiatric Hospital*, pp. 111–122.

Granzotto, A. and Zatta, P. (2014) ‘Resveratrol and Alzheimer’s disease: Message in a bottle on red wine and cognition’, *Frontiers in Aging Neuroscience*. doi: 10.3389/fnagi.2014.00095.

Gremer, L. *et al.* (2017) ‘Fibril structure of amyloid- $\beta$ (1–42) by cryo–electron microscopy’, *Science*. doi: 10.1126/science.aao2825.

Groen, J. *et al.* (2015) ‘Associative interactions in crowded solutions of biopolymers counteract depletion effects’, *Journal of the American Chemical Society*. doi: 10.1021/jacs.5b07898.

Gruenert, G. *et al.* (2010) ‘Rule-based spatial modeling with diffusing, geometrically constrained molecules’, *BMC Bioinformatics*, 11. doi: 10.1186/1471-2105-11-307.

- Gsponer, J. and Vendruscolo, M. (2006) 'Theoretical approaches to protein aggregation.', *Protein and peptide letters*, 13(3), pp. 287–293. doi: 10.2174/092986606775338407.
- Van Gunsteren, W. F. and Berendsen, H. J. C. (1988) 'A Leap-frog Algorithm for Stochastic Dynamics', *Molecular Simulation*. Taylor & Francis Group, 1(3), pp. 173–185. doi: 10.1080/08927028808080941.
- Haake, A. *et al.* (2020) 'An update on the utility and safety of cholinesterase inhibitors for the treatment of Alzheimer's disease', *Expert Opinion on Drug Safety*, 19(2), pp. 147–157. doi: 10.1080/14740338.2020.1721456.
- Halgren, T. A. (1996) 'Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94', *Journal of Computational Chemistry*, 17(5–6), pp. 490–519. doi: 10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.
- Hampel, H. *et al.* (2010) 'Biological markers of amyloid  $\beta$ -related mechanisms in Alzheimer's disease', *Experimental Neurology*. doi: 10.1016/j.expneurol.2009.09.024.
- Han, W. and Schulten, K. (2014) 'Fibril elongation by A $\beta$ 17-42: Kinetic network analysis of hybrid-resolution molecular dynamics simulations', *Journal of the American Chemical Society*. doi: 10.1021/ja507002p.
- Hardy, J. and Selkoe, D. J. (2002) 'The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics.', *Science (New York, N.Y.)*, 297(2002), pp. 353–356. doi: 10.1126/science.1072994.
- Harmeier, A. *et al.* (2009) 'Role of Amyloid- $\beta$  Glycine 33 in Oligomerization, Toxicity, and Neuronal Plasticity', *Journal of Neuroscience*, 29(23).
- Hartl, F. U., Bracher, A. and Hayer-Hartl, M. (2011) 'Molecular chaperones in protein folding and proteostasis.', *Nature*, 475(7356), pp. 324–332. doi: 10.1038/nature10317.
- Hayden, E. Y. and Teplow, D. B. (2013) 'Amyloid  $\beta$ -protein oligomers and Alzheimer's disease.', *Alzheimer's research & therapy*, 5(6), p. 60. doi: 10.1186/alzrt226.
- Hazra, M. K., Roy, S. and Bagchi, B. (2014) 'Hydrophobic hydration driven self-assembly of curcumin in water: Similarities to nucleation and growth under large metastability, and an analysis of water dynamics at heterogeneous surfaces', *Journal of Chemical Physics*. doi: 10.1063/1.4895539.

Heinig, M. and Frishman, D. (2004) 'STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins.', *Nucleic acids research*, 32(Web Server issue), pp. W500-2. doi: 10.1093/nar/gkh429.

Hess, B. *et al.* (1997) 'LINCS: A Linear Constraint Solver for molecular simulations', *Journal of Computational Chemistry*. doi: 10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

*HistoricaPollack, L. Klaus Schulten Reflects on his 2015 National Lecture Series* (no date). Available at: <http://www.ks.uiuc.edu/events/NationalLecture2015/reflection/> (Accessed: 3 March 2020).

Hoover, W. G. (1985) 'Canonical dynamics: Equilibrium phase-space distributions', *Physical Review A*, 31(3), pp. 1695–1697. doi: 10.1103/PhysRevA.31.1695.

Horobin, R. W. and James, N. T. (1970) 'The staining of elastic fibres with Direct Blue 152. A general hypothesis for the staining of elastic fibres', *Histochemie*. doi: 10.1007/BF00277460.

Hou, L. *et al.* (2004) 'Solution NMR Studies of the A $\beta$ (1-40) and A $\beta$ (1-42) Peptides Establish that the Met35 Oxidation State Affects the Mechanism of Amyloid Formation', *Journal of the American Chemical Society*. doi: 10.1021/ja036813f.

Hoyer, W. *et al.* (2002) 'Dependence of  $\alpha$ -synuclein aggregate morphology on solution conditions', *Journal of Molecular Biology*, 322(2), pp. 383–393. doi: 10.1016/S0022-2836(02)00775-1.

Hoyer, W. *et al.* (2008) 'Stabilization of a  $\beta$ -hairpin in monomeric Alzheimer's amyloid- $\beta$  peptide inhibits amyloid formation', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0711731105.

Huang, J. and Mackerell, A. D. (2013) 'CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data', *Journal of Computational Chemistry*. doi: 10.1002/jcc.23354.

Huang, T.-C. *et al.* (2011) 'Resveratrol protects rats from A $\beta$ -induced neurotoxicity by the reduction of iNOS expression and lipid peroxidation.', *PloS one*, 6(12), p. e29102. doi: 10.1371/journal.pone.0029102.

Humphrey, W., Dalke, A. and Schulten, K. (1996) 'VMD: visual molecular dynamics.', *Journal of molecular graphics*, 14(1), pp. 33–8, 27–8. Available at:

- <http://www.ncbi.nlm.nih.gov/pubmed/8744570> (Accessed: 16 May 2017).
- Iannuzzi, C., Irace, G. and Sirangelo, I. (2015) ‘The Effect of Glycosaminoglycans (GAGs) on Amyloid Aggregation and Toxicity’, *Molecules*. Multidisciplinary Digital Publishing Institute, 20(2), pp. 2510–2528. doi: 10.3390/molecules20022510.
- Jahn, T. R. and Radford, S. E. (2005) ‘The Yin and Yang of protein folding’, *FEBS Journal*, 272(23), pp. 5962–5970. doi: 10.1111/j.1742-4658.2005.05021.x.
- Jana, A. K. and May, E. R. (2020) ‘Structural and dynamic asymmetry in icosahedrally symmetric virus capsids’, *Current Opinion in Virology*. doi: 10.1016/j.coviro.2020.06.002.
- Javed, I. *et al.* (2019) ‘Inhibition of amyloid beta toxicity in zebrafish with a chaperone-gold nanoparticle dual strategy’, *Nature Communications*. doi: 10.1038/s41467-019-11762-0.
- Jha, S. *et al.* (2011) ‘Mechanism of amylin fibrillization enhancement by heparin’, *Journal of Biological Chemistry*, 286(26), pp. 22894–22904. doi: 10.1074/jbc.M110.215814.
- Johansson, J. (2005) ‘Amyloid fibrils’, *FEBS Journal*. doi: 10.1111/j.1742-4658.2005.05026.x.
- Jones, J. E., Jones and E., J. (1924) ‘On the determination of molecular fields. —II. From the equation of state of a gas’, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 106(738), pp. 463–477. doi: 10.1098/rspa.1924.0082.
- Jorgensen, W. L. and Tirado-Rives, J. (1988) ‘The OPLS Potential Functions for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin’, *Journal of the American Chemical Society*. doi: 10.1021/ja00214a001.
- Kahler, A., Sticht, H. and Horn, A. H. C. (2013) ‘Conformational Stability of Fibrillar Amyloid-Beta Oligomers via Protofilament Pair Formation - A Systematic Computational Study’, *PLoS ONE*. doi: 10.1371/journal.pone.0070521.
- Kamiya, N., Higo, J. and Nakamura, H. (2009) ‘Conformational transition states of a  $\beta$ -hairpin peptide between the ordered and disordered conformations in explicit water’, *Protein Science*. doi: 10.1110/ps.0213102.

- Karr, J. R. *et al.* (2012) ‘A whole-cell computational model predicts phenotype from genotype’, *Cell*, 150(2), pp. 389–401. doi: 10.1016/j.cell.2012.05.044.
- Kästner, J. (2011) ‘Umbrella sampling’, *Wiley Interdisciplinary Reviews: Computational Molecular Science*. doi: 10.1002/wcms.66.
- Ke, P. C. *et al.* (2017) ‘Implications of peptide assemblies in amyloid diseases’, *Chemical Society Reviews*. doi: 10.1039/c7cs00372b.
- Khan, S. *et al.* (2010) ‘Semi-Rigid Solution Structures of Heparin by Constrained X-ray Scattering Modelling: New Insight into Heparin-Protein Complexes’, *Journal of Molecular Biology*, 395(3), pp. 504–521. doi: 10.1016/j.jmb.2009.10.064.
- Kholina, E. G. *et al.* (2020) ‘Cationic Antiseptics Facilitate Pore Formation in Model Bacterial Membranes’, *The Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcc.0c07212.
- Klimov, D. K. and Thirumalai, D. (2003) ‘Dissecting the assembly of A $\beta$ 16-22 amyloid peptides into antiparallel  $\beta$  sheets’, *Structure*. doi: 10.1016/S0969-2126(03)00031-5.
- Koronyo-Hamaoui, M. *et al.* (2011) ‘Identification of amyloid plaques in retinas from Alzheimer’s patients and noninvasive in vivo optical imaging of retinal plaques in a mouse model’, *NeuroImage*. doi: 10.1016/j.neuroimage.2010.06.020.
- Koronyo, Y. *et al.* (2017) ‘Retinal amyloid pathology and proof-of-concept imaging trial in Alzheimer’s disease’, *JCI insight*. doi: 10.1172/jci.insight.93621.
- Krishna Deepak, R. N. V. and Sankararamkrishnan, R. (2016) ‘Unconventional N-H...N Hydrogen Bonds Involving Proline Backbone Nitrogen in Protein Structures’, *Biophysical Journal*. doi: 10.1016/j.bpj.2016.03.034.
- Kumar, Y., Singh, H. and Patel, C. N. (2020) ‘In silico prediction of potential inhibitors for the main protease of SARS-CoV-2 using molecular docking and dynamics simulation based drug-repurposing’, *Journal of Infection and Public Health*. doi: 10.1016/j.jiph.2020.06.016.
- Kumari, R., Kumar, R. and Lynn, A. (2014) ‘G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations’, *Journal of Chemical Information and Modeling*, 54(7), pp. 1951–1962. doi: 10.1021/ci500020m.
- Kung, H. F. (2012) ‘The  $\beta$ -Amyloid Hypothesis in Alzheimer ’ s Disease : Seeing Is

- Believing', *ACS Medicinal Chemistry Letters*, (Figure 1), pp. 265–267.
- Ladiwala, A. R. A. *et al.* (2012) 'Conformational differences between two amyloid  $\beta$  oligomers of similar size and dissimilar toxicity', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M111.329763.
- Laio, A. and Gervasio, F. L. (2008) 'Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science', *Reports on Progress in Physics*. doi: 10.1088/0034-4885/71/12/126601.
- Laio, A. and Parrinello, M. (2002) 'Escaping free-energy minima', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.202427399.
- Latshaw, D. C., Cheon, M. and Hall, C. K. (2014) 'Effects of macromolecular crowding on amyloid beta (16-22) aggregation using coarse-grained simulations', *Journal of Physical Chemistry B*. doi: 10.1021/jp508970q.
- Lee, M., Yoon, J. and Shin, S. (2019) 'Computational Study on Structure and Aggregation Pathway of A $\beta$ 42 Amyloid Protofibril', *Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcc.9b07195.
- Lee, W.-H. *et al.* (2013) 'Curcumin and its derivatives: their application in neuropharmacology and neuroscience in the 21st century.', *Current neuropharmacology*, 11(4), pp. 338–78. doi: 10.2174/1570159X11311040002.
- Lemkul, J. A. *et al.* (2012) 'Molecular Modeling of the Amyloid  $\beta$  -Peptide : Understanding the Mechanism of Alzheimer ' s Disease and the Potential for Therapeutic Intervention Molecular Modeling of the Amyloid  $\beta$  -Peptide : Understanding the Mechanism'.
- Lendrum, A. C., Slidders, W. and Fraser, D. S. (1972) 'Renal hyalin. A study of amyloidosis and diabetic fibrinous vasculosis with new staining methods.', *Journal of clinical pathology*. doi: 10.1136/jcp.25.5.373.
- Leone, V. *et al.* (2010) 'Targeting biomolecular flexibility with metadynamics', *Current Opinion in Structural Biology*. doi: 10.1016/j.sbi.2010.01.011.
- Li, C. and Pielak, G. J. (2009) 'Using NMR to distinguish viscosity effects from nonspecific protein binding under crowded conditions', *Journal of the American Chemical Society*. doi: 10.1021/ja808428d.

- Liang, C. *et al.* (2014) 'Kinetic intermediates in amyloid assembly', *Journal of the American Chemical Society*, 136(43), pp. 15146–15149. doi: 10.1021/ja508621b.
- Liang, L., Wang, L. W. and Shen, J. W. (2016) 'The self-assembly mechanism of tetrapeptides from the motif of  $\beta$ -amyloid peptides: A combined coarse-grained and all-atom molecular dynamics simulation', *RSC Advances*. doi: 10.1039/c6ra18204f.
- Lii, J. -H and Allinger, N. L. (1991) 'The MM3 force field for amides, polypeptides and proteins', *Journal of Computational Chemistry*, 12(2), pp. 186–199. doi: 10.1002/jcc.540120208.
- Lim, G. P. *et al.* (2005) 'A diet enriched with the omega-3 fatty acid docosahexaenoic acid reduces amyloid burden in an aged Alzheimer mouse model.', *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(12), pp. 3032–3040. doi: 10.1523/JNEUROSCI.4225-04.2005.
- Lin, W. H. *et al.* (2008) 'Amyloid- $\beta$  peptide (A $\beta$ ) neurotoxicity is modulated by the rate of peptide aggregation: A $\beta$  dimers and trimers correlate with neurotoxicity', *Journal of Neuroscience*. doi: 10.1523/JNEUROSCI.3916-08.2008.
- Liu, G. *et al.* (2010) 'Mechanistic studies of peptide self-assembly: Transient  $\alpha$ -helices to stable  $\beta$ -sheets', *Journal of the American Chemical Society*. doi: 10.1021/ja1069882.
- Liu, Hongli *et al.* (2020) 'The misfolding mechanism of the key fragment R3 of tau protein: A combined molecular dynamics simulation and Markov state model study', *Physical Chemistry Chemical Physics*. doi: 10.1039/c9cp06954b.
- Liu, T. and Bitan, G. (2012) 'Modulating Self-Assembly of Amyloidogenic Proteins as a Therapeutic Approach for Neurodegenerative Diseases: Strategies and Mechanisms', *ChemMedChem*. doi: 10.1002/cmdc.201100585.
- Luhrs, T. *et al.* (2005) '3D structure of Alzheimer's amyloid- $\beta$ (1-42) fibrils.', *Proc.Natl.Acad.Sci.Usa*, 102, pp. 17342–17347. doi: 16293696.
- Lührs, T. *et al.* (2005a) '3D structure of Alzheimer's amyloid-beta(1-42) fibrils.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(48), pp. 17342–7. doi: 10.1073/pnas.0506723102.
- Lührs, T. *et al.* (2005b) '3D structure of Alzheimer's amyloid- $\beta$ (1-42) fibrils', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0506723102.

- Luiken, J. A. and Bolhuis, P. G. (2015) 'Primary Nucleation Kinetics of Short Fibril-Forming Amyloidogenic Peptides', *The Journal of Physical Chemistry B*, 119(39), pp. 12568–12579. doi: 10.1021/acs.jpcc.5b05799.
- M.J. Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the G. and Team, D. (2015) 'GROMACS User Manual version 5.1'.
- Ma, B. and Nussinov, R. (2002) 'Stabilities and conformations of Alzheimer's  $\beta$ -amyloid peptide oligomers (A $\beta$ 16-22, A $\beta$ 16-35, and A $\beta$ 10-35): Sequence effects', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.212206899.
- Ma, L. L. *et al.* (2014) 'Association between cancer and Alzheimer's disease: Systematic review and meta-analysis', *Journal of Alzheimer's Disease*. doi: 10.3233/JAD-140168.
- Ma, Q. *et al.* (2007) 'Heparin Oligosaccharides as Potential Therapeutic Agents in Senile Dementia', *Curr Pharm Des.*, 13(15), pp. 1607–1616. doi: 10.1016/j.micinf.2011.07.011.Innate.
- Madine, J. *et al.* (2013) 'Heparin Promotes the Rapid Fibrillization of a Peptide with Low Intrinsic Amyloidogenicity', *Biochemistry*, 52(50), pp. 8984–8992. doi: 10.1021/bi401231u.
- Maiti, P. *et al.* (2016) 'A comparative study of dietary curcumin, nanocurcumin, and other classical amyloid-binding dyes for labeling and imaging of amyloid plaques in brain tissue of 5 $\times$ -familial Alzheimer's disease mice', *Histochemistry and Cell Biology*. doi: 10.1007/s00418-016-1464-1.
- Maiti, P. and Dunbar, G. L. (2016) 'Rationale for Curcumin Therapy in Alzheimer's Disease', *ARC Journal of Neuroscience*, 1(3), pp. 10–16. doi: 10.20431/2456-057x.0103002.
- Maiti, P. and Dunbar, G. L. (2018) 'Use of curcumin, a natural polyphenol for targeting molecular pathways in treating age-related neurodegenerative diseases', *International Journal of Molecular Sciences*. doi: 10.3390/ijms19061637.
- Malde, A. K. *et al.* (2011) 'An Automated force field Topology Builder (ATB) and repository: Version 1.0', *Journal of Chemical Theory and Computation*. doi: 10.1021/ct200196m.

- Man, V. H., Nguyen, P. H. and Derreumaux, P. (2017) 'Conformational Ensembles of the Wild-Type and S8C A $\beta$ 1-42 Dimers', *Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcc.7b00267.
- Marcus, C., Mena, E. and Subramaniam, R. M. (2014) 'Brain PET in the diagnosis of Alzheimer's disease', *Clinical Nuclear Medicine*. doi: 10.1097/RLU.0000000000000547.
- Maroli, N. and Kolandaivel, P. (2020) 'Comparative study of stability and transport of molecules through cyclic peptide nanotube and aquaporin: a molecular dynamics simulation approach', *Journal of Biomolecular Structure and Dynamics*. doi: 10.1080/07391102.2019.1570341.
- Marrink, S. J. *et al.* (2007) 'The MARTINI force field: Coarse grained model for biomolecular simulations', *Journal of Physical Chemistry B*. doi: 10.1021/jp071097f.
- Martin, T. D. *et al.* (2019) 'Computational Study of the Driving Forces and Dynamics of Curcumin Binding to Amyloid- $\beta$  Protofibrils', *Journal of Physical Chemistry B*. doi: 10.1021/acs.jpcc.8b09185.
- Martyna, G. J., Klein, M. L. and Tuckerman, M. (1992) 'Nosé-Hoover chains: The canonical ensemble via continuous dynamics', *The Journal of Chemical Physics*. doi: 10.1063/1.463940.
- Marzinek, J. K., Huber, R. G. and Bond, P. J. (2020) 'Multiscale modelling and simulation of viruses', *Current Opinion in Structural Biology*. doi: 10.1016/j.sbi.2019.12.019.
- McCammon, J. A., Gelin, B. R. and Karplus, M. (1977) 'Dynamics of folded proteins', *Nature*. doi: 10.1038/267585a0.
- McKenzie, K. J. *et al.* (1996) 'Is  $\beta$ -APP a marker of axonal damage in short-surviving head injury?', *Acta Neuropathologica*. doi: 10.1007/s004010050568.
- McLaurin, J. *et al.* (1999) 'Interactions of Alzheimer amyloid-beta peptides with glycosaminoglycans effects on fibril nucleation and growth.', *European journal of biochemistry*, 266(3), pp. 1101–10. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/10583407> (Accessed: 7 May 2017).
- McShane, R. *et al.* (2019) 'Memantine for dementia', *Cochrane Database of Systematic Reviews*. doi: 10.1002/14651858.CD003154.pub6.

- Meacham, K. and Green, D. (1996) 'Parallelization of the GROMOS87 molecular dynamics code: An update', in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. doi: 10.1007/3-540-61142-8\_544.
- Minton, A. P. (2001) 'The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media', *Journal of Biological Chemistry*. doi: 10.1074/jbc.R100005200.
- Miravalle, L. *et al.* (2000) 'Substitutions at codon 22 of Alzheimer's A $\beta$  peptide induce diverse conformational changes and apoptotic effects human cerebral endothelial cells', *Journal of Biological Chemistry*, 275(35), pp. 27110–27116. doi: 10.1074/jbc.M003154200.
- Morley, J. E. and Farr, S. A. (2014) 'The role of amyloid-beta in the regulation of memory', *Biochemical Pharmacology*. doi: 10.1016/j.bcp.2013.12.018.
- Motamedi-Shad, N. *et al.* (2012) 'Rapid oligomer formation of human muscle acylphosphatase induced by heparan sulfate.', *Nature structural & molecular biology*, 19(5), pp. 547–54, S1-2. doi: 10.1038/nsmb.2286.
- Motamedi-Shad, N., Monsellier, E. and Chiti, F. (2009) 'Amyloid formation by the model protein muscle acylphosphatase is accelerated by heparin and heparan sulphate through a scaffolding-based mechanism', *Journal of Biochemistry*, 146(6), pp. 805–814. doi: 10.1093/jb/mvp128.
- Mucke, L. and Selkoe, D. J. (2012) 'Neurotoxicity of amyloid  $\beta$ -protein: synaptic and network dysfunction.', *Cold Spring Harbor perspectives in medicine*, 2(7), p. a006338. doi: 10.1101/cshperspect.a006338.
- Munter, L. M. *et al.* (2007) 'GxxxG motifs within the amyloid precursor protein transmembrane sequence are critical for the etiology of A $\beta$ 42', *EMBO Journal*, 26(6), pp. 1702–1712. doi: 10.1038/sj.emboj.7601616.
- Nag, S. *et al.* (2011) 'Nature of the amyloid- $\beta$  monomer and the monomer-oligomer equilibrium', *Journal of Biological Chemistry*. doi: 10.1074/jbc.M110.199885.
- Nespovitaya, N. *et al.* (2017) 'Heparin acts as a structural component of  $\beta$ -endorphin amyloid fibrils rather than a simple aggregation promoter', *Chem. Commun.*, 53(7), pp. 1273–1276. doi: 10.1039/C6CC09770G.

- Ngo, S. T. and Li, M. S. (2012) 'Curcumin binds to A $\beta$ 1-40 peptides and fibrils stronger than ibuprofen and naproxen', *Journal of Physical Chemistry B*. doi: 10.1021/jp302506a.
- Nguyen, H. D. and Hall, C. K. (2004) 'Molecular dynamics simulations of spontaneous fibril formation by random-coil peptides.', *Proceedings of the National Academy of Sciences of the United States of America*, 101, pp. 16180–16185. doi: 10.1073/pnas.0407273101.
- Nguyen, H. D. and Hall, C. K. (2006) 'Spontaneous fibril formation by polyalanines; discontinuous molecular dynamics simulations', *Journal of the American Chemical Society*. doi: 10.1021/ja0539140.
- Nguyen, P. H., Li, M. S. and Derreumaux, P. (2011) 'Effects of all-atom force fields on amyloid oligomerization: Replica exchange molecular dynamics simulations of the A $\beta$ 16-22 dimer and trimer', *Physical Chemistry Chemical Physics*, 13(20), pp. 9778–9788. doi: 10.1039/c1cp20323a.
- Nikl, K. *et al.* (2019) 'Alzheimer's disease: Current treatments and potential new agents', *U.S. Pharmacist*.
- Noborn, F. *et al.* (2011) 'Heparan sulfate/heparin promotes transthyretin fibrillization through selective binding to a basic motif in the protein.', *Proceedings of the National Academy of Sciences of the United States of America*, 108(14), pp. 5584–5589. doi: 10.1073/pnas.1101194108.
- NosÉ, S. (2002) 'A molecular dynamics method for simulations in the canonical ensemble', *Molecular Physics*. doi: 10.1080/00268970110089108.
- Okumura, H. and Itoh, S. G. (2016) 'Structural and fluctuational difference between two ends of A $\beta$  amyloid fibril: MD simulations predict only one end has open conformations', *Scientific Reports*. doi: 10.1038/srep38422.
- Okur, A. *et al.* (2003) 'Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins', *Journal of Computational Chemistry*. doi: 10.1002/jcc.10184.
- Ono, K. *et al.* (2004) 'Anti-amyloidogenic activity of tannic acid and its activity to destabilize Alzheimer's ??-amyloid fibrils in vitro', *Biochimica et Biophysica Acta - Molecular Basis of Disease*, 1690(3), pp. 193–202. doi: 10.1016/j.bbadis.2004.06.008.

- Pandey, S. *et al.* (2020) 'An in silico analysis of deleterious single nucleotide polymorphisms and molecular dynamics simulation of disease linked mutations in genes responsible for neurodegenerative disorder', *Journal of Biomolecular Structure and Dynamics*. doi: 10.1080/07391102.2019.1682047.
- Pandi-Perumal, S. R. *et al.* (2008) 'Physiological effects of melatonin: role of melatonin receptors and signal transduction pathways.', *Progress in neurobiology*, 85(3), pp. 335–53. doi: 10.1016/j.pneurobio.2008.04.001.
- Pant, R., Mukhopadhyay, S. and Lakshmayya, B. (2014) 'review article', 3(2), pp. 1–19.
- Paravastu, A. K. *et al.* (2008) 'Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0806270105.
- Parrinello, M and Rahman, A. (1981) 'Polymorphic transitions in single crystals : A new molecular dynamics method Polymorphic transitions in single crystals : A new molecular dynamics method', *Journal of Applied Physics*, 52(12), pp. 7182–7190. doi: 10.1063/1.328693.
- Parrinello, M. and Rahman, A. (1981) 'Polymorphic transitions in single crystals: A new molecular dynamics method', *Journal of Applied Physics*. doi: 10.1063/1.328693.
- Pauling, L., Corey, R. B. and Branson, H. R. (1951) 'The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain.', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.37.4.205.
- Periole, X. *et al.* (2009) 'ELNEDYN: Keep your proteins stable', *J. Chem. Th. Comp.*
- Periole, X. and Marrink, S. J. (2013) 'The martini coarse-grained force field', *Methods in Molecular Biology*. doi: 10.1007/978-1-62703-17-5\_20.
- Petrov, D. and Zagrovic, B. (2014) 'Are Current Atomistic Force Fields Accurate Enough to Study Proteins in Crowded Environments?', *PLoS Computational Biology*, 10(5), p. e1003638. doi: 10.1371/journal.pcbi.1003638.
- Phillips, J. C. *et al.* (2005) 'Scalable molecular dynamics with NAMD', *Journal of Computational Chemistry*. doi: 10.1002/jcc.20289.
- Di Pierro, M., Elber, R. and Leimkuhler, B. (2015) 'A Stochastic Algorithm for the

Isobaric-Isothermal Ensemble with Ewald Summations for All Long Range Forces’, *Journal of Chemical Theory and Computation*. doi: 10.1021/acs.jctc.5b00648.

Pimplikar, S. W. (2014) ‘Neuroinflammation in Alzheimer’s disease: from pathogenesis to a therapeutic target.’, *Journal of clinical immunology*, 34 Suppl 1, pp. S64-9. doi: 10.1007/s10875-014-0032-5.

Pita, R. (2015) ‘The EU Regulatory Landscape of Non-Biological Complex Drugs’, in, pp. 357–380. doi: 10.1007/978-3-319-16241-6\_11.

Pol-Fachin, L. *et al.* (2012) ‘GROMOS 53A6<sub>GLYC</sub>, an Improved GROMOS Force Field for Hexopyranose-Based Carbohydrates’, *Journal of Chemical Theory and Computation*. American Chemical Society, 8(11), pp. 4681–4690. doi: 10.1021/ct300479h.

Pol-Fachin, L., Fernandes, C. L. and Verli, H. (2009) ‘GROMOS96 43a1 performance on the characterization of glycoprotein conformational ensembles through molecular dynamics simulations’, *Carbohydrate Research*. doi: 10.1016/j.carres.2008.12.025.

Pol-Fachin, L. and Verli, H. (2008) ‘Depiction of the forces participating in the 2-O-sulfo- $\alpha$ -l-iduronic acid conformational preference in heparin sequences in aqueous solutions’, *Carbohydrate Research*, 343(9), pp. 1435–1445. doi: 10.1016/j.carres.2008.04.016.

Portillo, A. *et al.* (2015) ‘Role of monomer arrangement in the amyloid self-assembly’, *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1854(3), pp. 218–228. doi: 10.1016/j.bbapap.2014.12.009.

Price, D. J. and Brooks, C. L. (2002) ‘Modern protein force fields behave comparably in molecular dynamics simulations’, *Journal of Computational Chemistry*. doi: 10.1002/jcc.10083.

Priyadarsini, K. I. (2014) ‘The chemistry of curcumin: From extraction to therapeutic agent’, *Molecules*, pp. 20091–20112. doi: 10.3390/molecules191220091.

Qi, R. *et al.* (2018) ‘Replica exchange molecular dynamics: A practical application protocol with solutions to common problems and a peptide aggregation and self-assembly example’, in *Methods in Molecular Biology*. doi: 10.1007/978-1-4939-7811-3\_5.

Quittot, N., Sebastiao, M. and Bourgault, S. (2017) ‘Modulation of amyloid assembly

- by glycosaminoglycans: from mechanism to biological significance', *Biochemistry and Cell Biology*, 95(3), pp. 329–337. doi: 10.1139/bcb-2016-0236.
- Rahman, M. M. *et al.* (2020) 'Virtual screening, molecular dynamics and structure–activity relationship studies to identify potent approved drugs for Covid-19 treatment', *Journal of Biomolecular Structure and Dynamics*. doi: 10.1080/07391102.2020.1794974.
- Ramachandran, G. and Udgaonkar, J. B. (2011) 'Understanding the kinetic roles of the inducer heparin and of rod-like protofibrils during amyloid fibril formation by Tau protein.', *The Journal of biological chemistry*. American Society for Biochemistry and Molecular Biology, 286(45), pp. 38948–59. doi: 10.1074/jbc.M111.271874.
- Ramaker, J. M. *et al.* (2016) 'Amyloid precursor proteins are dynamically trafficked and processed during neuronal development', *Frontiers in Molecular Neuroscience*. doi: 10.3389/fnmol.2016.00130.
- Rambaran, R. N. and Serpell, L. C. (2008) 'Amyloid fibrils: abnormal protein assembly.', *Prion*. doi: 10.4161/pri.2.3.7488.
- Reinke, A. A. and Gestwicki, J. E. (2007) 'Structure-activity relationships of amyloid beta-aggregation inhibitors based on curcumin: Influence of linker length and flexibility', *Chemical Biology and Drug Design*. doi: 10.1111/j.1747-0285.2007.00557.x.
- Rivière, C. *et al.* (2007) 'Inhibitory activity of stilbenes on Alzheimer's  $\beta$ -amyloid fibrils in vitro', *Bioorganic and Medicinal Chemistry*. doi: 10.1016/j.bmc.2006.09.069.
- Rogawski, M. A. and Wenk, G. L. (2003) 'The neuropharmacological basis for the use of memantine in the treatment of Alzheimer's disease.', *CNS drug reviews*, 9(3), pp. 275–308.
- Rosenblum, W. I. (2014) 'Why Alzheimer trials fail: Removing soluble oligomeric beta amyloid is essential, inconsistent, and difficult', *Neurobiology of Aging*. Elsevier Ltd, 35(5), pp. 969–974. doi: 10.1016/j.neurobiolaging.2013.10.085.
- Ryu, E. K. *et al.* (2006) 'Curcumin and dehydrozingerone derivatives: Synthesis, radiolabeling, and evaluation for  $\beta$ -amyloid plaque imaging', in *Journal of Medicinal Chemistry*. doi: 10.1021/jm0607193.
- Sanders, D. W. *et al.* (2016) 'Prions and Protein Assemblies that Convey Biological

- Information in Health and Disease', *Neuron*. doi: 10.1016/j.neuron.2016.01.026.
- Santini, S. *et al.* (2004) 'Pathway complexity of Alzheimer's  $\beta$ -amyloid A $\beta$ 16-22 peptide assembly', *Structure*. doi: 10.1016/j.str.2004.04.018.
- Santos, L. M. *et al.* (2016) 'Resveratrol administration increases transthyretin protein levels, ameliorating AD features: The importance of transthyretin tetrameric stability', *Molecular Medicine*. doi: 10.2119/molmed.2016.00124.
- Sarkar, M., Li, C. and Pielak, G. J. (2013) 'Soft interactions and crowding', *Biophysical Reviews*. doi: 10.1007/s12551-013-0104-4.
- Di Scala, C. *et al.* (2016) 'Common molecular mechanism of amyloid pore formation by Alzheimer's  $\beta$ -amyloid peptide and  $\alpha$ -synuclein', *Scientific Reports*. doi: 10.1038/srep28781.
- Schmidt, M. *et al.* (2009) 'Comparison of Alzheimer A $\beta$ (1-40) and A $\beta$ (1-42) amyloid fibrils reveals similar protofilament structures', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0905007106.
- Schmidt, M. *et al.* (2015) 'Peptide dimer structure in an A $\beta$ (1-42) fibril visualized with cryo-EM', *Proceedings of the National Academy of Sciences*, 112(38), pp. 11858–11863. doi: 10.1073/pnas.1503455112.
- Schmit, J. D., Ghosh, K. and Dill, K. (2011) 'What drives amyloid molecules to assemble into oligomers and fibrils?', *Biophysical journal*, 100(2), pp. 450–8. doi: 10.1016/j.bpj.2010.11.041.
- Schuler, L. D., Daura, X. and Van Gunsteren, W. F. (2001) 'An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase', *Journal of Computational Chemistry*, 22(11), pp. 1205–1218. doi: 10.1002/jcc.1078.
- Schüttelkopf, A. W. and Van Aalten, D. M. F. (2004) 'PRODRG: A tool for high-throughput crystallography of protein-ligand complexes', *Acta Crystallographica Section D: Biological Crystallography*, 60(8), pp. 1355–1363. doi: 10.1107/S09074444904011679.
- Schwierz, N. *et al.* (2016) 'Dynamics of Seeded A $\beta$ 40-Fibril Growth from Atomistic Molecular Dynamics Simulations: Kinetic Trapping and Reduced Water Mobility in the Locking Step', *Journal of the American Chemical Society*. doi: 10.1021/jacs.5b08717.

- Scott, W. R. P. *et al.* (1999) 'The GROMOS biomolecular simulation program package', *Journal of Physical Chemistry A*. doi: 10.1021/jp984217f.
- Selivanovitch, E. and Douglas, T. (2019) 'Virus capsid assembly across different length scales inspire the development of virus-based biomaterials', *Current Opinion in Virology*. doi: 10.1016/j.coviro.2019.02.010.
- Sengupta, U., Nilson, A. N. and Kaye, R. (2016) 'The Role of Amyloid- $\beta$  Oligomers in Toxicity, Propagation, and Immunotherapy', *EBioMedicine*, 6, pp. 42–49. doi: 10.1016/j.ebiom.2016.03.035.
- Serrano-Pozo, A. *et al.* (2011) 'Neuropathological alterations in Alzheimer disease', *Cold Spring Harbor Perspectives in Medicine*. doi: 10.1101/cshperspect.a006189.
- Seyrek, E. *et al.* (2003) 'Ionic strength dependence of protein-polyelectrolyte interactions', *Biomacromolecules*, 4(2), pp. 273–282. doi: 10.1021/bm025664a.
- Shao, H. *et al.* (1999) 'Solution structures of micelle-bound amyloid  $\beta$ -(1-40) and  $\beta$ -(1-42) peptides of Alzheimer's disease', *Journal of Molecular Biology*. doi: 10.1006/jmbi.1998.2348.
- Shaw, D. E. *et al.* (2008) 'Anton, a special-purpose machine for molecular dynamics simulation', *Communications of the ACM*. doi: 10.1145/1364782.1364802.
- Shi, Q., Izvekov, S. and Voth, G. A. (2006) 'Mixed atomistic and coarse-grained molecular dynamics: Simulation of a membrane-bound ion channel', *Journal of Physical Chemistry B*. doi: 10.1021/jp062700h.
- Shivu, B. *et al.* (2013) 'Distinct  $\beta$ -Sheet Structure in Protein Aggregates Determined by ATR-FTIR Spectroscopy', *Biochemistry*, 52(31), pp. 5176–5183. doi: 10.1021/bi400625v.
- Siwy, C. M., Lockhart, C. and Klimov, D. K. (2017) 'Is the Conformational Ensemble of Alzheimer's A $\beta$ 10-40 Peptide Force Field Dependent?', *PLoS Computational Biology*. doi: 10.1371/journal.pcbi.1005314.
- Sjögren, M. and Blennow, K. (2005) 'The link between cholesterol and Alzheimer's disease.', *The world journal of biological psychiatry: the official journal of the World Federation of Societies of Biological Psychiatry*, 6(2), pp. 85–97.
- Smaoui, M. R. *et al.* (2013) 'Computational assembly of polymorphic amyloid fibrils

reveals stable aggregates’, *Biophysical Journal*, 104(3), pp. 683–693. doi: 10.1016/j.bpj.2012.12.037.

Smith, M. D. *et al.* (2015) ‘Force-Field Induced Bias in the Structure of A $\beta$ 21-30: A Comparison of OPLS, AMBER, CHARMM, and GROMOS Force Fields.’, *Journal of chemical information and modeling*, 55(12), pp. 2587–95. doi: 10.1021/acs.jcim.5b00308.

So, M. *et al.* (2017) ‘Heparin-induced amyloid fibrillation of  $\beta$ 2-microglobulin explained by solubility and a supersaturation-dependent conformational phase diagram’, *Protein Science*, 26(5), pp. 1024–1036. doi: 10.1002/pro.3149.

Sofronova, A. A. *et al.* (2017) ‘Protein-polyelectrolyte complexes: Molecular dynamics simulations and experimental study’, *Polymer*, 113, pp. 39–45. doi: 10.1016/j.polymer.2017.02.047.

Sørensen, J. *et al.* (2011) ‘Protofibrillar Assembly Toward the Formation of Amyloid Fibrils’, *The Journal of Physical Chemistry Letters*, 2(19), pp. 2385–2390. doi: 10.1021/jz2010094.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, Alan E., *et al.* (2005) ‘GROMACS: fast, flexible, and free.’, *Journal of computational chemistry*, 26(16), pp. 1701–18. doi: 10.1002/jcc.20291.

Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, Alan E., *et al.* (2005) ‘GROMACS: Fast, flexible, and free’, *Journal of Computational Chemistry*, pp. 1701–1718. doi: 10.1002/jcc.20291.

Spreiter, Q. and Walter, M. (1999) ‘Classical Molecular Dynamics Simulation with the Velocity Verlet Algorithm at Strong External Magnetic Fields’, *Journal of Computational Physics*, 152(1), pp. 102–119. doi: 10.1006/jcph.1999.6237.

Stathopoulos, P. B. *et al.* (2008) ‘Sonication of proteins causes formation of aggregates that resemble amyloid’, *Protein Science*, 13(11), pp. 3017–3027. doi: 10.1110/ps.04831804.

Stefani, M. and Dobson, C. M. (2003) ‘Protein aggregation and aggregate toxicity: New insights into protein folding, misfolding diseases and biological evolution’, *Journal of Molecular Medicine*, pp. 678–699. doi: 10.1007/s00109-003-0464-5.

Stewart, K. L. and Radford, S. E. (2017) ‘Amyloid plaques beyond A $\beta$ : a survey of the

diverse modulators of amyloid aggregation’, *Biophysical Reviews*, 9(4), pp. 405–419. doi: 10.1007/s12551-017-0271-9.

Streets, A. M. *et al.* (2013) ‘Simultaneous Measurement of Amyloid Fibril Formation by Dynamic Light Scattering and Fluorescence Reveals Complex Aggregation Kinetics’, *PLoS ONE*. doi: 10.1371/journal.pone.0054541.

Suárez, D. and Díaz, N. (2020) ‘SARS-CoV-2 Main Protease: A Molecular Dynamics Study’, *Journal of Chemical Information and Modeling*. doi: 10.1021/acs.jcim.0c00575.

Sugita, Y. and Okamoto, Y. (1999) ‘Replica-exchange molecular dynamics method for protein folding’, *Chemical Physics Letters*. doi: 10.1016/S0009-2614(99)01123-9.

Suk, J. Y. *et al.* (2006) ‘Heparin accelerates gelsolin amyloidogenesis’, *Biochemistry*, 45(7), pp. 2234–2242. doi: 10.1021/bi0519295.

Swope, W. C. *et al.* (1982) ‘A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters’, *The Journal of Chemical Physics*. doi: 10.1063/1.442716.

Szala-Mendyk, B. and Molski, A. (2020) ‘Clustering and fibril formation during GNNQQNY aggregation: A molecular dynamics study’, *Biomolecules*. doi: 10.3390/biom10101362.

Tafrishi, H. *et al.* (2020) ‘Investigation of tetracosane thermal transport in presence of graphene and carbon nanotube fillers—A molecular dynamics study’, *Journal of Energy Storage*. doi: 10.1016/j.est.2020.101321.

Takase, H. *et al.* (2016) ‘Structural requirements of glycosaminoglycans for facilitating amyloid fibril formation of human serum amyloid A’, *Amyloid*, 23(2), pp. 1350–6129. doi: 10.3109/13506129.2016.1168292.

Tarasoff-Conway, J. M. *et al.* (2015) ‘Clearance systems in the brain - Implications for Alzheimer disease’, *Nature Reviews Neurology*. doi: 10.1038/nrneurol.2015.119.

Teodoro, G. *et al.* (2009) ‘Coordinating the use of GPU and CPU for improving performance of compute intensive applications’, in *Proceedings - IEEE International Conference on Cluster Computing, ICC*. doi: 10.1109/CLUSTER.2009.5289193.

Thomas, M. and Schwartz, R. (2017) ‘Quantitative computational models of molecular

self-assembly in systems biology’, *Physical Biology*, 14(3). doi: 10.1088/1478-3975/aa6cdc.

Tofoleanu, F. and Buchete, N.-V. (2012) ‘Molecular Interactions of Alzheimer’s A $\beta$  Protofilaments with Lipid Membranes’, *Journal of Molecular Biology*. Elsevier Ltd, 421(4–5), pp. 572–586. doi: 10.1016/j.jmb.2011.12.063.

Tomaselli, S. *et al.* (2006) ‘The  $\alpha$ -to- $\beta$  conformational transition of Alzheimer’s A $\beta$ -(1-42) peptide in aqueous media is reversible: A step by step conformational analysis suggests the location of  $\beta$  conformation seeding’, *ChemBioChem*. doi: 10.1002/cbic.200500223.

Verli, H. and Guimarães, J. A. (2004) ‘Molecular dynamics simulation of a decasaccharide fragment of heparin in aqueous solution’, *Carbohydrate Research*, 339(2), pp. 281–290. doi: 10.1016/j.carres.2003.09.026.

Verma, M., Vats, A. and Taneja, V. (2015) ‘Toxic species in amyloid disorders: Oligomers or mature fibrils’, *Annals of Indian Academy of Neurology*, 18, pp. 138–145. doi: 10.4103/0972-2327.144284.

Vilasi, S. *et al.* (2011) ‘Heparin induces harmless fibril formation in amyloidogenic W7FW14F apomyoglobin and amyloid aggregation in wild-type protein In Vitro’, *PLoS ONE*, 6(7). doi: 10.1371/journal.pone.0022076.

Vivekanandan, S. *et al.* (2011) ‘A partially folded structure of amyloid-beta(1-40) in an aqueous environment’, *Biochemical and Biophysical Research Communications*. doi: 10.1016/j.bbrc.2011.06.133.

Walker, D. and Lue, L.-F. (2007) ‘Anti-inflammatory and immune therapy for Alzheimer’s disease: current status and future directions.’, *Current neuropharmacology*, 5(4), pp. 232–43. doi: 10.2174/157015907782793667.

Walker, L. C. and LeVine, H. (2000) ‘The cerebral proteopathies: Neurodegenerative disorders of protein conformation and assembly’, *Molecular Neurobiology*.

Walsh, D. M. *et al.* (1997) ‘Amyloid  $\beta$ -protein fibrillogenesis–Detection of a protofibrillar intermediate’, *Journal of Biological Chemistry*, 272(35), pp. 22364–22372. doi: 10.1074/jbc.272.35.22364.

Walton, E. B. and Vanvliet, K. J. (2006) ‘Equilibration of experimentally determined protein structures for molecular dynamics simulation.’, *Physical review. E, Statistical,*

*nonlinear, and soft matter physics*, 74(6 Pt 1), p. 061901. Available at: <http://www.ncbi.nlm.nih.gov/pubmed/17280090> (Accessed: 30 September 2015).

Wang, R. G., Zhang, H. X. and Zheng, Q. C. (2020) ‘Revealing the binding and drug resistance mechanism of amprenavir, indinavir, ritonavir, and nelfinavir complexed with HIV-1 protease due to double mutations G48T/L89M by molecular dynamics simulations and free energy analyses’, *Physical Chemistry Chemical Physics*. doi: 10.1039/c9cp06657h.

Watanabe-Nakayama, T. *et al.* (2016) ‘High-speed atomic force microscopy reveals structural dynamics of amyloid  $\beta$ 1-42 aggregates’, *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1524807113.

Watts, C. R. *et al.* (2018) ‘Effects of force fields on the conformational and dynamic properties of amyloid  $\beta$ (1-40) dimer explored by replica exchange molecular dynamics simulations’, *Proteins: Structure, Function and Bioinformatics*, 86(3), pp. 279–300. doi: 10.1002/prot.25439.

Webb, B. and Sali, A. (2016) ‘Comparative protein structure modeling using MODELLER’, *Current Protocols in Bioinformatics*. doi: 10.1002/cpbi.3.

Weber, O. C. and Uversky, V. N. (2017) ‘How accurate are your simulations? Effects of confined aqueous volume and AMBER FF99SB and CHARMM22/CMAP force field parameters on structural ensembles of intrinsically disordered proteins: Amyloid- $\beta$  42 in water’, *Intrinsically Disordered Proteins*. Taylor & Francis, 5(1), p. e1377813. doi: 10.1080/21690707.2017.1377813.

Weber, P. C. (1991) ‘Physical Principles of Protein Crystallization’, *Advances in Protein Chemistry*. Academic Press, 41, pp. 1–36. doi: 10.1016/S0065-3233(08)60196-5.

Wei, G., Mousseau, N. and Derreumaux, P. (2007) ‘Computational simulations of the early steps of protein aggregation.’, *Prion*. doi: 10.4161/pri.1.1.3969.

Wei, Y. *et al.* (2014) ‘Design of novel cell penetrating peptides for the delivery of trehalose into mammalian cells’, *Biochimica et Biophysica Acta - Biomembranes*. doi: 10.1016/j.bbamem.2014.02.011.

Westermarck, G. T. *et al.* (2018) ‘Noncerebral amyloidoses: Aspects on seeding, cross-seeding, and transmission’, *Cold Spring Harbor Perspectives in Medicine*. doi:

10.1101/cshperspect.a024323.

*What's New in the Alzheimer's Treatment Pipeline? | BrightFocus Foundation-2020* (no date). Available at: <https://www.brightfocus.org/alzheimers/article/whats-new-alzheimers-treatment-pipeline> (Accessed: 28 March 2020).

Wu, C. *et al.* (2008) 'The Binding of Thioflavin T and Its Neutral Analog BTA-1 to Protofibrils of the Alzheimer's Disease  $\text{A}\beta_{16-22}$  Peptide Probed by Molecular Dynamics Simulations', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2008.09.062.

Wu, C., Bowers, M. T. and Shea, J. E. (2011) 'On the origin of the stronger binding of PIB over Thioflavin T to protofibrils of the Alzheimer Amyloid- $\beta$  peptide: A molecular dynamics study', *Biophysical Journal*. doi: 10.1016/j.bpj.2011.01.058.

Wu, C. and Shea, J. E. (2011) 'Coarse-grained models for protein aggregation', *Current Opinion in Structural Biology*. doi: 10.1016/j.sbi.2011.02.002.

Xiao, Y. *et al.* (2015a) ' $\text{A}\beta_{(1-42)}$  fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease.', *Nature structural & molecular biology*, 22(6), pp. 499–505. doi: 10.1038/nsmb.2991.

Xiao, Y. *et al.* (2015b) ' $\text{A}\beta_{(1-42)}$  fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease', *Nature Structural & Molecular Biology*, 22(6), pp. 499–505. doi: 10.1038/nsmb.2991.

Xiong, Z. *et al.* (2011) 'Curcumin mediates presenilin-1 activity to reduce  $\beta$ -amyloid production in a model of Alzheimer's disease', *Pharmacological Reports*. doi: 10.1016/s1734-1140(11)70629-6.

Xu, L., Chen, Y. and Wang, X. (2014) 'Assembly of amyloid  $\beta$  peptides in the presence of fibril seeds: One-pot coarse-grained molecular dynamics simulations', *Journal of Physical Chemistry B*. doi: 10.1021/jp505551m.

Xu, Y. *et al.* (2005) 'Conformational transition of amyloid  $\beta$ -peptide', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0501218102.

Xu, Y. *et al.* (2012) 'Effect of heparin on protein aggregation: Inhibition versus promotion', *Biomacromolecules*, 13(5), pp. 1642–1651. doi: 10.1021/bm3003539.

- Yakupova, E. I. *et al.* (2019) 'Congo Red and amyloids: History and relationship', *Bioscience Reports*. doi: 10.1042/BSR20181415.
- Yan, Y. and Wang, C. (2006) 'A $\beta$ 42 is More Rigid than A $\beta$ 40 at the C Terminus: Implications for A $\beta$  Aggregation and Toxicity', *Journal of Molecular Biology*. doi: 10.1016/j.jmb.2006.09.046.
- Yang, F. *et al.* (2005) 'Curcumin Inhibits Formation of Amyloid Oligomers and Fibrils, Binds Plaques, and Reduces Amyloid in Vivo', *Journal of Biological Chemistry*, 280(7), pp. 5892–5901. doi: 10.1074/jbc.M404751200.
- Yang, G. *et al.* (2006) 'Crystal shape control by manipulating supersaturation in batch cooling crystallization', *Crystal Growth and Design*. doi: 10.1021/cg0603873.
- Yeh, I. C. and Hummer, G. (2002) 'Peptide loop-closure kinetics from microsecond molecular dynamics simulations in explicit solvent', *Journal of the American Chemical Society*. doi: 10.1021/ja025789n.
- Yun, S. *et al.* (2007) 'Role of Electrostatic Interactions in Amyloid  $\beta$ -Protein (A $\beta$ ) Oligomer Formation: A Discrete Molecular Dynamics Study', *Biophysical Journal*, 92(11), pp. 4064–4077. doi: 10.1529/biophysj.106.097766.
- Zhang-Nunes, S. X. *et al.* (2006) 'The Cerebral  $\beta$ -Amyloid Angiopathies: Hereditary and Sporadic', *Brain Pathology*, 16(1), pp. 30–39. doi: 10.1111/j.1750-3639.2006.tb00559.x.
- Zhang, G. *et al.* (2014) 'Towards understanding the roles of heparan sulfate proteoglycans in Alzheimer's disease.', *BioMed research international*, 2014, p. 516028. doi: 10.1155/2014/516028.
- Zhang, R. *et al.* (2009) 'Interprotofilament interactions between alzheimer's  $\alpha\beta$  1-42 peptides in amyloid fibrils revealed by cryoEM', *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.0901085106.
- Zhang, S. *et al.* (2000) 'The Alzheimer's peptide A $\beta$  adopts a collapsed coil structure in water', *Journal of Structural Biology*. doi: 10.1006/jsbi.2000.4288.
- Zhang, Z. *et al.* (2007) 'Molecular Dynamics Simulations on the Oligomer-Formation Process of the GNNQQNY Peptide from Yeast Prion Protein Sup35', *Biophysical Journal*, 93(5), pp. 1484–1492. doi: 10.1529/biophysj.106.100537.

Zhao, H. F. *et al.* (2015) 'Resveratrol decreases the insoluble A $\beta$ 1-42 level in hippocampus and protects the integrity of the blood-brain barrier in AD rats', *Neuroscience*. doi: 10.1016/j.neuroscience.2015.10.006.

Zhao, L. N. *et al.* (2012) 'The Effect of Curcumin on the Stability of A $\beta$  Dimers', *The Journal of Physical Chemistry B*. American Chemical Society, 116(25), pp. 7428–7435. doi: 10.1021/jp3034209.

Zhao, Y. and Zhao, B. (2013) 'Review Article Oxidative Stress and the Pathogenesis of Alzheimer ' s Disease', 2013.

Zheng, J. *et al.* (2007) 'Modeling the Alzheimer A $\beta$ 17-42 fibril architecture: Tight intermolecular sheet-sheet association and intramolecular hydrated cavities', *Biophysical Journal*. doi: 10.1529/biophysj.107.110700.

Zhou, H. X. and Bates, P. A. (2013) 'Modeling protein association mechanisms and kinetics', *Current Opinion in Structural Biology*. doi: 10.1016/j.sbi.2013.06.014.

Zilka, N. and Novak, M. (2006) 'The tangled story of Alois Alzheimer.', *Bratislava Medical Journal*, 107, pp. 343–345.

# 9 Appendices

Appendix A.....	188
Appendix B.....	197
Appendix C.....	205

## Appendix A

**Table 9.1 Summary of A $\beta$  structures simulated in this study**

Structure (PDB code)	Peptide	Sequence	Starting conformation (stride(Heinig and Frishman, 2004)) Main secondary structure
Monomer (1IYT)(Orlando Crescenzi, Tomaselli, Guerrini, Salvadori, Anna M D'Ursi, <i>et al.</i> , 2002)	A $\beta$ (1-42)	DAEFRHDSGYEVHHQKLVFFA EDVGSNKGAIIGLMVGGVVIA	71% helical
5-mer fibril (2BEG)(Lührs <i>et al.</i> , 2005a)	A $\beta$ (17-42)	LVFFAEDVGSNKGAIIGLMVG GVVIA	47% $\beta$ -sheet
12-mer fibril (2MXU)(Xiao <i>et al.</i> , 2015b)	A $\beta$ (11-42)	EVHHQKLVFFAEDVGSNKGAIIGLMVGGVVIA	54% $\beta$ -sheet
Ring (M3Q_3)	A $\beta$ (1-40)	MDAEFRHDSGYEVHHQKLVFF AEDVGSNKGAIIGLMVGGVV	Turn- $\beta$ -sheet-turn ( )

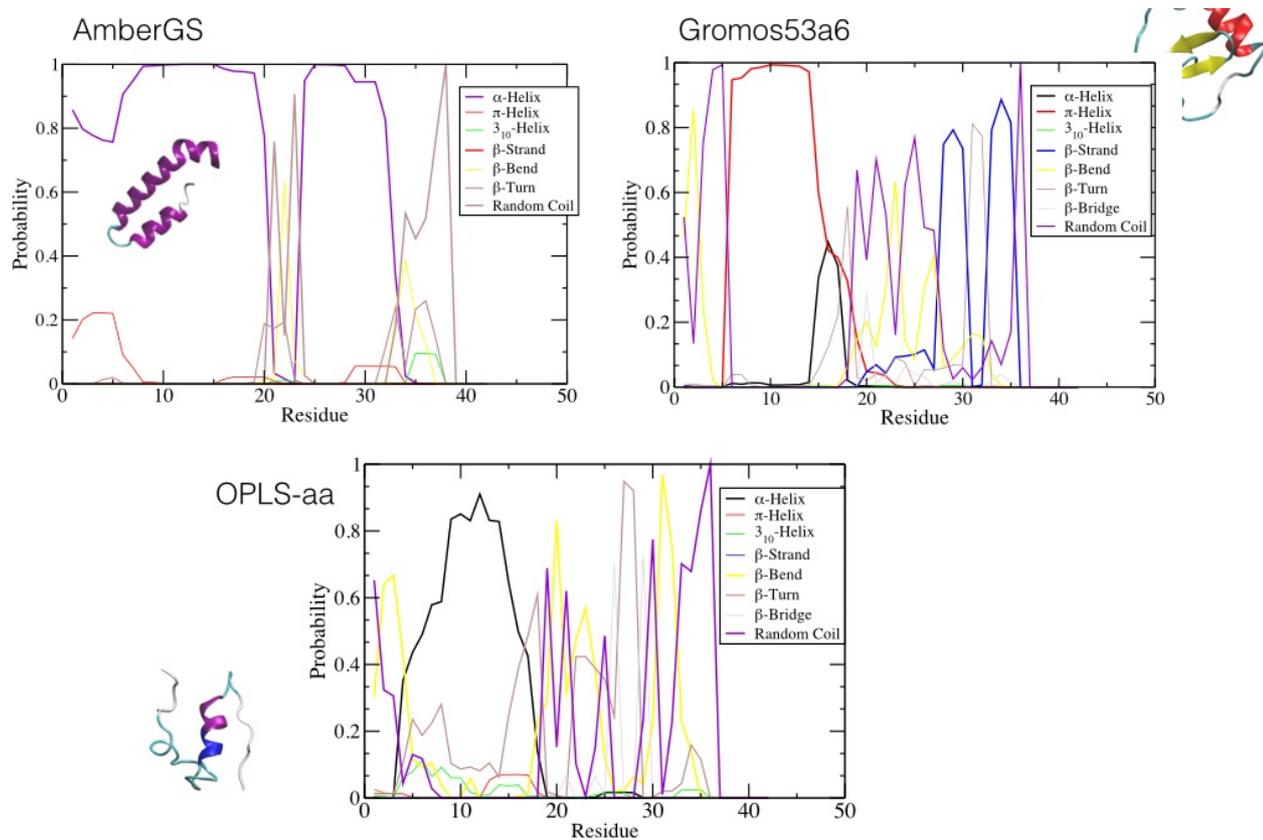
**Table 9.2 Summary of simulation details**

PDB	No. of atoms (Chains)	No. of water molecules	Box size x y z (nm)	No. of counterions
1IYT	637 (1)	3865	5x 5 x5	3
2BEG	1860(5)	6493	6 x6 x6	5
2MXU	5724(12)	31218	10x10x10	12
M3Q_3	5670(18)	69101	13x13x13	54

## 9.1 A $\beta$ monomer

**Table 9.3** The details of secondary structure content of A $\beta$  monomer

Forcefield	$\alpha$ -helix (%)	$\beta$ -sheet (%)	Coil (%)	Bend (%)	Turn (%)
AmberGS	73	0	15	4	4
Gromos53a6	3	14	35	13	8
OPLS-aa	22	0	33	21	18
CD	3	12	59	-	26



**Figure 9.1.** The averaged secondary-structure probability per residue in terms of regular secondary structure characteristics i-e  $\alpha$ -helices,  $\beta$ -sheets, turns and coils etc. The structure in AmberGS shows an elbow-shaped structure containing two  $\alpha$ -helices separated by a turn region at residue VAL24-ASN27, Gromos53aa6 favors the formation of  $\beta$ -sheets while OPLS-aa forms structure that is a mixture of helix and turns

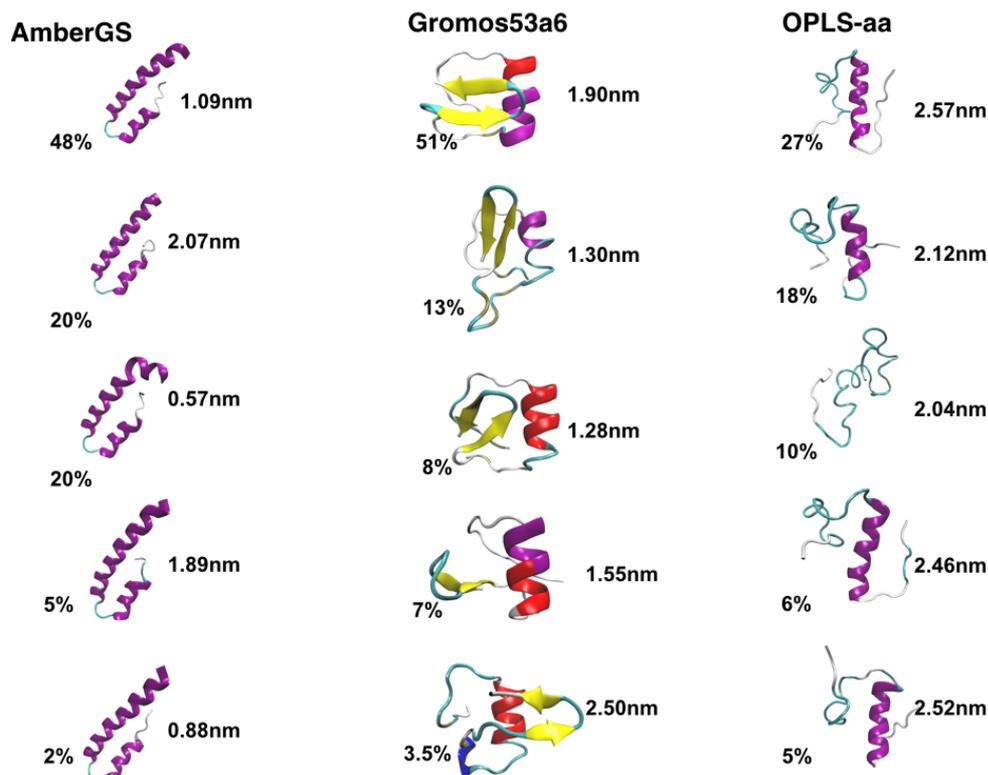


Figure 9.2 The five most populated clusters for each simulated force field, the percentage of snapshots each cluster represents. (RMSD cut-off 0.25nm) and the end-to-end distance of the  $A\beta_{42}$  monomer showing a change in the structure of various conformations.

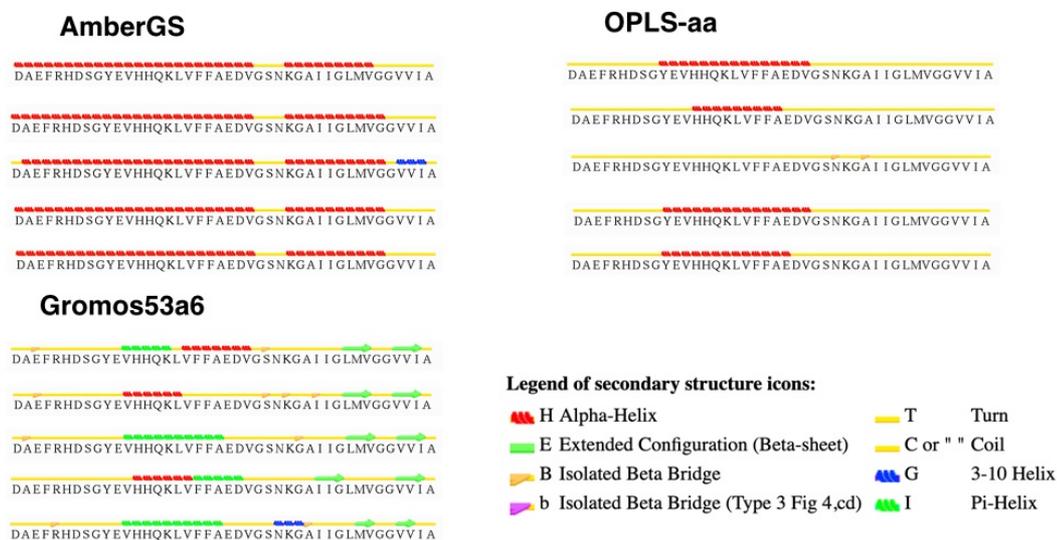


Figure 9.3. Secondary structure of five most populated clusters of  $A\beta_{42}$  monomer showing change in the structure of various conformations for each simulated force field, calculated from stride\*

## 9.2 U-shaped protofibril

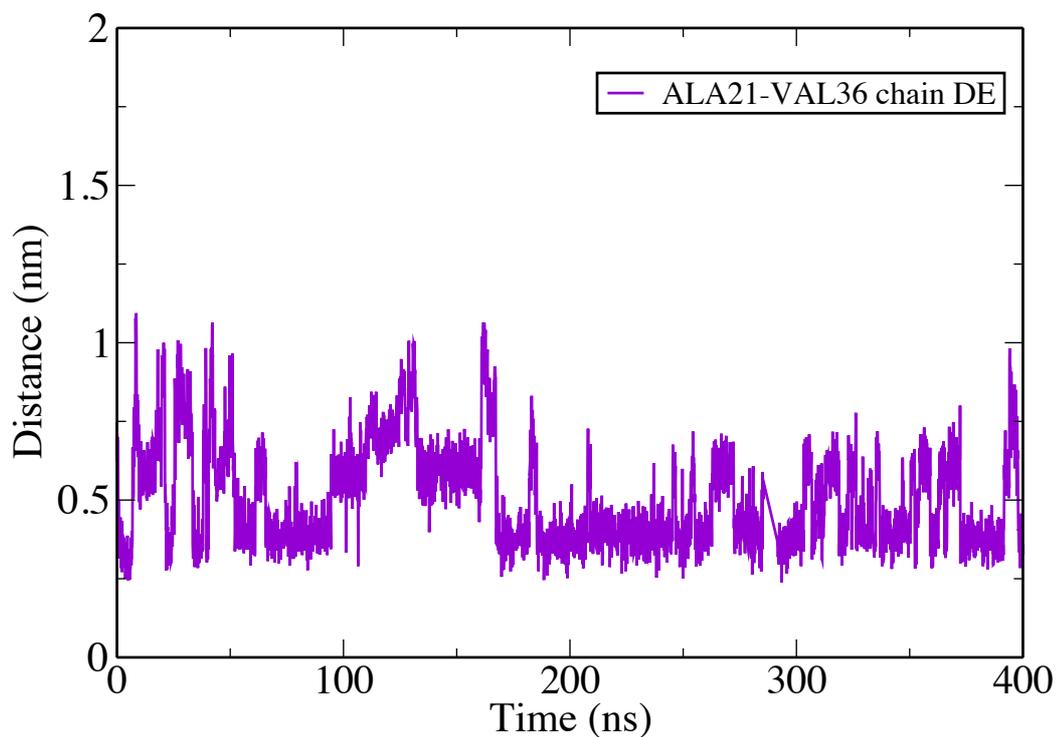
**Table 9.4. Average interpeptide distances U-shaped protofibril (nm  $\pm$  Std. Dev.)<sup>1</sup>**

Forcefields	PHE19/GLY38 (nm)	ALA21/VAL36 (nm)
AmberGS		
Chain AB	0.63 $\pm$ 0.04	0.62 $\pm$ 0.09
Chain BC	0.70 $\pm$ 0.03	0.68 $\pm$ 0.04
Chain CD	0.87 $\pm$ 0.05	0.80 $\pm$ 0.01
Chain DE	0.82 $\pm$ 0.03	0.50 $\pm$ 0.15
Gromos53a6		
Chain AB	0.98 $\pm$ 0.1	0.74 $\pm$ 0.06
Chain BC	1.02 $\pm$ 0.04	0.70 $\pm$ 0.03
Chain CD	1.00 $\pm$ 0.06	0.69 $\pm$ 0.04
Chain DE	0.98 $\pm$ 0.08	0.68 $\pm$ 0.05
OPLS-aa		
Chain AB	1.22 $\pm$ 0.1	1.06 $\pm$ 0.12
Chain BC	1.18 $\pm$ 0.11	0.99 $\pm$ 0.08
Chain CD	1.12 $\pm$ 0.10	0.98 $\pm$ 0.06
Chain DE	1.10 $\pm$ 0.10	0.96 $\pm$ 0.07

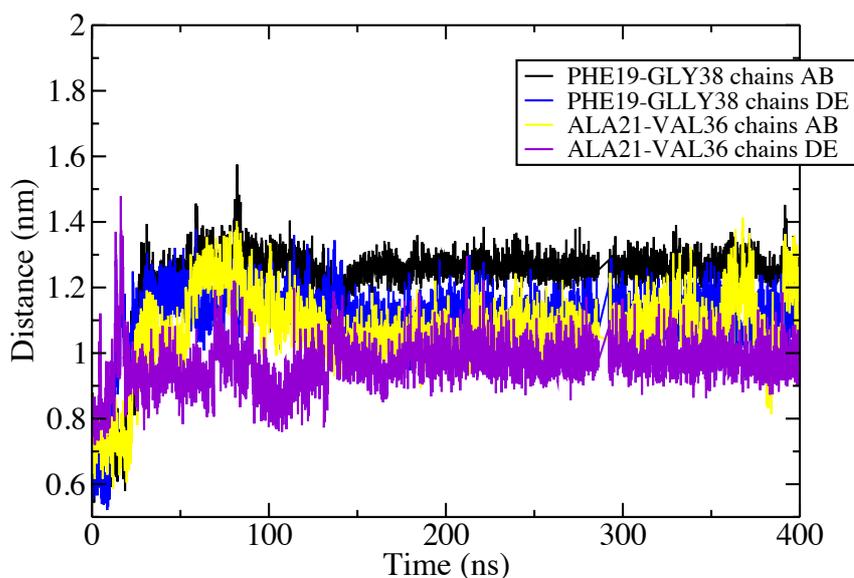
<sup>1</sup> The distance between the side chains of PHE19-GLY38 and ALA21-VAL36 was calculated to study their hydrophobic interaction. The distances reproduced by AmberGS are very close to NMR structures i.e. ( $\approx$ 0.8nm)

**Table 9.5. Average interpeptide Asp23 – Lys28 NH<sub>3</sub><sup>+</sup> – COO<sup>-</sup> distances (nm  $\pm$  Std. Dev.)**

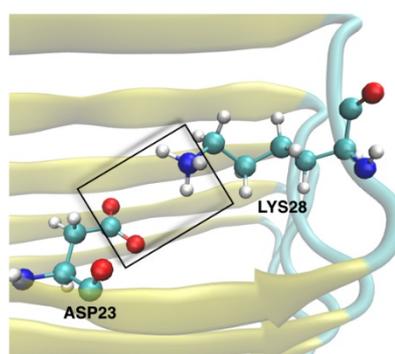
Forcefields	Chain AB	Chain BC	Chain CD	Chain DE
AmberGS	0.38 $\pm$ 0.05	0.37 $\pm$ 0.02	0.38 $\pm$ 0.02	0.31 $\pm$ 0.01
Gromos53a6	0.41 $\pm$ 0.1	0.37 $\pm$ 0.07	0.40 $\pm$ 0.08	0.58 $\pm$ 0.2
OPLS-aa	0.37 $\pm$ 0.07	0.35 $\pm$ 0.02	0.36 $\pm$ 0.01	0.35 $\pm$ 0.03



**Figure 9.4.  $C\alpha$ - $C\alpha$  distance between ALA21 and VAL36 in the last two chains of U-shaped protofibril in AmberGS showing large fluctuations in the lower end of the protofibril.**



**Figure 9.5.**  $\text{Ca-Ca}$  distance between PHE19 and GLY38 and ALA21 and VAL36 in the top and bottom chains of U-shaped protofibril in OPLS-aa showing fluctuations in the beginning of the simulations but later converge to an equilibrium distance.



**Figure 9.6.** Asp23 – Lys28 ( $\text{COO}^- - \text{NH}_3^+$ ) forming the salt bridge responsible for maintaining the U-shape of the U-shaped protofibril.

### 9.3 S-shaped fibril

**Table 9.6** Percentage of the secondary structure content S-shaped fibril calculated from the DSSP program from 400ns long simulation for AmberGS, Gromos53a6, and OPLS-aa forcefields.

Forcefields	Coil (%)	$\beta$ -sheet (%)	Bend (%)	Turn (%)
AmberGS	25	57	11	3
Gromos53a6	12	80	4	0
OPLS-aa	36	42	16	1

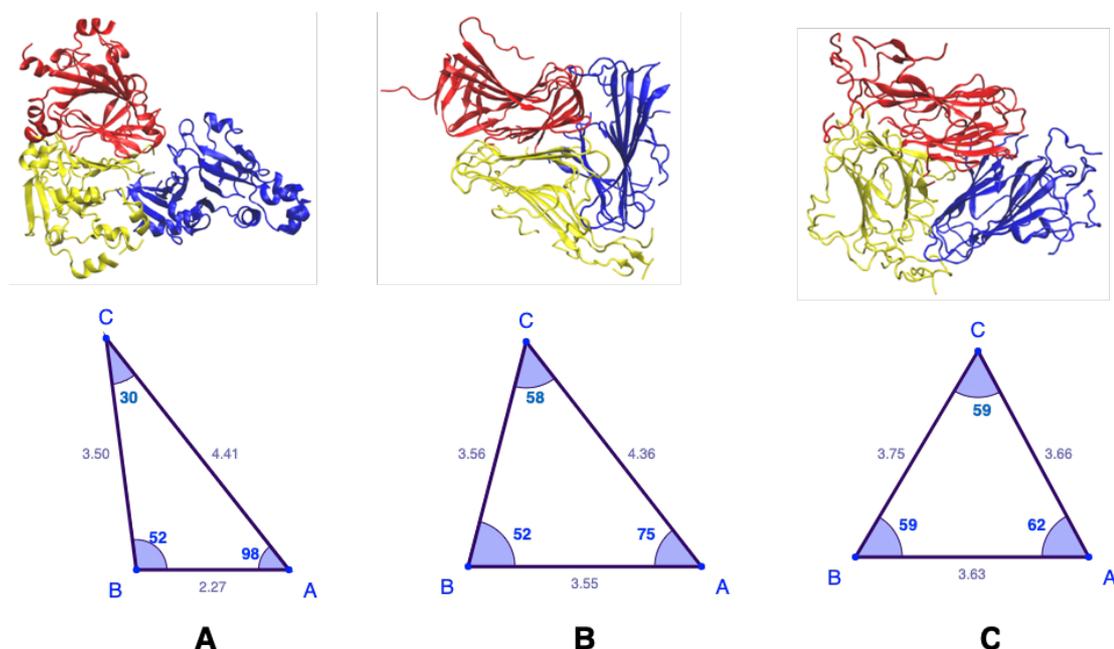
**Table 9.7 Interchain distances between HIS14, LYS17, PHE19 & 20, ASN27, ALA30 and LEU34 of two adjacent chains of S-shaped fibril.**

Forcefields	HIS14	LEU17	PHE19	PHE20	ASN27	ALA30	LEU34
(C $\alpha$ -C $\alpha$ )nm	-	-	-	-	-	-	-
	HIS14	LEU17	PHE19	PHE20	ASN27	ALA30	LEU34
AmberGS	0.48	0.49	0.50	0.51	0.48	0.48	0.51
Gromos53a6	0.49	0.49	0.48	0.53	0.50	0.49	0.49
OPLS-aa	9.3	0.51	0.50	0.50	0.49	0.62	0.48

**Table 9.8. Interchain salt bridges in S-shaped fibril ( Average interpeptide Asp23 – Lys28 NH3+ \_ COO– distances (nm  $\pm$  Std. Dev.)**

Forcefields	Chain AB	Chain BC	Chain CD	Chain DE	Chain EF	Chain FG	Chain GH	Chain HI	Chain IJ	Chain JK	ChainKL
AmberGS	1.09 $\pm$ 0.16	0.83 $\pm$ 0.19	0.85 $\pm$ 0.12	0.83 $\pm$ 0.08	0.83 $\pm$ 0.08	0.84 $\pm$ 0.06	0.89 $\pm$ 0.09	0.90 $\pm$ 0.07	0.93 $\pm$ 0.12	1.14 $\pm$ 0.18	2.91 $\pm$ 0.54
Gromos53a6	1.51 $\pm$ 0.19	1.34 $\pm$ 0.16	1.33 $\pm$ 0.14	1.34 $\pm$ 0.10	1.34 $\pm$ 0.98	1.35 $\pm$ 0.07	1.35 $\pm$ 0.07	1.36 $\pm$ 0.06	1.36 $\pm$ 0.08	1.42 $\pm$ 0.11	1.45 $\pm$ 0.19
OPLS-aa	1.01 $\pm$ 0.17	0.84 $\pm$ 0.10	0.78 $\pm$ 0.06	0.78 $\pm$ 0.06	0.79 $\pm$ 0.07	0.77 $\pm$ 0.09	0.75 $\pm$ 0.05	0.75 $\pm$ 0.06	0.76 $\pm$ 0.06	0.78 $\pm$ 0.10	3.24 $\pm$ 0.41

## 9.4 Ring fibril



**Figure 9.7. The three segments of this 3-fold symmetry fibril originally at 60° to each other but undergo a change especially for AmberGS and Gromos53a6 which give rise to distorted geometry. Out of three structures, (c) shows the least deviation from the original structure that is produced in OPLS-aa**

**Table 9.9. Average intrasheet distances ring fibril (nm  $\pm$  Std. Dev.)<sup>2</sup>**

Forcefields	PHE19/GLY38 (nm)	ALA21/VAL36 (nm)	ASP23/LEU34(nm)
<b>AmberGS</b>			
Chain A	1.80 $\pm$ 0.37	0.76 $\pm$ 0.13	0.85 $\pm$ 0.28
Chain B	1.93 $\pm$ 0.52	1.70 $\pm$ 0.48	1.25 $\pm$ 0.35
Chain C	1.98 $\pm$ 0.27	1.57 $\pm$ 0.14	1.43 $\pm$ 0.21
Chain D	3.78 $\pm$ 0.59	1.75 $\pm$ 0.26	1.54 $\pm$ 0.20
Chain E	2.24 $\pm$ 0.24	1.48 $\pm$ 0.19	1.42 $\pm$ 0.09
Chain F	3.44 $\pm$ 0.65	3.43 $\pm$ 0.77	1.57 $\pm$ 0.27
Forcefields	PHE19/GLY38 (nm)	ALA21/VAL36 (nm)	ASP23/LEU34 (nm)
<b>Gromos53a6</b>			
Chain A	1.64 $\pm$ 0.23	1.56 $\pm$ 0.14	1.51 $\pm$ 0.18
Chain B	1.92 $\pm$ 0.19	1.95 $\pm$ 0.16	2.02 $\pm$ 0.19
Chain C	2.00 $\pm$ 0.11	1.98 $\pm$ 0.08	1.98 $\pm$ 0.16
Chain D	2.22 $\pm$ 0.11	2.12 $\pm$ 0.11	2.07 $\pm$ 0.18
Chain E	2.46 $\pm$ 0.14	2.35 $\pm$ 0.12	2.30 $\pm$ 0.12
Chain F	2.51 $\pm$ 0.21	2.55 $\pm$ 0.18	2.26 $\pm$ 0.20
Forcefields	PHE19/GLY38 (nm)	ALA21/VAL36 (nm)	ASP23/LEU34 (nm)
<b>OPLS-aa</b>			
Chain A	1.47 $\pm$ 0.16	1.20 $\pm$ 0.09	1.25 $\pm$ 0.10
Chain B	1.69 $\pm$ 0.14	1.77 $\pm$ 0.15	1.85 $\pm$ 0.09
Chain C	2.09 $\pm$ 0.08	1.79 $\pm$ 0.09	1.95 $\pm$ 0.09
Chain D	2.11 $\pm$ 0.11	2.05 $\pm$ 0.08	1.92 $\pm$ 0.14
Chain E	2.21 $\pm$ 0.12	2.21 $\pm$ 0.10	2.05 $\pm$ 0.12
Chain F	2.47 $\pm$ 0.23	2.65 $\pm$ 0.15	2.24 $\pm$ 0.23

<sup>2</sup> Average intra-sheet distances between PHE19-GLY38, ALA21-VAL36 and ASP23-LEU34 shows that in all the three force fields the inter-sheet distances are well-maintained in the upper half of the fibril but as we move down the inter-sheet distance increases leading to an opened structure at the end.

## 9.5 References

- Crescenzi, O. et al. (2002) 'Solution structure of the Alzheimer amyloid beta-peptide (1-42) in an apolar microenvironment. Similarity with a virus fusion domain.', *European journal of biochemistry / FEBS*, 269(22), pp. 5642–8.
- Heinig, M. and Frishman, D. (2004) 'STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins.', *Nucleic acids research*, 32(Web Server issue), pp. W500-2. doi: 10.1093/nar/gkh429.
- Lührs, T. et al. (2005) '3D structure of Alzheimer's amyloid-beta(1-42) fibrils.', *Proceedings of the National Academy of Sciences of the United States of America*, 102(48), pp. 17342–7. doi: 10.1073/pnas.0506723102.
- Xiao, Y. et al. (2015) 'A $\beta$ (1–42) fibril structure illuminates self-recognition and replication of amyloid in Alzheimer's disease', *Nature Structural & Molecular Biology*, 22(6), pp. 499–505. doi: 10.1038/nsmb.2991

## Appendix B

### 9.6 Methodology

The details of simulations containing various concentrations  $A\beta_{42}$  and curcumin carried out during this work:

The interaction of curcumin with  $A\beta_{42}$  monomer

Three simulations were carried out: (i)  $A\beta_{42}$  monomer alone in aqueous solution; (ii) an  $A\beta_{42}$  monomer and a single curcumin molecule in aqueous solution and (iii) An  $A\beta_{42}$  monomer and 4 curcumin molecules in aqueous solution.

Effect of curcumin on self-assembly of  $A\beta_{42}$  monomers

The simulations investigated the effect of increasing concentration of curcumin on the self-assembly of  $A\beta_{42}$  monomers. A total of six simulations were carried out each containing 24 monomers of  $A\beta_{42}$ , with a varying number of curcumin molecules namely 2, 5, 19, 77, and 308, with one of the simulations being a control without curcumin. These molar ratios correspond to 5 mM of  $A\beta_{42}$  monomers, and 0.25, 1, 4, 16, and 64 mM of curcumin relative to water. Note that the simulated concentrations are in the milli-molar range and hence are markedly higher than experimental concentrations that are in the micromolar range to enhance the driving force for phase separation so as to make the system evolve quicker.

Interaction of curcumin with  $A\beta_{42}$  fibril

The fibril comprised a 25-mer unit of  $A\beta$  monomers of residues (17-42). Two simulations in a set of three were carried out with curcumin concentrations of 5 and 30 molecules.

The summary of simulations carried out and the models used are given in the Table 9.10

**Table 9.10. Summary of A $\beta$ <sub>42</sub> models and simulations systems**

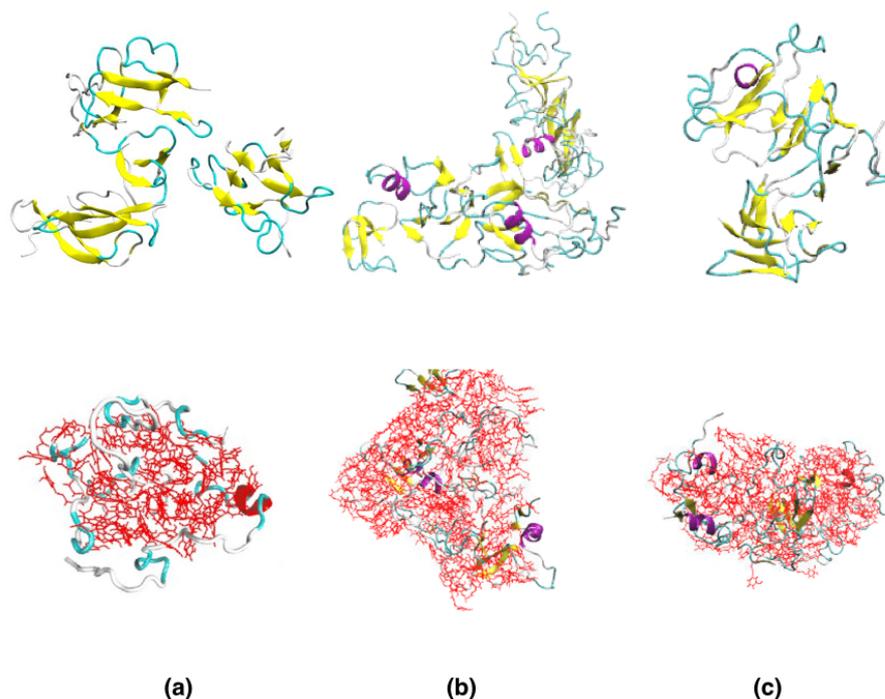
Model	System A $\beta$ <sub>42</sub> :Curcumin	Simulation time
A $\beta$ <sub>42</sub> monomer	1:0	200ns
	1:1	
	1:4	
A $\beta$ <sub>42</sub> monomer (Self- assembly)	24:0	100ns
	24:2	
	24:5	
	24:19	
	24:77	
	24:308	
A $\beta$ <sub>42</sub> fibril	1:4	100ns
	1:30	

## 9.7 Technical details

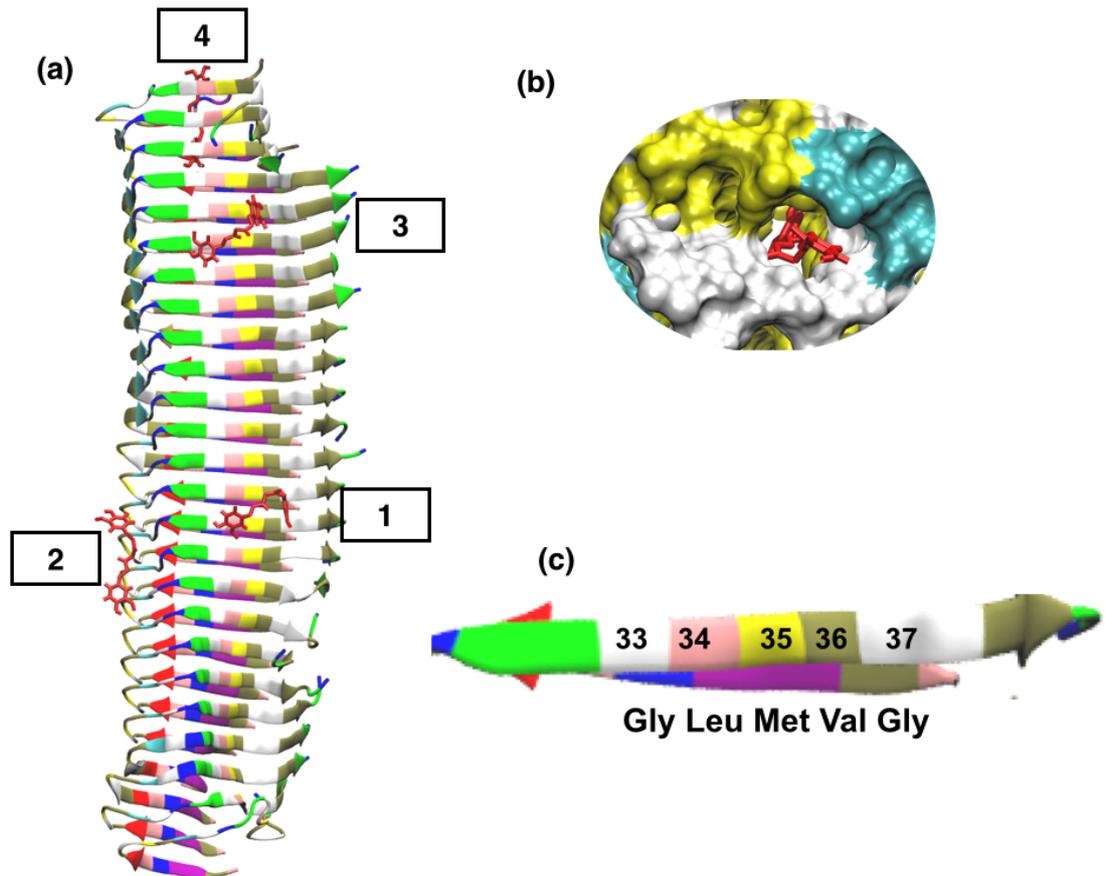
The binding of curcumin with A $\beta$  was explored using explicit-solvent, atomistic simulations on the nanosecond time scale. A $\beta$  monomer (PDB id 1IYT), (Orlando Crescenzi, Tomaselli, Guerrini, Salvadori, Anna M. D'Ursi, *et al.*, 2002) protofibril (PDB id 2BEG) (Lührs *et al.*, 2005a) and fibril (generated using CreateFibril v 2.5) (Smaoui *et al.*, 2013) were chosen to serve as models. The fibril comprised a 25-mer unit of A $\beta$  monomers of residues (17-42). The optimized structure and charges of curcumin diketone were taken from the work done by Son Tung Ngo *et al.* in 2012 (Ngo and Li, 2012). The simulations were carried out using Gromacs 5.1 package with parameters from the Gromos96 53A6 force field (Van Der Spoel, Lindahl, Hess, Groenhof, Alan E. Mark, *et al.*, 2005) (Schuler, Daura and Van Gunsteren, 2001) coupled with the SPCE water model. Long range electrostatic interactions were calculated using Particle-Mesh Ewald (PME) (Darden, York and Pedersen, 1993). The van der Waals interaction cutoff was 1.4nm, as was the cutoff for the real space Ewald

interaction. All of the systems were subjected to energy minimization to remove any bad contacts using 5000 steps of the steepest descent algorithm and then equilibrated for 500ps using the NVT ensemble followed by NPT with the peptides positions restrained. The simulations were carried out at 360K and 0.001 kbar using the Nose-Hoover thermostat and the Parrinello-Rahman barostat (isotropic mode) (Martyna, Klein and Tuckerman, 1992),(M. Parrinello and Rahman, 1981). The higher temperature was used to accelerate the system evolution given that standard MD simulations can only access limited timescale. For trajectory analysis we used the analysis utilities in the Gromacs package, and VMD (Visual Molecular Dynamics) for visualisation (Biarnés *et al.*, 2012). The binding energy of curcumin with the amyloid structures was calculated using the MM-PBSA(Kumari, Kumar and Lynn, 2014) method implemented in Gromacs 5.1.

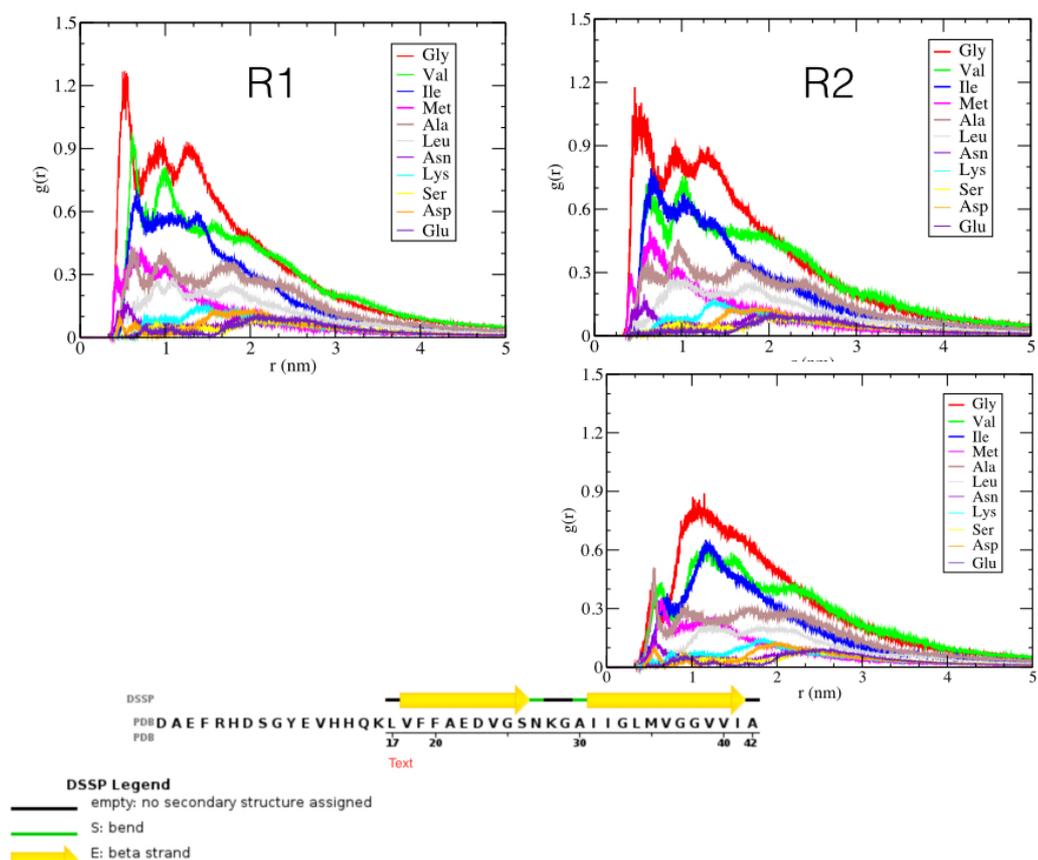
## 9.8 Results



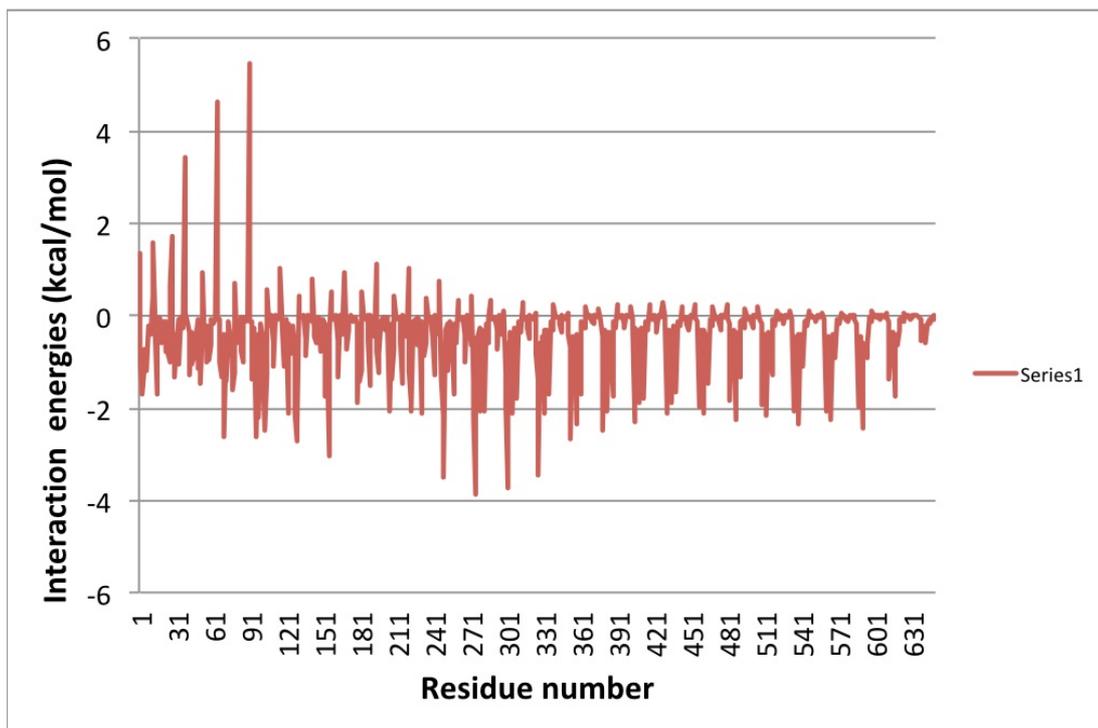
**Figure 9.8. Upper panel: A typical oligomer formed at 0mM curcumin (control); Lower panel, 64mM curcumin system. At higher concentration curcumin destroys  $\beta$ -sheets and stabilizes off-pathway oligomers; Snapshot from a 100ns long simulation shows integrated curcumin- $A\beta$  clusters via non-specific interactions.  $A\beta$  are represented by secondary structures while curcumin molecules are represented by red lines. (Water not shown for simplicity). (a), (b) & (c) represent the three replicate studies.**



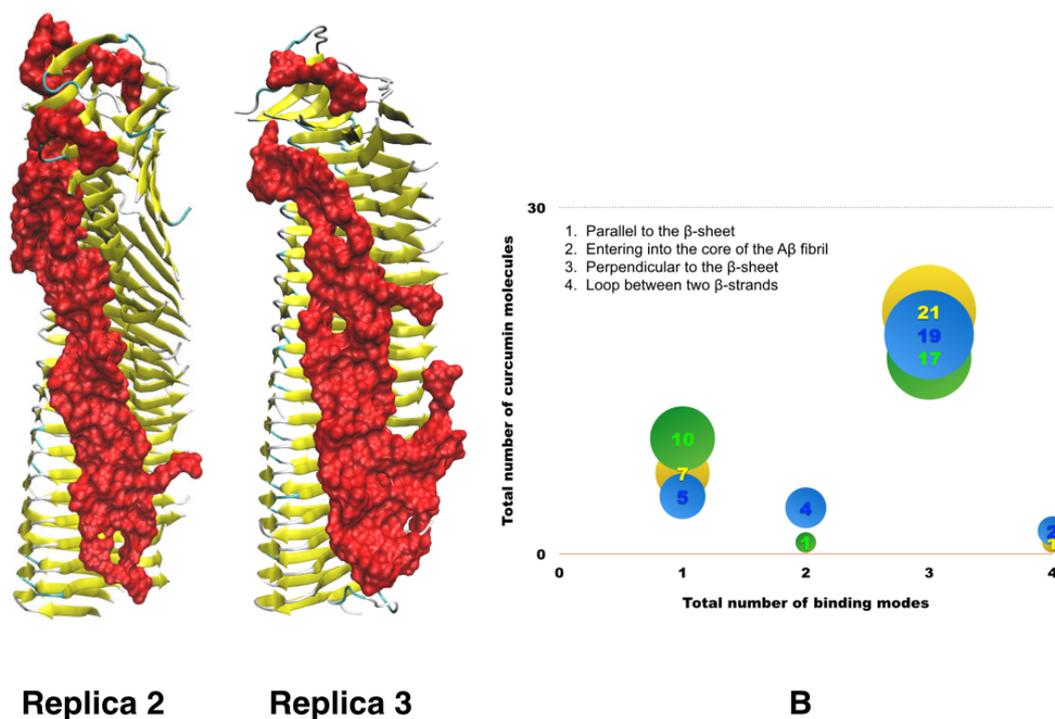
**Figure 9.9. Interaction of curcumin with  $A\beta$  fibril at lower concentration of curcumin.**  
 (a) Figure shows the binding of curcumin with  $A\beta$  fibril in three different modes. (1 &3) Curcumin parallel to the beta strands ( $BE = -30$  &  $-25$  kcal mol $^{-1}$ ). (2) Curcumin interacting with the loop region between the  $\beta$ -strands parallel to fibril axis ( $BE = 26$  kcal mol $^{-1}$ ). (4) Curcumin going into the core of the fibril ( $BE = -27$  kcal mol $^{-1}$ ); Curcumin is represented as line structure in red (b) Top view of curcumin entering the central cavity.  
 (c) The G33XXXG37 motif showing the main residues involved in binding with curcumin.



**Figure 9.10. Rdf plots generated for curcumin molecules to track the dynamics of curcumin– A $\beta$  fibril interaction. R1, R2 and L show the two aromatic and a linker region respectively. At the bottom of the figure, the sequence of A $\beta$  fibril is given (from protein data bank) (Lührs et al., 2005b). The sharp peaks for residues GLY, VAL, ILE and MET show the specificity of curcumin for these residues.**



**Figure 9.11. Per-residue analysis of interaction energies of all the residues of A $\beta$  fibril with curcumins calculated by MM-PBSA method. The periodicity of peaks represents the specific residues that are involved in the binding of curcumin to A $\beta$  including methionine, glycine, valine and isoleucine. Negative values indicate preferable interaction between a residue and curcumin molecule. The highest negative peaks representing glycines imply that Curcumin prefers to interact with GLY with both of its aromatic rings with a very strong binding.**



**Figure 9.12. Binding of curcumin to Aβ fibril in two replica studies showing highly-specific binding of curcumin to Aβ fibril, (b) A plot of the population of curcumin molecules oriented and located at various positions on the Aβ fibril in all the three replica studies showing that mode III i-e curcumin molecules lying perpendicular to the β-sheets is the most dominant mode of binding.**

## 9.9 References

Biarnés, X. *et al.* (2012) ‘METAGUI. A VMD interface for analyzing metadynamics and molecular dynamics simulations’, *Computer Physics Communications*. Elsevier B.V., 183(1), pp. 203–211. doi: 10.1016/j.cpc.2011.08.020.

Crescenzi, O. *et al.* (2002) ‘Solution structure of the Alzheimer amyloid β-peptide (1-42) in an apolar microenvironment: Similarity with a virus fusion domain’, *European Journal of Biochemistry*. doi: 10.1046/j.1432-1033.2002.03271.x.

Darden, T., York, D. and Pedersen, L. (1993) ‘Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems’, *The Journal of Chemical Physics*. doi: 10.1063/1.464397.

Kumari, R., Kumar, R. and Lynn, A. (2014) ‘G-mmpbsa -A GROMACS tool for high-

throughput MM-PBSA calculations’, *Journal of Chemical Information and Modeling*, 54(7), pp. 1951–1962. doi: 10.1021/ci500020m.

Lührs, T. *et al.* (2005) ‘3D structure of Alzheimer’s amyloid-beta(1-42) fibrils.’, *Proceedings of the National Academy of Sciences of the United States of America*, 102(48), pp. 17342–7. doi: 10.1073/pnas.0506723102.

Martyna, G. J., Klein, M. L. and Tuckerman, M. (1992) ‘Nosé-Hoover chains: The canonical ensemble via continuous dynamics’, *The Journal of Chemical Physics*. doi: 10.1063/1.463940.

Ngo, S. T. and Li, M. S. (2012) ‘Curcumin binds to A $\beta$ 1-40 peptides and fibrils stronger than ibuprofen and naproxen’, *Journal of Physical Chemistry B*. doi: 10.1021/jp302506a.

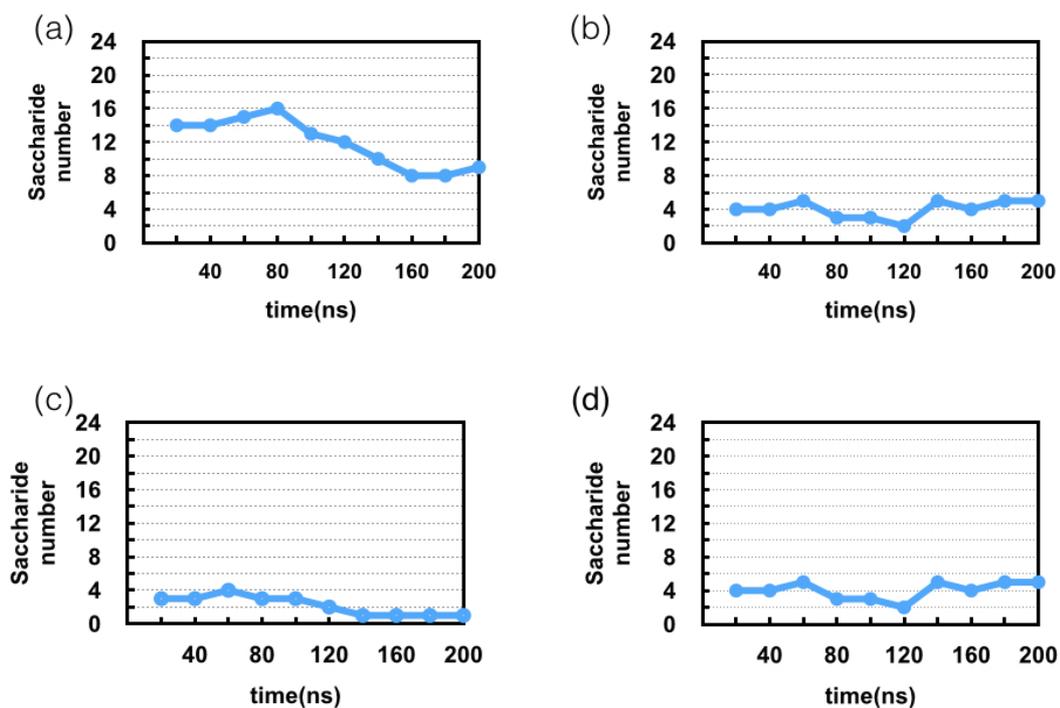
Parrinello, M. and Rahman, A. (1981) ‘Polymorphic transitions in single crystals: A new molecular dynamics method’, *Journal of Applied Physics*. doi: 10.1063/1.328693.

Schuler, L. D., Daura, X. and Van Gunsteren, W. F. (2001) ‘An improved GROMOS96 force field for aliphatic hydrocarbons in the condensed phase’, *Journal of Computational Chemistry*, 22(11), pp. 1205–1218. doi: 10.1002/jcc.1078.

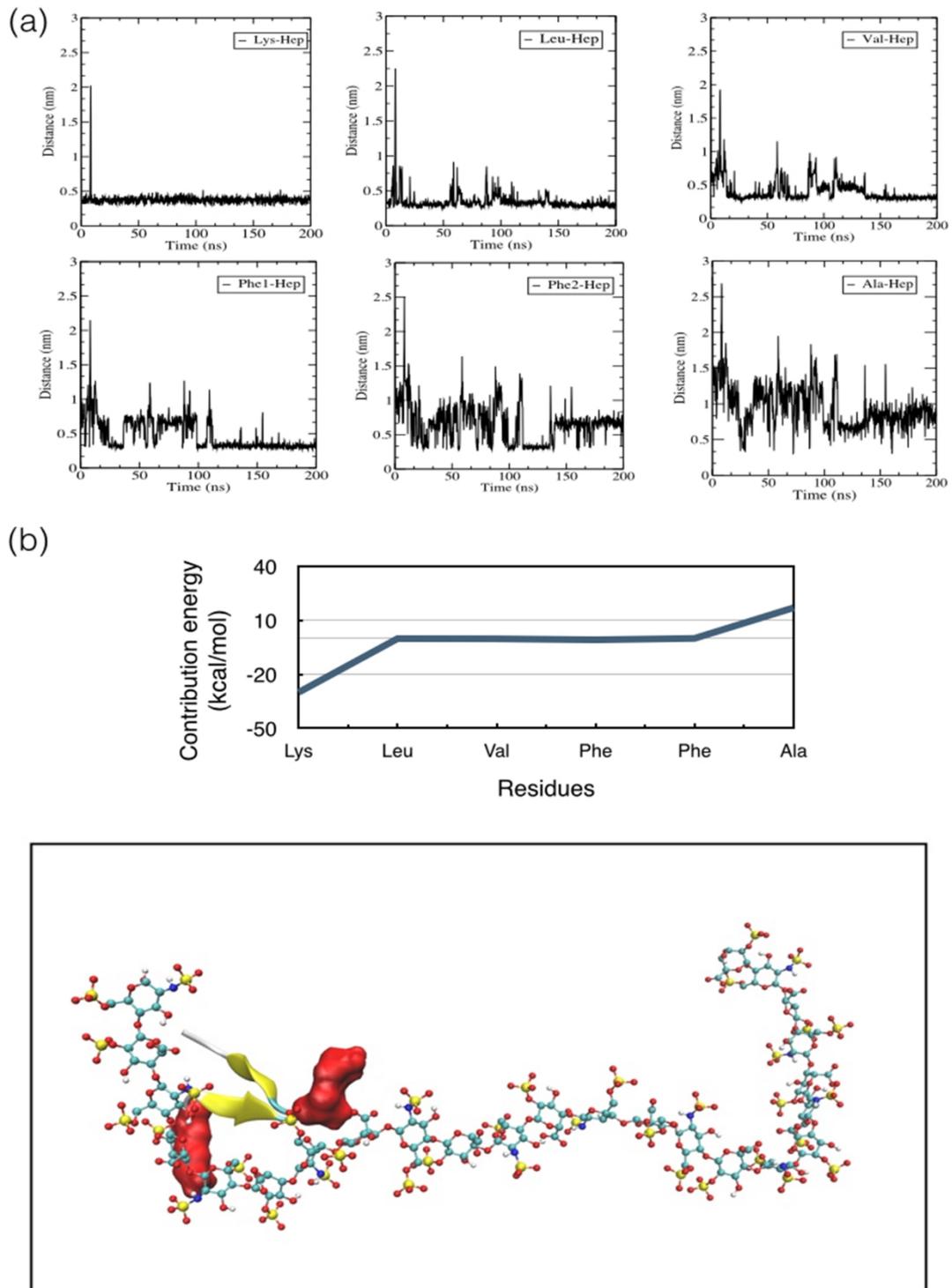
Smaoui, M. R. *et al.* (2013) ‘Computational assembly of polymorphic amyloid fibrils reveals stable aggregates’, *Biophysical Journal*, 104(3), pp. 683–693. doi: 10.1016/j.bpj.2012.12.037.

Van Der Spoel, D. *et al.* (2005) ‘GROMACS: Fast, flexible, and free’, *Journal of Computational Chemistry*, pp. 1701–1718. doi: 10.1002/jcc.20291.

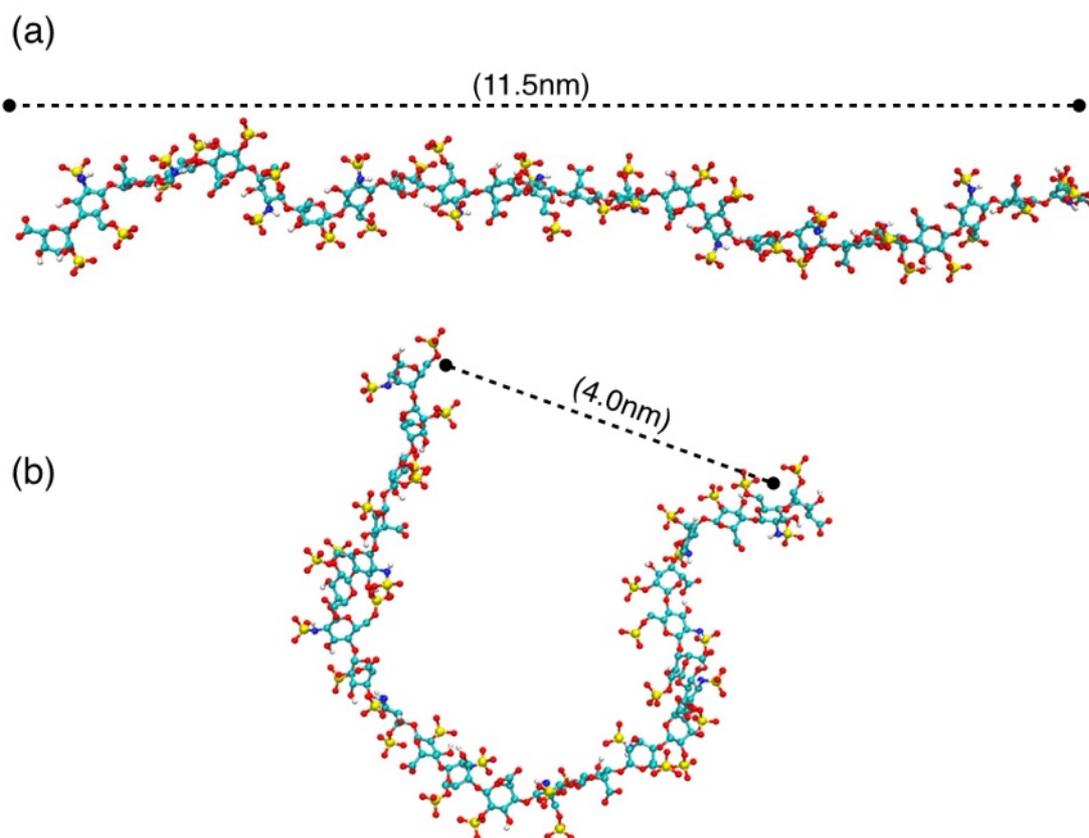
## Appendix C



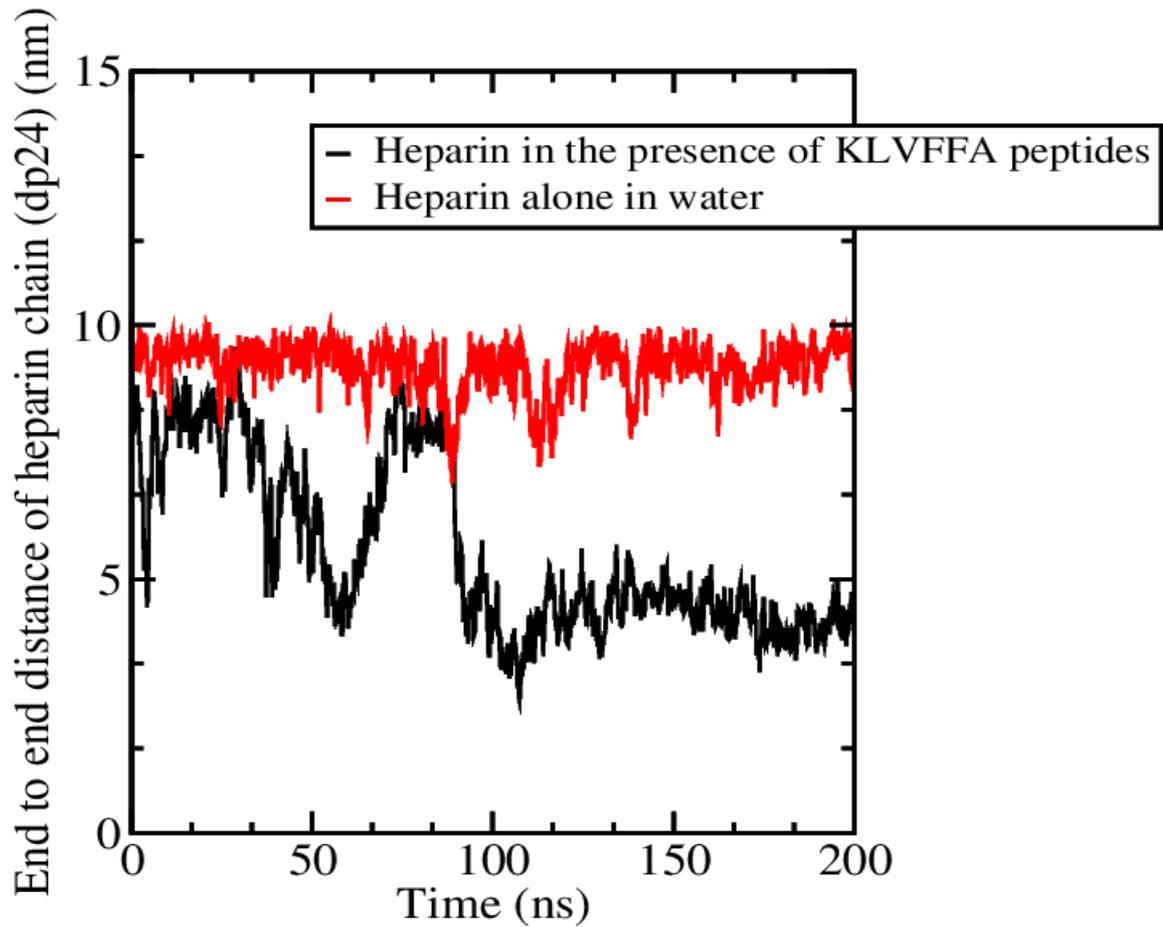
**Figure 9.12.** Comparison of peptide (monomer) position on heparin's saccharide units at various time intervals. (a) Fully flexible heparin (dp24) at low ionic strength. (b). Position restrained heparin (dp24) at low ionic strength. (c) Fully flexible heparin (dp24) at high ionic strength. (d) Position restrained heparin (dp24) at high ionic strength. It is clear from the plots that the movement of peptides is more restricted for the position-restrained heparin.



**Figure 9.13. (a) Minimum distance between various residues of KLVFFA and negatively charged sulpho groups of heparin (cut-off = 0.6nm). The results indicate that the lysine residues of the KLVFFA peptides are involved in binding. Once the contacts are formed, they remain conserved throughout the simulation time(200ns). (b) The major contribution to the binding includes electrostatic interactions between cationic sites (lysine) on the peptides and anionic sites on heparin. (c) Peptide klvffa (dimer) binding to heparin, the interacting residues (lysines in both the strands) are shown in surface presentation.**



**Figure 9.14.** The effect of different heparin chain lengths on heparin-facilitated assembly of KLVFFA peptides at 150mM. As the chain length increases, aggregation is also enhanced. Shorter heparin chain lengths (dp=4 and dp=6).are less effective as they form multiple heparin-peptide complexes that aggregate only slowly. (a) dp2, (b) dp4, (c) dp6, (d) dp8 and (e) dp24. Heparin and peptides are shown in CPK and cartoon (secondary structure) representation respectively



**Figure 9.15. (a) The extended structure of a heparin molecule in neutral aqueous solution. The end-to-end distance of heparin is 11.5nm. (b) Snapshot of the heparin molecule in the presence of peptides (pre-dimerized) (not shown) where the end-to-end distance reduces to around 4nm. (c) Comparison of end to end distance of heparin in water and in the presence of peptides showing that it reduces significantly when heparin interacts with the peptides (dimers).**