ABSTRACT

Title of Dissertation: EXPLORING THE EFFECTS OF PHYSIOLOGICAL ENVIRONMENT ON AMYLOID AGGREGATION

Abhilash Sahoo
Doctor of Philosophy, 2022

Dissertation Directed by: Dr. Silvina Matysiak
Department of Bioengineering

Molecular level self-assembly/aggregation processes are common in biomolecular systems. Specifically, aggregation of protein molecules results in formation of amyloid deposits, that has been associated with neuronal dysfunction leading up to neurodegeneration. The protein aggregation is often influenced by several external physiological features, which can modulate this pathological process in a specific or non-specific manner. This thesis aims to elucidate the role of such factors in amyloid aggregation in the context of neurodegeneration. As test cases, we have focused on different fragments of Amyloid-β peptide and Huntingtin protein and explored common interaction schemes in the presence of phospholipid membranes, solvated glucose molecules and added trailing sequences.

Phospholipid membranes, composed of a heterogeneous distribution of lipid molecules, serve as packaging envelopes in cellular systems. But several studies have suggested a role of cellular membranes in abetting protein aggregation in neurodegenerative diseases. The first

section of this thesis explores Aβ 16-22 aggregation in the presence of membranes. Lipid membranes have been shown to modulate peptide aggregation in a charge dependent manner with anionic membranes promoting faster peptide aggregation into ordered fibrillar structures compared to zwitterionic membranes. In this work, we evaluate the role of this electrostatic membrane headgroup charge on Aβ 16-22 peptide aggregation with model lipid membranes composed of POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) and POPS (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoserine) lipids. Beyond, membrane charge, membrane's physical organization can also affect peptide-peptide and peptide-membrane interactions. Here, we have curated the effects of applied surface-tension, as a proxy for membrane curvature, on peptide fibrillation propensities.

Apart from ordered structures such as membranes, solvated small molecules are a large class of molecules that can affect aggregation patterning by affecting peptides through both specific and non-specific interactions. The second section of this thesis explores Aβ 16-22 aggregation in varying hyperglycemic conditions, to draw correlations between Alzheimer's disease and type 2 diabetes. Here, we discovered that the glucose prefers partitioning onto the aggregate-water interface in a specific manner, leading to a loss in rotational entropy that propels peptide aggregation.

In the final section, we discuss the case of pathological peptide aggregation in the case of Huntington's disease. Broadly, Huntinting protein's N-terminal region which consists of 17-residue N-terminal domain (N17) and the following Glutamine repeat tract (Poly-Q) are our objects of interest and associated with pathological aggregation. The aggregation landscape of N17 is analyzed in presence of added different lengths of trailing Poly-Q tract and the presence of curved membranes.

We have approached our research through a computational lens using molecular dynamics simulations. To address the relevant concerns of large spatio-temporal scales necessary to study peptide aggregation systems with molecular simulations, we have developed a coarse-grained forcefield (ProMPT: Protein Model with Polarizability and Transferability) that uses reduced spatial resolution to accelerate phase-space exploration. The forcefield can capture secondary and tertiary folding of protein structures with minimal constraints, and is transferable across biomolecular systems without a need for re-parametrization.

My dissertation presents a holistic picture of peptide aggregation and various physiological factors that affect it, with biomolecular simulation across multiple scales.

EXPLORING THE EFFECTS OF PHYSIOLOGICAL ENVIRONMENT ON AMYLOID AGGREGATION

by

Abhilash Sahoo

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Dr. Silvina Matysiak, Chair/Advisor
Dr. Pratyush Tiwary
Dr. Jinwoo Lee
Dr. Yanxin Liu
Dr. Jeffery B. Klauda, Dean's representative

Acknowledgments

First, I would like to thank Dr. Silvina Matysiak for her supervision, patience and support throughout my doctoral studies. She has helped me understand how to navigate academic and scientific research, and I consider myself incredibly lucky to be a part of her lab for five years.

I am grateful to my committee members Dr. Jeffery Klauda, Dr. Pratyush Tiwary, Dr. Yanxin Liu and Dr. Jinwoo Lee for their time and constructive feedback, that was really influential in shaping my dissertation.

I would like to thank my research collaborators Pei-yin Lee, Neha Nanajkar and Dr. Hongcheng Xu, for all your support and sticking with me over these years as a graduate student. Also, I am grateful to other members of Dr. Matysiak's lab — Riya Samanta, Gregory Custer, Suhas Gotla, Meenal Jain and Neel Sanghvi for all the discussions and overall making my graduate school enjoyable.

My wife, Dr. Alisha Pradhan has been my psychological and intellectual support system during these six years of graduate school. I am grateful to have her as my partner and completing this Ph.D. journey together. Thank you for always motivating me, your unwavering belief, being my sounding board and my best friend. I owe my deepest thanks to my parents and brother for their unending love and support through graduate school as an international student. Thank you for being there for me. Lastly, this journey would not have been possible without the wonderful friends I made over these years — Spandan, Pramod, Biswak, Vishnu, Akshita, Vikram, Utkarsh and Teja, who made my time in college park memorable.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Abeta	Amyloid-Beta
AFM	Atomic force microscopy
AWSEM	Associative Memory, Water Mediated, Structure and Energy Model
CD	Circular Dichorism
CG	Coarse-Grained
DOPC	1,2-Dioleoyl-sn-glycero-3-phosphocholine
DOPG	1,2-dioleoyl-sn-glycero-3-phospho-(1'-rac-glycerol)
DPC	Dodecyl Phosphocholine
EM	Electron Microscopy
FG	Fine Grained
GUV	Giant Unilamellar Vesicles
htt	Huntingtin
LUV	Large Unilamellar Vesicle
MD	Molecular Dynamics
N17	Nterminal-17
NMR	Nuclear Magnetic Resonance
OPEP	Optimized Potential for Efficient Protein structure prediction
Poly-P	Poly-Proline
Poly-Q	Poly-Glutamine
POPC	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine
POPS	1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoserine
ProMPT	Protein Model with Polarizability and Transferability
SUV	Small Unilamellar Vesicle
TEM	Transmission Electron Microscopy
TIRF	Total Internal Reflection Fluorescence
UV	Ultra Violet
WEPCGM	Water Explicit Polarizable Coarse Grained Model
WEPMEM	Water Explicit Polarizable Membrane Model
WEPPRO	Water Explicit Polarizable Protein Model

Chapter 1: Introduction

1.1 Objective of Thesis

Peptide misfolding and accumulation of aberrant proteinaceous amyloid-like aggregates is a recurrent theme in numerous diseases associated with neuronal dysfunction such as Alzheimer's disease (AD), Huntington's disease (HD) and Parkinson's disease (PD) [1]. In particular, with progressive increase in the aging population, neurodegenerative diseases present a significant current social and economic challenge.

In this thesis, I have focused on fragments of Amyloid-beta (Aβ) peptide associated with Alzheimer's disease and Huntingtin protein associated with Huntington's disease as test cases to understand interaction schemes that drive this self-association behavior. A detailed understanding of the molecular mechanisms and pathological event pathways of peptide's conformational changes and self-association can aid towards development of rational therapeutics. But such aggregation processes do not occur in isolation and are often affected by the presence of several physiological factors such as biomembranes. Previous research has implicated the importance of environment in protein folding and aggregation pathways [2–5]. Moreover, common pathways for amyloid aggregation related cytotoxicity and inhibitory mechanisms involve biomolecular structures, such as lipid membranes and the extracellular matrix [6–9]; and mutations/modifications. This thesis explores the impact of membranes, varying concentration of glucose molecules and addition of

trailing lengths of glutamines on aggregation behavior of fragments of Aβ and Huntingtin protein.

While, wet-lab experiments are essential for our understanding of biological systems, they can suffer from some specific disadvantages. Experimental investigations into structural features of amyloid - environment interaction is limited due to extensive structural heterogeneity and complex competing interactions. Moreover, the disordered nature of these peptides is highly susceptible to local environmental alterations and generate experimental artifacts which often remain unaccounted for, leading to controversial results. This has also led to diverging interpretations to same/similar experimental results.

Computer simulations and molecular modeling, aided by statistical thermodynamics can provide insights on how complex bio-molecular systems behave beyond what theory and experiments can deliver separately. Molecular dynamics (MD) simulations involve analyzing the spatial and temporal motion of atoms derived from a system's model Hamiltonian (forcefield), to generate both atomic-scale structural and kinetic insights. It can be particularly ideal for characterizing transient peptide structures which are difficult to study experimentally and has been widely used to evaluate heterogeneous ensemble of structures generated by amyloid peptides, with appropriate controls for environmental factors [10–13]. While traditional atomistic simulations provide higher resolution and more detailed insights about the peptide-based biomolecular systems, the spatio-temporal scales to study conformational transitions and peptide/aggregate - environment interactions cannot be reliably achieved by present-day computational machineries. This necessitates the need for specialized tools to access such time scales and length scales. Towards this direction, for advanced sampling protocols — both biased and unbiased simulations have been developed. Biased simulation methods such as metadynamics and umbrella sampling, improve exploration of phase space along particular reaction coordinates, whereas unbiased methods such as replica

exchange (REMD, H-REMD) seek for a more general enhancement of sampling. But these methods require significant computational resources and can be particularly unfeasible for large spatiotemporal processes such as peptide aggregation [14–17]. On the other hand, due to fewer number of particles – resulting in lowered resolution and a smoother free-energy landscape, coarse grained (CG) MD can provide a more holistic picture for such multi-agent phenomena in larger biomolecular systems [18].

In my research, I have developed coarse-grained models and used molecular dynamics with these CG models to curate how the presence of physiological features — membranes, local enhancement of cosolutes (glucose molecules) and added trailing polar sequences can affect the structural properties and kinetics of peptide aggregation. This thesis address three primary objectives.

- Explore the impact of membrane charge and membrane curvature on aggregation of Amyloid-Beta (Aβ) peptide's central hydrophobic core (Aβ 16-22).
- Understand the effects of hyperglycemic conditions on aqueous aggregation of Aβ CHC.
- Study the impact of membrane curvature and the presence of polyglutamine (Qn) repeats on aggregation behavior of Huntingtin protein's N-terminal domain (N17).

1.1.1 Effects of Membrane charge and structure on aggregation of Amyloid-Beta (Aβ) peptide's central hydrophobic core (Aβ 16-22)

In the first section of the thesis, I focused on how membranes can affect aggregation patterning and kinetics of Amyloid-beta aggregation. Amyloid plaques and neurofibrillary tangles,

contributing to progressive cognitive decline have been established as hallmarks for Alzheimer's disease [1]. The amyloid cascade hypothesis, has been widely accepted by neuropathologists as the primary model of AD pathogenesis. According to this hypothesis, oligomerization of Aβ peptides initiates a cascade of events, culminating in neuronal dysfunction and dementia. Pathogenic Aβ peptides are about 39–43 residue long intrinsically disordered peptide in aqueous solution and ordered alpha helix rich structures in apolar environments, formed from successive incisions by β-secretase and γ-secretase, which can aggregate into β-sheet rich aggregates.

Mammalian cell membranes are dynamic and complex microsystems of significant physiological relevance. The lipid membranes are self-assembled biological structures composed of lipid molecules — amphiphilic molecule with a hydrophilic head and a hydrophobic alkyl tail. These molecules are highly diverse in chemical specificity due to variations in the head and tail types and therefore impart varying physico-chemical properties to membranes. Phospholipids are the primary component of cellular membranes, marked by the presence of a phosphate group at the "head". The net charge on a lipid molecule has been shown to modulate membrane properties and function. Therefore, although zwitterionic lipid molecules particularly 1-palmitoyl-2-oleoyl-sn-glycero- 3-phosphocholine (POPC) forms a significant fraction of lipid molecules in mammalian cells [19], anionic lipids such as 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-l-serine (POPS) [20, 21], which predominantly exist in the inner leaflet, plays a more significant role in cellular signaling pathways and interactions with peptides. The production of Aβ peptides occurs in a membranous environment, exposing the peptides to a number of lipid–peptide interactions. Also, the perturbation of cellular membranes, followed by ion-dysregulation due to oligomeric forms of Aβ peptides is hypothesized to be a central part of Aβ assisted AD pathology [22–24]. Therefore, a mechanistic understanding of bio-mechanical interactions of Aβ peptides

with cellular membranes is necessary to gain insights into the amyloid cascade pathway. Aβ peptides have been shown to exhibit varying aggregation patterns on lipid bilayers depending on their structure and composition. CD and Thioflavin T assay studies of unilamellar vesicles have revealed an accelerated aggregation of Aβ 16–28 peptides into ordered beta sheets on an anionic bilayer (DPPG) as compared to a zwitterionic bilayer (DPPC) [25]. In addition, an AFM experiment using supported bilayers has shown that the disruption of zwitterionic membranes (DOPC) is higher than that of anionic membranes (DOPG) in presence of Aβ peptides [26]. Recent evidence from imaging studies using TEM, AFM and total internal reflection fluorescence microscopy has suggested that small unilamellar vesicles (SUVs) with a larger curvature promotes amyloid fibril formation when compared to large unilamellar vesicles [27, 28]. Therefore, the interaction of Aβ peptides with cellular membrane can result in complex changes in energetics and kinetics of structural transitions and can also result in significant membrane disruptions. The interactions are highly heterogeneous with significant dependencies on membrane composition, oligomer structure, peptide/lipid ratio and cellular environment.

Experimental characterization of peptide aggregates by X-ray diffraction has revealed common structural features such as a cross beta sheet architecture [29, 30]. Structural studies of Aβ peptides by solid state NMR, hydrogen–deuterium exchange and electron microscopy have also shown the presence of similar cross beta sheet patterns [31]. The central hydrophobic core (CHC), residues 17–21 (L–V–F–F–A) of the complete Aβ, is crucial for fibrillation [32–36]. In addition, solid state NMR studies have confirmed that Aβ 16–22 (K–L–V–F–F–A–E) is one of the the smallest peptide sequences capable of forming highly ordered, stable beta sheet rich fibrils at neutral pH [37]. Therefore studies of the structural and kinetic properties of a simpler tailorable model peptide, Aβ 16–22, can provide a better understanding of the molecular forces

responsible for fibril formation/elongation [38–40].

A morphological characterization of Aβ oligomers is difficult due to its transient and soluble nature. On the other hand, computational studies, particularly molecular dynamics (MD) can be an ideal alternative to access this small time scale, transient behaviour. Atomistic simulations often coupled with advanced sampling techniques have been extensively implemented to study small-scale peptide oligomerization in solution. Some recent atomistic studies on peptide–lipid interactions have investigated pre-formed membrane-inserted oligomers and the very initial phases of oligomer–lipid interactions [11, 13, 41]. Due to high computational costs and sampling issues, atomistic simulations have not been used to study peptide aggregation on lipid bilayers starting from a solvated monomeric configuration. Coarse grained molecular dynamics (CG-MD), which provides a reduced resolution description of a system and significantly improved sampling of a protein conformational landscape is an effective tool to study complex systems with extended spatio-temporal scales, specifically peptide aggregation starting from monomeric peptides.

With our in-house developed coarse grained model (WEPPRO/WEPCGM) we were able to study the imapct of membrane charge through model membranes created with zwitterionic (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine: POPC) and anionic (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine: POPS) membranes. We also curated the effects of membrane-curvature (expressed through changes in applied surface-tension) on Aβ fragment's aggregation behavior.

1.1.2 Effects of hyperglycemic conditions on aqueous aggregation of Aβ 16-22

In the second section, I have focused on more disordered and non-specific interactions associated with the presence of co-solvated sugar molecules. Several clinical and molecular studies have outline pathological correlations between type II diabetes (T2D) and Alzheimer's disease [42–45]. Insulin resistance and improper sugar metabolism are the established outcomes of T2D, which can result in increasing the blood sugar content (hyperglycemia). This has engendered studies to understand this effect of hyperglycemia on Alzheimer's disease. Recent structural evidences have established that chemical crosslinks generated by glycation of lysines, thereby preventing the peptide aggregates from dissolution as the primary correlation between increased severity of Alzheimer's in T2D patients [46–48]. But this idea was recently challenged with experiments reporting a time-scale separation between increased aggregation in presence of sugar molecules and formation of chemical crosslinks, which happen at a later point [49]. This points to a possible alternate *thermodynamic* mechanism. In this section, we aim to investigate this pathological correlation from a thermodynamic perspective with coarse-grained molecular simulations.

1.1.3 Impact of membrane curvature and the presence of polyglutamine (QN) repeats on aggregation behavior of Huntingtin protein's N-terminal domain (N17)

Another neurodegenerative disease that involves progressive amyloid deposition and associated membrane disruption is Huntington's disease [50–52]. The pathogenesis of dominantly inherited

Huntigton's disease is linked to fibrillar nano-scale deposits of Huntingtin protein (htt). A mutant htt gene (with multiple CAG repeats) encodes variants of htt protein with an anomalous expanded homopolymeric Poly-Q sequence that aids the aggregation process. While the flanking amino acid residues, particularly, the first 17 N-terminal amino acid residues (N17) modulate aggregation behavior and lipid binding by formation of amphipathic alpha helix, the length of Poly-Q tract directly participates in aggregation and generation of a variety of aggregate species — oligomers and larger fibrillar structures. Several reports indicate htt protein interacts with membrane, either through intracellular vesicular transport, or by association with Endoplasmic reticulum and Golgi apparatus [53]. In addition, the pathogenesis of Huntington's disease is hypothesized to proceed through mitochondrial dysfunction. But the membrane interactions of htt and polyQ deposits have not been fully characterized.

Huntington's disease is among a larger section of diseases marked by the presence of expanded poly-glutamine tract [53–55]. While, some global common themes have been suggested for structural details of N17, the stuctures at the poly-Q tract and the impact of poly-Q tract on N17-poly-Q aggregation remains ill defined. As such, in this section, we focus on the impact of poly-Q on the conformational landscape and aggregation of Huntingtin protein's N-terminal domain.

Moreover, similar to Aβ, both membranes and oligomers are affected by htt-membrane association. Solution AFM studies have reported oligomeric and fibrillar deposits over mica surface [56, 57]. Studies have demonstrated an alteration in htt alpha helical content in presence of POPC and POPC/POPS SUVs [58]. AFM studies with Total Brain Lipid Extract (TBLE) suggests local alterations of bilayer compressibility on interaction with htt oligomers [59]. Investigations by Chaibva *et al.* suggested an enhanced peptide binding and aggregation propensities with

increase in membrane curvature [60].

The intrinsically disordered nature of N17, similar to Aβ peptides, in solution makes experimental characterization difficult. The transience of peptide conformations has necessitated the need for computational investigations. While scarce in number, atomistic molecular simulations keep reiterating the importance of N17 in driving membrane interactions. Due to the associated structural transitions and complexity at capturing accurate environmental effects (such as local dielectrics), coarse-grained simulations of N17-membrane interaction has not been attempted yet. A coarse grained forcefield capable of encoding environmental features and generating structural transitions can be an essential tool to study N17's behaviour in membranous environments.

The relation between neurodegeneration and membrane curvature has been established for several peptide sequences — Tau, Aβ, α-synuclein and huntingtin [61, 62]. One recurring theme among such peptides is the presence of a membrane binding amphipathic helical (AH) motif. The AH motifs can bind to membranes and significantly retard peptide diffusion, which could in turn promote peptide aggregation [63]. Due to shortness of its sequence, the AH motif of the Huntingtin protein (Huntington's disease) can be reliably used as a model to investigate this behavior, and provide a general outline how neurodegenerative AHs can sense membrane curvature. In addition, the amphipathic nature of this model N17 can be treated as a window into the biophysical properties of similar amphipathic helical regions, often prevalent in neurodegenerative peptides such as tau and α synuclein [61].

1.2 Outline of Thesis

In chapter 2, I introduce the coarse grained model for proteins (Water Explicit Polarizable PROtein Model: WEPPRO) and membrane (Water Explicit Polarizable MEMbrane Model: WEPMEM) that has been developing in our lab over years. Here, I present an overview of the forcefield and discuss its benefits and potential applications.

In chapter 3, I explored differential patterns of membrane-induced Aβ 16–22 (K–L–V–F–F–A–E) aggregation from the microscopic perspective of molecular interactions. In agreement with experimental observations, anionic POPS molecules promote extended configurations in Aβ peptides that contribute towards faster emergence of ordered β-sheet-rich peptide assemblies compared to POPC, suggesting faster fibrillation. In addition, lower cumulative rates of peptide aggregation in POPS due to higher peptide–lipid interactions and slower lipid diffusion result in multiple distinct ordered peptide aggregates that can serve as nucleation seeds for subsequent Aβ aggregation.

In chapter 4, I introduce the effects of POPC membrane curvature (using applied surface-tension as proxy) on Aβ 16-22 aggregation. This work presents a mechanistic and morphological overview of the relationships between the biomembrane local structure and organization, and Aβ peptide aggregation.

In chapter 5, I adopt coarse-grained molecular dynamics simulations to investigate protein aggregation behavior with Aβ 16-22 peptides in near-pathological hyperglycemia. Here, we uncovered entropic effects that direct glucose concentration-induced increased aggregation of peptides.

In chapter 6, I discuss the development of a new transferable CG forcefield — ProMPT

with an explicit representation of the environment for accurate simulations with proteins. The forcefield consists of a set of pseudo-atoms representing different chemical groups that can be joined/associated together to create different biomolecular systems. This preserves the transferability of the forcefield to multiple environments and simulation conditions. We have added electronic polarization that can respond to environmental heterogeneity/fluctuations and couple it to protein's structural transitions. The non-bonded interactions are parametrized with physics-based features such as solvation, and partitioning free energies determined by thermodynamic calculations and matched with experiments and/or atomistic simulations. The bonded potentials are inferred from corresponding distributions in non-redundant protein structure databases. We present validations of the CG model with simulations of well-studied aqueous protein systems with specific protein fold types- TRP-cage, TrpZip4, Villin, WW-domain and β-α-β. We also explore the applications of the forcefield to study aqueous aggregation of Aβ 16-22 and dimerization of Glycophorin A in presence of Dodecylphosphocholine (DPC) micelles.

In chapter 7, we examine the impact of pathological glutamine repeats on peptide structure and aggregation, with the coarse-grained model developed in chapter 6 — ProMPT. Here, we curated a list of structural features for N17 with different lengths of poly-glutamine repeats, both as a single peptide, and the peptide in its aggregate form. Increase in the length of trailing polyglutamine tract resulted in a reverse-micelle architecture with the aggregate core dominated by fibrillar poly-glutamine tract, and a disruption of the traditional hydrophobic core formed by bulky hydrophobic groups.

Moreover, we also studied the structural features of the POPC membrane and N17 peptide, that drives the curvature sensing of the Huntingtin protein. Our simulations delineate gradual progression of N17 from a dominantly unstructured peptide in solution, to a more structured

α-helical patch on the membrane. In presence of a curved membrane, we observed that N17 peptides preferentially interact with the curved region. Here, we noted that while the polar and charged groups drive the initial membrane-peptide interaction, the membrane curvature sensing is dominantly controlled by the bulky hydrophobic groups.

Chapter 2: Water-Explicit Polarizable Coarse Grained Model - WEPCGM

2.1 Overview

The application of classical molecular dynamics (MD) simulations at atomic resolution (fine-grained level - FG), to most biomolecular processes, remains limited because of the associated computational complexity of representing all the atoms. To address this, coarse-grained models have been developed. In the past decade, the development of various coarse-grained models has provided key insights into the driving forces in folding and aggregation. But, models in literature are either implicit-solvent, non-transferable or cannot study structural transitions. Here, we discuss the coarse-grained model that was developed by our lab to address these shortcomings.

2.2 Introduction

While, wet-lab experiments are essential for our understanding of biological systems, they suffer from some disadvantages. Specifically for complex biological systems, wet-lab experiments are complicated and isolating individual effects/interactions can be very difficult. This can be crucial for biological systems with many inter-dependent entities. Moreover, experimental results can have diverging interpretations and are susceptible to environmental fluctuations. As an alternative, computer simulations can be used to explain and direct experimental research. It

can be used to characterize molecular biophysics with theoretical precision, while capturing the impacts of all the interactions in the biological system. Classical molecular dynamics involves direct numerical integration of Newton's equations, aided by mathematical functions and parameters — forcefield, that describe the dependence of system's potential energy on individual atomic positions to generate time evolution of molecules. The efficiency and accuracy of bio-molecular simulations is dependent on mathematical forcefields and packaged molecular dynamics simulation programs (GROMACS [**?**], CHARMM [64], NAMD [65] and DLPoly [66], etc.). Some of the popular chemically specific peptide–lipid forcefield families are Assisted Model Building and Energy Refinement [67, 68] (AMBER), Chemistry at HARvard Molecular Mechanics [69–71] (CHARMM), GROningen MOlecular Simulation [72, 73] (GROMOS), Optimized Potential for Liquid Simulations [74, 75] (OPLS). Improvements in parallel computing architecture and use of graphical processing units have enabled millisecond level atomistic simulations to study protein folding and unfolding in an unbiased manner.

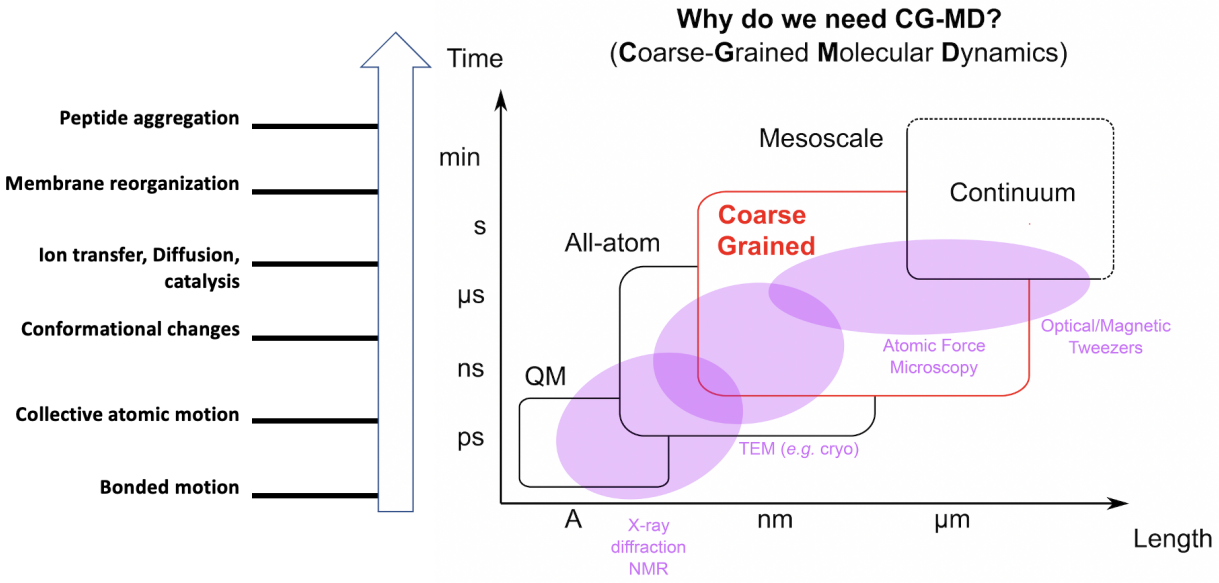

Figure 2.1: Spatio-temporal scales accessible to different biomolecular methods

Under rather typical experimental conditions, the time scale of biological events can be orders of magnitude larger than the time scale achieved by classical molecular dynamics simulations(Fig. 2.1). To address this time-sale issue, many novel techniques have been developed. One such technique — spatial coarse-graining involves local averaging of atomic positions, which can smooth out the free energy landscape, and provide computational speed-up. On the basis of particle-based resolution, molecular dynamics simulations can be broadly classified into all-atom/atomistic (AA), united-atom (UA) and coarse grained (CG). Coarse grained molecular dynamics (CG-MD) simulations can be applied to access relevant length and timescales to study biophysical features such as membrane remodeling, peptide folding and aggregation and membrane-assisted peptide aggregate's structural changes. Some novel and popular coarse grained forcefields include PRIME-20 [76], AWESEM [77], OPEP [78] and MARTINI [79]. In current literature, the transferable coarse grained forcefields are often either limited by the variations of biomolecule species they can represent or by the ability to capture secondary/tertiary peptide structures from primary sequence.

In this chapter, I elaborate on the coarse grained models for proteins (Water-Explicit Polarizable Protein Model - WEPPROM) and lipid membrane (Water-Explicit Polarizable Membrane Model - WEPMEM) that I have used in the upcoming chapters [80–82]. This family of coarse grained forcefields (Water-Explicit Polarizable Coarse Grained Models - WEPCGM) address the current shortcomings in the coarse-grained forcefields present in literature. These models have an explicit representation of environment that allows studies with environment-effected biomolecular motion. With the protein model, we can study structural transitions into secondary structures. Finally, these coarse-grained models are transferable and do not require re-parametrizations in order to study different biomolecular systems.

2.3 Water-Explicit Polarizable Protein Model - WEPPROM

Our protein model was first introduced by Ganesan *et al.* and has been going through constant updates. The model can capture structural transitions from a primary sequence of amino-acids without any added bias. The CG protein can consist of three types of beads - charged (+/-), hydrophobic (H) and polar (P) mapped in an atomistic amino acid sequence (Fig. 2.2). Each amino acid's peptide backbone has been mapped into a single polarizable backbone bead (BB) with structural polarization added through dummy positive/negative charges (BBp/BBm) of equal absolute magnitude. These dummy charges are tethered to the central interaction site (BB) through harmonic restraints. The dummies can interact with the environment through electrostatic forces and that can result in an induced-dipole effect, adding directionality to peptide backbone–backbone interactions, and generating secondary and super-secondary structures. The side-chains are amino-acid length, hydrophobicity and charge specific. The current version of our coarse-grained model does not account for chirality of protein backbones. The model has been parametrized with Yesylevskyy *et al.* 's [83] polarizable MARTINI water model. Please refer to Ganesan *et al.* for more details on model parametrization and finer details of the model [80, 81]. A description of the non-bonded interactions between coarse-grained interaction sites is provided in Table 2.2

In comparision to the previous implementation, the backbone dummy (D) angle (D-BB-D) was modified from 0 degree to 180 degrees to mimic the structure of typical peptide bonds. The modification increases the dipole moment of polarizable peptide backbone beads and stabilize the formation of secondary structure as in agreement with quantum mechanical calculations . The original BB-BB-S1 (Backbone-Backbone-first sidechain) angles were removed from the model

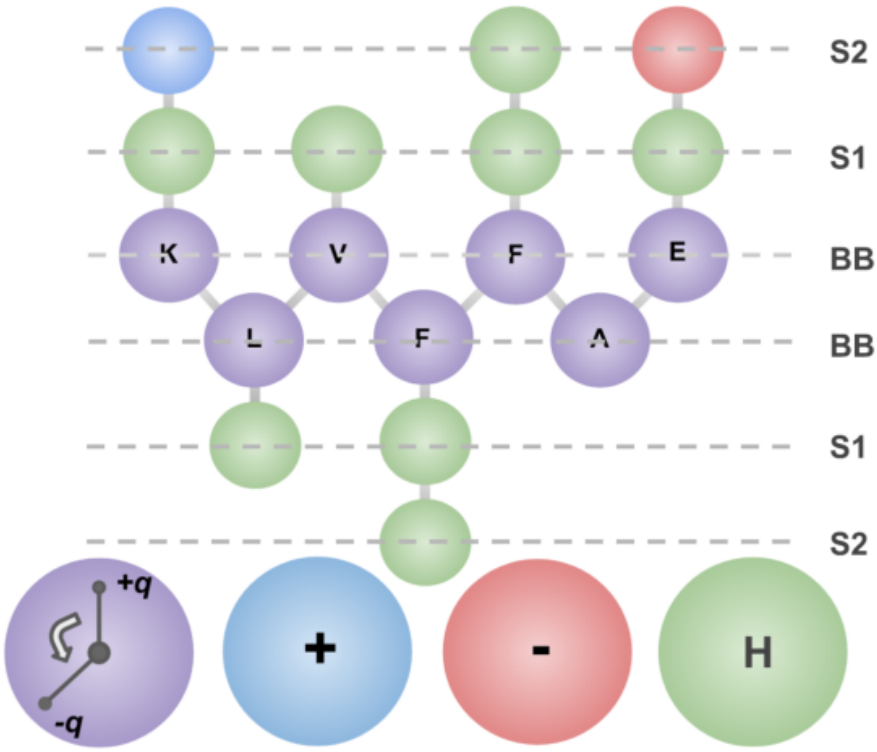

Figure 2.2: A schematic description of peptide coarse-grained model for K-L-V-F-F-A-E.

and replaced by non-bonded interactions between backbone beads to sidechain beads on adjacent amino acids. Similar to the original implementation of this model, no external dihedral potential has been used to maintain the secondary structure of the peptide.

2.4 Water-Explicit Polarizable Membrane Model - WEPMEM

Our lipid forcefield complements the protein model and uses same/similar interaction-center types. Similar to the original MARTINI forcefield, lipids - POPC and POPS in our lipid model — adapted from WEPMEM [82, 84], are modelled by 13 CG beads each through a 4:1 mapping scheme (Fig. 2.3). The non-bonded interactions between interaction sites is presented in Table 2.2. At the lipid-headgroup, phosphate (PO4) and choline (NC3) are mapped to a charged

Table 2.1: Bond and angle parameters in WEPPRO model. BB: Backbone bead, D: dummy particle, S1: Sidechain 1 bead, S2: Sidechain 2 bead.

bonds	R_{bond} (nm)	K_{bond} (kJ mol^{-1} nm^{-2})
BB-BB	0.385	7500
BB-D	0.14	5000
BB-S1	0.25	5000
S1-S2	0.28	5000
angles	$\theta_0(deg)$	K$_{angle}$ (kJ mol^{-1})
D-BB-D	180	7.2
BB-BB-BB	109	75
BB-S1-S2 (LYS, PHE)	151	25
BB-S1-S2 (GLU)	180	25

Table 2.2: Non-bonded Lennard-Jones (LJ) interaction strengths in WEPCGM model. Unit of interaction strength (ϵ) is in kJ/mol. The radius (σ) of all LJ interactions is 4.7Å.

Beads	BB (P5)	H1 (C1)	H2 (C3)	C$^+$ (Qd)	C$^-$ (Qa)	Water (POL)
BB (P5)	5.0	2.0	2.7	5.32	5.32	4.75
H1 (C1)	2.0	3.5	3.5	2.3	2.3	1.0
H2 (C3)	2.7	3.5	3.5	2.7	2.7	2.7
C$^+$ (Qd)	5.32	2.3	2.7	3.5	4.0	5.0
C$^-$ (Qa)	5.32	2.3	2.7	4.0	3.5	5.0
Water (POL)	4.75	1.0	2.7	5.0	5.0	4.0

coarse grained bead, whereas glycerol-esters (GL1 and GL2) and serine (CNO) are represented by polarizable beads with structural polarization generated through two dummy charges similar to the peptide model. Lipid oleoyl tails are designed with five hydrophobic beads, whereas the palmitoyl tails with four.

In comparision to previous implementation of WEPMEM, the LJ interactions between peptide backbone bead type P5 (BB) and charged beads (Qd/Qa) were changed from 4 kJ/mol to 5.32 kJ/mol to avoid overbinding. Also, if the interaction between hydrophobic beads and water beads are too low, the hydrophobic beads tend to aggregate in the aqueous environment. To maintain the intricate balance between apolar-water and apolar-apolar interactions, the interactions between hydrophobic sidechain bead type C1 and water bead type POL were modified from 0.2 kJ/mol to 1.0 kJ/mol. Please refer to Ganesan *et. al.* for a more detailed validation of the

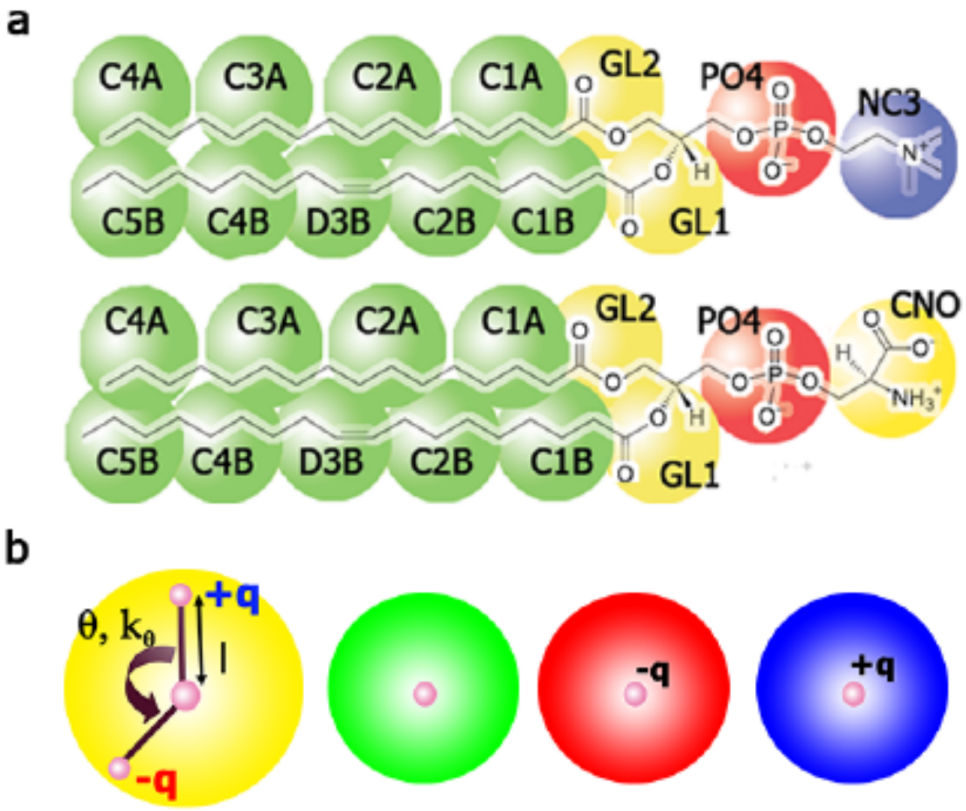

Figure 2.3: A schematic description of lipid coarse-grained model

forcefield and other model descriptions [81, 84]. As of now, we have representations for POPC and POPS only.

2.5 Validations

Our membrane model was validated through several comparisions with atomistic simulations and experimental reports [81, 82]. With simulations starting from a random distribution of the lipid molecules in a simulation box, the lipid molecules conformed into well ordered lipid bilayers, exhibiting reasonable descriptions of lipid behavior. The model WEPMEM membrane matched the structural, dynamic and electrostatic properties, such as the cross-bilayer density

profile, area per lipid, bilayer thickness, mean squared displacement, line tension, dipole potential, dielectric profile, head group orientation and formation of lipid clusters, compared to experiments and/or all-atom simulations. Our model could also capture the right trend in membrane-interfacial electric potential compared to the MARTINI model.

Similarly, our protein model could achieve realistic α/β content with patterns created from hydrophobic-hydrophilic residues [80]. This helped in charecterization of dipole–dipole and dipole–charge interactions in shaping the secondary and supersecondary structure of proteins. Formation of helix bundles and β-strands. Moreover, the model could also achieve folding into ordered β-sheets for elastin like model octa-peptides in presence of a hydrophilic-hydrophobic interface [85]. In another publication from the group, folding of anti-cancer peptide SVS1 was studied in presence of model POPC and POPS membrane. Our model could reproduce the expected folding of SVS1 in presence of anionic membrane into a β hairpin structure [81].

2.6 Limitations

In the current form, WEPCGM suffers from several limitations. While the model has been shown to reproduce common secondary structures of smaller peptides, it has not been validated with larger protein structures with complex tertiary packing. In addition, we don't have representation for *all* amino acids with this model. We will address these limitations in the new version of our CG model that we introduce in Chapter 6 — Protein Model with Polarizability and Transferability (ProMPT). Similarly, the repository of lipid molecules is also limited with current representation for POPC and POPS only. Ongoing work in our lab aims to expand this set to include POPG, POPE and sterols.

2.7 Conclusion

In this chapter, we introduce the coarse-grained models developed by our lab that I have contributed to and applied in my research. The coarse-grained model features explicit representation of the environment, that allows for studying environment assisted structural changes. The presence of explicit solvent also aids in creating a transferable forcefield without a need for reparametrization in order to study another biomolecular system. In the upcoming chapters we will use WEPPROM in conjuction with WEPMEM to study the aggregation of a fragment Aβ peptide in presence of the membrane and hyperglycemia.

Chapter 3: Effects of membrane headgroup charge on Amyloid-β 16-22 aggregation

3.1 Overview

This chapter is based on the author's publication: Pathways of amyloid-beta absorption and aggregation in a membranous environment. Abhilash Sahoo, Hongcheng Xu and Silvina Matysiak. Physical Chemistry Chemical Physics, 2019.

Aggregation of misfolded oligomeric amyloid-beta (Aβ) peptides on lipid membranes has been identified as a primary event in Alzheimer's pathogenesis. However, the structural and dynamical features of this membrane assisted Aβ aggregation have not been well characterized. In this chapter, we explore differential patterns of membrane headgroup -induced Aβ 16–22 (K–L–V–F–F–A–E) aggregation from the microscopic perspective of molecular interactions. Physics-based coarse-grained molecular dynamics simulations were employed to investigate the effect of lipid headgroup charge – zwitterionic (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine: POPC) and anionic (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine: POPS) – on Aβ 16–22 peptide aggregation. Our analyses present an extensive overview of multiple pathways for peptide absorption and biomechanical forces governing peptide folding and aggregation. In agreement with experimental observations, anionic POPS molecules promote extended configurations in Aβ peptides that contribute towards faster emergence of ordered β-sheet-rich peptide assemblies compared to POPC, suggesting faster fibrillation. In addition, lower cumulative rates of peptide

aggregation in POPS due to higher peptide–lipid interactions and slower lipid diffusion result in multiple distinct ordered peptide aggregates that can serve as nucleation seeds for subsequent Aβ aggregation. This study provides an *in-silico* assessment of experimentally observed aggregation patterns, presents new morphological insights and highlights the importance of lipid headgroup chemistry in modulating the peptide absorption and aggregation process.

3.2 Introduction

Aberrant aggregation of peptides and proteins on cellular membranes has been associated with the pathogenesis of a number of neurodegenerative diseases such as Alzheimer's (AD), Parkinson's (PD) and Huntington's (HD) disease [86–88]. Alzheimer's disease, characterized by extra-cellular amyloid plaques [89–91] and intra-cellular neurofibrillary tangles [92, 93], is a significant social challenge which has affected 5.7 million people in the United States [94]. The amyloid cascade hypothesis for AD presents aggregation of a 39-43 residue long intrinsically disordered peptide-amyloid beta (Aβ) as the trigger for a cascade of events culminating in neuronal deaths and dementia [95, 96]. Multiple alloforms of Aβ peptides with an intrinsic tendency to form fibrillar aggregates are created by successive excisions of Amyloid precursor proteins (APP) by beta secratase in the endosomal pathway and gamma secretase in the plasma membrane [97–99]. Recent evidences have implicated soluble, low molecular weight Aβ oligomers as the primary cytotoxic agents, correlating strongly with cognitive defects [4, 5, 100, 101]. The structural diversity of polymorphic Aβ oligomers contributes towards multiple pathways for Aβ -induced neuronal toxicity [102].

A broad range of proteins and peptides, regardless of large variations in the amino acid

sequence, have been shown to form amyloid fibril at high concentrations [103–105]. Experimental characterization of peptide aggregates through X-ray diffraction has revealed common structural features such as the cross beta sheet architecture [30, 106]. Structural studies of Aβ peptides through solid state NMR [31], hydrogen-deuterium exchange [107] and electron microscopy [108–110] have also shown the presence of similar cross beta sheet patterns. The central hydrophobic core (CHC), residues 17-21 (L-V-F-F-A) of the complete Aβ is crucial for fibrillation [32–36]. In addition, solid state NMR studies have confirmed that Aβ 16-22 (K-L-V-F-F-A-E) is one of the the smallest peptide sequence capable of forming highly ordered, stable beta sheet rich fibrils at neutral pH [37]. Therefore studies of the structural and kinetic properties of a simpler tailorable model peptide, Aβ 16-22, can provide a better understanding of the molecular forces responsible for fibril formation/elongation [38–40].

The production of Aβ peptides occurs in a membranous environment, exposing the peptides to a number of lipid-peptide interactions [22]. Also, the perturbation of cellular membranes, followed by ion-dysregulation due to oligomeric forms of Aβ peptides is hypothesized to be a central part of Aβ assisted AD pathology [22–24]. Therefore, a mechanistic understanding of bio-mechanical interactions of Aβ peptides with cellular membranes is necessary to gain insights on the Aβ cascade pathway. Aβ peptides have been shown to exhibit varying aggregation patterns on lipid bilayers depending on their structure and composition [25,26,111–116]. CD and Thioflavin T assay studies on unilamellar vesicles have revealed an accelerated aggregation of Aβ 16-28 peptides into ordered beta sheets on anionic bilayer (DPPG) as compared to zwitterionic bilayer (DPPC) [25]. In addition, an AFM experiment using supported bilayers, has shown that disruption of zwitterionic membranes (DOPC) is higher than anionic membranes (DOPG) in presence of Aβ peptides [26]. Recent evidences from imaging studies using TEM, AFM and total

internal reflection fluorescence microscopy have suggested that small unilamellar vesicle (SUVs) with a larger curvature promotes amyloid fibril formation when compared to large unilamellar vesicles [27, 28].

A morphological characterization of Aβ oligomers is difficult due to its transient and soluble nature [117–119]. On the other hand, computational studies, particularly molecular dynamics (MD) can be an ideal alternative to access this small time scale, transient behaviour [120]. Atomistic simulations, often coupled with advanced sampling techniques have been extensively implemented to study small-scale peptide oligomerization in solution [121–128]. Some recent atomistic studies on peptide-lipid interactions have investigated pre-formed membrane-inserted oligomers and the very initial phases of oligomer-lipid interaction [11, 13, 41]. Due to high computational costs and sampling issues, atomistic simulations have not been used to study peptide aggregation on lipid bilayers starting from a solvated monomeric configuration. Coarse grained molecular dynamics (CG-MD), which provides a reduced resolution description of a system and significantly improved sampling of protein conformational landscape, is an effective tool to study complex systems with extended spatio-temporal scales [129], specifically peptide aggregation starting from monomeric peptides. Many novel coarse grained force-fields (In-lattice and Off-lattice) have successfully characterized ordered amyloid aggregation [76, 130–134]. PRIME 20, an intermediate resolution unbiased peptide coarse graining scheme has been implemented with discontinuous molecular dynamics (DMD) on many amyloidogenic sequences including Aβ peptides to study fibril formation in solution [135]. Zheng, et. al. explored the aggregation free energy landscape of Aβ peptides using AWSEM-MD — a predictive coarse grained force-field [136]. A minimalistic, phenomenological model (Clafisch model) was implemented with a variable dihedral term to generate peptide aggregation on model vesicles [131]. Recently,

another phenomenological coarse grained three bead-per-residue, amyloidogenic peptide model (Shea model) was employed with implicit solvent to demonstrate spontaneous peptide aggregation into beta sheets on model lipid membranes [132]. These coarse graining techniques, either have been designed only for peptides (PRIME20-Hall model) or do not provide peptide sequence/lipid type specificity (Clafisch and Shea models).

Here, we present a CG-MD study of partitioning, folding and aggregation dynamics of Aβ 16-22 peptides (K-L-V-F-F-A-E), in presence of model lipid bilayers composed with zwitterionic-POPC (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine) and anionic-POPS (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine) starting from their solvated monomeric state. We designed the peptide using a modified version of prior-developed coarse grained model - Water-Explicit Polarizable PROtein Model (WEPPROM) [80,81], which generated secondary structures of small peptides from primary amino acid sequences without any built-in bias. Lipids were modeled by another slightly altered variant of recently created Water-Explicit Polarizable MEMbrane (WEPMEM) model that could accurately reproduce dielectric properties at the lipid-headgroup (interface) region [81,84]. Physically, peptide aggregation involves an interplay between hydrophobic and hydrophilic effects, indicating the importance of precise modeling of electrostatic interactions [80, 85, 137]. The novelty of these models is the introduction of structural polarization at the peptide-backbone and lipid-headgroup which is necessary for studying peptide-lipid interaction and membrane induced peptide folding [81]. Both these models have been parameterized to use Yesylevskyy's polarizable water model [83].

Outer neuronal cell membranes are extremely diverse with a wide variety of glycerophospholipids and ceramides [138]. Aβ peptides show an enhanced propensity to aggregate into ordered beta sheets on negatively charged membranes composed with anionic lipids [25, 137, 139–142]. Also,

many experimental studies to understand the effect of lipid type on peptide aggregation have used POPS as one of the anionic components [143–145]. This prompts our choice for using POPS as the model anionic lipids to study the oligomerization of Aβ peptides in membraneous environments. In this chapter we explore the differences in the morphology and kinetics of membrane-assisted beta sheet formation by Aβ 16-22 on model lipid membranes using coarse grained MD simulations. The chapter also presents a mechanistic explanation to experimentally observed kinetic behaviour. To our knowledge, this is the first coarse grained simulation study that captures the complex process of peptide aggregation in membraneous environment, from peptides solvated in their monomeric state to formation of ordered secondary structures, while maintaining peptide sequence and lipid type specificity.

3.3 Methods

3.3.1 Peptide Model

The Aβ peptide model is primarily derived from WEPPROM — described in details in Chapter 2 and previous publications [80–82]. The forcefield has been successfully applied to study interfacial folding and aggregation of small peptides without any externally added bias [80, 81]. The mapping scheme for the peptide Aβ 16-22 is provided as Fig. 2.2.

The peptide model was validated using experimental and computational evidences of Aβ 16-22 aggregation into ordered cross beta structures in aqueous solution [30,38,106,146]. Simulations with 12 Aβ monomers, solvated in water (0.14 M) aggregated into stable beta sheet rich oligomers (Fig. 3.1).

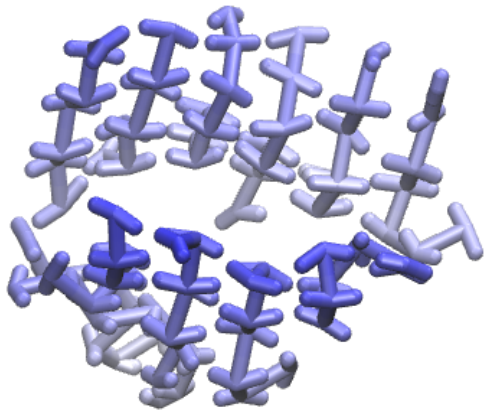

Figure 3.1: Cross-beta structure of Aβ 16-22 in solution.

3.3.2 Lipid Model

The lipid molecules in this work were modeled from WEPMEM forcefield, described in detail in Chapter 1 and previous publications [81, 82, 84]. The forcefield could capture accurate structural, dynamic and electrostatic features of the lipid membrane. Please refer to Fig. 2.3 for details of the mapping scheme.

3.3.3 Simulation Protocol

The model POPC bilayer, composed with 240 lipids and POPS bilayer with 242 lipds were solvated with CG water (lipid to water ratio fixed approximately at 20). The simulations were performed on GROMACS 4.5.4 [147]. After the preliminary energy minimization, the bilayers and counter-ions (in case of POPS bilayer) were equilibrated for 10 ns (time step, dt = 10 fs) using NPT ensemble. Temperature was maintained at 300K through a Noose-Hoover thermostat [148, 149] with a time constant of 1 ps. Parinello-Rahman barostat [150] with a time

constant of 1 ps and compressibility of 3×10^{-5}/bar was used alongside with semi-isotropic pressure coupling to maintain a pressure of 1 bar. Particle mesh Ewald (PME) [151, 152] with a relative dielectric constant of 2.5 and cutoff distance of 1.6 nm was used to compute long range electrostatics. The Lennard-Jones interactions were modified starting from 0.9 nm to 0 at 1.2 nm by the GROMACS shift scheme.

After creation of an equilibrated bilayer, 48 peptides (peptide to lipid ratio of about 0.2) are randomly added into the solution. The peptide to lipid molar ratio of 1:5 has been previously investigated in small (Aβ 25-35) peptide-membrane experiments [153]. The composite system is then energy minimized and re-equilibrated for 50 ns with position restraints placed on 4th residue backbone (on F-19 BB in x, y an z directions) in peptides and phosphate (on PO4 in z direction) in lipids. The bilayers were simulated with a fixed surface tension (Berendsen barostat) to mimic an enhanced area per lipid due to the outer membrane curvature of SUVs. The area per lipid of both POPC and POPS bilayers were fixed at 95 $\mathring{A}^2$ to simulate outer membrane of a SUV with 13.4 nm diameter. These values were obtained through a WEPMEM simulation of a small unilamellar vesicle system composed of 877 lipids and 61113 CG-water molecules and also verified against other coarse grained molecular dynamics simulations with comparable vesicle sizes [154, 155]. All other simulation parameters were kept similar to the first equillibration step. This second step allows us to equilibrate the monomeric peptides in the solution in presence of a lipid bilayer.

Finally, position restraints were removed and a production run of 1.5 μs was carried out using the remaining simulation parameters from the second equillibration step. To verify the statistics presented in this chapter, we also simulated a replica of bilayer-peptide systems with different initial states.

3.3.4 Analysis

Built-in functions of GROMACS, analysis modules of Visual Molecular Dynamics [156] (VMD) and in-house developed scripts were used to analyze molecular simulation trajectory.

3.3.4.1 Peptide absorption

The relative position of two F-S2 beads from the lipid bilayer surface, described by locally close (six nearest neighbors from center of mass) phosphate (PO4) beads has been used as metrics to differentiate between absorbed and unabsorbed peptides. To accommodate for the local curvature of lipid bilayers, the height of bilayer surface (h) is determined by the average heights of six nearest PO4 beads to each individual peptide. The peptides are then categorized into three classes - *completely absorbed* (CA), *partially absorbed* (PA) and *unabsorbed* (UA) based on positions of Phenylalanine S2 (F-19/F-20 S2) on a single peptide chain. If the position of both F-S2 along bilayer normal (z) is less or equal to h, the peptide is classified as completely absorbed (CA) whereas if $z > h$, the peptides are considered unabsorbed (UA).

3.3.4.2 Peptide-aggregate Clusters

Peptide aggregation in the simulated systems was quantified by clustering connected peptides. In this analysis, the peptides are described by main beads - BB, S1 and S2. This connectedness is established using distance cutoffs determined from the first peak of radial distribution functions ($g(r)$) between these beads. Two peptides are connected if they have at-least one pair of beads within their respective cutoff distance. These cutoffs are detailed in the Table 1.

The number of peptide clusters, representing peptide aggregates at a particular time is

Table 3.1: Non-bonded Lennard-Jones (LJ) interaction strengths in WEPPRO model. Unit of interaction strength (ϵ) is in kJ/mol. The radius (σ) of all LJ interactions is 4.7Å.

Interaction	Cutoff (nm)
BB - BB	0.45
BB - S1	0.40
BB - S2	0.61
S1 - S1	0.65
S1 - S2	0.42
S2 - S2	0.70

computed by determining connected components. In addition, the peptide aggregates is also categorized into three groups based on absorption classes of component peptides - *completely absorbed aggregates*, *partially absorbed aggregates* and *unabsorbed aggregates*. A peptide-aggregate has been listed as a completely absorbed aggregate, if *all* of its component peptides can be classified as either completely absorbed or partially absorbed. On the other hand, an aggregate is designated as a partially absorbed aggregate, if *at-least* one of its component peptide can be classified as partially absorbed and *at-least* one another as unabsorbed. Finally, if *all* of the component peptides are marked as unabsorbed, then the peptide-aggregate is categorized as unabsorbed aggregate.

3.3.4.3 Beta Sheet Content

A backbone contact between two peptides is defined by alignment of back-bond dipoles (BBm–BBp), characterized using a distance cut off of 2.5 nm between two oppositely charged BB-dummies. This distance was determined from the interaction peak of radial distribution function, $g(r)$, between opposite charged dummy particles. We considered two peptides as beta sheets, if they have at least five (71.5 $\%$) such backbone-backbone contacts and end-to-end distance greater than 1.2 nm, similar to beta sheet determination technique used by Lu et.al. [146].

The end-to-end length is the distance between the back-bone (BB) beads of flanking amino acids - K and E. The fraction of peptides in mutual beta sheets constitute the total beta sheet content.

The relative position of two F-S2 beads from the lipid bilayer surface, described by locally close (six nearest neighbors from center of mass) phosphate (PO4) beads has been used as metrics to differentiate between absorbed and unabsorbed peptides. To accommodate for the local curvature of lipid bilayers, the height of bilayer surface (h) is determined by the average heights of six nearest PO4 beads to each individual peptide. The peptides are then categorized into three classes - *completely absorbed* (CA), *partially absorbed* (PA) and *unabsorbed* (UA) based on positions of Phenylalanine S2 (F-19/F-20 S2) on a single peptide chain. If the position of both F-S2 along bilayer normal (z) is less or equal to h, the peptide is classified as completely absorbed (CA) whereas if $z > h$, the peptides are considered unabsorbed (UA).

3.4 Results and discussion

Fig. 3.2a shows the number of unabsorbed peptides over time whereas Fig. 3.2a-*inset* presents the variation in the number of partially absorbed aggregates with time. The peptides can partition into lipid membranes as individual monomers (Fig. 3.2b), as small oligomeric aggregates (Fig. 3.2c) or by slow dissociation of larger peptide aggregations (Fig. 3.2d) bound to the lipid membrane. Previous AFM studies with full length Aβ peptide have also reported some of these pathways — absorption as monomers and oligomers into supported bilayers [26]. There was a sudden significant decrease in the number of unabsorbed peptides within about first 15 ns for both anionic-PS and zwitterionic-PC simulations (Fig. 3.2a). Following that, the absorption slowed down due to formation of a number of partially absorbed aggregates. These aggregates

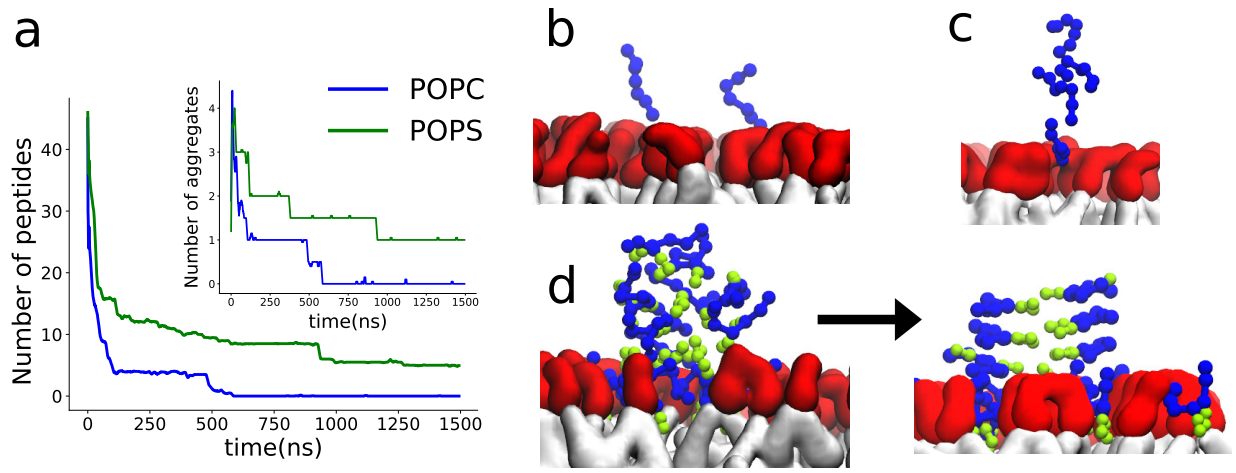

Figure 3.2: a) Variation of number of "unabsorbed peptides" with time, averaged over two replica-simulations. The variation of the number of partially absorbed aggregates has been provided as *inset*. b,c,d) Different pathways for peptide absorption into lipid bilayer. b) Single (monomeric) peptide absorption. c) Peptide absorption as oligomeric aggregates. d) Peptide aggregation through dissociation and rearrangement of partially absorbed aggregates. Coloring scheme: Light green beads - Sidechains of Phenylalanines (F); Blue beads - Peptide backbones; Red region - Polar/charged lipid headgroup; White region - Hydrophobic alkyl tails (Lipids).

arranged themselves to protect the hydrophobic cores of their component peptides (Fig. 3.2d), which afforded some stability and reduced absorption rates. Over time, these partially absorbed aggregates continued to slowly lose peptides to the membrane and decrease in total number and size. Particularly, in contrast to POPC where all inter-facial aggregates were completely absorbed, one such aggregate on POPS rearranged and conformed into a stable layered beta sheet structure (Fig. 3.2d) attached to the membrane, similar to structures observed in membrane-free systems.

The backbones of the absorbed peptides had variable residue-wise insertion into a lipid membrane modulated by their side chain hydrophobicity and relative position along the peptide chain (Fig. 3.3). Throughout the 1.5 μs simulation, the absorbed peptides remained close to the bilayer headgroup region (Fig. 3.2d, S6). Once absorbed into the membrane, the peptides (monomers and aggregates) start diffusing laterally on bilayer surface as shown in the supplementary

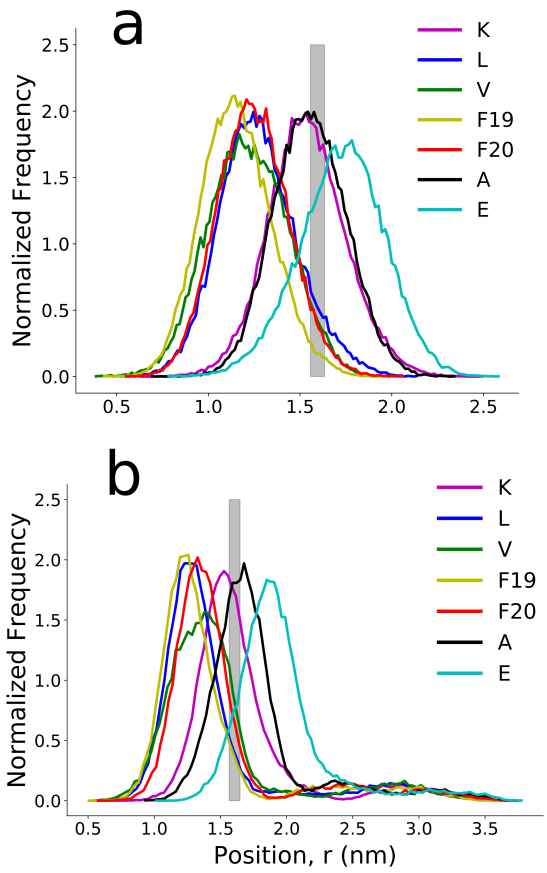

Figure 3.3: Residue-wise insertion of Backbone beads (BB) into different membranes. a)POPC bilayer; b)POPS bilayer. The gray region describes the average location of bilayer headgroup (PO4). The presented results have been averaged over both replica-simulations.

movie. In addition, as the peptides are pre-dominantly hydrophobic (71.43 %), the aggregation on the membrane occurs primarily by pushing away polar/charged lipid heads thereby exposing the hydrophobic tail region as is evident in Fig. 3.4a-d, where peptides have pre-dominantly aggregated on top of the hydrophobic alkyl tails (white region).

3.4.1 Rate of peptide aggregation

The variation in cumulative number of peptide clusters (ordered + disordered) over time is presented in Fig. 3.4e. Over time, due to continued aggregation, the number of peptide clusters

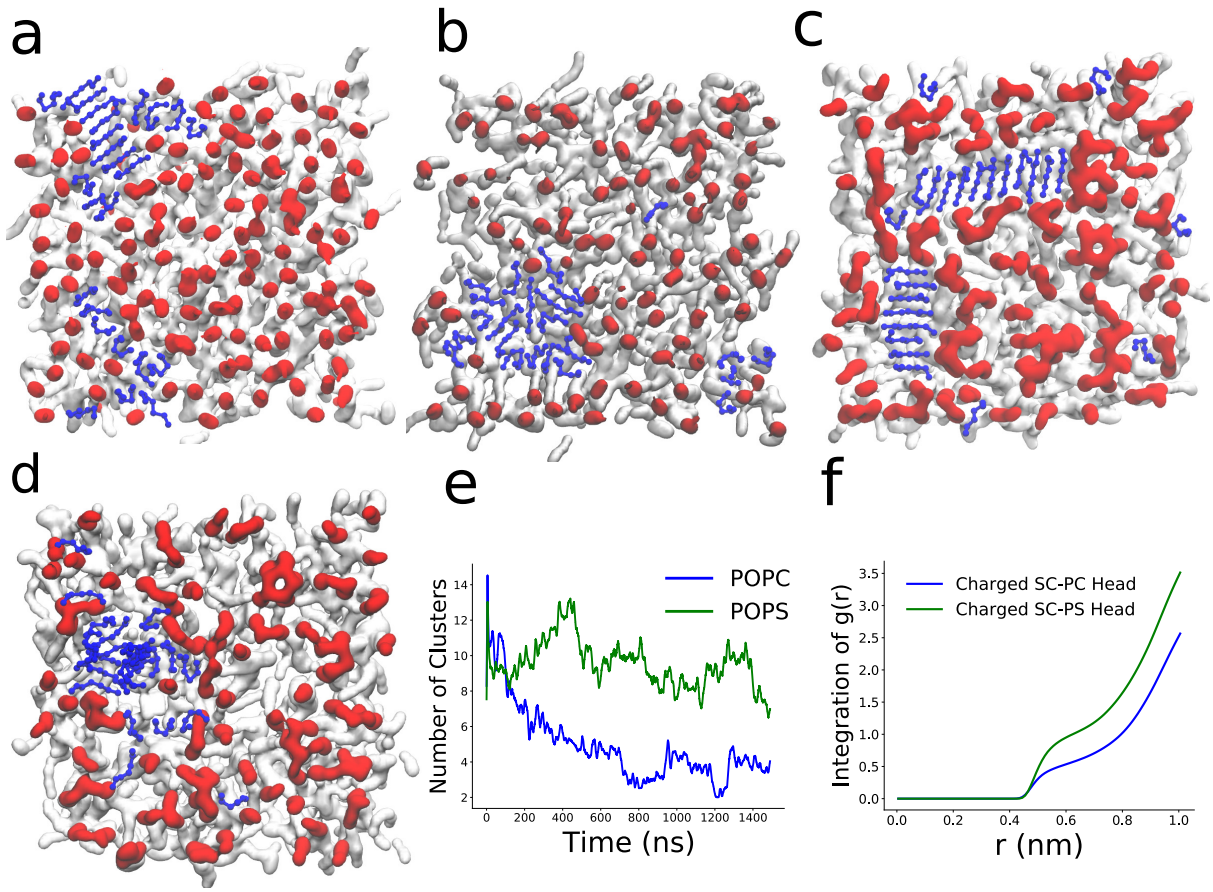

Figure 3.4: a,b) Last frame snapshot of peptide aggregates on two opposing leaflets of POPC lipid membrane in simulation 1. c,d) Last frame snapshot of peptide aggregates on two opposing leaflets of POPS lipid membrane in simulation 1. The red part in this representation corresponds to polar headgroup and the white part corresponds to hydrophobic tails. The blue connected beads represent peptide backbone. e) Variation in the number of Aβ 16-22 aggregates over time, averaged over both replica-simulations. Even connected components of size one (monomers) have been designated as a single cluster. f) Integration of radial distribution function between charged peptide sidechains (E/K-S2) and lipid headgroup (POPC:NC3/PO4, POPS:CNO/PO4), averaged over both replica-simulations.

on POPC steadily reduced in number (Fig. 3.4e) to about three clusters and increased in size — number of peptides (Fig. 3.5). In contrast, on POPS membrane, the number of peptide aggregates/clusters continued to be relatively high, featuring significant variations in aggregate sizes, pointing to a comparatively slower aggregation rate. This disparity in aggregation patterns can be explained through difference in diffusion rates of PC and PS. Lateral diffusion of lipids is

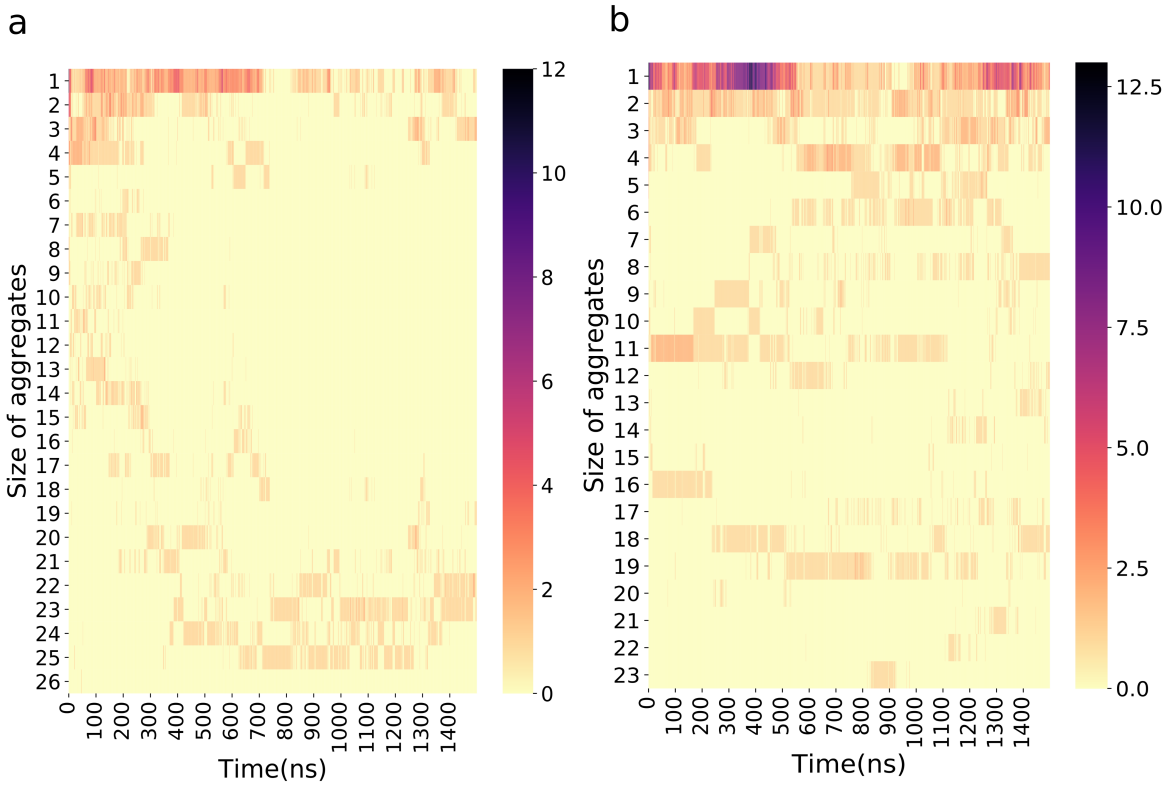

Figure 3.5: Variation of size of peptide aggregates with time. The colors of heatmap correspond to frequency of particular sized aggregate. a)POPC bilayer; b)POPS bilayer. The presented results have been averaged over both replica-simulations.

relatively slower in POPS with a lateral coarse grained diffusion constant of 0.0369 x 10ˆ-5 cm^2/s as compared to 0.0899 x 10ˆ-5 cm^2/s in POPC. While coarse grained diffusion constant is not directly comparable to experimental and atomistic results, it can capture relative trends. Similar qualitative trend has also been observed in other reported values for lateral diffusion constants [157]. The rigid headgroup of POPS due to intra-molecular pseudo-hydrogen bonds captured by CNO dipole-dipole interactions in this coarse grained scheme, prevents faster diffusion for lipids. This restricts the effective exclusion of lipid molecules [81], preventing peptide aggregation. In addition, increased interaction between lipid head group (CNO/PO4) and peptides in POPS can further result in better mixing of peptides and lipids, compared to POPC (Fig. 3.4f).

3.4.2 Beta Sheet Content

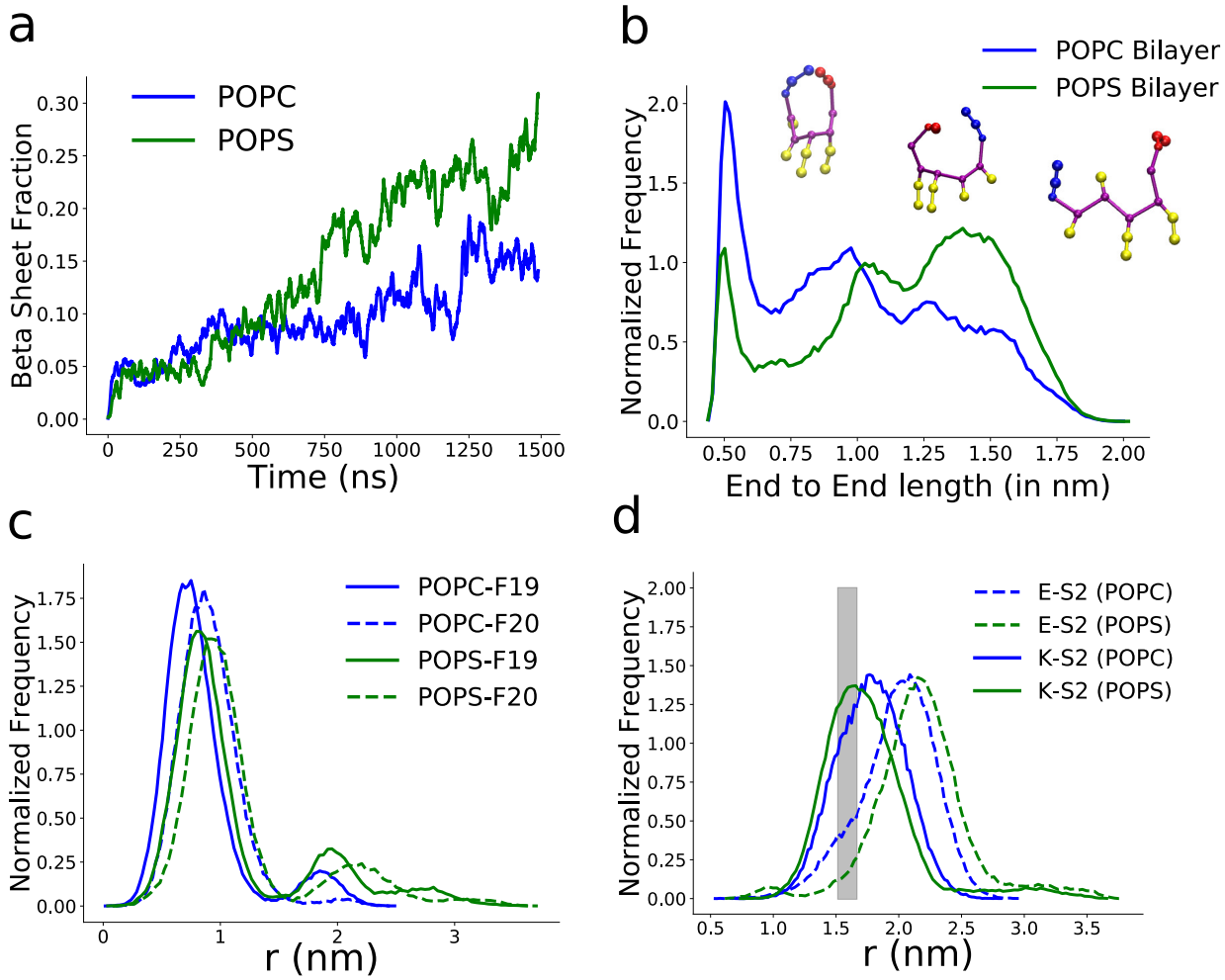

Figure 3.6: a) Time evolution of beta sheet fraction. b) Distribution of end-to-end length of peptides over the last 200 ns. The gray line shows the end-to-end distance criteria used to determine beta sheets. (*inset*)-single peptide representative snapshots describing end-to-end lengths of peptides. Peptide backbone of Aβ 17-21 (LVFFA) is presented in magenta, whereas residues K and E are represented by blue and red respectively. The connected blue beads represent hydrophobic sidechains. c) Density distribution of E-S2/K-S2 (POPC/POPS) along bilayer normal from bilayer center over the last 200 ns of simulation time. The gray region describes the average location of bilayer headgroup (PO4). d) Density distribution of F19-S2/F20-S2 (POPC/POPS) beads along bilayer normal from bilayer center over last 200 ns of simulation time. All the results presented here have been averaged over all replica-simulations.

Over time, peptides rearranged into aggregates with high beta sheet content on both bilayers.

Fig. 3.6a presents time evolution of beta sheet content over time. Similar to reported CD and

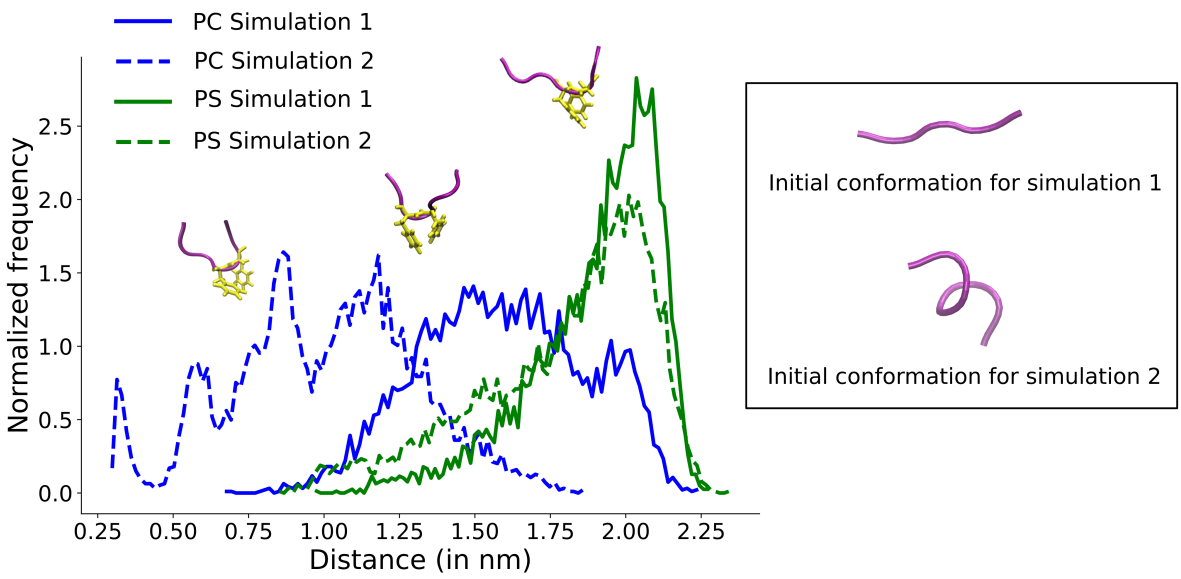

Figure 3.7: Distribution of peptide end-to-end distances over different atomistic simulations. The end-to-end distance is defined as the distance between terminal nitrogen (N) of K which is a part of peptide backbone and terminal carbon (C) of E which is also a part of peptide backbone. Color scheme of VMD snapshots: Purple - Peptide backbone, Yellow - F (Phenylalanine)

Thioflavin T assay studies [25] with Aβ 16-28, the beta sheet content in peptide aggregates was significantly larger (approximately double) for POPS as compared to POPC. Even with larger sized aggregates, peptides on PC bilayer were more unstructured. The higher beta sheet content on PS bilayer can be explained by the distribution of end-to-end distance of peptides. Peptides on PS are in general more elongated (Fig. 3.6b) which exposes their backbone to more peptide backbone-backbone interactions, thereby increasing the overall beta sheet content. In agreement with our coarse grained simulations, atomistic molecular dynamics simulations of single peptide-membrane systems also reproduce a trend towards more elongated peptides in PS than PC (Fig. 3.7). This lower end-to-end distance of peptides in PC can be reasoned in terms of membrane compressibility and peptide-lipid insertion. The higher compressibility of POPC membranes as compared to POPS, that has been previously reported [158] and also captured by our CG model [81], increases the relative penetration of peptides into POPC membranes. Fig. 3.6c

presents the distribution of F-S2, the "highest inserted" side chain bead into the membrane. This increased insertion of F-S2 on POPC membranes, coupled with charged residues at the ends — K and E which prefer to stay close to bilayer surface, distorts the shape of the peptide into a sharper "U/O - like" shape (Fig. 3.6b) thereby decreasing the average end-to-end distances and backbone-bone contacts. The higher membrane disruption for some zwitterionic lipids like DOPC compared to anionic DOPG has also been recorded in AFM experiments by Hanes et. al [26]. Moreover, similar to FTIR experiments on Aβ peptides in membranous environment by Yu *et. al.* [159], a relatively higher insertion of F-19 as compared to F-20 is observed. In addition, as apparent from Fig. 3.6d the sidechain of Glutamate (E-S2) in POPS is positioned slightly away from the bilayer-headgroup as compared to POPC because of the interaction between positively charged choline (NC3) and negatively charged E-S2 which is absent in POPS. This further increases possibility of intra-peptide K-E interactions which results in a less expanded peptide conformation.

In addition, due to higher number of total inter-peptide interactions in large amorphous peptide aggregations on POPC, it is difficult for peptides to effectively rearrange into ordered beta sheets necessary for efficient fibrillation. On the other hand, peptides on PS form numerous slow diffusing smaller oligomeric aggregates which can reorganize comparatively easily into beta sheet rich structures before coalescing into larger aggregates.

Oligomeric deposits on membranes can act as nucleation-seeds for subsequent peptide aggregation from solution. To test this hypothesis, after the initial production run of 1.5 μ s, we added 48 more peptides into both bilayer systems (with pre-existing peptide aggregates) and recorded their dynamics for 500 ns. Similar to previous experimental observations [160], initial pre-existing oligomers acted as nucleating sites, recruiting either individual monomers or smaller

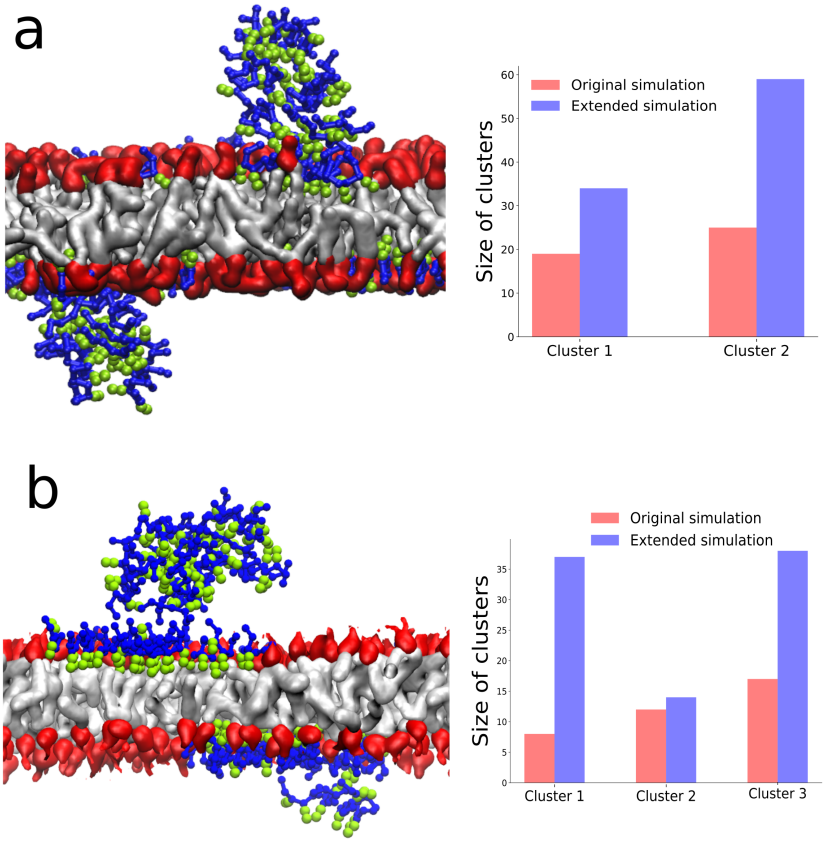

Figure 3.8: a) Sanpshot of peptide aggregation on POPC bilayer at the end of extended simulation. b) Snapshot of peptide aggregation on POPS bilayer at the end of extended simulation. Coloring scheme of VMD snapshots: Light green beads - Sidechains of Phenylalanines (F); Blue beads - Peptide backbones; Red region - Polar lipid headgroup; White region - Hydrophobic alkyl tails (Lipids). *Right*-Increase in size of clusters by addition of 48 new peptides and extension of simulation for POPC (a) and POPS (b). The size of initial cluster increased due to recruitment of peptides during 500 ns of extended simulation.

oligomers to create large fibrillar aggregates (3.8). This presents a possibility of multiple ordered aggregates on POPS (larger than POPC) to act as nucleation sites that can progressively increase subsequent aggregation rates. These fibrils are attached to the membranes primarily through inter-peptide interactions with embedded peptides and extend into solution. The aggregate sizes of fibrils in solution, attached to the bilayers is on average larger than oligomers lying flat on membrane surfaces.

Studies with central hydrophobic core (CHC) provides insights about how structures develop

in peptide aggregates [161–163]. Faster development of ordered beta sheet rich structures in Aβ 16-22 can be corroborated with faster ordered fibrillation and resulting toxicity of full length Aβ peptides in POPS. Similar to our observations, Lindberg et. al. [164] also found about two fold increase in fibrillation of Aβ 1-42 in an anionic membrane (DOPS). This effect of lipid headgroup chemistry has also been corroborated in previous work with surface plasmon resonance and magnetic bead assay on Aβ 1-40 [165].

3.5 Conclusion

Our coarse grained molecular dynamics simulations reproduce and provide a mechanistic explanation to a broad spectrum of experimental results [25, 26, 159, 164, 165]. We characterized multiple pathways for peptide absorption into membranes composed of POPC and POPS. Both lipid molecules have distinct effects on aggregation patterns of absorbed peptides. While rapid cumulative aggregation (ordered + disordered) was observed in zwitterionic PC bilayer, the emergence of ordered beta sheets and by extension, fibrillation was faster in presence of anionic POPS lipids. The results are in agreement with previous experimental studies [25, 137, 139–142, 164, 165] that had observed faster growth of amyloid fibrils in presence of anionic lipids. The discrepancy in the cumulative aggregation rates is a consequence of faster lateral diffusion of POPC lipid molecules compared to POPS. On the other hand, increased beta sheet content in POPS membranes is due to the differences in membrane compressibility. Higher membrane compressibility of POPC membrane compared to POPS, results in a relatively higher peptide insertion into the bilayer. This distorts the geometry of individual peptide molecules which hinders their participation in beta sheet formation. Some of the morphological aspects of membrane

assisted Aβ aggregation reported in this study such as relatively higher membrane insertion of F19 compared to F20, have been supported by previous experimental evidences [159]. We also revealed the propensity of initial oligomeric deposits to act as nucleation seeds to enhance further fibrillation. Considering the presence of multiple, ordered aggregates in POPS due to slow cumulative aggregation and peptide aggregates operating as nucleation seeds, POPS membranes can have an increased progressive peptide aggregation rates. This work unravels how lipid headgroup driven biochemical interactions in homogeneous model membranes shape peptide absorption and aggregation.

Chapter 4: Effects of Applied Surface-tension on Membrane-assisted Aβ Aggregation

4.1 Overview

This chapter is based on the author's publication: Effects of Applied Surface-tension on Membrane-assisted Aβ Aggregation; Abhilash Sahoo, and Silvina Matysiak. Physical Chemistry Chemical Physics, 2021.

In previous chapter, we focused on how lipid headgroups can impact Aβ aggregation. Beyond, membrane charge, membrane's physical organization can also affect peptide-peptide and peptide-membrane interactions.

Studies connecting biological membrane organization and Aβ aggregation are limited due to experimental and computational challenges. While experiments have suggested that an increased membrane curvature results in faster Aβ peptide aggregation in the context of Alzheimer's disease, a mechanistic explanation for this relation is missing. In this work, we are leveraging molecular simulations with a physics-based coarse grained model to address and understand relationships between lipid packing in membranes and aggregation of a model template peptide Aβ 16-22. In agreement with experimental results, our simulations also suggest a positive correlation between increased peptide aggregation and lipid packing defects. More curved membranes have higher lipid packing defects that engage peptide's hydrophobic groups and promote faster diffusion leading to the peptide's fibrillar structures. In addition, we curated the effects of peptide aggregation

on membrane's structure and organization. Interfacial peptide aggregation results in heterogeneous headgroup-peptide interactions and an induced crowding effect at the lipid headgroup region, leading to a more ordered headgroup region and disordered lipid-tails at the membrane core. This work presents mechanistic and morphological overview of relationships between biomembrane's local structure and organization, and Aβ peptide aggregation.

4.2 Introduction

Alzheimer's disease is characterized by extracellular self-assembly of Aβ protein segments into ordered fibrillar aggregations [89–91]. Recent reports have pointed to amyloid-membrane interactions as a crucial step in modulating the aggregation kinetics and associated disease pathogenesis [22,111,116,166]. In particular, biophysical aspects of membrane organization such as membrane order and charge can affect peptide aggregation kinetics. In addition, peptide aggregation can also effect changes to membrane's structural properties. Three experiment-derived hypothesis detailing effects of peptide aggregation on membrane organization have been suggested — carpeting model, membrane-pore model and detergent model [167]. But the exact biophysical mechanism characterizing this interaction is still missing. While significant research efforts have been applied to understanding peptide aggregation in aqueous solvent, studies in presence of membranous environments are rather limited. Research in this direction can impact our understanding of event-pathways in Alzheimer's disease.

Recent experimental reports have suggested that Aβ peptide can have differential aggregation patterns when interacting with anionic membranes compared to zwitterionic/cationic ones [25, 26]. Here, peptide insertion and peptide aggregation in presence of zwitterionic membranes

were correlated to the increased membrane compressibility. Experimental studies have also characterized the dynamic role of membrane curvature in driving the formation of and interaction with amyloids [27, 28]. Peptide aggregation was studied with Phosphatidylcholine vesicles of varying sizes using Thioflavin T fluorescence assay to track fibrillation. This study observed a reduced lag time with smaller vesicles compared to larger ones. The authors also used isothermal titration calorimetry to report endothermic binding between Aβ and small lamellar vesicles compared to exothermic binding with large unilamellar vesicles. Such curvature associated modulations in aggregation thermodynamics and kinetics has also been observed in the case of α-Synuclein (Parkinson's disease) and Huntingtin protein (Huntington's disease), suggesting a fundamental nature in protein aggregate-membrane interactions [168,169]. This presents a need for biophysical studies into membrane curvature/lipid-packing driven peptide aggregation.

While most studies of membrane-curvature induced peptide behaviour using atomistic simulation have been limited to studies with a single peptide, some emerging work have reported the correlated effects of membrane curvature and peptide aggregation [170, 171]. Physics-based transferable coarse grained forcefields like MARTINI have enabled molecular simulations with exact amino-acid sequences that have coupled protein aggregations to membrane-curvature [79, 172, 173]. But MARTINI involves restraining the secondary structure of each protein units, thereby preventing studies of conformational transitions.

In this work, we studied the inter-dependent effects of membrane lipid packing and peptide aggregation with 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and Aβ 16-22. Here, we have used POPC to create our model lipid membranes, as phosphatidylcholines are one of the most abundant lipid molecules in mammalian cells and previous experimental studies have used such membranes for peptide-aggregation studies [27].

4.3 Methods

The coarse grained simulation systems were created with Water-Explicit protein and membrane model outlined in chapter 2 (Fig. 2.2, Fig. 2.3). Both the membrane and peptide models feature local grouping of atoms to generate functional coarse-grained atom-types (CG-beads) that are representative of the physics and chemistry of the atoms they represent. Previous research has explored the use of these CG models to study membrane-induced peptide folding, calcium ion driven lipid demixing and peptide aggregation in presence of membranes [81, 82, 84, 85, 85, 174]. We validated our forcefield by simulations with Aβ 16-22 peptide in aqueous solution and in presence of membranes composed of POPC and POPS lipids, where peptides were allowed to interact with the membrane from solution and form ordered aggregates [82] (Described in Chapter 3, Fig.3.1).

Similar to Terakawa *et al.* 's experimental study on the effect of membrane-curvature on peptide aggregation, our model lipid membranes are also composed of POPC lipids [27]. The initial frame was prepared with Aβ 16-22 peptides (with peptide to lipid ratio of 1:10, similar to experimental systems outlined in Kandel *et al.* [153]) solvated randomly away from the membrane. We instituted a hardcore repulsion with c_6 and c_{12} terms set to 0 and 0.00247 respectively between lipid headgroups (phosphate and choline) in the lower leaflet and the peptide backbone. This forces the peptides to exclusively interact with the upper leaflet. As previous simulations have shown that this peptide fragment primarily stays at the membrane-water interface, such applied repulsion between the lower leaflet and the peptide backbone would not affect the peptide aggregation and adsorption properties. Each simulation system was composed with 242 POPC lipid molexules, 24 Aβ 16-22 peptides solvated in 5400 coarse grained water molecules.

The membranes were held with different values of surface-tension to simulate conditions with varying curvature and membrane packing defects. Such an approach of using applied surface-tension as proxies for membrane curvature has been previously used in atomistic simulations to study membrane insertion of anti-microbial peptides [175]. Furthermore, Aβ 16-22 stays at the membrane-interface, where membrane-curvature primarily manifests as an increased area-per-lipid [82]. Fig. 4.1 shows the relationship between the area-per-lipid and the applied surface-tension values in standalone simulations with POPC membranes in absence of peptides. The area per lipid is computed simply by dividing the total membrane lateral area with the number of lipid molecules in one leaflet. The highest surface-tension in our case corresponds to a small unilamellar vesicle with a diameter of 13.4 nm, which was verified with a 100 ns molecular simulation of complete vesicle (877 lipids) [82]. These values are similar to those reported with coarse-grained and large all-atom simulations [154, 155, 176]. We also simulated a larger membrane with 512 POPC lipid molecules, 50 Aβ 16-22 peptides and 9973 coarse-grained water molecules at the highest value of surface-tension for 1.5 microseconds, to verify that our system does not suffer from finite size effects. We note here that membrane curvature is a complex process involving several structural features of lipid headgroup and tails. Here, we are presenting a minimalist view of membrane curvature by focusing on the membrane packing defects arising from increased area-per-lipid in membranes at high surface-tension.

The simulation protocol is identical to that mentioned in Sahoo *et al.* [82]. A brief overview is presented here. The initial peptide-membrane system was created with peptides dispersed randomly in solution, away from the membrane. The molecular system was then energy minimized and equilibrated for a period of 50 ns, with position restraints on the peptides. After that, the position restraints were removed and the peptides were allowed to interact among themselves and

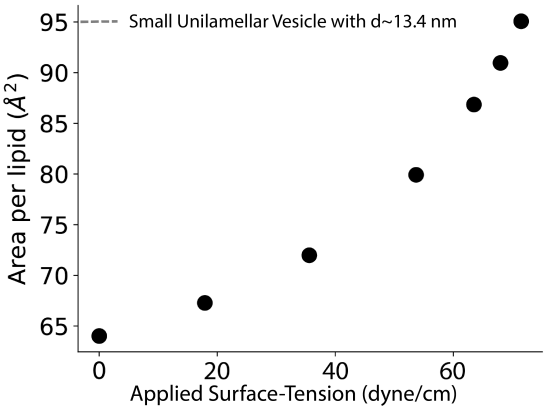

Figure 4.1: Relationship between applied surface-tension and the area-per-lipid. Each point denotes a simulation system.

with the membrane during the production phase of 1.5 microseconds. All the simulations were performed with GROMACS 4.5.4 [177], with temperature maintained at 300K with a Nose-Hoover thermostat [148] and pressure by Berendsen barostat [178]. Two independent replicas were simulated for each simulation condition. The long range electrostatics is determined through Particle-Mesh Ewald method with relative dielectric constant of 2.5.

4.3.1 Analysis

We have used visual molecular dynamics [156] and in-house developed scripts to analyze our molecular simulations. Different geometric features of peptide and lipid organization was used to define peptide absorption and ordered aggregation. We used the position of two second sidechain beads (S2) of PHE to determine absorption of a peptide into the membrane. First, six nearest lipid phosphate groups closest to the peptide center-of-mass were used to create a local membrane plane for each peptide, thereby allowing accommodations for local membrane undulations. Then, peptides were classified as absorbed if at least one of the PHE-S2 of the peptides moved beyond that local plane in the z-direction. Here, we classified peptide aggregates

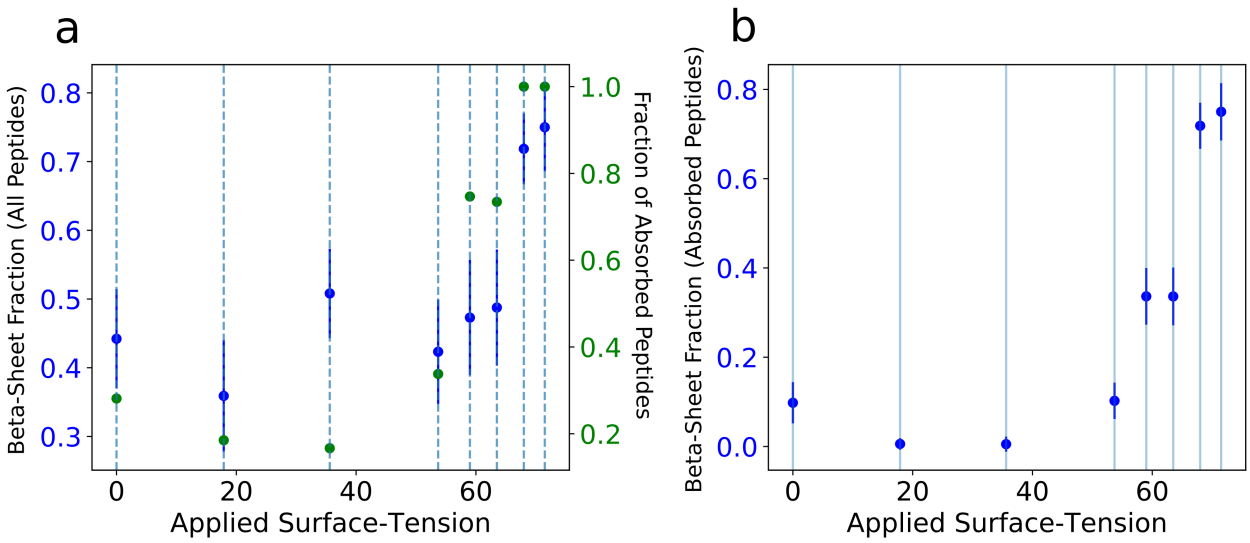

Figure 4.2: Variation of peptide absorption and ordered aggregation with increasing surface-tension. a) Absorption (Green) and ordered aggregation (Blue) among all the peptides (in-solution + on-membrane). b) Ordered aggregation (Blue) among peptides absorbed on the membrane only. These values are averaged over the last 200 ns of two independent replicates.

as "ordered" if each peptide in that aggregate participates in at least three backbone-driven dipole orientations. Such dipole-dipole interactions can be construed as equivalent to hydrogen bonds (as observed in atomistic representation of a β sheet) in a coarse-grained sense.

4.4 Results and Discussion

4.4.1 Impact of induced curvature on peptide aggregation

To determine the effect of membrane curvature and lipid packing defects on peptide aggregation, we characterized the structure of peptide aggregates in simulation with varying applied surface-tension values. Fig. 4.2 records the overall behavior (in-solution + on-membrane) of peptide aggregates, with measures for total ordered-aggregate fraction (Fig. 4.2a) and ordered-aggregate fraction among the peptides absorbed on the membrane (Fig. 4.2b). Peptides can interact and aggregate both in solution and on top of membranes. This competition between peptide-peptide

and peptide-membrane interaction shapes the peptide aggregation behavior. At surface-tension values greater than 35 dyne/cm, peptide absorption progressively increases. This increase in peptide absorption, in turn leads to an increased arrangement of the absorbed peptides into ordered macro-structures on the membrane surface (Fig. 4.2b), which increases the overall ordered content. It is worth noting that such increased fibrillation in more-curved membranes are in line with experimental observation of increased beta-sheets in small unilamellar vesicles compared to larger ones [27]. The computed fraction of absorbed peptides (1.0) and fraction of total ordered aggregation (0.68, standard deviation 0.02) from a larger simulation system at the highest surface-tension agrees with the values from the smaller system, suggesting no finite size effects. Therefore all further analyses presented in this article have focused on the smaller peptide-membrane systems.

To understand of this packing-induced peptide fibrillation, we used geometric and physical metrics that quantify membrane's structure and organization. Fig. 4.3a shows the distribution of hydrophobic solvent-accessible-surface-area (HSASA) for membranes at different values of surface-tension. An increased HSASA is indicative of more hydrophobic defects that can affect membrane-peptide interactions. These evidences suggest, peptide absorption is driven by increased hydrophobic defects in membranes with higher surface-tension. This is apparent in the molecular snapshot (Fig. 4.3b) showing these peptide aggregations on top of an exposed hydrophobic patch of the membrane at the highest surface-tension. Furthermore, increased area compressibility of these membranes allow for easy rearrangement of lipid molecules to sustain larger absorbed aggregates. Also, due to reduced crowding, peptide diffusion is faster in membranes with higher surface-tension. To capture that, we simulated a single absorbed peptide in a membrane at several different surface-tension values. Then we tracked the mean squared deviation of the peptide

center of mass with time lag (Fig. 4.4), which shows faster lateral diffusion with increasing applied surface-tension.

Now, as we are working with a dominantly hydrophobic fragment of the Aβ peptide, the hydrophobic side-chains of the absorbed peptides primarily interact with the membrane's acyl core. With increasing surface-tension values, these absorbed peptides with their side-chains engaged can diffuse around faster. The interaction between the peptides are then driven by backbone dipole alignments leading to ordered beta-sheet like structures. A parallel with Glycine-Valine repeats at hexadecane-water interface can be drawn here, with fibrillation at the interface region driven by buried hydrophobic sidechains into apolar media [80]. Therefore, this increased fibrillation fraction at later surface-tension values can be attributed to the cumulative effects of peptide absorption, higher membrane compressibility and faster peptide diffusion. It is worth noting here that similar to earlier solid-state NMR experiments, our simulations also report some ordered aggregations into cross-beta like structures in solution phase on membranes with lower surface-tension values (Fig. 4.3 - c, d) [37]. Here the peptide aggregates primarily in solution, tethered to the membrane (Fig. 4.3c). Such cross-beta structures are also noted for our simulations in solution, in absence of any membranes (Chapter 3). But, the prevalence of ordered aggregation is higher on "more curved" membranes because of the specific topological aspects (hydrophobic defects) that promote more absorption and backbone-driven inter-peptide interactions.

4.4.2 Effects of peptide aggregation on membrane structure

Amyloid depositions can alternatively affect membrane's structure which can in turn result in functional changes. Several studies have reported Aβ assisted membrane damage with formation of heterogeneous pores or detergent-like effects [166]. A correlation between membrane fluidity and Aβ self-assembly was suggested in experiments using dynamic light scattering, fluorescence and electron microscopy techniques [179]. This disruption of membranous environments is also noted by Lemkul *et al.* for single Aβ peptide with atomistic simulations [180]. The authors reported several structural impacts of Aβ-membrane interactions including local disordering of lipid groups, and increased tail-tilts.

In this work, we also analyzed induced changes in membrane order and headgroup tilt due to peptide aggregation. While the reported behavior exist in all simulation systems, we have shown results from membranes with the highest surface-tension as that system has the highest number of absorbed peptides, and therefore can be used to report statistically significant results. The coarse-grained tail order parameter, $\frac{1}{2}\left(3\left\langle cos^2\theta\right\rangle - 1\right)$, quantifies disorder in acyl tails. Here θ is the angle made by each bond vector in the acyl tail with the bilayer normal. The order parameter varies from -0.5 to 1, reaching the maximum value when bond vectors are aligned parallel to membrane normal. We classified lipid molecules in contact (any CG-bead within 7 angstroms) from the peptide aggregate as "close-to-aggregate" and others as "away-from-aggregate". Fig. 4.5a reports the tail order parameter at the interaction-sites of the acyl tail starting from the region near the headgroups to the center of the membrane. As compared to lipid molecules that are away from the aggregate, the lipid molecules closer to the aggregate are more ordered at the interfacial region, and more disordered deeper inside the membrane. This observation can be explained in

terms of the location of the peptide aggregate along the Z-axis (Fig. 4.5b). Due to the presence of peptide aggregates, the interfacial regions get more crowded, inducing increased order closer to the aggregates (Fig. 4.5a). Farther inside the membrane, a hollow defect is generated by pushing the headgroups away at the interface (Fig. 4.5b), which contributes to more disorder near the bilayer center, closer to the aggregates. These observations align with the reported "aggregate induced order" in previous-mentioned experimental study [179]. We observed this order at the interfacial region, and a more disordered acyl tail inside the membrane core as the short peptides reside at the membrane interface.

Finally, we also looked at how the lipid headgroups behave closer to the aggregate as compared to away from it. Differential arrangement of lipid headgroups can alter local membrane potential and have implications in cellular signalling processes. The distribution of headgroup ($\vec{PN}$) tilt with respect to the bilayer normal is plotted in Fig 4.5c. $\vec{PN}$ tilt is more dispersed closer to the aggregate. This can be understood in terms of increased heterogeneity induced by peptide-lipid interactions at the headgroups, closer to the aggregate.

4.5 Conclusions

Peptide aggregation into structured deposits have been associated with many neurodegenerative diseases, particularly through interaction with cellular membranes. Experimental results have suggested that both membranes and peptide can influence each other's structure and dynamics. Lipid packing defects associated with membrane curvature can promote extensive fibrillation, with faster formation of ordered structures in presence of small unilamellar vesicles, compared to larger ones. On the other hand, peptide aggregates can alter membrane's structure and packing. In

this work we have investigated these intertwined effects with coarse-grained molecular simulations. Our results agree with previous observations of curvature-driven peptide aggregation into ordered structures, and suggest possible biophysical mechanisms for it [27,28]. Membranes with increased curvatures lead to increased peptide absorption, due to more exposed hydrophobic defects. The absorbed peptides can then laterally diffuse around, and interact through the peptide backbone to form ordered fibrillar structures. The presence of such peptide aggregates also affects the lipid membrane's local structure. The lipid groups closer to the aggregates are more ordered at the crowded interfacial region due to interfacial presence of peptides; and less ordered deep inside the membrane core. These observations align with previous experiments and molecular simulations that also highlighted membrane disruption by peptides [179,180]. Our results support the previously suggested "carpeting model" of membrane disruption, by increasing membrane fluidity inside the membrane core [166, 181]. These locally close lipids also have a broad distribution in $P\vec{N}$ vector tilt from the membrane normal, suggesting the heterogeneous nature of peptide-lipid interactions. Such local variations in $P\vec{N}$ vector tilt can manifest in changes to the membrane's electrostatic potential, ion distributions and alter electrostatics-assisted membrane signalling [182,183]. This study unravels the effects of lipid packing in aiding peptide aggregation, and the effects of peptide aggregation in reshaping membrane's local attributes.

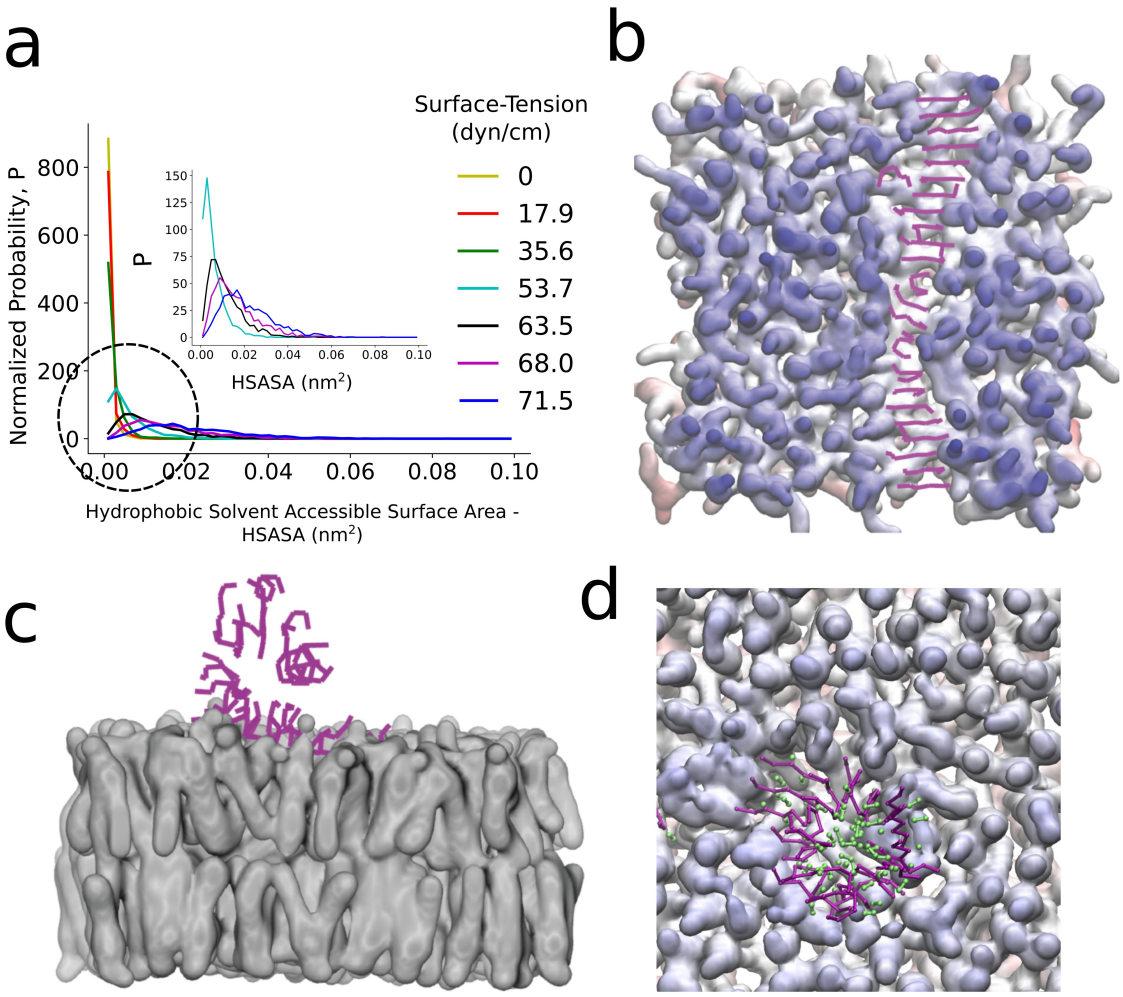

Figure 4.3: (a) Hydrophobic solvent accessible surface-area of membranes in absence of peptides. Snapshots of membranes with peptide aggregate. b - Simulation with surface-tension of 71.5 dyne/cm. Coloring Scheme - Membrane components are colored by their position along Z, from red to blue; Peptides: Magenta. c - Simulation without surface-tension (Lateral view). Coloring scheme - Membranes: Grey. d - Simulation without surface-tension (Top view). Coloring Scheme - Membrane components are colored by their position along Z, from red to blue; Peptides: Magenta; Hydrophobic groups: Lime

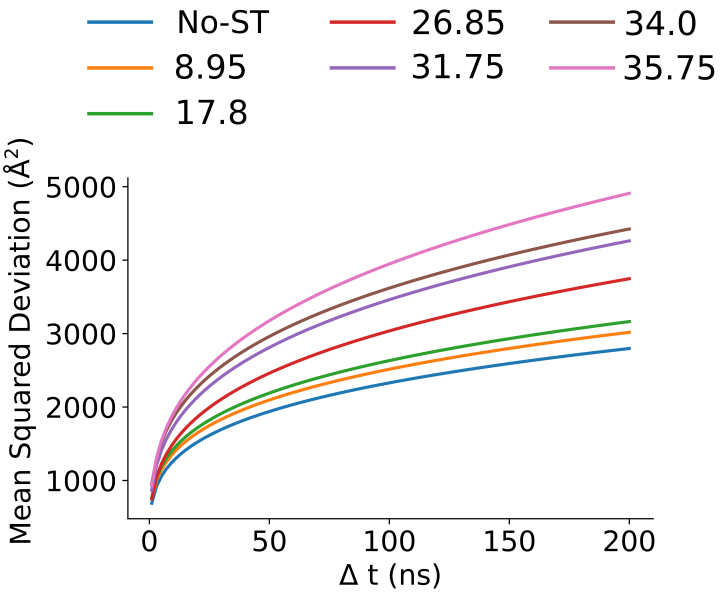

Figure 4.4: Mean squared deviation of a single peptide as a function of time-lag.

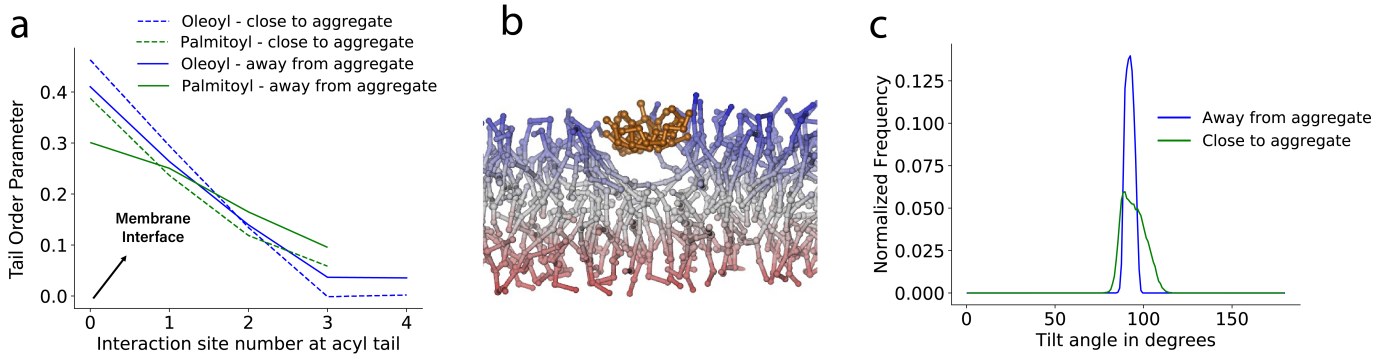

Figure 4.5: a - Lipid tail order with respect to interaction site at acyl tail, numbered starting from the membrane interface. b - A snapshot of peptide aggregate on the membrane for simulation with surface-tension of 71.5 dyne/cm. Coloring Scheme - Membrane components are colored by their position along Z, from red to blue; Peptides: Orange. c - Distribution of headgroup (PN-vector) tilt with-respect-to the bilayer normal.

Chapter 5: Aggregation of Aβ 16-22 in Hyperglycemic Conditions

5.1 Overview

Over chapter 3 and 4, we have been focusing on peptide membrane interactions and how that affects specific details of peptide aggregation. Beyond cell membranes, there are several other factors in the physiological environment that can shape peptide aggregation behavior. Co-solutes are a large class of molecules that can affect aggregation patterning by affecting peptides through both specific and non-specific interactions. In this chapter we explore how solvated glucose molecules affect Aβ 16-22 aggregation in a concentration dependent manner.

Previous reports have drawn out correlations between pathogenesis of Alzheimer's disease and hyperglycemic conditions associated with type 2 diabetes. Previous research has proposed chemical crosslinking, in presence of glucose molecules as responsible for increased toxicity. But the time-scale separation between increased aggregation and formation of chemical cross-links suggests the presence of an alternate thermodynamic mechanism, which we aim to explore in this research. In this work, we apply coarse-grained molecular dynamics simulations to study the how glucose molecules can drive Aβ 16-22 aggregation. Our analyses suggest that increasing the concentration of solvated glucose molecules can result in faster aggregation of Aβ 16-22, without any appreciable change in β-sheet content. This change in aggregation rates can be explained in terms of relative orientations of interfacial glucose molecules. The glucose molecules at the

peptide-water interface features a preferential orientation, resulting in loss of rotational entropy in a concentration dependent manner. This can assist in faster aggregation of the peptide, to reduce the cumulative availability of solvent accessible surface-area.

5.2 Introduction

Several studies have suggested pathological correlations and epidemiological linkages between a very common metabolic disease, type II diabetes (T2D) with Alzheimer's disease [42–45]. T2D involves development of cellular insulin resistance, which leads to improper sugar metabolism and high blood sugar concentration. Recent fMRI studies have suggested a declined cognitive performance of patients suffering from T2D [45]. Another study has suggested that patients with T2D have a 50% higher chance of suffering from AD [42]. But, the mechanistic event pathways of this correlation is still not clear. While, numerous experimental and computational studies have investigated Aβ conformational changes and aggregation in aqueous environments, investigations into Aβ-sugar aggregate structures and sugar-induced peptide aggregation, at elevated levels of physiological sugar concentrations in the case of T2D is limited [184].

Previous research efforts have postulated the post-translational glycation (Advanced glycation end products - AGE) induced increased aggregation as possible pathological pathway correlating AD and T2D [46–48]. These covalent modifications stabilize peptide aggregates against degradation, linking it to neuronal dysfunction. But, *in-vitro* studies have reported that formation of AGE can take upto a month with incubation in very high concentrations of glucose. At early stages of Aβ aggregation formation of AGE is not reported [49]. Experiments with cellular culture has revealed the upstream effects of glucose in production of Aβ peptides and preventing degradation

of Amyloid Precursor Protein. Glucose can also effect structural changes on peptide aggregates and alter its interaction with physiological structures. Recent experiment by Kedia *et. al.* have suggested that the presence of sugar molecules can lead to faster formation of toxic, unstructured and membrane-active oligomers that can be taken up by the cells and can interfere with mitochondrial activity [49].

The presence of co-solutes can modulate protein folding and peptide aggregation pathways, prompting the use of simple saccharides as macromolecular crowding agents [185–187]. These are often used to simulate cell-like crowded conditions in *in-vitro* experiments. Several experimental, computational and theoretical studies have outlined the effect of such crowded conditions — from altering the thermodynamic phase space to modulating the kinetics of processes [188–191]. This poses the question whether there are thermodynamic pathways that can explain the behavior of Aβ peptides in hypoglycemic conditions, beyond the current understanding of covalent modifications.

Computational studies on Aβ peptides in hyperglycemic conditions is limited. A recent atomistic molecular dynamics with beyond-physiological (0.8 M) glucose and full-length Aβ 1-42 revealed a caging effect of glucose molecules on peptide aggregates, which led to an increased hydration of protein structures [192]. A complete biophysical picture of amyloid aggregation in physiological/pathological hyperglycemic condition is missing.

In this work, we adopt coarse-grained molecular dynamics simulations to investigate protein aggregation behavior with Aβ 16-22 peptides in near-pathological hyperglycemia. Aβ 16-22 peptides are recognized as one of the smallest fragments capable of forming fibrillar structures, which has prompted research studies with this segment as templates [37–40]. With coarse-graining, we reduce the number of interaction-centers in our simulation-system to access spatio-

temporal scales not typically accessible by traditional atomistic molecular dynamics simulations. Here, we used the in-house developed coarse-grained protein model, Water-explicit Polarizable Coarse-Grained Model (WEPCGM, Chapter 2) that can capture protein's secondary structural transitions starting from a primary sequence of amino acids, specifically in presence of external stimulus. The coarse grained model can reproduce cross-beta like Aβ 16-22 aggregate structures in aqueous solution, and ordered membrane-adsorbed beta sheet aggregates in presence of model membranes (Chapter 3, Chapter 4). In this chapter, we present unique structural features of Aβ 16-22 and glucose co-aggregates and discuss mechanistic perspectives of glucose-accelerated Aβ aggregation.

5.2.1 Peptide and Glucose Model

We refer readers to previous chapters [Chapter 2] and previous publications for details of the CG model [80–82]. Please refer to the schematic figure 2.2 for details of the mapping scheme.

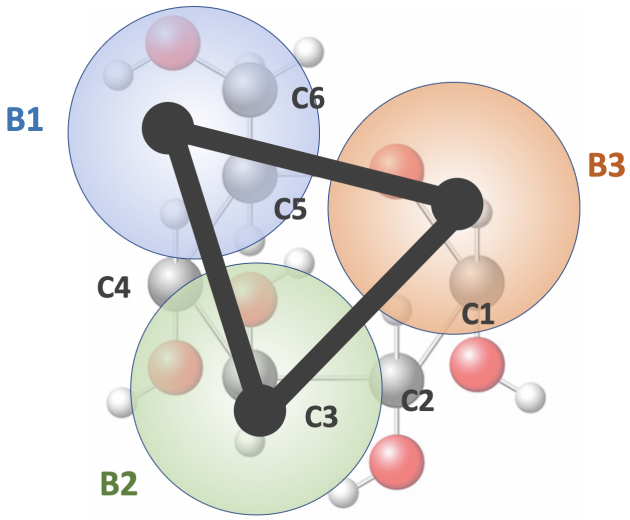

Figure 5.1: Geometry of the Glucose Molecule. The atomistic numbering is in black. B1: Blue; B2: Green; B3: Orange

The geometry and cross-interactions terms for glucose molecules are borrowed directly from the MARTINI model [Fig. 5.1] [193]. We verified the fidelity of the glucose coarse grained model — glucose-glucose and glucose-water interactions in polarizable water by computing the osmotic second virial coefficient, B_{22}, of the osmotic pressure that can be a measure of deviation from ideality of solutions and matching with previous recorded values in experiments and atomistic simulations. Here, we followed the simulation protocol outlined by Schmalhorst *et al.*. A simulation box of $(19)^3 nm^3$ was created with 420 randomly placed glucose molecules and then solvated with Martini polarizable water [83]. Here, we followed an approximate strategy to derive B_{22} using the radial-distribution-function, *g(r)* and cumulative number distribution function *N(r')* developed by Schmalhorst *et al.* [194]. We obtained an average B_{22} value of 0.119 (+/- 0.02) L/mol, close to the reported experimental B_{22} for glucose (0.117 L/mol).

Glucose molecules have carbon numbered from C1-C6 starting from the aldehyde (in open conformation) group. Here carbon and the associated oxygen atoms for C1 and C2 are mapped to CG interaction site B3; C3 and C4 mapped to B2; and C5 and C6 are mapped to B1. The sites B1, B2 and B3 are all linked together forming a planar, triangular shape representing glucose in its cyclic closed form.

5.2.2 Simulation Setup

The mixed systems — glucose molecules, Aβ 16-22 and water molecules were created by placing the groups randomly in a simulation box. We created three different compositions (peptide-to-sugar ratios of 1.66, 0.33 and 0.16) for analyzing the effect of increasing sugar concentration on Aβ aggregation. The concentration of glucose molecules were correspondingly

maintained at 1.98 mM, 9.8 mM and 19.8 mM respectively. The regular blood sugar concentration in healthy individuals can range from 2.5-6.5 mM, and hyperglycemia associated with diabetes has a concentration of 7-10 mM [195]. Emergency hospitalization is recommended for concentrations higher than 16.7 mM.

After the preliminary energy minimization, the systems were equilibrated for 5 ns (time step, dt = 10 fs) using NVT ensemble. Temperature was maintained at 300K through a Noose-Hoover thermostat [148, 149] with a time constant of 1 ps. Following this, production simulation was run with constant-pressure ensemble. Parinello-Rahman barostat [150] with a time constant of 1 ps and compressibility of 3×10^{-5}/bar was used alongside with isotropic pressure coupling to maintain a pressure of 1 bar. Particle mesh Ewald (PME) [151, 152] with a relative dielectric constant of 2.5 and cutoff distance of 1.6 nm was used to compute long range electrostatics. The Lennard-Jones interactions were modified starting from 0.9 nm to 0 at 1.2 nm by the GROMACS shift scheme.

5.2.3 Analysis

To calculate peptide aggregate sizes, we used an interaction cut-off of 7 $\mathring{A}$ from any CG interaction site to determine if a peptide is part of the aggregate. We also quantified the structure of these peptide aggregates through a β-sheet metric. Peptides were considered as part of β sheet if greater than 57 $\%$ of the backbone dipoles aligned. These explicit dipole-dipole interactions can be used as surrogates in our CG representation for hydrogen-bonds between peptides in β sheets. This analysis metric is specific to our coarse-grained forcefield, and we have used similar metric in our previous publications to quantify ordered peptide aggregation.

Finally we also report the relative enrichment of glucose molecules as a function of distance from the peptide aggregate.

$$P_i(r) = \frac{n_i(r)\,[N_i + N_{ow}]}{N_i\,[n_i(r) + n_{ow}(r)]}$$

Here, N_i and N_{ow} are the total number of groups of species i and water molecules respectively. This metric has been applied to study solvation effect of proteins and protein aggregation [196–199]. Here, we compared this relative enrichment of different parts of the glucose molecule (B1, B2 and B3) near the peptide aggregate.

5.3 Results and Discussion

5.3.1 Impact of Glucose on Aβ 16-22 Aggregation

To characterize the effect of glucose on peptide aggregation, we analyzed the peptide aggregate size in simulations with varying glucose concentrations. Fig.5.2 a-d shows the time-series of aggregate sizes with different near-physiological sugar concentrations. In most (except one replica for no glucose case) of the simulations, we recorded aggregation into larger aggregates with all peptides involved (Snapshots - Fig. 5.2e, f). The scatter plots (Fig. 5.2a-d) have data aggregated from two replica simulations for each peptide-sugar ratio. In our simulations, due to high concentration of Aβ peptides (33mM), we do not observe the expected lag phase in self-assembly processes, rather there is a fast initial evolution of aggregate sizes in all simulation systems. We observed a concentration dependence in peptide aggregation kinetics, with faster *complete* aggregation at increased sugar content and slower evolution for reduced concentrations.

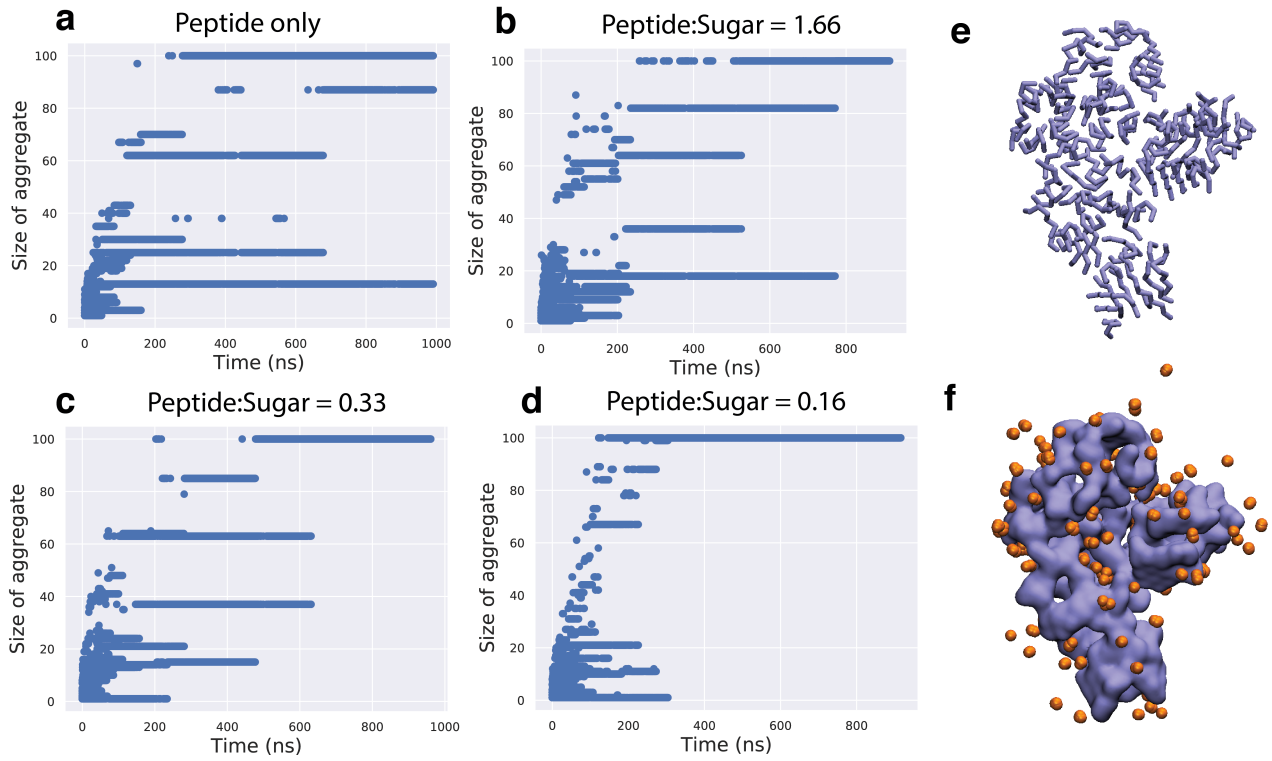

Figure 5.2: 1a-d: Evolution of peptide aggregate sizes over time at varying concentration of co-solvated glucose (1a) 0M; 1b) 1.98 mM; 1c) 9.8 mM and 1d) 19.8 mM) 1e-f: Structure of peptide aggregate (1e/1f - violet) and spatially close glucose molecules (1f - orange) created from the 19.8 mM glucose simulation at the final time-step

This correlations between increased hyperglycemia and Aβ fibrillation has been previously noted in experiments by Kedia *et al.* [49].

The glucose molecules were predominantly concentrated at the interface of the aggregate, forming a cage-like structure at about 5-7 $\mathring{A}$ distance, with fast glucose exchanges between interface and the bulk-solvent (Fig.5.2 f), similar to the organization suggested by Menon *et al.* in their atomistic study with Aβ 1-42. Such, organization of glucose molecules can influence local environmental alterations [192]. The authors here reported a decrease in protein dimerization because of co-solvated glucose. But this can be attributed to the non-physiological concentration of glucose molecules (0.8 M) in their simulation systems. This can be explained in terms of slower peptide diffusion in these crowded environments. Slower peptide, and aggregate diffusion can result in lower peptide-peptide interactions necessary for increase in aggregate size. Such diffusion-limited peptide aggregation behavior is noted in previous research on macromolecular crowding and aggregation [189]. Here, we need to note that physiological levels of glucose is about 10 mM in the case of type 2 diabetes. Previous experiments investigating the impact of glucose on Aβ aggregation have also adopted glucose concentrtions of 5 mM and 10 mM.

5.3.2 Secondary Structure in Protein Aggregates

Further, we analyzed β sheet content in our protein aggregates and tracked its growth over time (Fig. 5.3). Our simulations do not show any particular trend in the amount of ordered β sheet. There is a quick gain in ordered structures at the start of the simulation, followed by a rough close to β-sheet content of 0.4/0.5. The presence of glucose at near physiological levels led to *faster* aggregation, but did not affect the fibrillation levels. Previous experimental reports

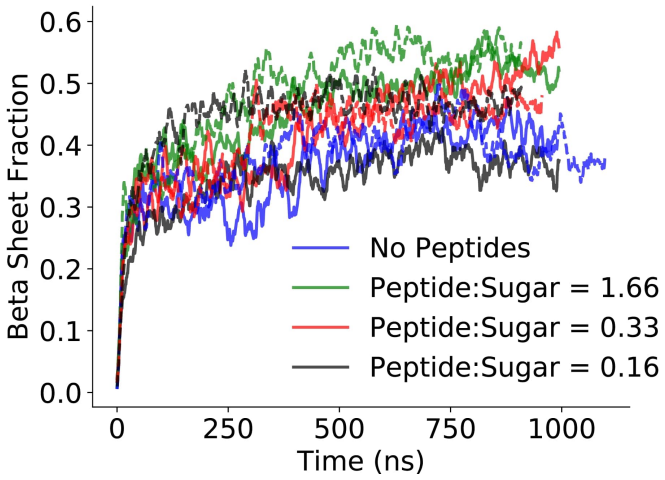

Figure 5.3: Evolution of β sheet content over time.

by Kedia *et. al.* also reported this increased aggregation in presence of glucose without any changes in circular dichorism [49].

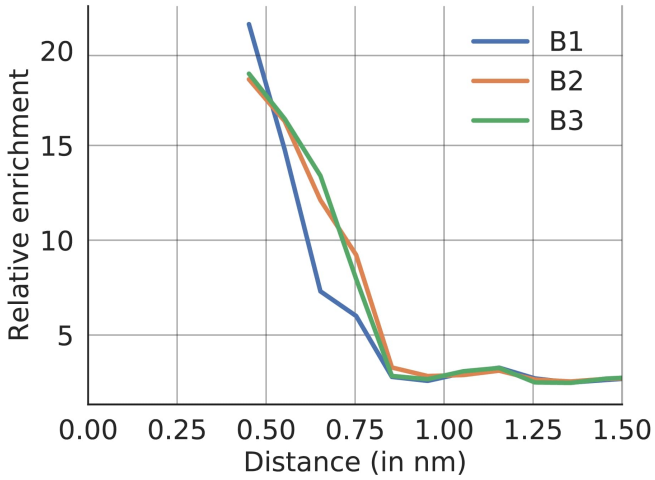

Figure 5.4: Relative enrichment of individual coarse-grained interaction site of glucose

5.3.3 Restricted rotation of interfacial glucose molecules

Finally, we quantified the local enhancement of glucose molecules close to the peptide aggregate. Fig. 5.4 is the relative enrichment (RE) of each CG interaction site of glucose, compared to water molecules as a function of distance from the protein aggregate. As glucose is an asymmetric molecule, here, we want to understand if there are any directional nature of glucose-peptide interaction.

We found that the C1 and C2 groups of glucose (CG interaction site - B3) interacted preferentially with the peptide, whereas B2 (representing C3 and C4) and B1 (representing C5 and C6) preferred interacting with water. Previous *ab initio* calculations of glucose's hydrogen-bonding patterns in water has noted a similar asymmetric set of interactions [200]. The authors found that the oxygens at C5 and C6 have lower hydrogen bonding capabilities compared to other oxygens. We observe a similar trend, with C5 and C6 partitioned closer to the peptide aggregate. This preferential orientation of glucose molecules will result in restricted rotations, and therefore a loss of rotational entropy. This can explain our concentration dependent increase in aggregation rates. With increase in size of aggregate (concurrent increase in aggregation rates), the aggregate surface-area accessible for interaction with glucose decreases, which in turn reduces this entropic penalty. Similar macromolecular crowding and entropy associated effects (folding/denaturation) are noted also for RNA and protein systems [201, 202].

While epidemiological studies have correlated AD and T2D, the primary reason for this has been attributed to covalent modification of arginine and lysines (AGE), resulting in cross-linking of amyloid aggregates. These covalent crosslinks prevent dissolution of peptide aggregates. Reports have suggested of a possible temporal mismatch between peptide aggregation and chemical

modifications [49]. Here we uncover a complementary thermodynamic pathway for this observed increase in peptide aggregation with glucose.

In the case of protein aggregation, research on macromolecular crowding effects has been generally limited to non-interacting (steric) crowders and very high concentrations to mimic physiological viscosity in biological environment [189, 203]. Latsaw *et. al.* reported that such crowders can lead to decrease in effective sizes of aggregates resulting in increased number of smaller oligomers. In some instances, addition of *ad-hoc* hydrophobic effect to the crowder could shape the structural features of the aggregate. Finally, adding soft non-specific crowder interactions have been shown to have unique implications, and reverse the effect of systems with *only* steric crowders [204, 205]. In this work, we report a non-specific, but directional interaction of glucose molecules, with a certain preferred orientation. This uncovers another possible mechanism how specific co-solutes can affect biologically relevant processes.

5.4 Conclusion

Self-assembly of Aβ peptides into specific deposits is noted as an essential upstream process in Alzheimer's disease. Several recent studies have suggested interlinks between high blood sugar levels and Alzheimer's disease. This has led to biophysical research to uncover connections between hyperglycemic conditions and Aβ aggregation. Several research studies have indicated the ability of glucose to chemically modify amino acids and create covalent crosslinks, which are resistant to dispersion. Currently, to our knowledge, there are limited studies aimed at uncovering thermodynamics-derived mechanisms that can aid peptide aggregation in presence of glucose molecules. In this work, we approach this research question from a computational lens — using

coarse-grained molecular simulations to understand the aggregation of Aβ 16-22 with varying concentrations of glucose (1.98 mM, 9.8 mM and 19.8 mM). Our analyses of coarse-grained simulation trajectories also suggest a glucose concentration dependent increase in aggregation rate. We did not see any changes in the amount of β-sheet rich ordered structures in presence of glucose molecules. The glucose molecules close to the peptide aggregate preferred a particular orientation, with C5 and C6 partitioning closer to the aggregate, while other CG interaction center associated with other carbon molecules preferring interaction with water. This restricted rotation of this ring molecule would result in a rotational entropy loss, that can force a reduction of accessible area for glucose molecules to interact. Therefore, hastening the peptide aggregation process. This could be a possible thermodynamic mechanism for this concentration dependent increase in aggregation rate of glucose.

Chapter 6: Transferable and Polarizable Coarse grained model for Proteins - ProMPT

6.1 Overview

This chapter is based on the author's publication: Transferable and Polarizable Coarse grained model for Proteins - ProMPT; Abhilash Sahoo, Pei-Yin Lee and Silvina Matysiak. A.S. and P.L. are co-first authors in this publication.

The application of classical molecular dynamics (MD) simulations at atomic resolution (fine-grained level - FG), to most biomolecular processes, remains limited because of the associated computational complexity of representing all the atoms. This problem is magnified in the presence of protein-based biomolecular systems that have a very large conformational space and MD simulations with fine-grained resolution have slow dynamics to explore this space. Current transferable coarse-grained (CG) force fields in literature are either limited to only peptides with the environment encoded in an implicit form or cannot capture transitions into secondary/tertiary peptide structures from a primary sequence of amino acids. To address these constaints, we initially came up with our *in-house* transferable coarse-grained model — water-explicit polarizable coarse grained model for proteins (WEPPRO) that I introduced in chapter 2 (methods). Although, this model afforded us access to several biomolecular systems; it has several shortcomings.

Firstly, WEPPROM does not have representation for all the essential amino-acids. Secondly, it has not been studied and validated for tertiary folding protein structures.

In this work, we present a transferable CG forcefield (Protein Model with Polarizability and Transferability - ProMPT) with an explicit representation of the environment for accurate simulations with proteins. The forcefield consists of a set of pseudo-atoms representing different chemical groups that can be joined/associated together to create different biomolecular systems. This preserves the transferability of the forcefield to multiple environments and simulation conditions. We have added electronic polarization that can respond to environmental heterogeneity/fluctuations and couple it to protein's structural transitions. The non-bonded interactions are parametrized with physics-based features such as solvation, and partitioning free energies determined by thermodynamic calculations and matched with experiments and/or atomistic simulations. The bonded potentials are inferred from corresponding distributions in non-redundant protein structure databases. We present validations of the CG model with simulations of well-studied aqueous protein systems with specific protein fold types- TRP-cage, TrpZip4, Villin, WW-domain and β-α-β. We also explore the applications of the forcefield to study aqueous aggregation of Aβ 16-22 peptides and dimerization of glycophorin A (GPA) in micellar environments.

With this model, we primarily address the shortcomings of WEPPRO. ProMPT now has access to all the amino acid representation. It has updated bonded and non-bonded interaction parameters with which we validated the conformational landscape of several small proteins and aggregates.

6.2 Introduction

Physiological functions of protein molecules are closely intertwined with their associated structure and dynamics [206,207]. This complex macro-organization is shaped through microscopic multi-body interactions that define a protein molecule's conformational landscape. In this direction, computer simulations, primarily molecular dynamics are being increasingly leveraged to understand protein biochemistry and biophysics [208–210].

With recent advancements in dedicated high performance computing architectures and graphical processing unit, it is now possible to run long simulations with small proteins for multiple microseconds, at conventional all-atom resolution [211–214]. Such long molecular simulations could capture a small number of protein folding-unfolding transitions in an unbiased manner. But such processes require significant amount of computing resources and are still not scalable to larger systems. Most biological systems of interest involve large biomolecules that interact over long spatiotemporal scales. To alleviate some of these shortcomings, several enhanced sampling approaches such as metadynamics and replica exchange molecular dynamics have been proposed to allow faster exploration of conformational space with limited resources [14–17]. But such approaches require extensive knowledge about the particular biomolecular system and/or can be infeasible for larger systems.

Coarse-grained molecular dynamics (CG-MD) involves creating a simplified representation or minimal model of biomolecules that can capture the essential biophysics [18]. This approach allows efficient access to long spatio-temporal scales by directly reducing computational complexity and allowing fast conformational transitions by smoothing the local free energy landscape. Early coarse-grained (CG) models with simplified phenomenological potentials were instrumental in

establishing the foundations of energy landscape theory of protein folding [215–217]. The coarse grained models require creating interaction sites that are representative of a particular molecule and defining interaction schemes (potentials) that allow these interaction sites to communicate. The interaction potentials, also known as *forcefields* can be variations of knowledge-based (developed from analysis of statistical databases) or physics-derieved potentials (created for example to fit free energies, partition coefficients, and biomolecular phases.) [**?**, 218–224]. Coarse grained molecular simulations of proteins have played a significant role in shaping our understanding of physiological processes such as protein folding, protein aggregation and membrane-protein interactions [217, 225–231].

CG models with varying molecular resolutions and diverse coarse graining strategies have been proposed. The levels of coarse-graining can vary from multiple-residues represented by a single interaction site, to models with representation for all the heavy atoms in a biomolecular system. Several of the CG models employ an implicit description of the solvent environment through interaction potentials [220, 221, 224, 232–237]. These models have been used to study several protein-based dynamical processes such as folding, adsorption, misfolding and aggregation. While these models provide a significant reduction in system complexity, an implicit representation of the environment cannot be used to study heterogeneous environmental effects that can be essential in crowded physiological systems. Moreover, the role of solvent in governing thermodynamics and kinetics of protein folding is well documented [238–240]. In some of the implicit-water CG models, the dynamics is biased towards native state, and cannot capture non-native states [216, 241–243].

MARTINI is a popular coarse grained forcefield with explicit description for solvents and a modular architecture [79, 244]. The interactions are directed through a set of interaction-

levels created by leveraging environment-dependant free energies (solvation, vaporization and partitioning free energies). Due to its reasonable accuracy and ease of use, particularly in heterogenous and crowded environment, this forcefield has been widely adopted by the natural science and engineering communities. In the case of proteins, MARTINI has been used to study ligand-binding, aggregation, surface-absorption and membrane translocation [245–250]. However the model relies on restraining protein's secondary structures through artificial potentials, and the global/tertiary structure also needs to be restrained to prevent spontaneous unfolding. This prevents any study of dynamical changes to protein's conformation over the simulation time using MARTINI. While G$\bar{o}$ model has been employed along with the MARTINI forcefield to study large-amplitude conformational changes, it can suffer from several shortcomings including loss of amino-acid identity, insensitive to point-mutations and environmental changes [225, 251–253].

In our previous publications, we had introduced an explict-solvent polarizable CG model for a selection of amino-acids to study secondary structure transitions starting from primary sequence in presence of different environmental stimulus such as hydrophobic media, interfaces and lipid bilayers [80–82, 85, 254]. In this work, we have formalized our approach to model parametrization, introduced new bonded and non-bonded potentials, updated residue geometry; and extended the representation to all natural amino acids to capture accurate protein tertiary structures.

Our Protein Model with Polarizability and Transferability (ProMPT) consists of two types of interaction sites (beads) — primary coarse grained beads and dummy beads. The basic biomolecule structure is created with primary interaction sites which feature modular architecture and geometry similar to the MARTINI model that allows for easy transferability; and are parametrized along the MARTINI scales [79, 255]. Through additional off-center dummy charges to the

primary sites that represent polarizable entities, we have introduced explicit local dipole moment, that can result in anisotropic coulombic interactions similar to hydrogen bonds in protein's secondary structure. The angle and dihedral potentials are derived from statistical distributions of these features from the protein data bank (PDB). ProMPT can capture secondary and tertiary conformational transitions, along with appropriate intermediate conformations and accurate folding free energy profiles. As such, this CG model can be applied to study spatiotemporally complex biological phenomena and processes involving proteins. In this paper, we present model validations using simulations of small protein structures in aqueous solvent, and aqueous protein aggregation. We have also discussed future applications, and current shortcomings of this CG potential.

6.3 Methods

The coarse grained interaction sites follow a basic *MARTINI-influenced* nomenclature, with atom-types grouped into polar, neutral, charged and hydrophobic beads [79,255]. The polar beads have explicit electrostatic dipoles added through charged dummy particles, constrained to the main interaction site (Fig. 6.2). For parametrization, we initially run a 50 ns atomistic simulations (using CHARMM36m/TIP3P force field [256, 257] and CHARMM-GUI [258] equilibration protocol) of relevant atoms that map to the particular polar CG interaction site tripeptides that have the corresponding mapped regions in chemical space. The parameters (charge on the beads - q and distance between the beads - r) are generated to match the dipole moment corresponding to the maximum in dipole moment distribution of the mapped region from the atomistic simulation. The charge and relevant geometry of these charged dummies are provided in Table 6.1.

The dummy charges interact with the local environmental through electrostatic interactions

Dummy Types	Charge	Bond-length	k_{angle}
Backbone	$\pm$ 0.340	0.14	7.2
SER-Sidechains	$\pm$ 0.144	0.14	7.2
THR-Sidechains	$\pm$ 0.153	0.14	7.2
GLN-Sidechains	$\pm$ 0.256	0.14	7.2
ASN-Sidechains	$\pm$ 0.256	0.14	7.2
TYR-Sidechains	$\pm$ 0.138	0.14	7.2
TRP-Sidechains	$\pm$ 0.136	0.14	7.2

Table 6.1: Charges and characteristic bonded potentials for dummy beads. Bond-length is the length of the tether from the primary interaction-center to the charged dummies. k_{angle} is the spring constant preventing deviation of the angle between charged dummies and the primary-interaction-center from 180 degrees.

(such as change in dielectric constant at membrane-water interface), which then directs protein's structural changes. As these charges are placed off-center to the main interaction site, they can introduce an asymmetry and directionality to interactions, and introduce spatial heterogeneity in local charge distribution. Here, a parallel can be drawn between the direct dipole-dipole and charge-dipole interactions in this CG model and electrostatic alignments in hydrogen bonds. Previous publications with a much simplified variant of this CG model, could generate appropriate sequence-specific secondary structures through alignment of these dipolar charges [80–82, 85, 254]. Several experiment and simulation-based studies have also underlined the importance of these molecular dipoles in protein folding.

ProMPT has been parametrized to work with the Yesylevsky *et al.*'s polarizable MARTINI water model [83]. This water model uses three coarse-grained (CG) interaction sites (beads) that map to four water molecules. One of the beads is the primary/central interaction site with the other two dummy interaction sites tethered to it [Fig. 6.1 - charges on dummies: +/- 0.46; Bond-length: 0.14 nm]. The authors could capture effective orientational polarizability of real water, that can modulate inhomogenous dielectric response. They applied the water model to study ion permeation in lipid membranes, and electroporation of lipid membranes and oil-slabs.

This model could naturally capture relative dielectric changes due to the local environment.

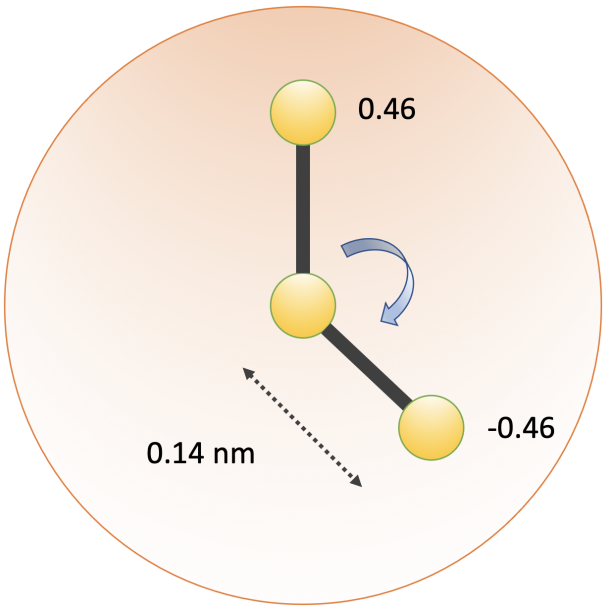

Figure 6.1: Schematic geometry of Martini polarizable water

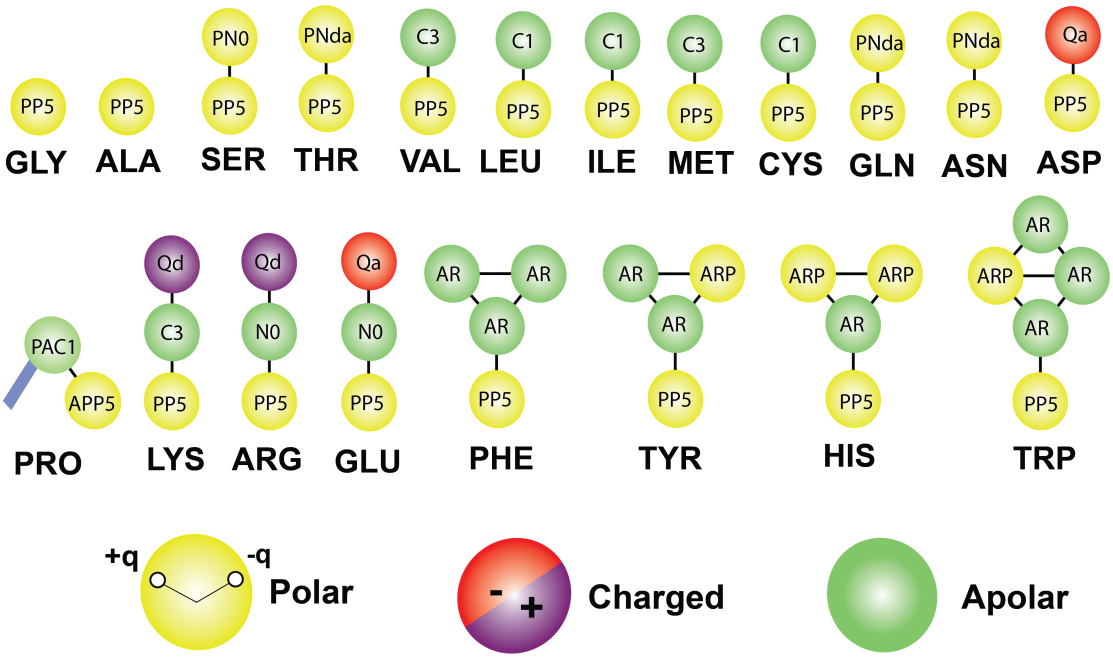

Figure 6.2: Schematic geometries of coarse-grained amino-acids

The amino-acid geometry defining the structures of the CG amino-acids are shown in

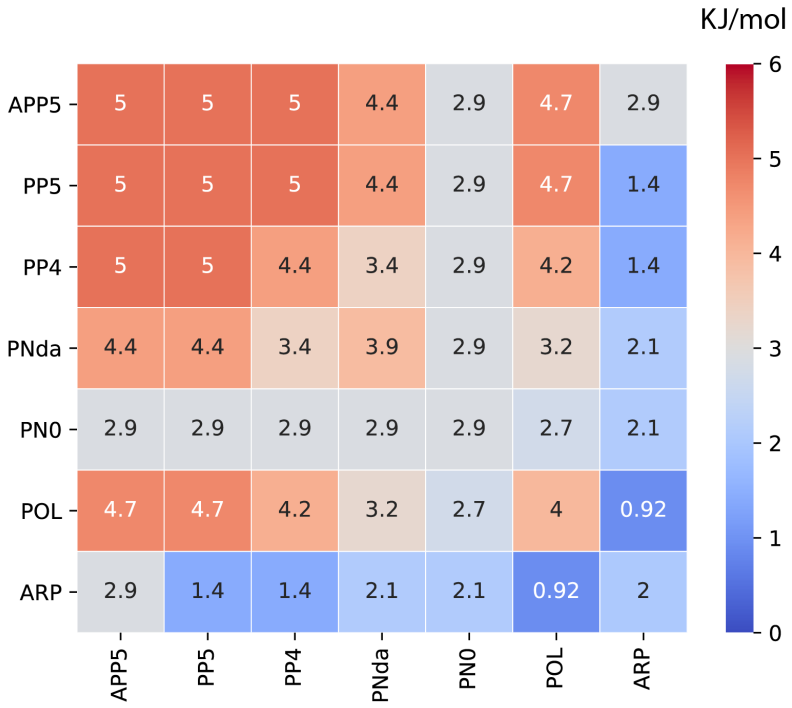

Figure 6.3: Non-Bonded Interactions [POL-POL]. Refer to Fig. 6.2 for naming nomenclature

Fig. 6.2. The interaction sites representing protein backbone and polar sidechains are assigned polarizable atom-types, with off-center dummy charges. The primary CG interaction sites that map to aromatic regions in the chemical space have a radius of 0.43 nm, compared to 0.47 nm for all other interaction sites.

6.3.1 Non-Bonded Interactions

All the non-bonded Lennard-Jones (LJ) type interactions between the atom-types are parametrized along the MARTINI interaction levels, which allows for easy transferability, and can therefore be used with all biomolecular environments that can be represented with the MARTINI forcefield. A complete description of the non-bonded interaction parameters is presented in through heatmaps (Fig. 6.3-6.8).

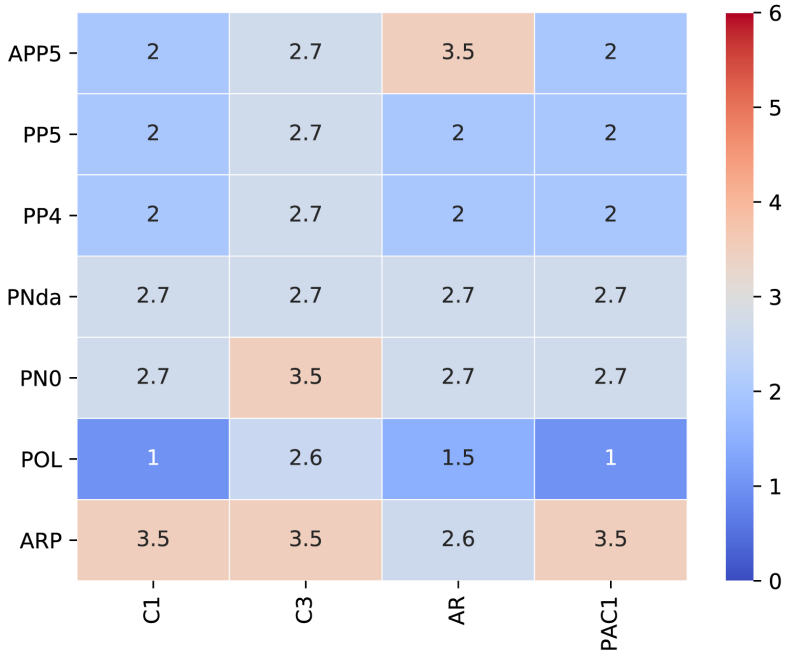

Figure 6.4: Non-Bonded Interactions [POL-HYD]. Refer to Fig. 6.2 for naming nomenclature

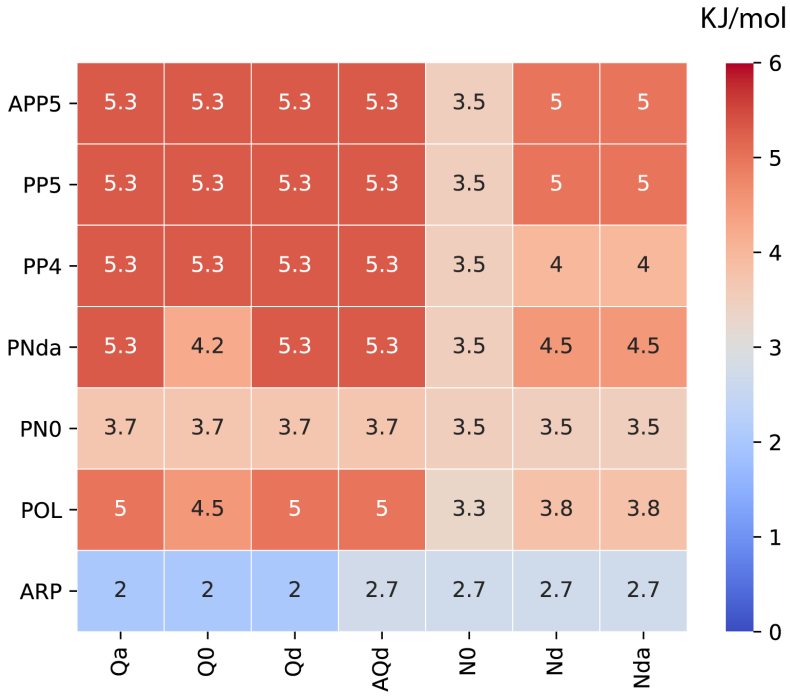

Figure 6.5: Non-Bonded Interactions [POL-Others]. Refer to Fig. 6.2 for naming nomenclature

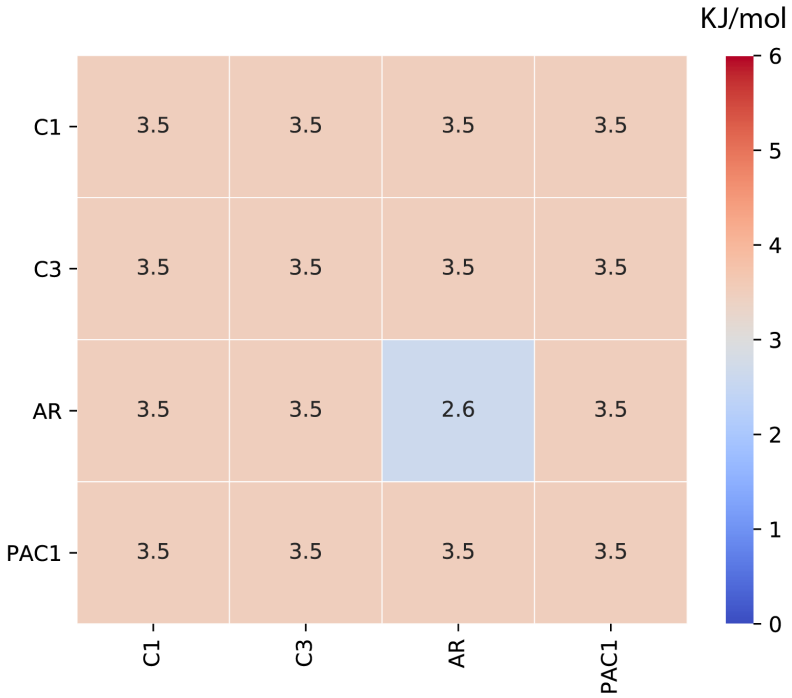

Figure 6.6: Non-Bonded Interactions [HYD-HYD]. Refer to Fig. 6.2 for naming nomenclature

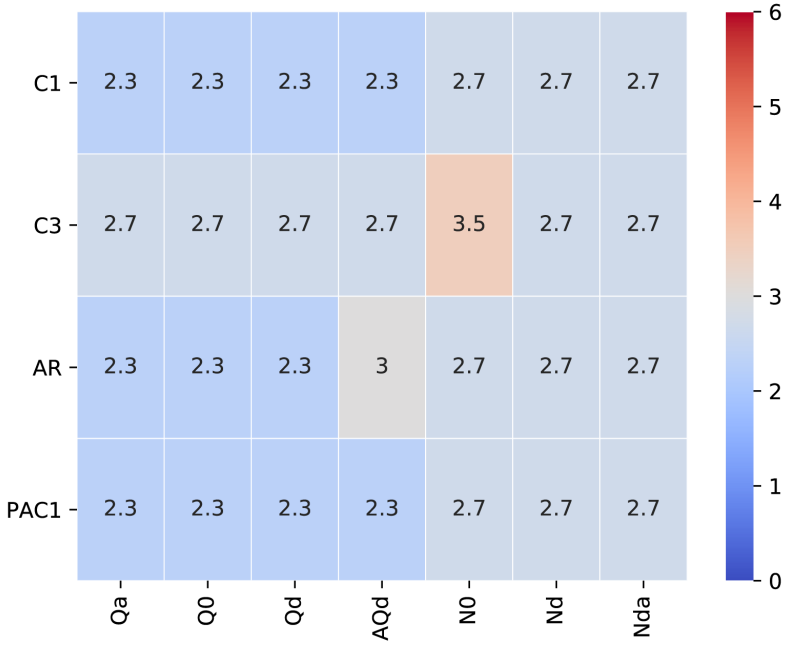

Figure 6.7: Non-Bonded Interactions [HYD-Other]. Refer to Fig. 6.2 for naming nomenclature

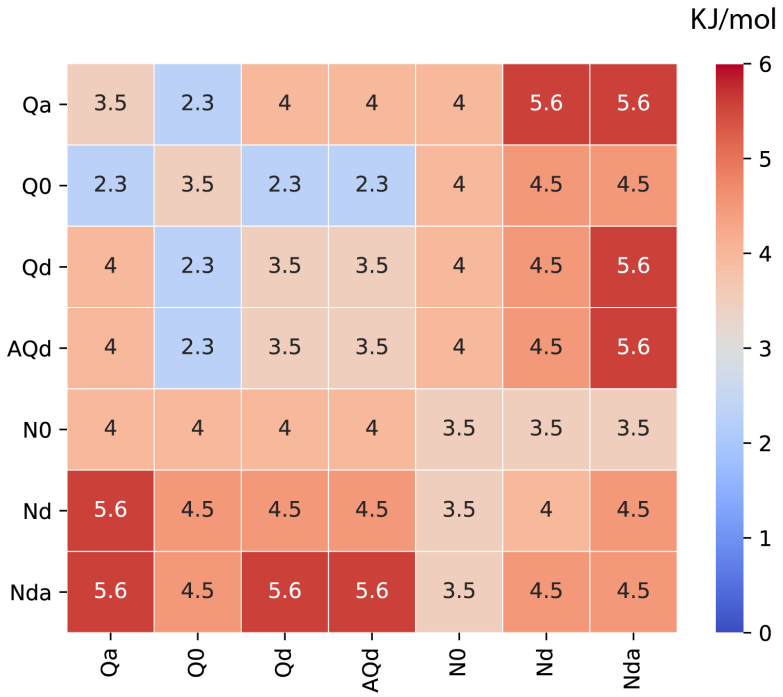

Figure 6.8: Non-Bonded Interactions [Other-Other]. Refer to Fig. 6.2 for naming nomenclature

While, most of the cross-interactions between non-polarizable interaction sites are directly borrowed from the MARTINI forcefield, the interactions between our polarizable groups had to be re-parametrized to balance out the added electrostatic interactions from the charged dummies. Similar to the MARTINI forcefield, the parametrization here aimed to fit non-bonded interaction parameters to reproduce accurate free energies of solvation and partitioning. The interaction level between hydrophobic groups and the interaction sites representing water is reduced by upto 50% from the MARTINI level to reproduce appropriate environment-induced structural transitions. These reductions balance out issues of over-polarization by backbone dipoles and allow conformational switching from helices to β-strands. We refer readers to the previous publication by Ganesan *et al.* and Sahoo *et al.* regarding further details about non-bonded parametrization [81, 82]. The electrostatic charges on specific dummy beads are enumerated

in the supplementary.

A set of special interactions were added in an *ad-hoc* manner to capture specific protein-protein interactions. We applied specific attraction (ϵ = 3.0 kJ/mol) between positively charged groups (such as ions and cationic amino acid sidechains) and the aromatic rings to mimic cation-π effects [259]. Similarly, interaction between aromatic rings, and proline sidechains with aromatic rings were made attractive (ϵ = 3.0 kJ/mol) to capture π-π stacking and CH-π interactions that have been highlighted in quantum mechanical calculations [260–262]. These electronic effects are ubiquitous in protein folding. Additional dummy-dummy and dummy-charged group hard-core repulsion ($c_{12} \sim 10^{-7}$ nm) were added to prevent over-interaction between CG groups similar to methods to prevent *polarization catastrophe* in polarizable forcefields [263, 264]. This allows for fast switching between conformations through binding/unbinding among charged dummies.

6.3.2 Bonded Interactions

Bonded interactions can be broadly categorized as bonds, angles and dihedrals. The features describing bonds in ProMPT — bond lengths and bond rigidity were borrowed directly from the MARTINI forcefield. The corresponding parameters are enumerated below (Table 6.2-6.6). Some angular (between backbone beads and first side-chain) and dihedral potentials (backbone only) were informed from a statistical distributions of protein structures.

Bond-Types	Bond-Length	Bond-Rigidity/Constraints
A. Between backbone interaction-centers	0.385	7500
B. Between BB-S1 of non-aromatic amino-acids	0.250	5000
C. Between S1-S2 of non-aromatic amino-acids	0.280	5000
D. Bonds in planar rings of aromatic residues	0.270	Constraints

Table 6.2: Bonded interaction potentials between different primary CG interaction-centers. BB: Backbone; S1: First Sidechain; S2: Second Sidechain

Type	Bond-length
Backbone (BB) - Backbone (BB)	0.385 nm
Backbone (BB) - Sidechain1 (S1)	0.28 nm
Sidechain1 (S1) - Sidechain2 (S2)	0.25 nm
Between aromatic rings	0.27 nm

Table 6.3: Bond lengths between primary coarse grained interaction sites.

Angle-Types	Angle (degrees)	Angular-Rigidity
A. Between backbone interaction-centers	109	75
B. Between BB-S1-S2 of non-aromatic amino-acids	151	25
C. Between BB-S1-S2/S3 for F Y	150	50
D. Between BB-S1-S2 for W	210	50
E. Between BB-S1-S3 for W	90	50

Table 6.4: Angular interaction potentials between different primary CG interaction-centers. BB: Backbone; S1: First Sidechain; S2/S3/S4: Second/Third/Fourth Sidechain

We used a non-redundancy p-value of 10^{-7} to create a database of about 14000 protein structures from the protein data bank. The angular and dihedral potentials were based on their corresponding normalized distributions from this database. The angle between the protein backbone sites was universally set to 109^o, restrained through a harmonic potential. The angular potential between the backbone and the sidechain was created specific to each amino-acid by applying Boltzmann inversion at 300K to each amino-acid specific distributions. These potential energy functions corresponding to each amino-acid have been reported in Fig. 6.9-6.11. Finally, the dihedral potentials between the backbone interaction-centers are secondary-structure specific, with different tabulated potentials for α-helix, 3-10 helix and β-sheets. The functional forms

Angle-Types	Dihedral (degrees)	Dihedral-Rigidity
A. Between backbone interaction-centers	Tabulated	NA
B. Between BB-S2-S3-S1 for F Y	Improper; 0	50
C. Between BB-S2-S3-S1 for W	Improper; 0	50
D. Between S1-S2-S4-S3 for W	Improper; 0	200

Table 6.5: Dihedral interaction potentials between different primary CG interaction-centers. BB: Backbone; S1: First Sidechain; S2/S3/S4: Second/Third/Fourth Sidechain

Secondary-Structure	A	B	C	D	ϕ_0
α-helix	1.799	1.969	-1.199	-1.03	1.227
3-10 helix	1.799	1.969	-1.199	-1.03	1.969

Table 6.6: Dihedral potential parameters for α-helix and β-sheet corresponding to equation 1 in the main text.

of α-helix and 3-10 helix potentials were taken from previous publications of Cα based coarse grained models [265].

$$
\begin{aligned}
V(\phi; A, B, C, D) & \\
&= A\left[1 + cos(\phi + \phi_0)\right] \\
&+ B\left[1 + cos(\phi - \phi_0)\right] \\
&+ C\left[1 + cos(3(\phi + \phi_0))\right] \\
&+ D\left[1 + cos(\phi + \phi_0 + \frac{\pi}{4})\right]
\end{aligned}
\tag{6.1}
$$

The fitted parameters are provided in Table 6.6. These periodic functions allow for introducing multiple local/global minimums, in contrast to single and deep global minima by Boltzmann inversion. We parametrized these functions to set the global functional minimum to match the maximum value in the dihedral probability distributions of relevant secondary structures, in addition to introducing several local minimas. These local extremas introduce additional frustrations to the protein's conformations and has been adopted in previous coarse grained models [265]. The tabulated β sheet potential is generated through a direct Boltzmann inversion (at 300K) of dihedral distributions from regions specific to beta sheets in the protein structure database. Fig. 6.12 shows the dihedral potentials employed in ProMPT.

6.3.3 Simulation setup

All simulations are performed with GROMACS 2019.4 [177]. The starting conformation and the protein topology are generated with in-house codes, taking the amino acid sequence

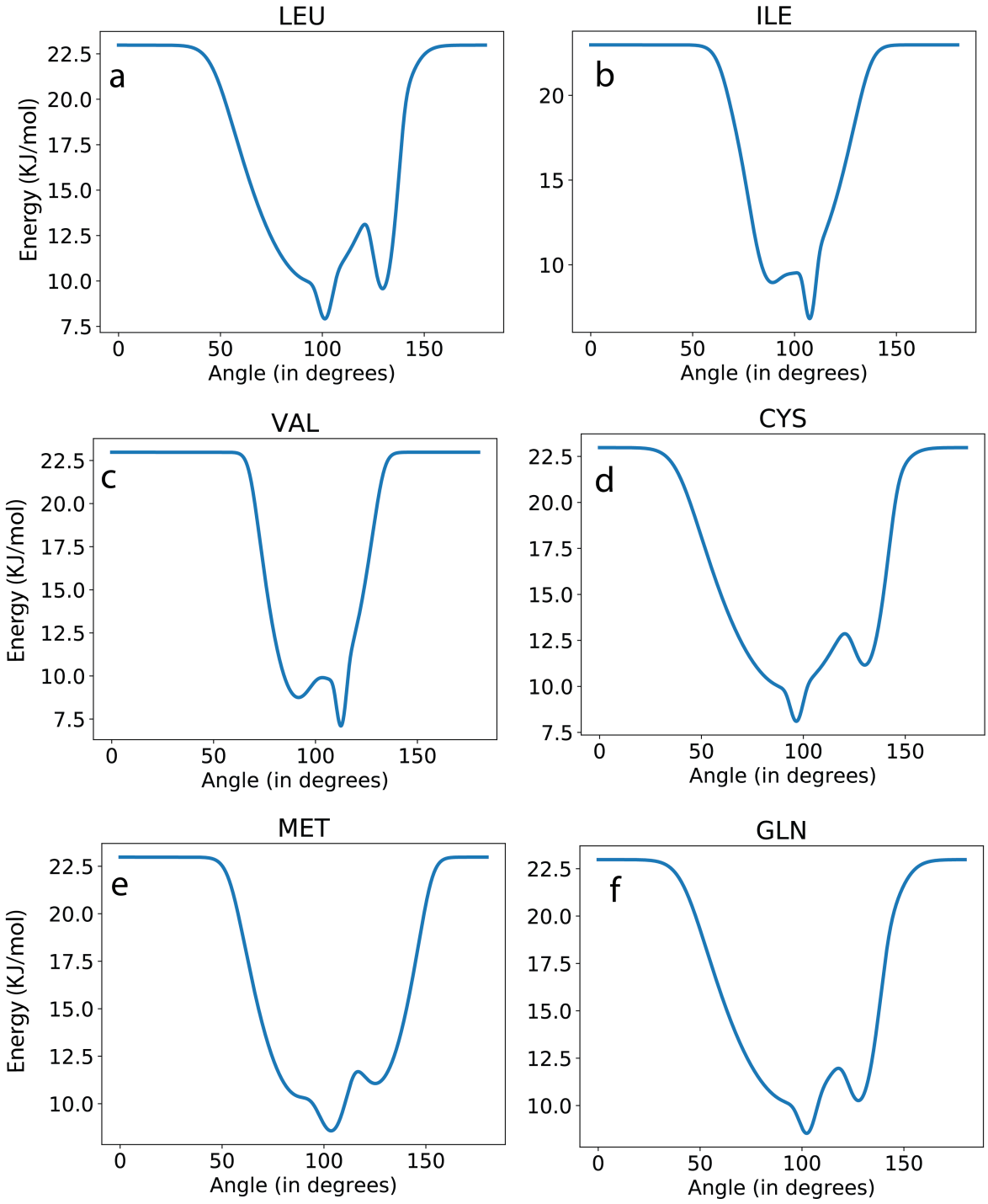

Figure 6.9: BB(Previous amino acid)-BB-S1 tabulated angular potentials (Set 1)

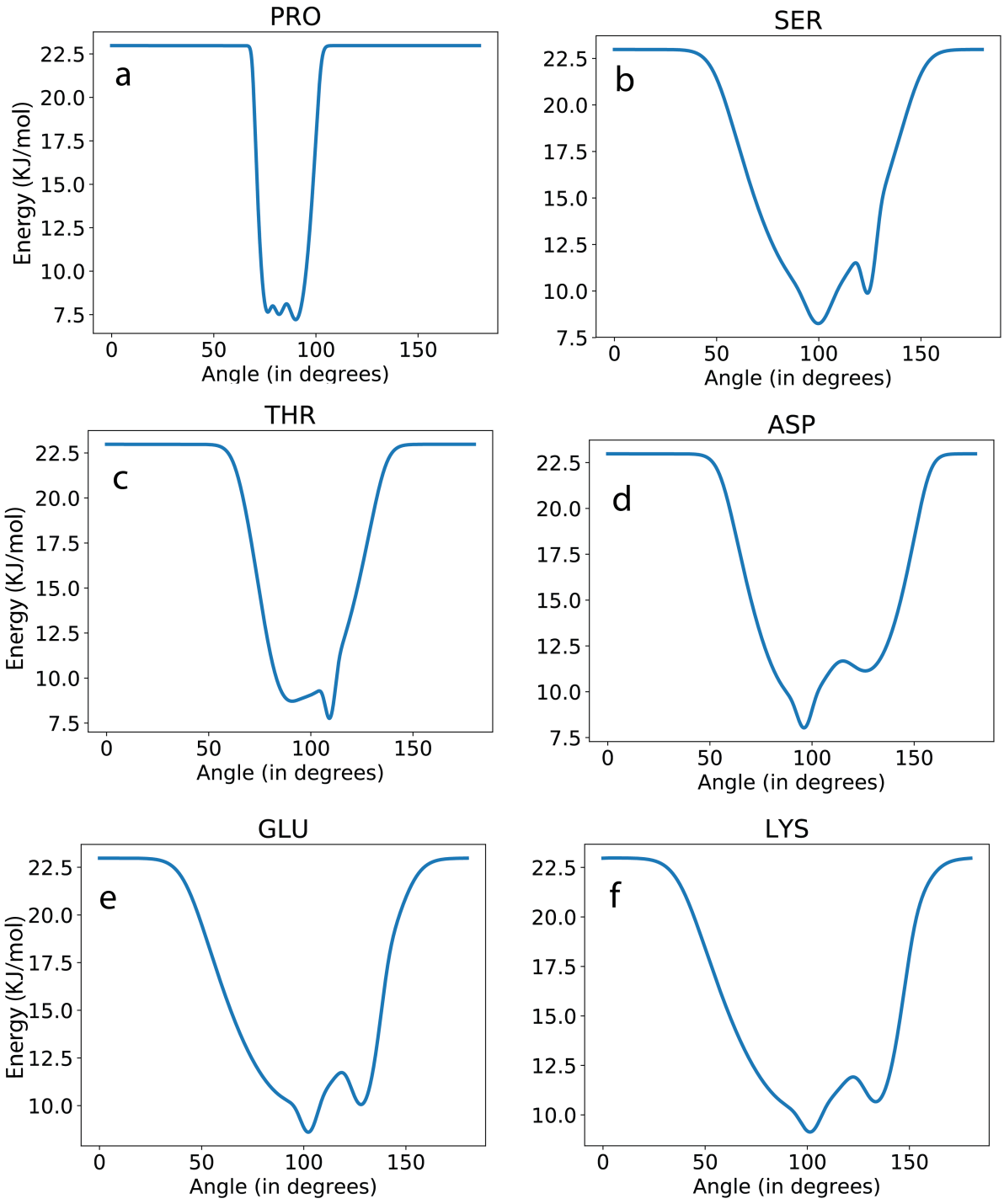

Figure 6.10: BB(Previous amino acid)-BB-S1 tabulated angular potentials (Set 2)

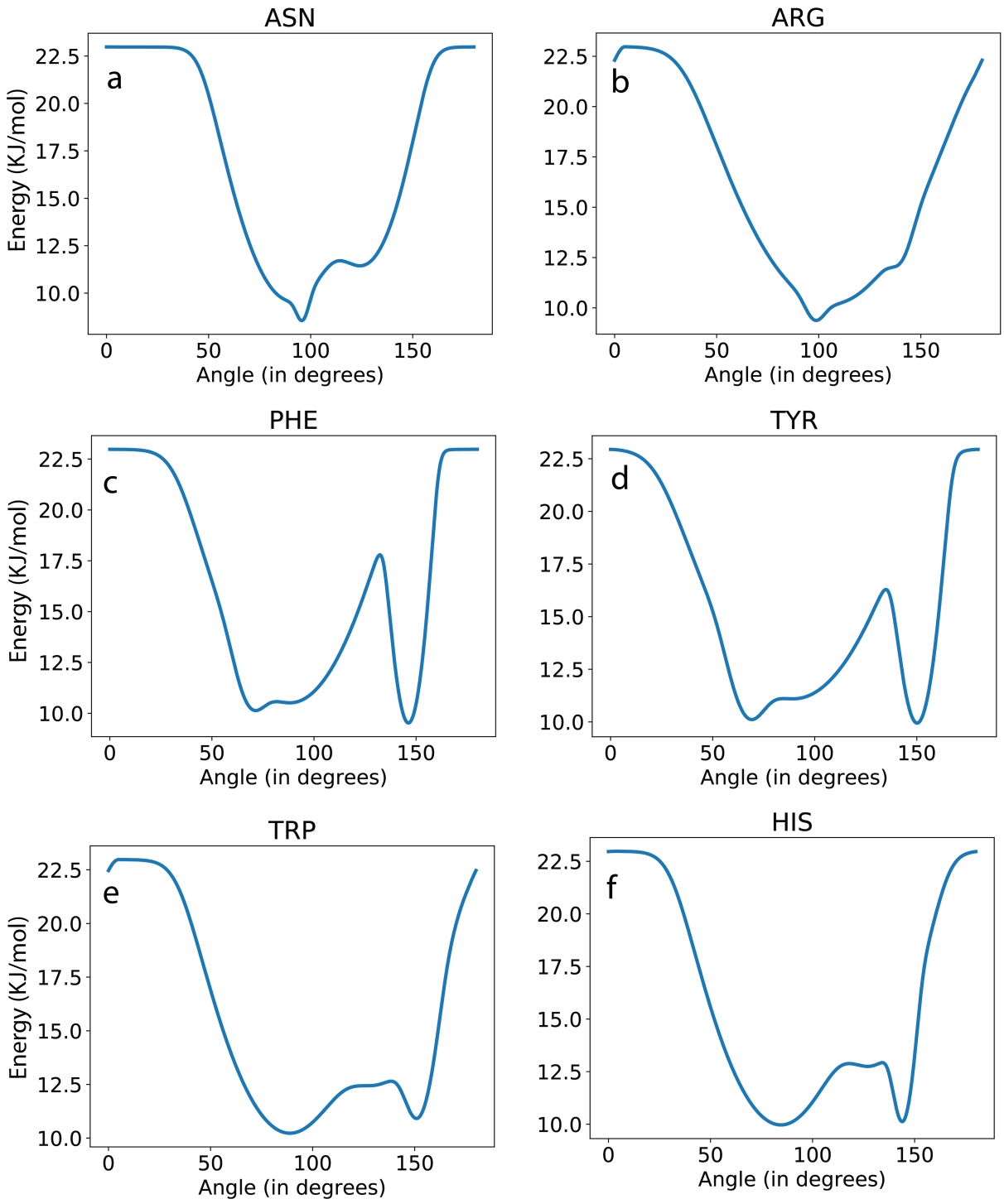

Figure 6.11: BB(Previous amino acid)-BB-S1 tabulated angular potentials (Set 3)

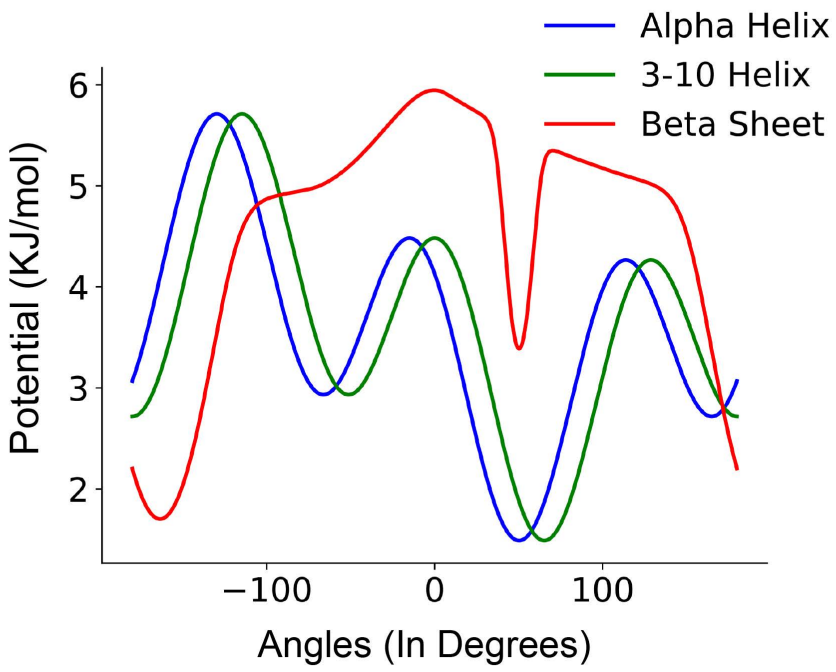

Figure 6.12: Secondary-structure specific dihedral potential used in the CG forcefield. The tabulated potentials are fitted (for α-helix and 3-10 helix) or derieved (β-sheet) to capture maximum value in the dihedral probability distributions

from the reference PDB structures listed in Table 6.7. The initial conformation for all proteins is a random unfolded extended conformation. The setup for each protein system including the number of protein beads, the number of polarizable water molecules, and temperature details are also listed in Table 6.7. Adequate number of ions are added to neutralize the system. We have used reduced temperature, $T^* = K_b T/\epsilon$ as our temperature scale, where ϵ is the highest interaction strength (5.6 kJ/mol) in our CG potential. Energy minimization is first conducted with steepest descent. An NPT equilibration run at T^*=0.52 for 50 ps is followed at 1 bar and with time step 0.001 ps. The canonical production run is then performed with time step 0.01 ps for the simulation time listed in Table 6.7 at different reduced temperatures. Since the box size is

constant, the solvent density remains the same across various temperatures. Electrostatic cutoff is 1.6 nm and particle-mesh Ewald (PME) method is applied for long range electrostatic interactions [151]. Nos$\tilde{e}$-Hoover thermostat is used to maintain the system at the desired temperature [266].

	Trp-cage	Trpzip4	Villin	WW-Domain	β-α-β	Aβ16-22
PDB code	1L2Y	1LE3	1YRF	1E0L	2KI0	-[a]
Number of proteins	1	1	1	1	1	8
Number of water	1957	1610	26391	30136	24998	2584
Simulation time per temperature (ns)	200	200	600	100	200	300
Temperature range performed	0.52-1.04	0.52-1.04	0.52-1.04	0.52,0.82	0.52,0.82	0.52
Total number of temperature simulated	15	15	19	2	2	1

Table 6.7: Simulation setup for each protein. Sequence: KLVFFAE.

To validate the use of ProMPT in heterogenous environmental conditions, we studied the dimerization of Glycophorin A (GpA) monomers starting from their solvated-unfolded state. Two CG GpA monomers are put into an 8 nm wide cubic box with a distance of 5.66 nm between the two monomers pointing to the same direction. The box is then solvated with 80 DPC detergents and 3000 CG water molecules. Energy minimization is first conducted with steepest descent. An NPT equilibration run at 350K for 5 ns is followed at 1 bar and with a time step of 0.001 ps. The compressibility is 3.50×10^5 bar^{-1}. The production run in an NPT ensemble is then performed with a time step of 0.01 ps for the simulation time of 500 ns at 350K. Electrostatic cutoff is 1.6 nm and particle-mesh Ewald (PME) method is applied for the long range electrostatic interactions [151]. Nos$\tilde{e}$-Hoover thermostat is used to maintain the system at the desired temperature [266]. Six replicas are run for each system.

6.3.4 Comparision with Atomistic Simulations - Replica Exchange with Solute Tempering

The replica exchange with solute tempering (REST) [267] simulations are performed for Trp-cage and Trpzip4 in water and the Amber99sb [268] atomistic force field is used. For Trp-cage, an unfolded coil is solvated with 2678 water molecules and neutralized with one chloride ion. The system then goes through energy minimization, 200 ps of NVT equilibration and 200 ps of NPT equilibration. After equilibration, the REST protocol is used for the production runs where 10 replicas are simulated spanning a temperature range from 290K to 540K. Each replica is run for 700 ns and the resulting exchange rates are between 22% to 36%. For Trpzip4, an unfolded coil is solvated with 3570 water molecules and neutralized with 3 sodium ions. The system undergoes the same equilibration process and the same REST protocol as those for Trp-cage. For Trpzip4, a longer simulation time is needed for convergence, each replica is run for 1400 ns and the resulting exchange rates are between 24% to 32%.

6.3.5 Analysis

We used potential of mean force (PMF) plots of Trp-cage, Trpzip4, and villin to validate the ability of our CG force field to capture a protein's conformational landscape. Backbone native contact and root-mean-squared-deviation (RMSD) of peptide fragments were used as reaction coordinates for PMF calculations. Backbone native contact measures the fraction of backbone contacts with a cutoff of 7Å that are replicated in our CG trajectories that are native to the PDB-converted-CG structure. RMSD backbone (BB) is a measure of relative deviation of our CG

backbone from the backbone beads in PDB-converted-CG structure. For all proteins, the first 80ns of trajectory-data is removed and the equilibrated data are collected at all temperatures simulated and re-weighted through Multistate Bennett Acceptance Ratio (MBAR) method [269]. Here, we present the final PMF at T^*=0.52. For villin, we used a cluster analysis tool from GROMACS with a RMSD cutoff of 0.3 nm to generate representative conformation corresponding to each basin in the PMF. For the aggregation simulations, two peptides are considered part of β-sheet if more than three dummy-dummy interaction pairs are formed. A fraction of one means that all the peptides (8 of them) are forming β-sheet.

In this study we used potential of mean force to capture the folding and dimerization landscape of our Glycophorin A. Here, we only used the trajectory after at least one CG interaction site contact (cutoff=6$\mathring{A}$). We computed helical content per monomer as number of native backbone contacts (cutoff=6$\mathring{A}$) compared to the PDB-converted-CG structure for the same sequence of the wide type. We used average helical content and the number of backbone (BB) contacts are used as the reaction coordinates for the PMF calculations.

6.4 Results and discussion

Mini proteins Trp-cage and Trpzip4 are selected to be our first test proteins because they are already well-studied both computationally and experimentally [270–274]. Trp-cage has an α-helix starting from the N-terminus and a 3-10 helix at the middle part, and an unstructured C-terminal tail, which collapse together forming a hydrophobic core (Fig. 6.13c). Trpzip4 has a β-hairpin structure with four Trp residues packing together. Fig. 6.13a shows the PMF for Trp-cage using native contact and RMSD helix BB as two reaction coordinates. A two-state

folding/unfolding path can be observed from the PMF, which agrees with experiments and REMD simulation results [270, 275]. The folded basin at around 0.1 nm RMSD helix BB indicates a perfect α-helix. The distribution of native contact is relatively broader, from 0.6 to 0.7. The unfolded basin has a a native contact of 0.4 to 0.5 and RMSD helix BB from 0.2 to 0.3 nm, indicating some partial helicity in agreement with previous simulation reports [276]. The PDB structure and a representative folded structure with our CG model for Trp-cage are shown in Fig. 6.13c. Both α-helix and 3-10 helix can be well reproduced at the correct positions. The hydrophobic core formed by Tyr3, Trp6, and three proline residues at the C-terminus is also well captured as seen in the experimental structure. Our estimation of ΔG for protein folding is 3.2 kJ/mol, close to the experimental estimate of 2.6 kJ/mol [277]. We also perform a set of replica exchange with solute tempering (REST) simulations to estimate the folding-unfolding free energy difference for Trp-cage, from the atomistic simulations, the ΔG is estimated to be around 2.81 kJ/mol at 290K (Fig. 6.14).

Fig. 6.13b shows the PMF for Trpzip4 with native contact and RMSD BB as the reaction coordinates. The PMF of Trpzip4 also shows a two-state folding/unfolding path, which agrees with experiments and atomistic simulations. The folded basin for Trpzip4 is located at 0.35 nm RMSD BB with a native contact of around 0.8. The unfolded basin has a broad RMSD BB distribution from 0.5 to 0.7 nm with a native contact of around 0.4. Fig. 6.13d shows both the PDB structure (left) and a representative folded state for Trpzip4 with our model (right). The β-sheet content can be well captured. Four tryptophan residues are interacting with each other and forming a hydrophobic core as observed in the experimental structure. Here, our CG model estimated a ΔG of about 7.3 kJ/mol for protein folding, compared to 12.3 kJ/mol reported in experiments [278]. From the REST simulations we run, a ΔG of 4.53 kJ/mol is estimated (Fig.

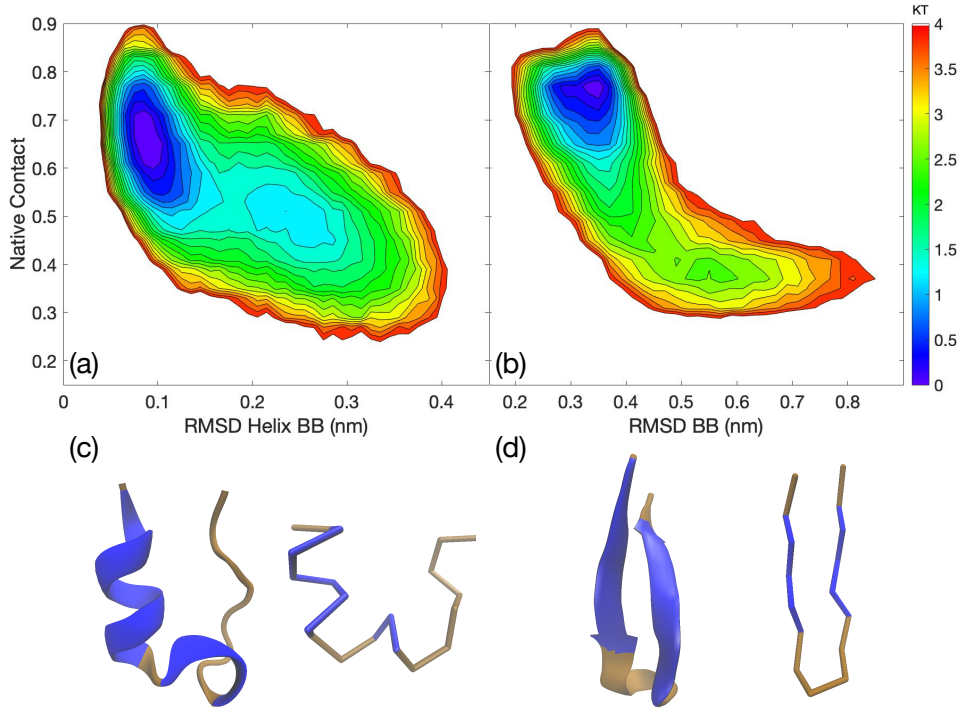

Figure 6.13: PMF for Trp-cage (a) with RMSD helix BB and native contact as the reaction coordinates. PMF for Trpzip4 (b) with RMSD BB and native contact as the reaction coordinates. Both PDB structure (left) and the representative structure (right) from our model for Trp-cage (c) and Trpzip4 (d) are shown. Blue indicates the specific secondary structure each protein exhibits.

6.15), which is lower than both the experimental data and the estimation from our CG model.

Our next target protein - villin has three helices with a hydrophobic core. Successful folding of villin validates our CG protein model for general applications to proteins with tertiary packing of homogeneous secondary structures. Fig. 6.16 shows the PMF for villin with RMSD strand 1 (S1) and RMSD strand 2 (S2) as two reaction coordinates, where the locations of the two strands are highlighted in red. Strand 1 contains two helices from the N-terminus (residue 2-18), while strand 2 only contains one long helix (residue 21-32) from the C-terminus. There are one native state and two intermediate states populated at T^*=0.52. At higher temperatures, the location of the unfolded basin shifts to more extended conformations. All of the states populated at T^*=0.52 have similar RMSD S1 values, but different RMSD S2 values, which correspond to different tertiary structures. Indeed, it has been reported in experiments that although there are fluctuations in the N-terminal part of villin, large-scale unfolding will not occur unless undergoing global unfolding [279]. In our model, global unfolding happens at higher temperatures, where disruption of strand 1. The native state is centered at RMSD S1 0.35 nm and RMSD S2 0.1 nm. The representive structure of the folded basin with our model is very similar to the experimental structure from PDB. For the intermediate states, one is centered at RMSD S1 0.35 nm and RMSD S2 0.2 nm while the other one is centered at RMSD S1 0.35 nm and RMSD S2 0.35 nm. Both of the intermediate basins show a decrease in helicity for strand 2 and losing of hydrophobic core, while the helicity for the basin with larger RMSD S2 is reduced further more. The intermediate states are similar to the states described in the experimental work from Reiner and co-workers, where the intermediate states have an unfolded strand 2 that is undocked from the hydrophobic core [279]. Both the native state and intermediate states are on the folded side of the folding-unfolding landscape.

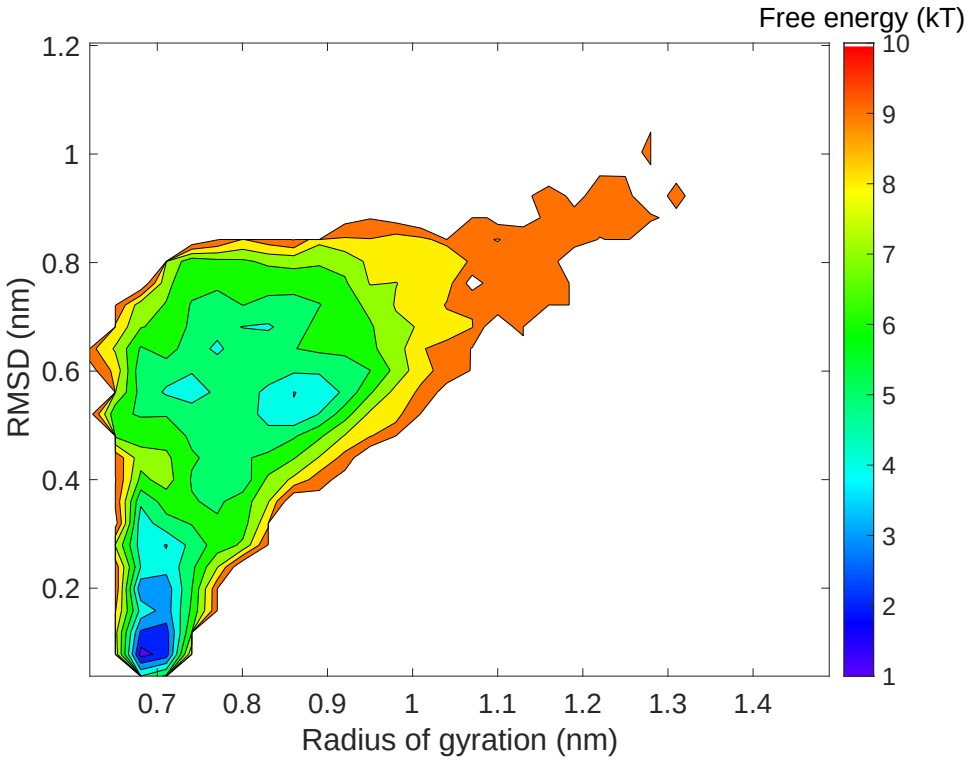

Figure 6.14: Free energy plot for Trp-cage at 290K from the REST simulations. The folded basin has a free energy of 2.82 kT and the unfolded basin has a free energy of 3.98 kT. The ΔG is estimated to be 1.17 kT (2.81 kJ/mol).

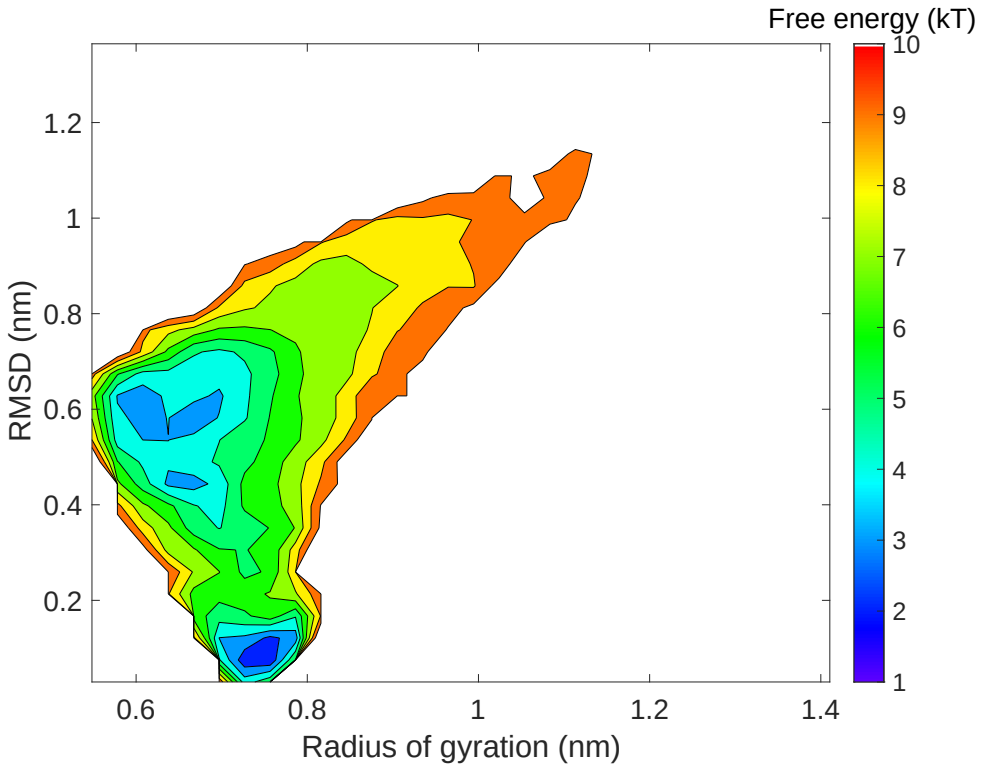

Figure 6.15: Free energy plot for Trpzip4 at 290K from the REST simulations. The folded basin has a free energy of 2.12 kT and the unfolded basin has a free energy of 4.00 kT. The ΔG is estimated to be 1.88 kT (4.53 kJ/mol).

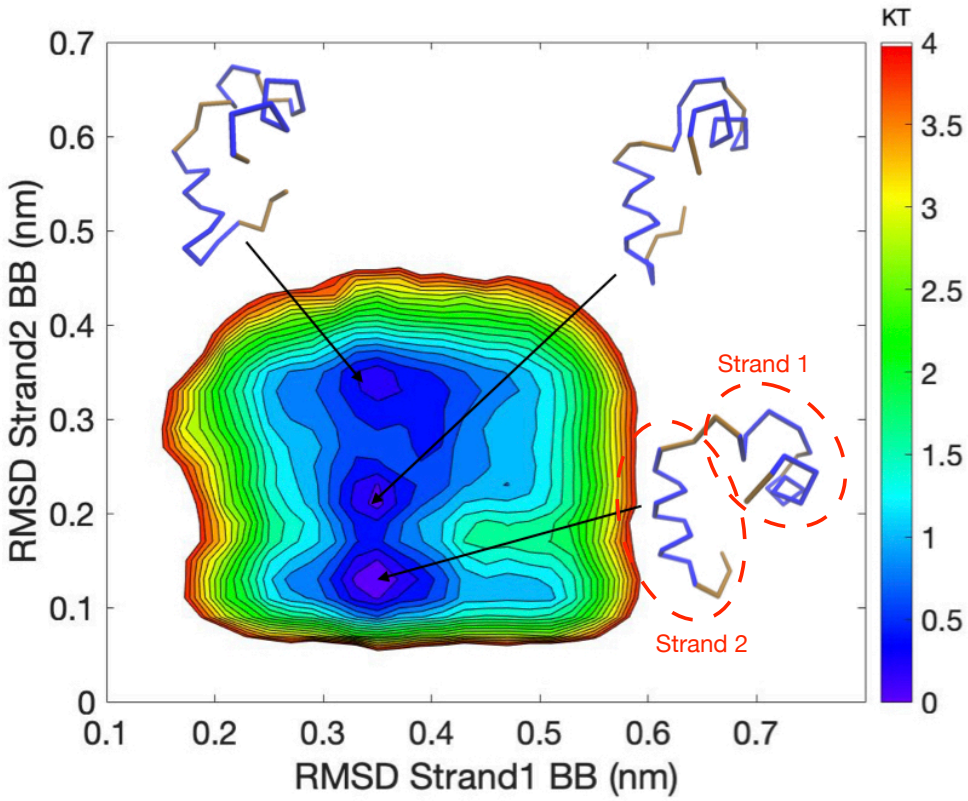

Figure 6.16: PMF for villin with RMSD S1 and RMSD S2 as the reaction coordinates at T^*=0.52. The representative structure for each basin is shown as insets.

For the proteins having a tertiary structure with homogeneous β-sheets, we used WW-domain as our benchmark protein. The PDB structure is shown in Fig. 6.17c with three antiparallel β-strands and two long tails at both ends. A representative folded structure with our model is shown in Fig. 6.17c. In Fig. 6.17a, the RMSD BB time series without considering the loop regions are shown at two temperatures. From the time series at T^*=0.52 we observe that the simulation converges at a value of RMSD BB around 0.3 nm. The simulation can reach the folded state in about 20 ns. We could note frequent transitions between folded and unfolded states at T^*=0.82. The last protein we focus on β-α-β (PDB code: 2KI0), which is a designed protein that has a mixture of secondary structures (partially α-helix and β-sheet) with a defined tertiary structure. Using DSSP [280] we assigned the secondary structures using one of the deposited NMR structures in the PDB database. The antiparallel β-strands only have four residues on each strand. Fig. 6.17d shows a representative folded structure with our model, the α-helix can be well formed, while the β-sheets are formed but shifted. Fig. 6.17b shows the RMSD BB time series for β-α-β at two temperatures, without considering the tails. The RMSD BB converges at around 0.5 nm at T^*=0.52. The time series at a higher temperature (T^*=0.82) is shown in red, which confirms the ability of ProMPT to capture frequent transitions between different states.

We have demonstrated that with our model we can fold a variety of proteins, which include pure β-sheets, pure helices, and proteins that have a defined tertiary structure composed with homogeneous secondary structures. For proteins that have mixed secondary structures such as β-α-β, the packing could not be perfectly reproduced. We also attempted simulating the folding of another well-studied protein with heterogeneous secondary structures - GB1 (PDB code: 2J52.pdb). GB1 is a slightly larger protein with four β-strands and one α-helix. Although our CG model could generate accurate secondary structure at appropriate regions, the overall tertiary

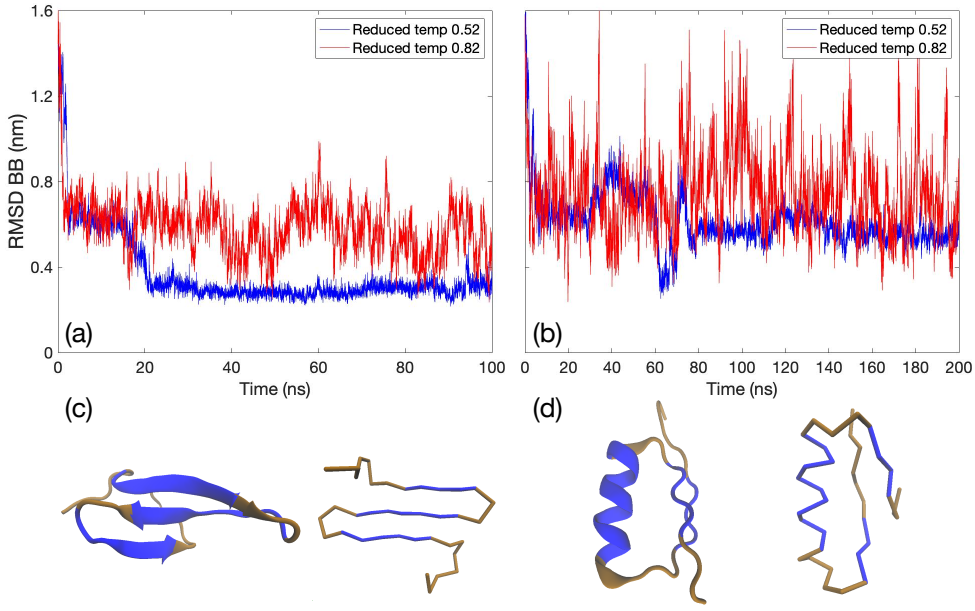

Figure 6.17: RMSD BB time series for (a) WW-domain and (b) β-α-β at T*=0.52 (blue) and T*=0.82 (red). The PDB structure (left) and the representative structure (right) are shown in (c) and (d) for WW-domain and β-α-β, respectively. Blue indicates the specific secondary structure each protein exhibits.

packing was not reproduced. We compared the molecular packing of amino-acid side-chains in ProMPT for GB1 to GB1 PDB native structure mapped into our CG description. ProMPT simulations, with backbone restrained to the native state, resulted in a protein core with packing defects and voids (Fig. 6.18). On the other hand, CG structure mapped directly from PDB native state showed no such anomalies. Therefore, we postulate that appropriate modifications to side-chain size, can achieve adequate protein-core packing required for correct tertiary structures.

Proteins with mixed folds will be our next target proteins and could potentially be solved by adjusting the size and shape of each amino acid in the CG model to better mimic the side-chain geometry. Recent parametrization of MARTINI 3 force field has adopted this strategy and included several smaller bead-types to capture the accurate geometrical shapes of biomolecules [244]. This resulted in significant improvement in molecular packing, and could capture correct

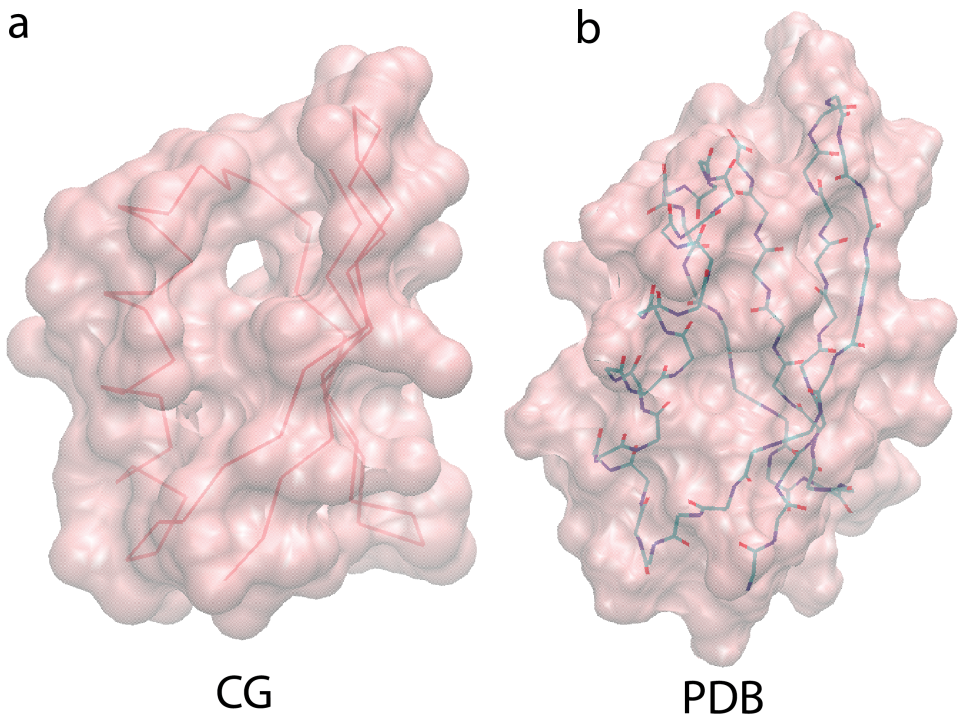

Figure 6.18: Surface Plot of GB1 with a solvent probe of 1.4 $\mathring{A}$. a- CG after 5 ns of simulation starting from the folded state, b- PDB. The backbone is traced to show the tertiary packing of the protein.

open and close conformations in transmembrane proteins.

Beyond studying protein folding, ProMPT can also be applied to study peptide aggregation. Here we demonstrate the aggregation of Aβ 16-22 peptides in water at T*=0.52. The protein concentration in the system is 0.0426 M. Fig. 6.19 shows the time series for β-sheet fraction. The β-sheet fraction fluctuates between 60% to 80%, indicating significant fibrillation. Occasionally the β-sheet fraction can reach 100%, where all the Aβ peptides are involved in the β-sheet formation. Similar to previous experimental and simulation studies, peptides organize into stable, layered β-sheet rich structures [189,281–284]. PHE residues are shown in yellow and a hydrophobic core formed by PHE packing can be observed from the aggregate snapshot (inset of Fig. 6.19). We observed anti-parallel β-sheets in our simulations, similar to the structures reported by experiments [285]. Nguyen *et al.* applied replica exchange MD on Aβ 16-22 monomer, dimer, and trimer to study the effects of different atomistic force fields on the formation of β-sheets [286]. While AMBER99 and OPLS generated a diversity of conformations, only GROMOS96 could successfully generate antiparallel β-sheets. With ProMPT, we observe aggregation into β-sheet rich, fibril-like structures by 30 ns and can also qualitatively study the kinetics of this aggregation process.

6.4.1 Simulations of Glycophorin-A and Mutants

Fig. 6.20a shows the PMF of wild type GpA, where the most populated basin is observed with a high average helical content and a broad range of number of BB contacts. A representative snapshot of GPA dimerization is also presented. The WT-GPA dimerizes as a parallel dimer. In addition, we could also reproduce Cα-H contacts prevalent in the G_{79}xxxG_{83} motif, along with T87-T87 contact that is noted in previous research [287].

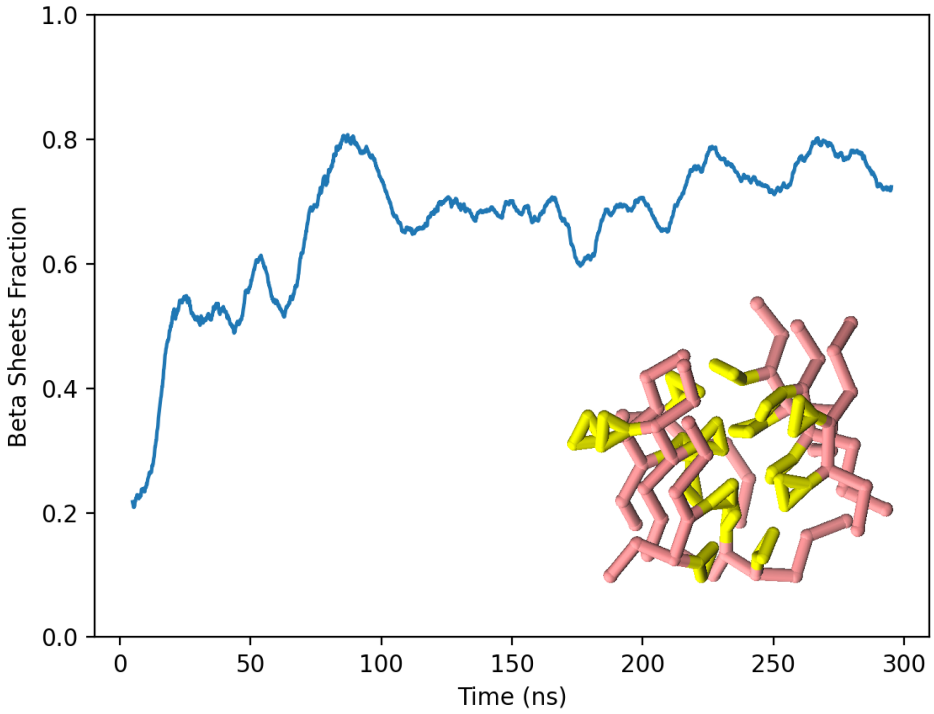

Figure 6.19: Time series for the number of Aβ 16-22 peptide forming β-sheets. An illustration of β-sheets aggregation is shown in the inset with yellow representing PHE residues.

The helical crossing angle is close to PDB deposited structures (marked by black horizantal lines), with a strong peak between -50 and 20 degrees for WT-GpA [Fig. 6.21]. Here, we also noted some anti-parallel orientations, which can be because of the less restrictive media of DPC micelles compared to the membrane. Anti-parallel dimers have been reported before in atomistic simulations in both implicit membrane and implicit cyclohexane [288].

The contact map shown in Fig. 6.22 shows that the majority of the wild type GpA dimers are parallel, which agrees with the crossing angle result shown earlier. In addition, from the residue-residue contacts it can be seen that the dimers are highly symmetric and specific. This specific contact surface is also observed in CG simulations in membrane where a mobile helix is found to be anisotropically distributed near the fixed helix [289] that is originated from the close contacts of the GxxxG motif.

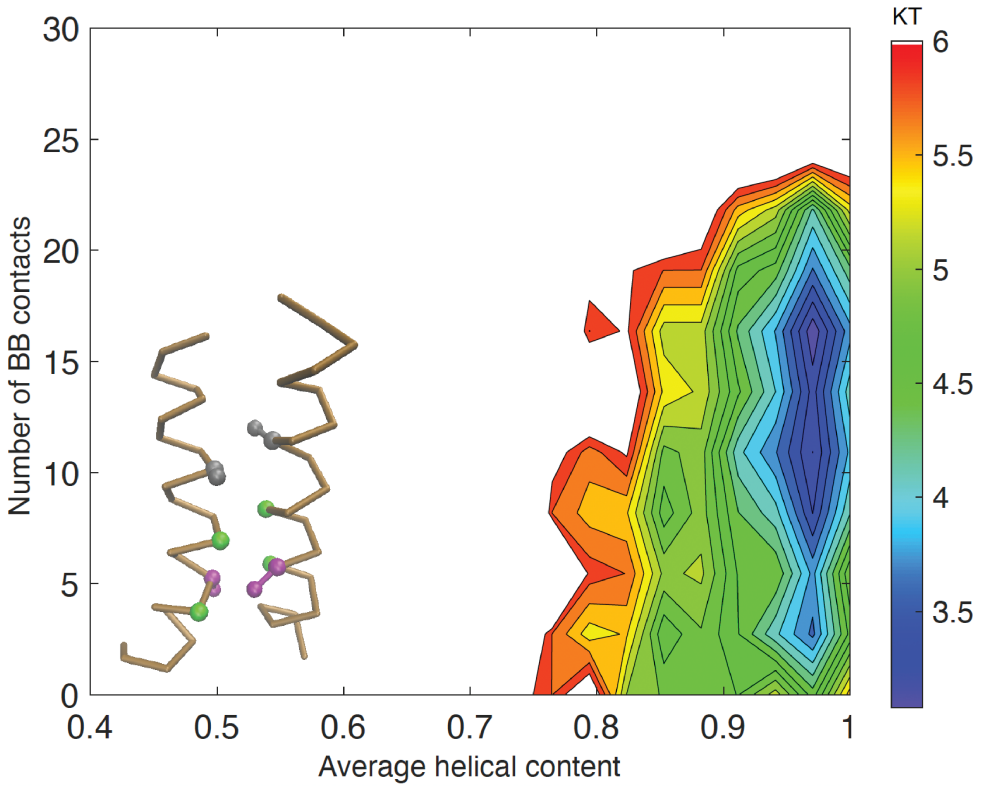

Figure 6.20: PMF for GpA with the average helical content and the number of BB contacts as the reaction coordinates. The representative conformation is shown as *inset*-figure. Color code: Thr (grey), Gly (green), Val (purple), Leu (orange).

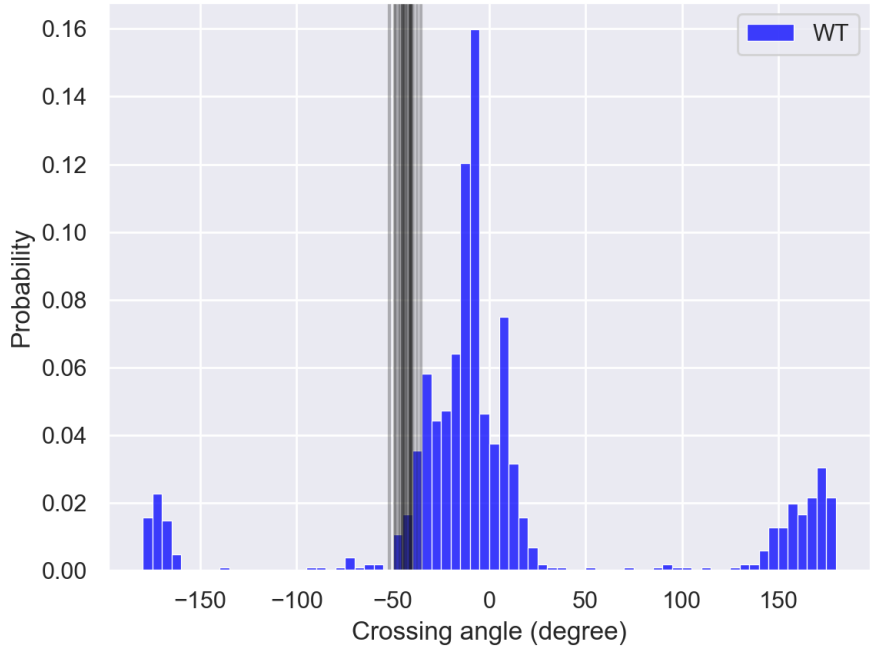

Figure 6.21: Helical crossing angle between the two helices of Glycophorin A. The horizantal lines reflect PDB values.

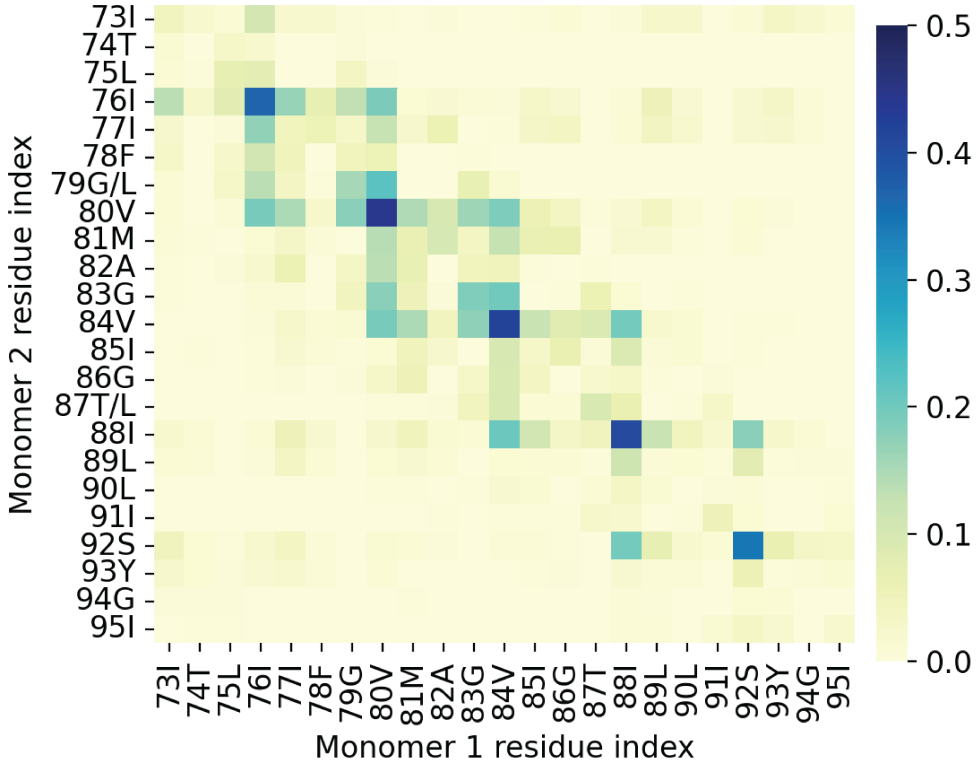

Figure 6.22: The residue-residue contact maps for GpA

6.4.2 Note on Computational Efficiency

To compare the computational efficiency gained by applying our CG model, we estimate the computational resource used for obtaining a converged PMF for Trp-cage via REST and our CG model. In order to get a converged PMF using REST atomistic simulations, 10 replicas at different temperatures are needed with 700 ns simulation time for each replica. With optimized simulation speed of 0.17 hours/ns and 20 cores-per-replica, the simulation required about 23.80 kSU (supercomputing units) of resources. For the converged PMF with our CG model, we used 15 replicas at different temperatures, with 200 ns simulation time per replica. With optimized simulation speed of 0.028 hour/ns and 27 cores-per-replica, the simulation required about 2.27 kSU. In order to get a converged PMF using REST, several parameters need to be tested (such as exchange frequency, number of replicas, range of temperatures), including the use of force field. These will result in a much higher computational cost in practice.

The applications and validations of our CG model in this work was focused on capturing accurate protein conformational landscape and structural diversity in *aqueous* environment. This explicit treatment of environment can capture local electrostatic changes such as dielectric fluctuations more effectively that can contribute towards protein's structural changes. This also allows for easy transferability of interactions, without a need for re-parametrization. Therefore, ProMPT can have special applications in studies involving chemically heterogeneous biomolecular systems due to its *explicit* description of environment (and solvent). Because of similar parametrization of non-bonded interactions, ProMPT can be patched with our *in-house* coarse grained membrane model — Water Explicit Polarizable MEMbrane model (WEPMEM) to study membrane and detergent assisted conformational transitions [81, 82, 84, 254]. Moreover, ProMPT can also

be patched with our CG model for divalent ions to study metal-assisted protein folding [174]. Finally, the modular architecture of the forcefield allows for easy parametrization of other biomolecules. Future work in our lab will be focused on studies to elucidate this environment-driven conformational changes in protein structure and dynamics with this forcefield.

6.5 Conclusion

In this article, we developed a polarizable coarse-grained model with explicit representation for the environment — ProMPT. The key objective here was to create a coarse-grained molecular dynamics forcefield to capture tertiary folding of protein structures with minimal constraints. The CG mapping scheme follows closely the MARTINI coarse grained forcefield geometry. In ProMPT, the polar CG beads have explicit drude-like charges that can couple a protein's environment to its structure and dynamics. We parametrized the non-bonded interactions between CG beads through free energies of their interaction with the environment. The bonded interaction potentials were generated through analysis of corresponding normalized bonded-feature distributions generated from a database of non-redundant protein structures from the protein data bank. This CG forcefield was validated by reproducing the conformational free energy landscape of several well-studied small protein systems. While ProMPT can accurately fold protein structures with homogeneous secondary structures, we observed some side-chain packing defects in folding of proteins composed with heterogenous secondary structures. Our future research efforts will be focused on fixing sidechain packing through accurate mapping of sidechain geometry. Due to its transferability, ProMPT can be patched with our previous *in-house* lipid and divalent ion models to study protein folding assisted by physiological environment.

Chapter 7: Effect of varying poly-Q tract and the presence of curved membranes on aggregation of Huntingtin protein's N-terminal domain

7.1 Overview

In the previous chapters (chapter 3-5), we characterized the aggregation behavior of a fragment of Aβ peptide. Along another direction in this chapter, we focus on Huntington's disease.

The initial 17 residues of the N-terminal and the following Glutamine repeats (poly-Q/GLN-repeats) of the Huntingtin protein have been associated with the pathogenesis of Huntington's disease (HD). A direct correlation has been established between the length of poly-Q tract and the severity of Huntington's disease. Moreover, HD belongs to the larger class of diseases whose pathogenesis involves poly-Q aggregation. While previous research have presented a consensus overview of the structure and dynamics of N17 fragment, a similar overview of ther poly-Q tract is missing. There have been competing evidences of secondary structure in the poly-Q tract. Moreover, since this is a fast aggregating peptide, it is difficult to effectively discern between individual peptides and peptides in their aggregate form through structural studies. In this chapter, we aim to explore the conformations of N17-Qn (where n $\in \{7, 15, 35, 45\}$), both as a single peptide and peptides in their aggregate form. For the single peptide, we noted a

trend towards more globular (decreasing asphericity) peptide structures with increasing lengths of polyglutamine tract. We also observed a significant specificity towards Q-Q interactions also observed in previous experimental reports. Finally, we also found a gradual change from helical-rich structures to increasing sheet-like features with increase in number of trailing GLN residues. For the case of peptide aggregates, we found that the GLN-repeats drove peptide aggregation. While, for smaller poly-Q tract, the hydrophobic core led by bulky Phenylalanine groups on N17 are formed, but a similar core is absent for longer poly-Q tract.

N17 often structures into amphipathic helices that can tether the protein to the membrane surface. Peptide anchoring to the membrane surface can have unique implications to disease pathogenesis, such as the peptide diffusion can get arrested resulting in increased aggregation. While some structural features of N17-membrane interaction are known, detailed biophysical reports of peptide-membrane interactions are still missing. In this chapter we also aim to understand these biomolecular details of peptide-membrane interaction and the molecular mechanism of curvature sensing. Our simulations suggest a gradual progression from a dominantly unstructured peptide in solution, to a more structured α-helical patch on the membrane. In presence of a curved membrane, we observed that N17 peptides preferentially interact with the curved region. Here, we discovered that while the polar and charged groups drive the initial membrane-peptide interaction, the membrane curvature sensing is dominantly controlled by the bulky hydrophobic groups such as PHE. Omission of these bulky residues from the sequence leads to loss of this curvature sensing ability of N17.

We will approach this research with the coarse grained protein model that we developed in our previous chapter — ProMPT.

7.2 Introduction

Huntington's disease (HD) is a hyperkinetic movement disorder and is associated with rapid neurodegeneration [290–293]. In recent years, this has emerged as a significant socio-economic concern as this disease is incurable and invariably lethal. Structural and functional alterations of Huntingtin protein is responsible the Huntington's disease [294–296].

Huntingtin protein is a large (348 Kda) soluble cytosolic protein often found in the central nervous system [294]. The wild-type Huntingtin protein consists of a 17 residue N-terminal domain, followed by 35 or less Glutamine repeats, a poly-proline rich segment and a cluster of HEAT repeats (α-helix-turn-α-helix unit) [294]. The final HEAT motifs are essential for several protein-protein interactions associated with this protein's functions. Abnormality in the Glutamine repeats has been associated with the pathology of Huntington's disease (HD). Fig. 7.1 is a schematic image highlighting these regions.

Studied have correlated the increase in polyglutamine lengths (mid 30s and higher) to increased disease severity [297–299]. Now, HD is a part of a large number of diseases such as spinocerebellar ataxia (SCA), and spinal and bulbar muscular atrophy (SBMA) are associated with glutamine repeats and share several common pathological signatures [53–55]. Indications of poly-Q diseases involve inclusion-bodies with several poly-Q stretches and progressive neuronal cell deaths. Studies with animal models have reported that *just* the presence of long poly-glutamine chains is enough for initiation of neuronal death and neurodegeneration [55].

For HD, clinical research found that mutants with expanded poly-glutamine tract can lead to pathological protein aggregation, often associated with neurodegeneration [290–296]. These findings have fueled increasing efforts to understand the biophysical properties of GLN that result

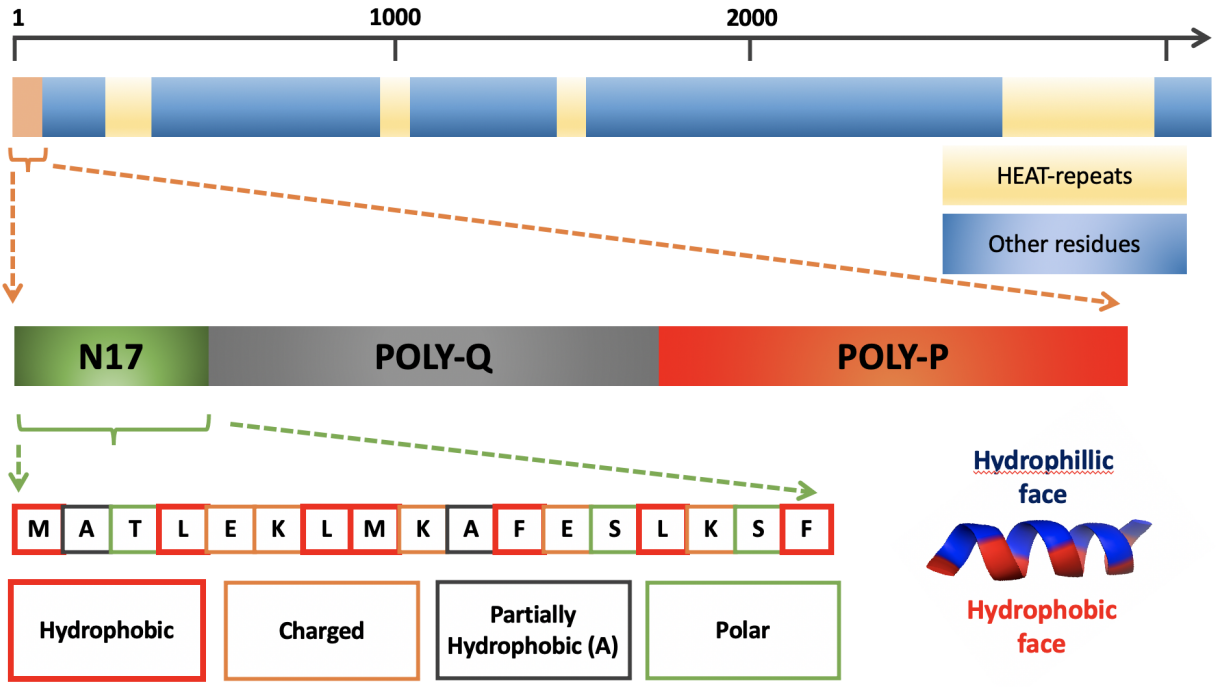

Figure 7.1: Schematic of Huntingtin protein with focus on the N-terminus

in aggregation of poly-Q. But, as early aggregation of poly-Q is fast and difficult to study by traditional experimental methods, computational studies have often been applied at those stages. Early computational and experimental studies found that water is a *poor* solvent for GLN and the poly-glutamines prefer to self-interact and assemble into disordered structures, which can over time evolve into fibrillar deposits [300–302]. Several *in-vitro* studies have also confirmed that poly-Q can aggregate by itself into ordered and insoluble β-sheet rich patches, among other more disordered structures [298]. These insoluble aggregates were also confirmed in cell cultures and mice models [55, 303]. Poly-Q, by itself, can be disordered in solution, and collapse into a globular shape with increase in Q-repeats [55, 304].

While poly-Q forms the unifying theme for these glutamine repeat diseases, other fragments can have unique implications, particularly in protein aggregation. In huntingtin protein, there are two more fragments associated with the disease pathology — the first 17 residues (N17) and the

38 residue poly-proline (polyP-C38) rich domain following poly-Q. While, polyP-C38 is known to retard protein aggregation rates, N17 aggravates it [305–307].

N17 is intrinsically disordered in solution, but can form amphipathic helix-like structures in hydrophobic and membranous environments [307]. Ampipathic helices (AH) are commonplace in several biological systems, and have functional relevance to many physiological processes [63, 308, 309]. Particularly, these helices can help in tethering the protein to the cellular membranes. This can help in protein co-localization and membrane permeation — a strategy that is employed by several anti-microbial peptides. Larger AHs such as Bar domain have been known to affect membrane remodeling, lipid phase separation and local lipid diffusion [309, 310].

In the case of Huntington's disease, recent biophysical evidences have suggested that N17 is the cis-acting switch for driving poly-Q aggregation [307, 311]. Particularly these experiments established the importance of hydrophobic residues. Point-mutations that broke the helix or reduced hydrophobicity, in turn reduced the aggregation propensity. Even in the absence of poly-Q, N17 is known to aggregate into peptide clusters.

Huntingtin protein is specifically associated with several membrane-related processes such as vesicle trafficking and has been found to localize at the membranes of endoplasmic reticulum, golgi apparatus and endosomal vesicles [53]. N17, being an AH forms the membrane binding domain of huntingtin protein. But, in the case of pathologically mutated htt, membrane binding can tether the peptide and arrest the dimensionality of diffusion, thus promoting aggregation. As N17 plays a central role in aggregation both in solution and in presence of membranes, it can be the key towards therapeutic intervention [307].

While several experimental research attempted to understand the biophysics of the N17 and N17-poly-Q domain, structural information remains limited due to highly dynamic and

transient (quick aggregating) nature of these biomolecules. Recent experimental characterization revealed that with a particular variant of N17-poly-Q, the peptide has a propensity towards β-sheet structures in Q-rich regions [312, 313]. Experimental data by Langen and coworkers with Time-Resolved Fluorescence Energy Transfer (TR-FRET) and Electron Paramagnetic Resonance (EPR) suggest both α-helix and β-sheets within the poly-Q segment [305]. Another structural study by Nagai *et al.* also reported an increased β-sheet fraction with increase in Q-repeats [314].

Recently Michaelek *et al.* used two dimensional solution NMR to identify the lowest energy structure of this N17 domain in micellar conditions [315]. In the same study, they also delineated the structure in presence of POPC membranes. In another structural investigation, Tao *et al.* used model small unilamellar vesicles to study the structure of N17-Q25 fragment [316]. Their study showed the propensity of this peptide to interact interfacially with the membrane, with stable helicity in the N17 part. Chaibva *et al.* used atomic force microscopy with POPC membranes restricted to a fixed curvature to record a interaction of N17-Q35-P10-KK fragment dominantly with regions with positive curvature [60].

Only a few computational studies have been attempted to characterize these systems because of the associated large system sizes and the intrinsic disordered nature of these peptides. Moreover, discrepancies can be noted in these computational findings. Initial results by Kelly *et al.* with simulated tempering suggested two stable populations of N17 structures — the more stable double helix (bent) and a single linear helical patch [317]. With Multiscale Coarse grained (MS-CG) simulations by Wang and Voth reported quick interchange between folded-unfolded states for lower Q-repeats ($\leq$28), and long stable helices for longer Q-repeats ($\geq$28) [318]. Similarly, investigations by Dlugosz and Trylyska with atomistic simulations and replica exchange molecular dynamics also suggested a helix-rich poly-Q (Q-55) domain with a solvent exposed N17 headpiece

[319]. On the other hand, Lakhani *et al.* with discrete molecular dynamics and an atomistic resolution forcefield found an increased tendency to form β sheet rich structures with expanded lengths of Q [320]. Both computational and experimental structural studies centered around understanding the morphology of N17-Poly-Q aggregates are scarce, and we have limited knowledge of aggregate architecture.

Zhang *et al.* investigated the folding of N17 domain through a consensus forcefield approach, using simulations from multiple different forcefields (OPLS-AA/L, CHARMM36, and AMBER99sb-ILDNP) in presence of dodecylphosphocholine (DPC) micelles. They found a relatively low barrier (3 kcal/mol) of transformation into a helical structure [321]. Cote *et al.* characterized the membrane interaction of N17 through a combination of molecular simulations and experiments [322]. Their analysis points to a single α-helical structure, more stable than the micellar NMR structure suggested by Michaelek *et al.* [315]. A more recent study elucidated correlation between peptide structure and the length of Q-repeats for a single peptide. Here the authors report of a super-compact and more globular structures with progressive increase in Q-repeats.

Studies on how N17 drives membrane interaction and curvature sensing is limited due to compounded complexity of the presence of membranes, in addition to the intrinsically disordered and highly transient peptide. While peptide aggregates are acknowledged as the primary pathological entity in Huntington's disease, computational research focus has been primarily on monomeric structure and monomer-membrane interaction. Moreover, N17 being an amphipathic helix can have preferential interaction with curved membrane as noted earlier by Chabaiva and coworkers with atomic force microscopy [60]. But, the details of peptide structure that enable such curvature sensing is still missing. In this work, we aim to expand our understanding to peptide aggregate structures and how they drive membrane-protein interaction, particularly curvature sensing. Our

recently developed coarse grained forcefield — ProMPT is uniquely poised to address this because of the explicit-solvent and transferable nature of this CG forcefield. Also, coarse-graining affords the access to length and time-scales beyond the current atomistic levels to study peptide-membrane interactions in the case of aggregates.

The objectives of this research are two fold.

1. Understand the impact of pathological glutamine repeats on peptide structure and aggregation - While previous research have presented a consensus overview of the structure and dynamics of N17 fragment, a similar overview of ther poly-Q tract is missing. There have been competing evidences of secondary structure in the poly-Q tract. Moreover, since this is a fast aggregating peptide, it is difficult to effectively discern between individual peptides and peptides in their aggregate form through structural studies. Therefore, in this work we aim to explore the conformations of N17-Qn (where n $\in \{7, 15, 35, 45\}$), both as a single peptide and peptides in their aggregate form.

2. Present a mechanistic view on how N17 peptides sense curvature on POPC membranes - Previous research on peptide-membrane interactions has been limited both from a experimental and computational perspective. While some structural features of N17-membrane interaction are known, detailed biophysical reports of peptide-membrane interactions are still missing. In this chapter we aim to understand these biomolecular details of peptide-membrane interaction and the molecular mechanism of curvature sensing.

To our knowledge, this is the first computational research that studies unbiased peptide aggregation and peptide-membrane interactions of N17/N17-Qn peptides starting from solvated unstructured peptides.

7.3 Methods

For this work, we used the coarse-grained forcefield developed in chapter 5 — ProMPT. Due to the explicit-solvent nature of this forcefield, it can capture local environmental fluctuations (such as local dielectrics) well. This can contribute towards environment-assisted structural changes in protein structure. Moreover, this explicit representation of environment allows for easy transferability across multiple biomolecular systems, without need for re-parametrization. As such, we used ProMPT (see chapter 5) to represent the protein, WEPMEM (see chapter 2) to represent membranes with the MARTINI polarizable water [81–83].

7.3.1 Single Peptide (N17-Qn) simulations in aqueous solution

To understand the folding landscape of N17 with multiple Q-repeats, we ran simulations with 7, 15, 35 and 45 trailing glutamines. All simulations are performed with GROMACS 2019.4 [177]. The starting conformation and the protein topology are generated with in-house codes. The initial conformation for all proteins is a random unfolded extended conformation. The simulation box is adequately solvated and neutralized through addition of ions. 9 simulations were run for each peptide system ranging from 300K (T$*$=0.445, previous chapter) to 460K (T$*$=0.682, previous chapter). This allows us to escape local free energy barriers and have a better exploration of conformational landscape.

Energy minimization is first conducted with steepest descent. An NPT equilibration run at T=300K for 50 ps is followed at 1 bar and with time step 0.001 ps. The canonical production run is then performed with time step 0.01 ps for 200ns at different temperatures. Since the box size is constant, the solvent density remains the same across various temperatures. Electrostatic cutoff is

1.6 nm and particle-mesh Ewald (PME) method is applied for long range electrostatic interactions [151]. Nos$\tilde{e}$-Hoover thermostat is used to maintain the system at the desired temperature [266].

7.3.2 Multi Peptide (N17-Qn) simulations in aqueous solution

We also did a similar exercise, but with multiple peptides. 6 peptides ($\sim$ 6.9 mM), each with trailing glutamine repeats of 7, 15, 35 and 45 were solvated in a polarizable water box, with ions added to neutralize the box. In the initial conformation, all copies of peptides are randomly distributed in a simulation box. The simulation procedure is identical to the single monomer simulation.

7.3.3 Single Peptide (N17) in presence of a planar membrane patch

In this section, we characterize the curvature sensing nature of the N17 part of Huntingtin protein. As membrane interaction is primarily driven by the N17 fragment of the peptide, here we have focused on just that segment. To understand how single N17 peptide interacts with the membrane, we created two replica simulations of a single N17 peptide in presence of a smaller patch of planar lipid membrane. These simulations can be used as validations against pre-existing results of N17-membrane interactions. We created the small membrane patch with 240 POPC lipids on each leaflet, and 12 nm (in the Z-direction) solvated box. After initial energy minimization and 1 ns of equilibration, the production simulation was run for 200 ns with constant-pressure ensemble. Temperature was maintained at 300K through a Noose-Hoover thermostat [148, 149] with a time constant of 1 ps. Parinello-Rahman barostat [150] with a time constant of 1 ps and compressibility of 3×10^{-5}/bar was used alongside with semi-isotropic

pressure coupling to maintain a pressure of 1 bar. Particle mesh Ewald (PME) [151, 152] with a relative dielectric constant of 2.5 and cutoff distance of 1.6 nm was used to compute long range electrostatics. The Lennard-Jones interactions were modified starting from 0.9 nm to 0 at 1.2 nm by the GROMACS shift scheme.

7.3.4 Multi Peptide (N17) simulations in presence of a curved membrane

Next, we pivot onto curvature sensing aspect of N17 peptides. Towards this, we created an artificial curved membrane with a diameter of 13 nm from POPC lipid molecules. The overall architecture consists of a semi-sphere and a planar region. Fig. 7.2 is provided as a reference. The curved membrane is created by first placing the POPC lipids as points, and growing them to their individual sizes through slow-growth method of Gromacs. We followed the approach outlined in BuMPY [323]. We also added dummy particles close to the lipid headgroup, and instituted a hardcore repulsion ($c_1 2 \sim 10^{-5}$) between lipid tails and the added dummies. This maintains the membrane curvature throughout the simulation time.

This curved membrane was energy-minimized and then equilibrated for 1ns with a dt of 0.001 ps with a Berendsen thermostat at 300K. We followed this up with another dt=0.01 ps equilibration for 1 ns with nose hoover thermostat. Finally, we ran a small production run for 20 ns with constant pressure ensemble. All the simulation protocol is identical to the planar membrane case.

After, this 20 ns run, we added 36 peptides ($\sim$ 8 mM), above the critical concentration for Huntingtin protein aggregation; and 125 mM of NaCl ions [324]. The peptides were initially held with position restraints at the backbone of residue 8, while initial minimization and equilibration.

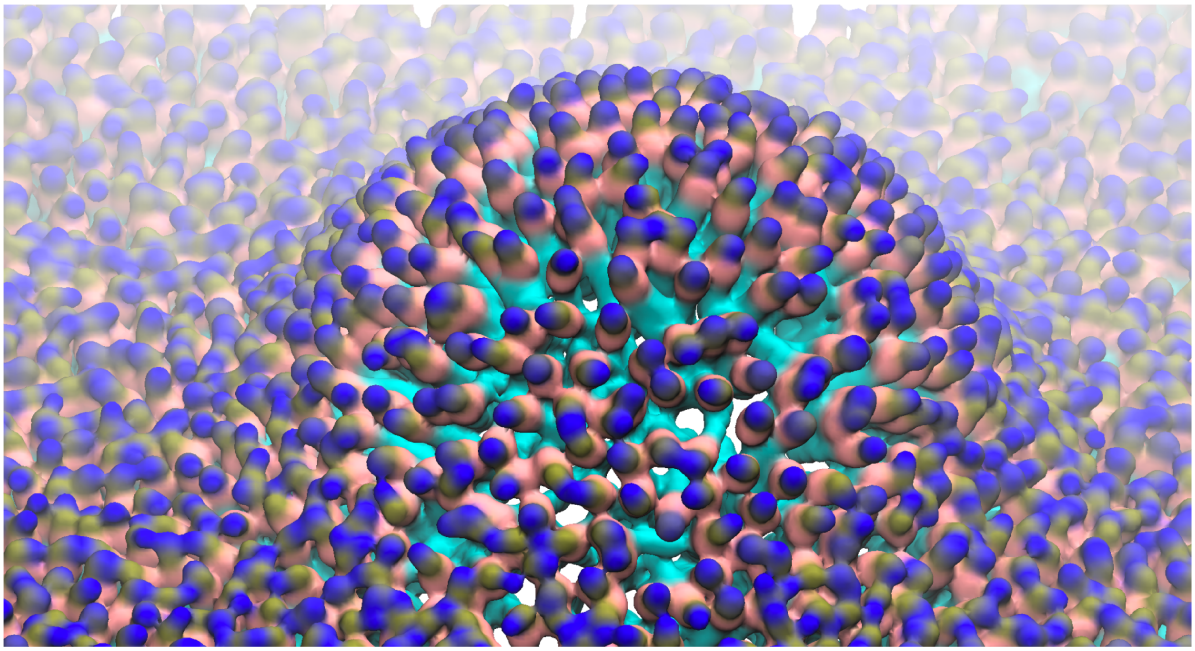

Figure 7.2: A snapshot of the created curved membrane

The initial equilibration steps were conducted with dt=0.001 for gradually increasing number of steps — 0.5 ns, 1 ns and 5 ns. All the other parameters for simulation control are similar to the curved bilayer creation stage. Finally, at the production stage the simulation was run for 400 ns. We ran two independent replicas of this simulation system. For an appropriate control, we also simulated a curved bilayer without any peptides. The amount of curved membrane surface area available to the peptide is about 265.46 nm^2, compared to 623.51 nm^2 for the planar region. So, without any preference for curved/planar region, the chance of interaction with curved region is 36 $\%$, compared to 64 $\%$ on the planar region. Moreover, to decode the effects of Phenylalanine (bulky hydrophobic group) from the other parts of the peptide, we also simulated (for 750 ns) a peptide system with the two Phenylalanines mutated to Valines (F11V/F17V).

7.3.5 Analysis

The re-weighting for all the trajectories in the case of solvated peptide systems (single peptide and peptide aggregate) were done with pyMBAR module [269]. This allows us to analyze a particular reaction coordinate in a defined ensemble. The results shown in this chapter has reweighted the reaction coordinate to distributions at 300K. Similarly, the contact plots shown in this work is the expectation-value of a particular contact at 300K, along the generated distribution of that contact.

Special care had to be taken for calculations on the curved membrane. This is because, there is no uni-directional normal as is the case for a planar lipid bilayer. Therefore, we created parallels for commonly used membrane metrics for our case, that takes into account the local membrane curvature.

7.3.6 Area per lipid on membranes with varying curvature

Traditionally, area per lipid is calculated through a voronoi tesselletion of the membrane surface, and ascribing each lipid with a defined area. But this calculation can be tricky for curved membranes, as creating a convex hull in two-dimensions would not be appropriate, and a three dimensional convex hull is often error-prone at our scales. Therefore, in this work, we apply a local normal approach outlined by Buchoux in their implementation of FATSLiM [325]. First, the locality set—A_i, for each lipid molecule, defined by the locally close phosphates to the phosphate of ith lipid molecule is created. Now, the principal component analysis is used to determine the local normal at ith point, and then identify the local tangent plane for ith lipid (T_i). The local neighborhood of the ith lipid, A_i is projected onto this T_i to create $A_i^{'}$. Finally, a voronoi

tesselletion is performed on $A_i^{'}$ to define the local surface area.

7.3.7 Fine-grained surface for density calculations

For the density calculation, we interpolated the lipid membrane to create a super fine-grained local grid of points along the membrane surface. Initially, we create a lateral grid of about 40000 points placed in the simulation box along X and Y. Now, the Z direction is added to this grid of points, by creating a set of weights based on two-dimensional distance from the lipid-phosphates. This two dimensional distance is penalized by exponent of squared distance. That way, the points closer get more weights and this reduces for points farther off. The approach of including weights from local points to ascribe Z position of the grid, generates smooth local curvature. Now, the distance for each specific species can be calculated by computing distance from locally present grid points.

7.4 Results and Discussion

7.4.1 Conformational landscape of N17-Qn in solution

First, we attempt an exploration of the conformational landscape and structural features for N17 with varying lengths of poly-Q tract. Towards this, we ran simulations with 7, 15, 35 and 45 trailing GLN in aqueous solution at varying temperatures. We reweighted the trajectory data (by Multi-state Benett Acceptance Ratio - MBAR) at 300K, and created potential of mean force plots using *asphericity* as the reaction coordinate [Fig. 7.3] [269]. Asphericity computes how spherically asymmetric a particular structure is, and can be used as a measure for globular nature of a structure. Asphericity, b can be measured in terms of principal moments of gyration tensor

(λ_x, λ_y and λ_z).

$$b = \lambda_z - \frac{1}{2}(\lambda_x + \lambda_y)$$

The values of asphericity can range from 0 being complete spherically symetric, and 1 being planar.

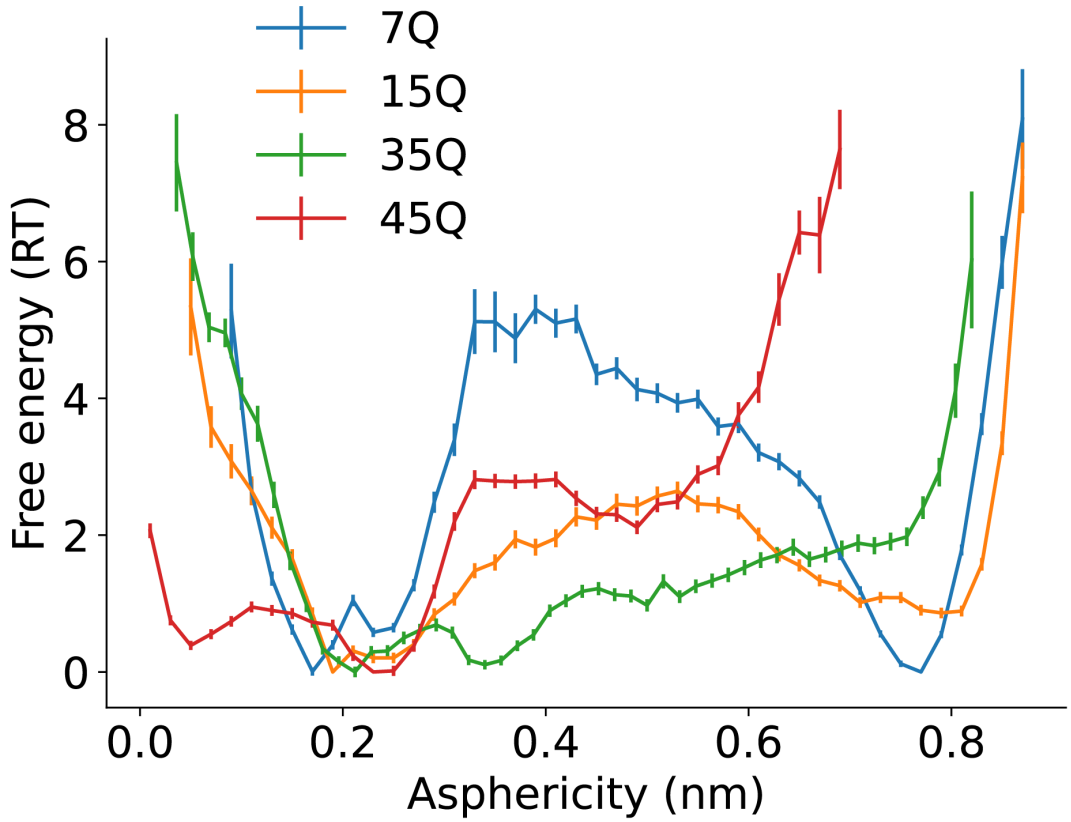

Figure 7.3: Free energy landscape of a single N17 with varying Qs, in solution with asphericity as the reaction coordinate

Our simulations suggest a trend towards more globular (decreasing asphericity) peptide structures with progressively increasing lengths of polyglutamine tract. This decreasing asphericity with increasing Q-repeats was also recently noted in an atomistic study coupled with replica exchange molecular dynamics [304].

Fig. 7.4 shows the overview of the conformations we capture with our simulations. Fig. 7.4-a presents the number of particular heavy atom contacts (cutoff = 7$\mathring{A}$) between each fragment of the peptide — N17-N17, N17-poly-Q and poly-Q-poly-Q for varying trailing poly-Qs. Here, the number of contacts between N17 and Poly-Q does not change with increasing number of Qs, while poly-Q-poly-Q contacts increase. This suggests a significant specificity towards Q-Q interactions that are also noted in previous experimental and computational reports [300–302].

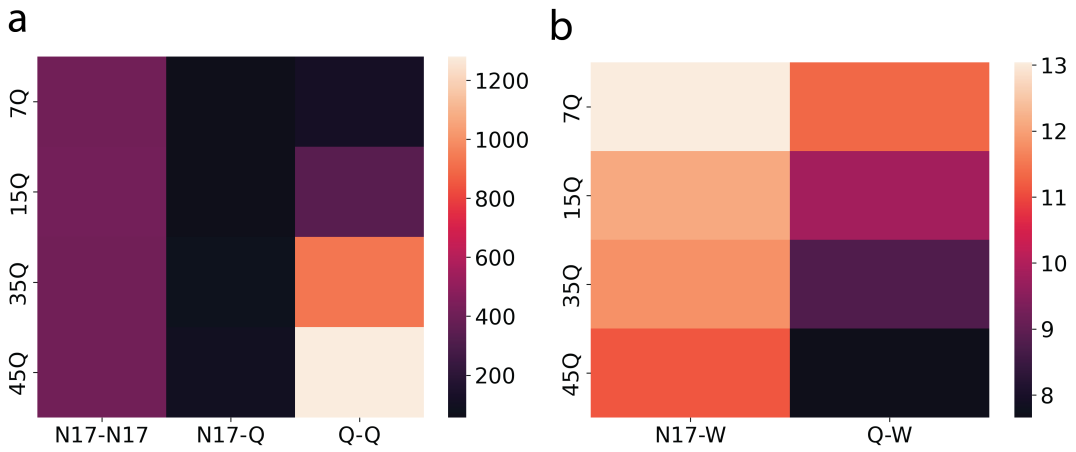

Figure 7.4: a - Number of contacts between different fragments of the peptide. b - Number of peptide-water contact per residue. The lighter shade is indicative of increase in number.

Fig. 7.4-b presents the average number of water contacts per *each residue* belonging to a particular fragment of the peptide. There is an effective decrease in how solvated particular regions (N17/poly-Q) of the peptide are, with the faster decrease for the poly-Q fragment. This corroborates with previous experimental and computational evidences that suggest water can behave as a poor solvent for GLN polymers [300–302].

For a more fine-grained perspective, we created contact maps of the backbone for each peptide (Fig. 7.5). Here, a contact is defined by primary backbone interaction center within 7 $\mathring{A}$ of each other. We created these heatmaps by reweighting trajectories across the temperature

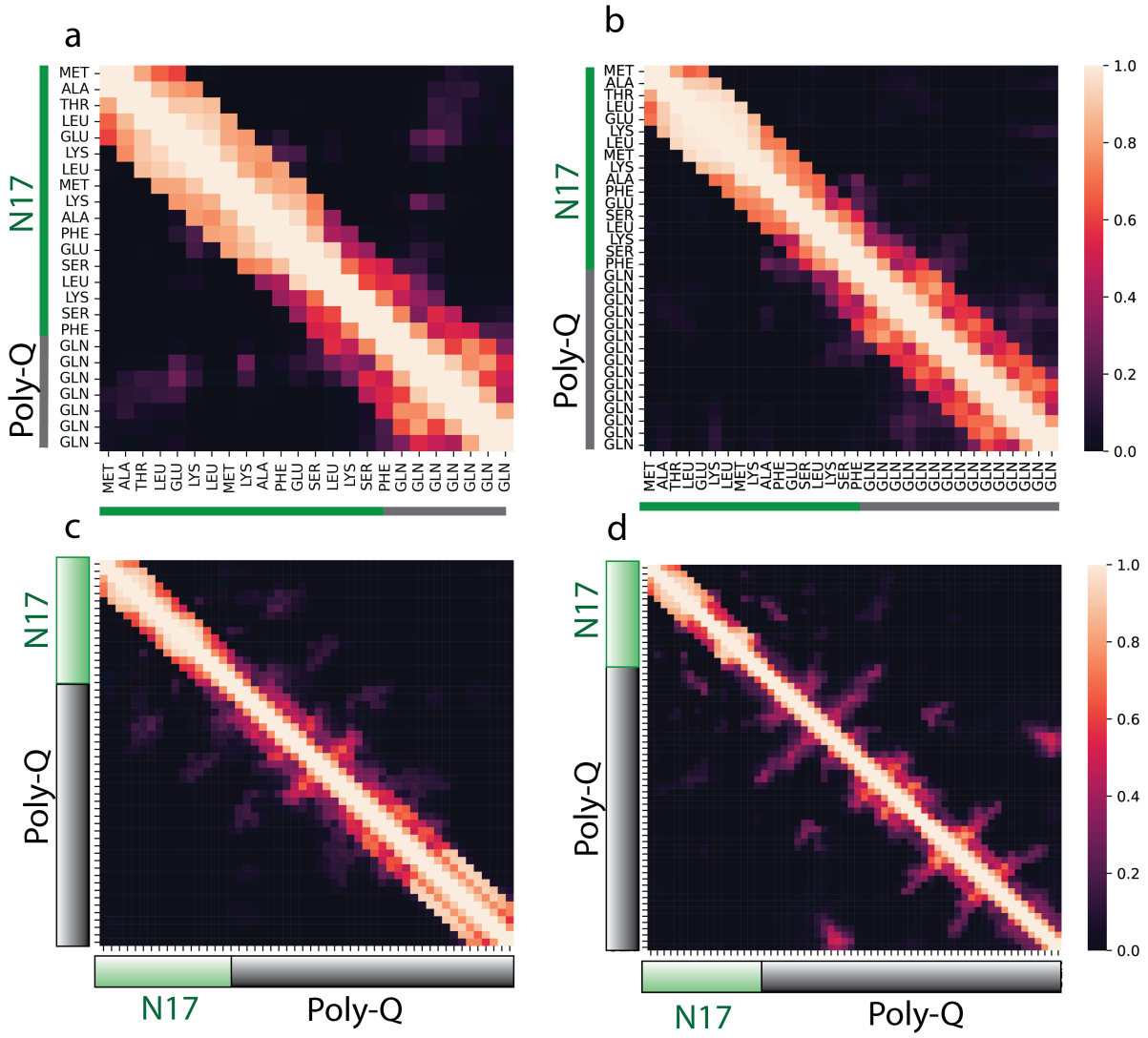

Figure 7.5: Reweighted contact (backbone CG interaction sites within 7 $\mathring{A}$) map for different peptides — N17-7Q (a), N17-15Q (b), N17-35Q (c) and N17-45Q (d). The N17 and poly-Q regions are marked in green and black respectively.

range through MBAR. Contiguous contacts with positive or negative slope (off the diagonal) marks sheet-like features. While near-diagonal contacts (i,i+4) are suggest α-helical structures. We noted a gradual change from helical-rich structures to increasing sheet-like features with increase in number of trailing GLN residues. Representative snapshots of single peptide structure is provided in Fig 7.6

While computational methods are well-poised to study secondary structure of these peptides, the results have been rather vague and contradictory. While some reports have pointed to primarily helical structures in the complete exon, others have noted some sheet factions. Recent experiments by Urbanec *et al.* used site-specifically labelled NMR samples to pose an ensemble view of htt-exon1 structure [326]. They suggested an increase in propensity for extended structures at later positions of the poly-Q tract, and a competition between helix and extended structures at initial positions. This propensity for increased β-sheets is also corroborated by a recent atomistic Replica-exchange Molecular Dynamics study by Kang *et al.* [304]. Our observations suggest support for the hypothesis of increase in sheet-like structures, following the addition of GLN residues.

7.4.2 Aggregation of N17-Qn in solution

Similarly, we also explored the conformational landscape of large protein aggregates, through simulations of six copies ($\sim$ 6.9 mM) of each peptide (N17-7Q, N17-15Q, N17-35Q and N17-45Q). The plots on Fig. 7.7 present inferences on aggregate's structure, with information re-weighted across multiple temperature to 300K by MBAR. Representative VMD snapshots are provided in Fig. 7.8.

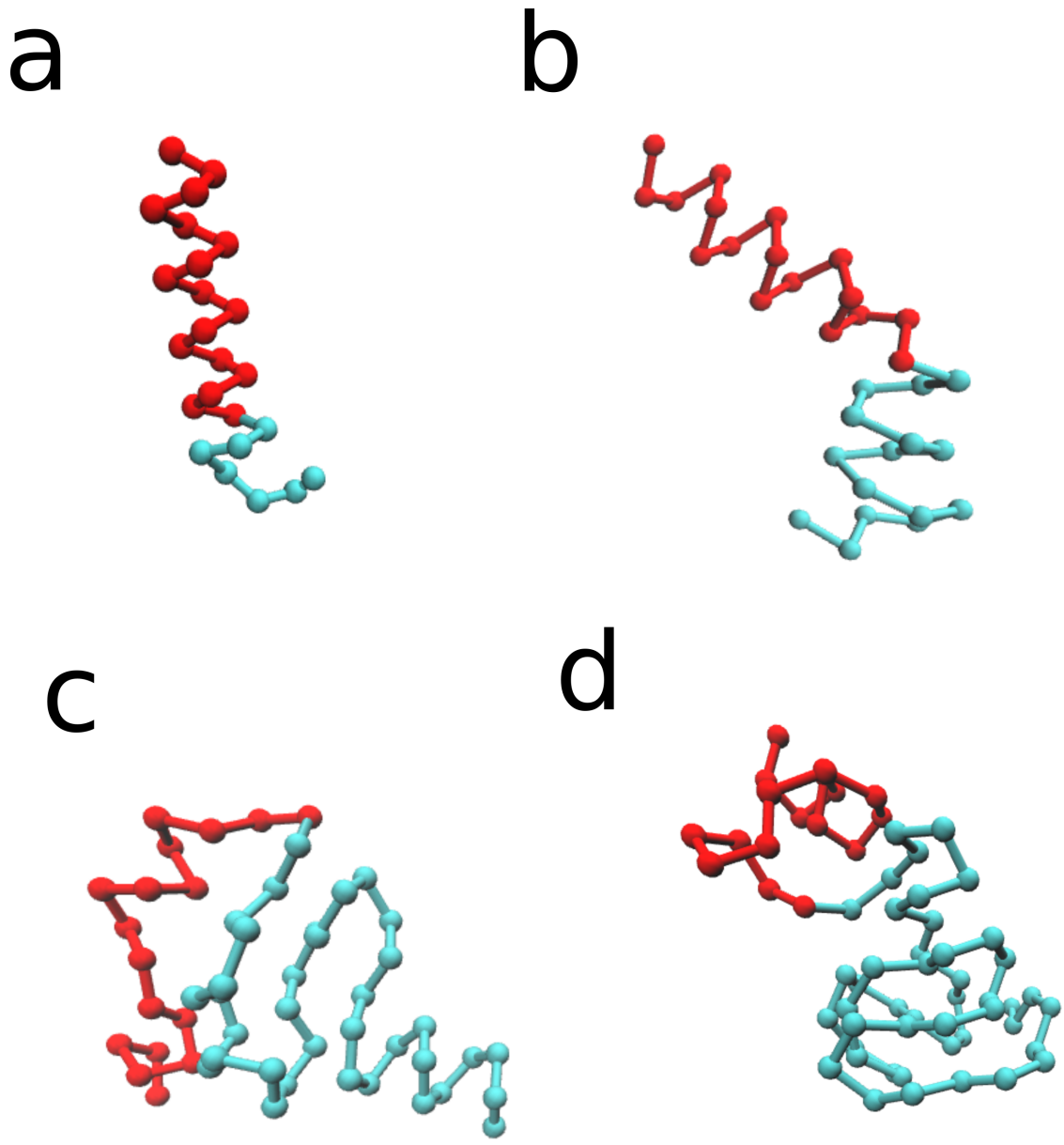

Figure 7.6: Representative VMD snapshots at 300K for a-N17-7Q, b-N17-15Q, c-N17-35Q, d-N17-45Q. Color scheme - N17 backbone: Red; Poly-Q backbone: cyan; PHE sidechain: Orange

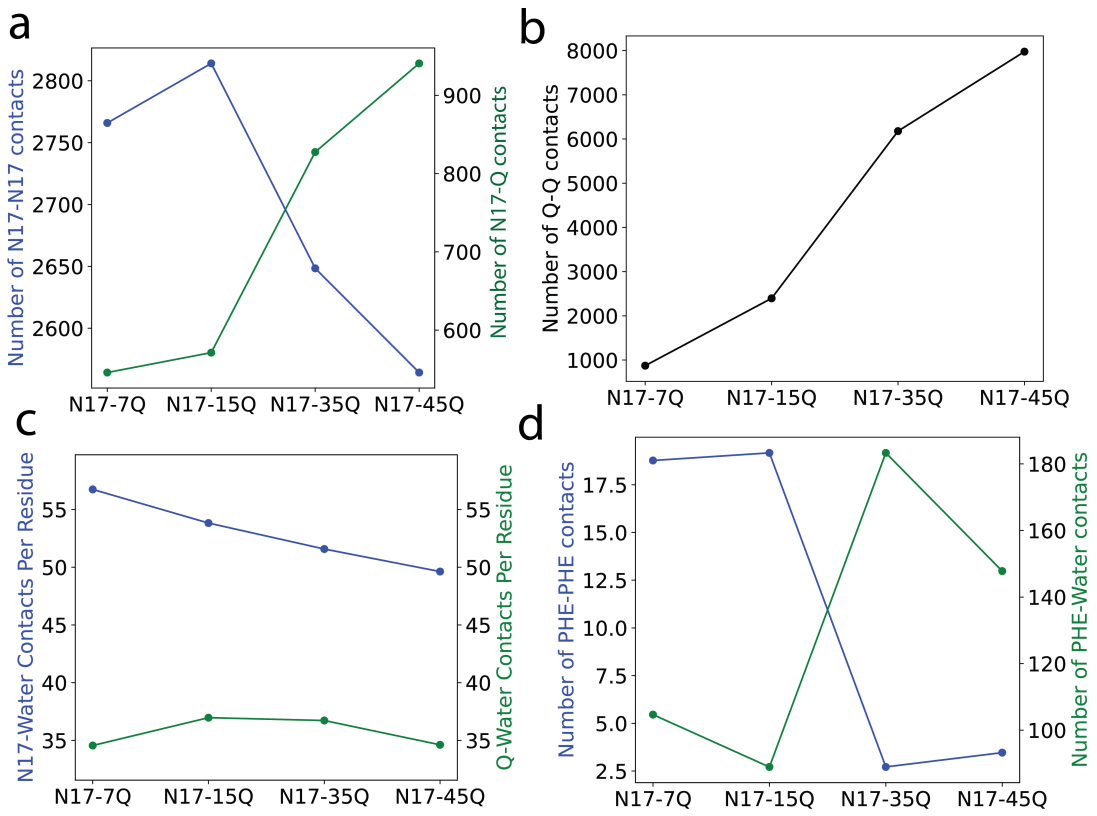

Figure 7.7: Contacts between different domains in the aggregate structure. a - Number of contacts between N17-N17 (blue) and N17-Poly-Q (green). b - Number of GLN-GLN contacts. c - Water solvation per residue of N17 (blue), poly-Q (green). d - PHE-PHE contacts (blue) and PHE solvation (green). All the plots here are a function of the length of polyglutamine tract.

Fig. 7.7-a shows the the number of N17-N17 (blue) and N17-Q (green) contacts with increase in the length of poly-Q. Here, we observed a significant decrease in the number of N17-N17 contacts, with progressive increase in poly-Q. This decrease is compensated by an increase in N17-Q contacts. Fig. 7.7-b shows the increase in the number of GLN-GLN contacts with increase in GLN lengths. We note an up to eight fold increase in the number of GLN-GLN contacts (also noted from VMD snapshots of Fig. 7.8), which similar to the single peptide case also suggests the specificity of interactions between glutamines. This specificity was also noted in simulations by Chen *et al.* [324]. Fig. 7.7-c shows the hydration per residue of N17 (blue) and poly-Q (green). While, N17 shows a slow dehydration with increase in the length pf poly-Q tract, the solvation of the poly-Q tract itself is not affected. The overall solvation is significantly higher for the N17 compared to the poly-Q, which is also an indication of water being a poor solvent, and the N17-repeats situated closer to the periphery compared to the larger core created by the poly-Q repeats. This feature is also evident in Fig. 7.7-d, which shows the number of Phenylalanine residue contacts (blue lineplot) in different simulations. We observe a significant drop in the number of PHE-contacts from N17-7Q/N17-15Q to more pathological N17-35Q/N17-45Q. Moreover the number of PHE-water contacts (Fig. 7.7-d, green) increases, which establishes N17 as an interfacial domain (also Fig. 7.8-d) in the case of aggregate structures for longer poly-Q tracts. These structural measures suggest that there is a reduction in the size of the hydrophobic core, with N17 distributed at the interfacial region of the aggregate. This is in contrast to the traditional view of hydrophobic cores for aggregates/micelles. Here, at longer lengths of the poly-Q tract, the glutamines drive aggregation and a polar core develops with GLN at the aggregate-center, similar to the case of reverse micelles [327]. This disrupts the number/size of the hydrophobic core led by bulky PHE groups and delegates N17 to the periphery of the peptide, where hydrophobic

groups are more water exposed.

To understand backbone-backbone interactions and get a fine-grained perspective of secondary structures in our protein aggregates, we created re-weighted backbone contact-maps (Fig 7.9). These contact maps show the expectation value for each possible inter and intra peptide backbone-backbone contact (cutoff 7 $\AA$) re-weighted to 300K. The contacts between groups which have bonds (i, i$\pm$1) have been removed.

With increasing length of poly-Q tract we note an increase in sheet-like features. Contiguous contacts with positive or negative slope (off the diagonal) marks sheet-like features. While near-diagonal contacts (i,i+4) are α-helical structures. Similar to the single peptide simulations, here we also note a change from helical-rich structures to sheet-like features with increase in number of trailing GLN residues. This fibrillation is particularly high for the pathological N17-45Q fragment, in line with the current expectation of increased fibrillar patches at longer and pathological poly-Q fragments [297–299].

7.4.3 Membrane interaction of a single N17 peptide

Next, we probe into peptide membrane interaction of Huntingtin protein. Since, N17 is the membrane binding domain, we would focus on these initial 17 residues only. As a first set of simulations, we analyze the interaction of a single N17 peptide with a planar membrane. These simulations are aimed to understand the pathways of peptide's partitioning onto the membrane, and peptide folding as a consequence of peptide-membrane interaction.

We created the dihedral angle box-plot for the peptide backbone with different slices of the trajectory [Fig. 7.10a/b - 0-10 ns; Fig. 7.10c/d - 40-50 ns, Fig. 7.10e/f - 90-100 ns; Fig.

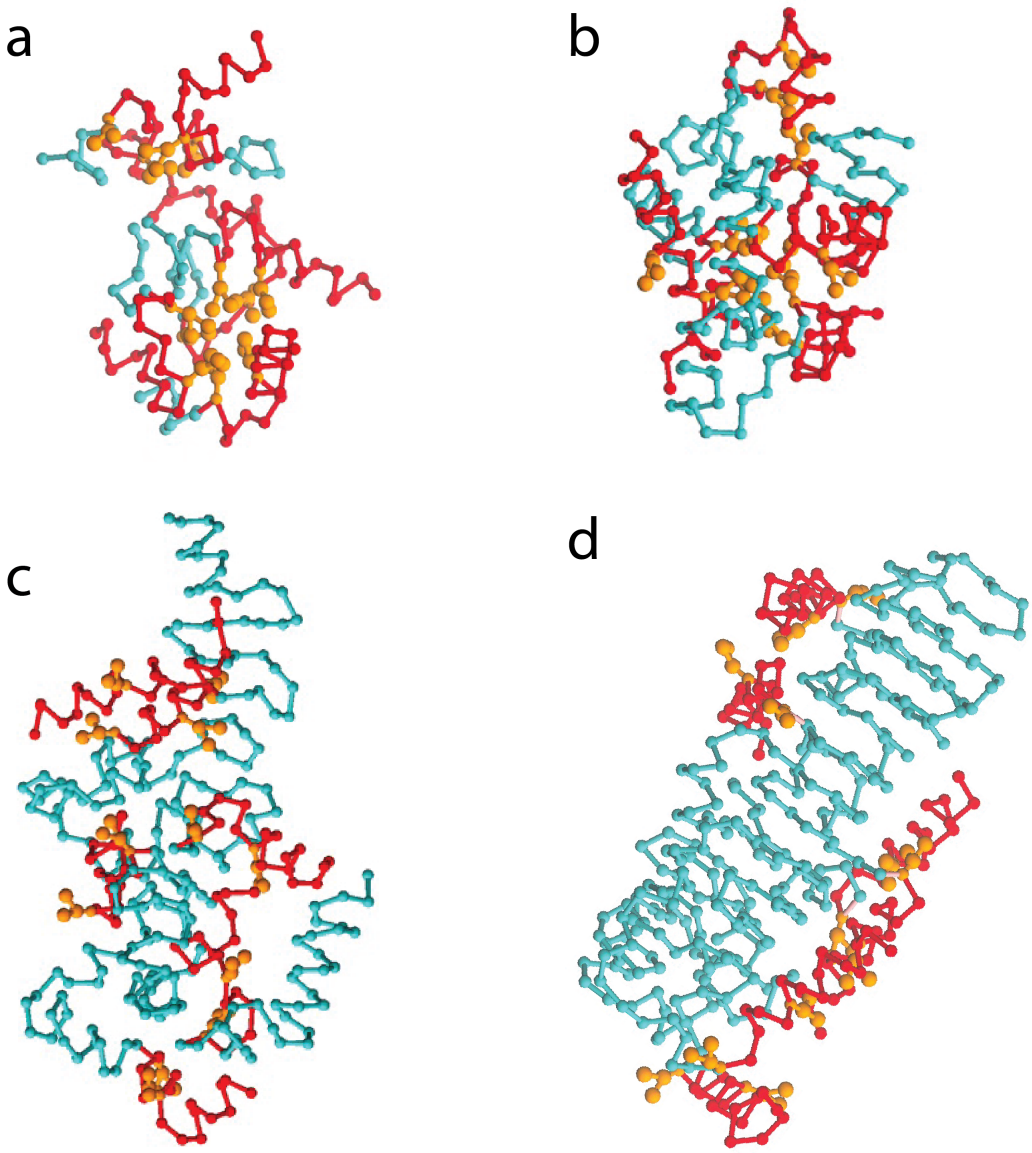

Figure 7.8: Representative VMD snapshots of peptide aggregate for a-N17-7Q, b-N17-15Q, c-N17-35Q, d-N17-45Q. Color scheme - N17 backbone: Red; Poly-Q backbone: cyan; PHE sidechain: Orange

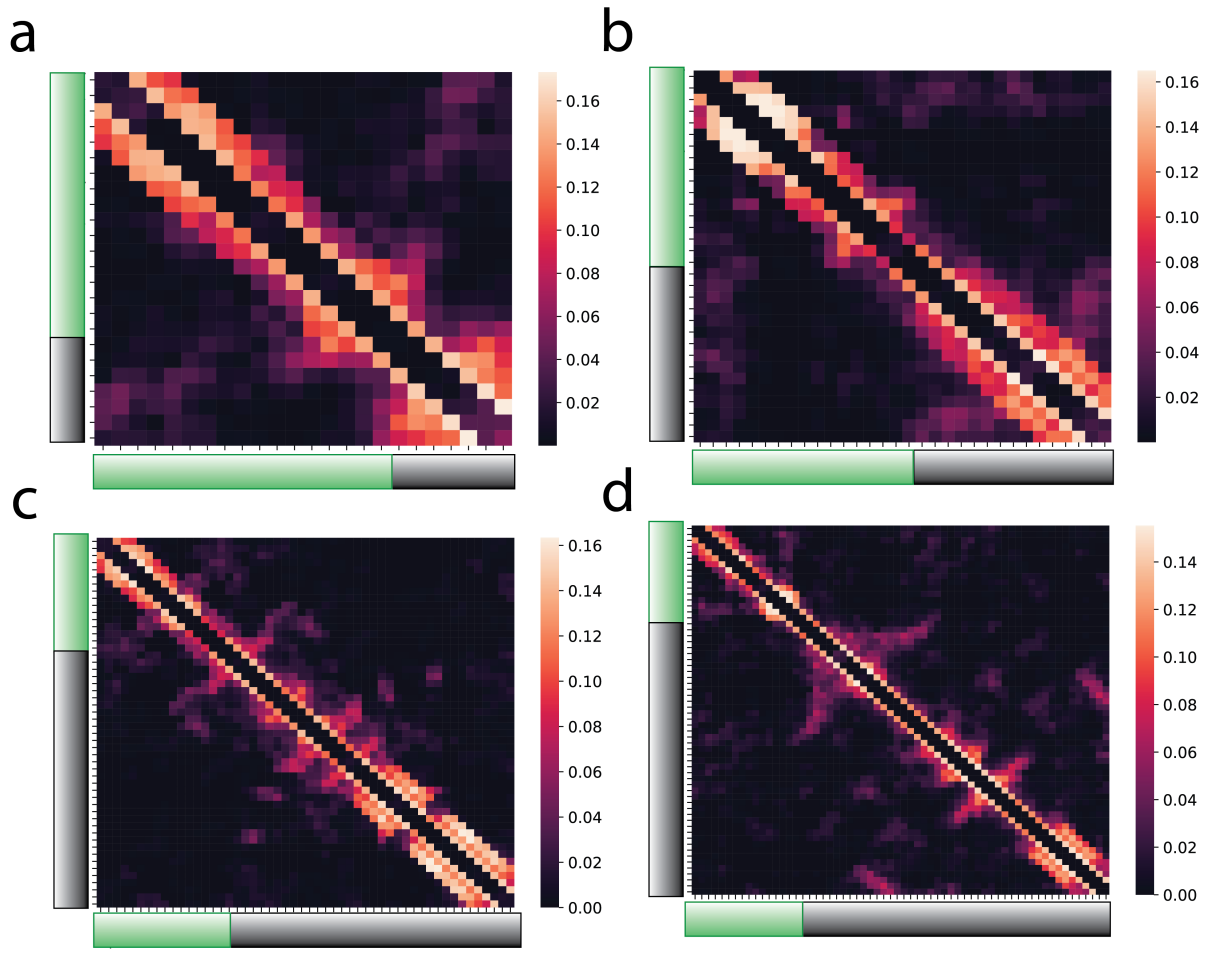

Figure 7.9: Re-weighted contact (backbone CG interaction sites within 7 $\mathring{A}$) map for different peptide aggregates — N17-7Q (a), N17-15Q (b), N17-35Q (c) and N17-45Q (d). The N17 and poly-Q regions are marked in green and black respectively.

7.11a/b - 140-150 ns, Fig. 7.11c/d - 190-200 ns]. The horizontal black line on the boxplots mark 50.4 degrees, which is the preferred backbone dihedral angle for α-helices as noted from the non-redundant protein structures we scraped from the protein data bank [see Chapter 5, section 6.3.2].

Our simulations show a gradual progression from a dominantly unstructured peptide in solution, to a more structured α-helical patch on the membrane. This can be noted from a gradual reduction in variation of backbone dihedral with time, moving closer to 50.4 degrees. The peptide forms fast initial contact with the membrane within 15-20 ns, followed by gradual membrane absorption by 45-50 ns [Fig. 7.12]. This initial contact is led by THR (cyan), MET (green) and to some extent LYS (yellow) co-located at the N-terminus of the peptide. The peptide, while disordered, forms a hydrophobic core in solution through the phenylalanine groups. Now, after the initial contact, between 20-40 ns, these PHE groups partition onto the membrane. This starts off a slow secondary structure re-adjustment phase, completing the helical transformation by 160-170 ns. Here, we record that significant rearrangement needs to be done at A10-L14 fragment to break the PHE-PHE interaction in favor of interaction between PHE and the acyl tails, that reshapes the peptide into a proper helix.

This preference towards structured α helices has been noted for N17 in both membranous and micellar environments. Michaelek *et al* noted a pronounced transition into helical conformation in presence of DPC micelle (2D solution NMR) and POPC membrane (solid state NMR) [315]. Stable membrane associated α-helix for N17 is also corroborated by several atomistic and coarse-grained computational work [317,318]. Our work, validates this observation of increased helicity in presence of membranes and points to a possible mechanism for this structural transition.

Fig. 7.13 shows the distribution of the angle that the regression line fitted to the peptide

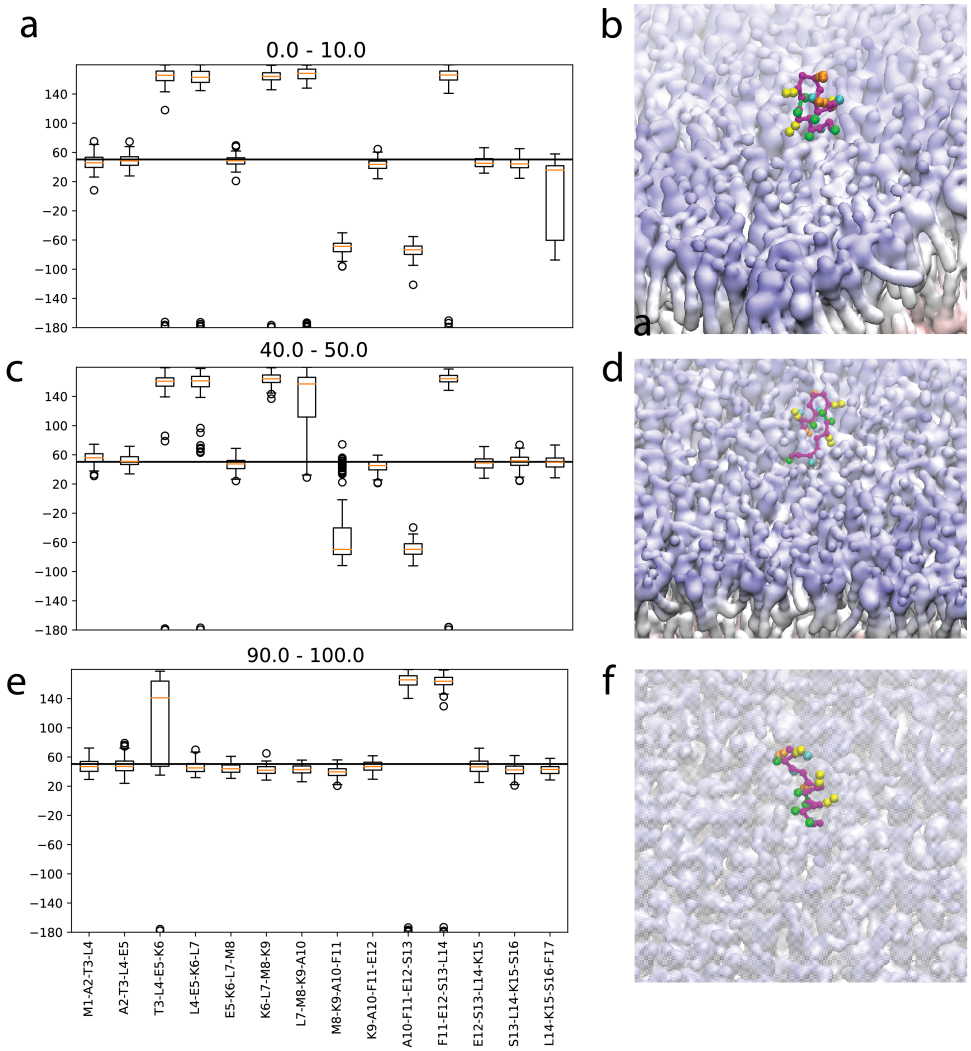

Figure 7.10: a, c, e - Boxplot for backbone dihedral angles, with median marked in orange for different trajectory slices. The black horizontal line corresponds to 50.4 degrees. (a) 0-10 ns; (c)40-50 ns; (e)90-100 ns; b - A representative snapshot corresponding to 0-10 ns; d - A representative snapshot corresponding to 40-50 ns; f - A representative snapshot corresponding to 90-100 ns Coloring scheme: Magenta:peptide backbone; Orange: Phenylalanine; sidechain; Green:Other hydrophobic groups sidechain; Cyan:Polar residue sidechains; Yellow: Lysine sidechains

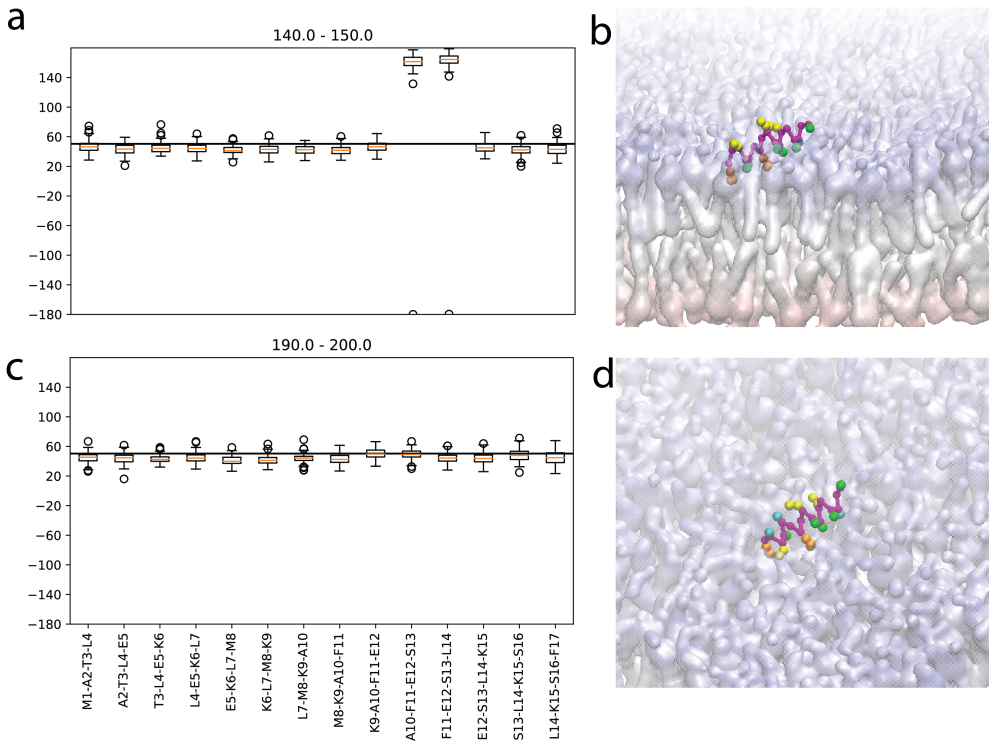

Figure 7.11: a, c - Boxplot for backbone dihedral angles, with median marked in orange for different trajectory slices. The black horizontal line corresponds to 50.4 degrees. (a) 140-150 ns; (c)190-200 ns; b - A representative snapshot corresponding to 140-150 ns; d - A representative snapshot corresponding to 190-200 ns Coloring scheme: Magenta:peptide backbone; Orange: Phenylalanine; sidechain; Green:Other hydrophobic groups sidechain; Cyan:Polar residue sidechains; Yellow: Lysine sidechains

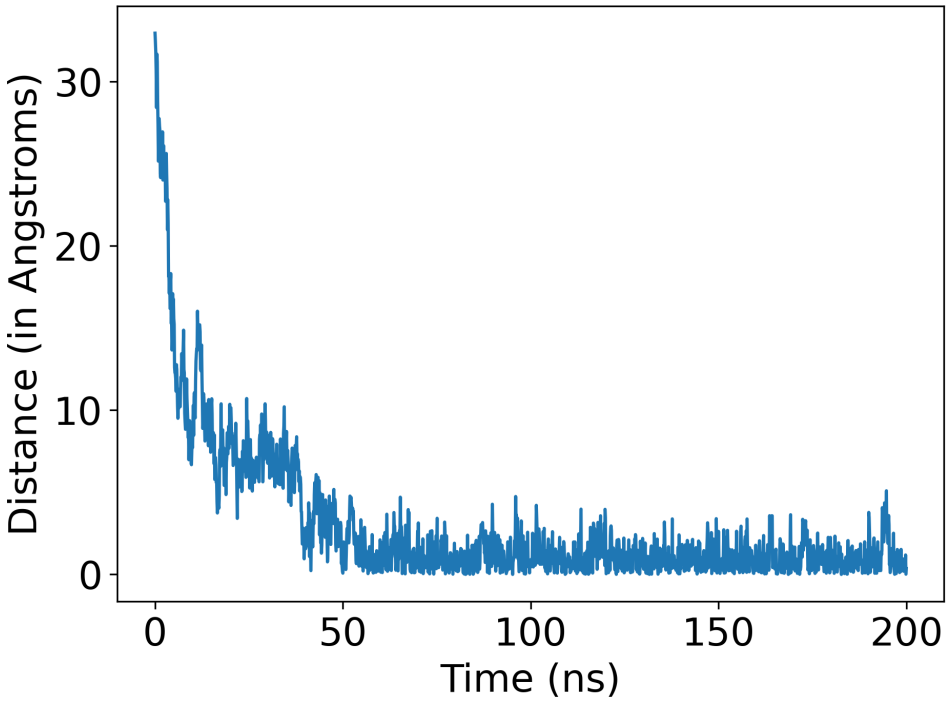

Figure 7.12: Normal distance between peptide center-of-mass and the center-of-mass of lipid headgroups.

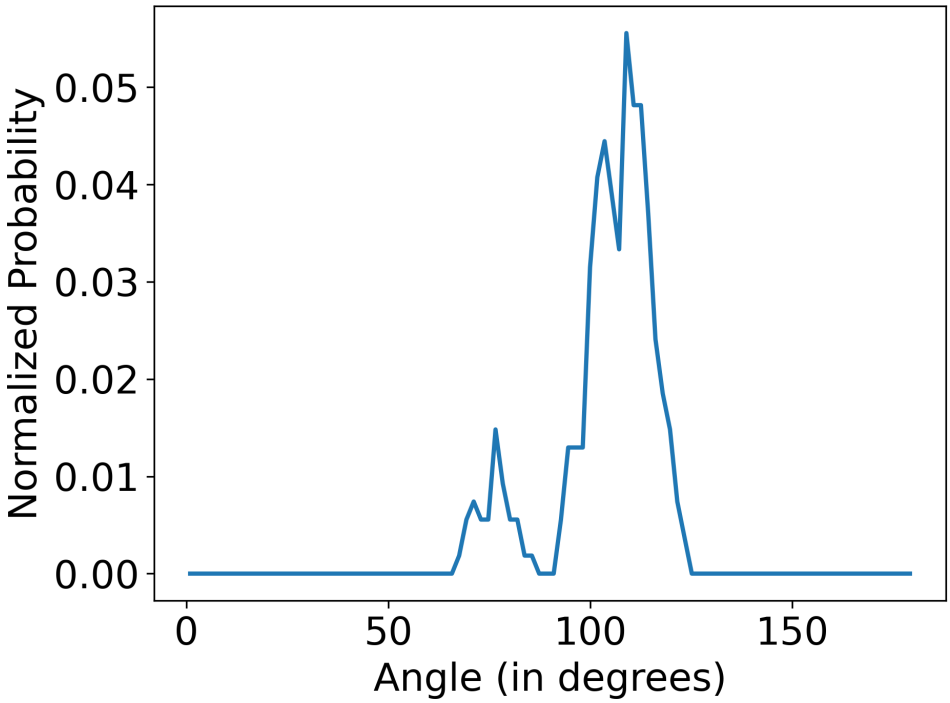

Figure 7.13: Distribution of helical-tilt with bilayer normal. Here the last 30 ns of trajectory has been used for calculation.

backbone makes with the bilayer normal, with the data from the last 30 ns (when the peptide has transformed into more structured α-helix). The peptide stays in an almost-lateral orientation with the lipid membrane, with hydrophobic groups facing the membrane and hydrophobic/polar groups more solvent exposed.

This more inserted C-terminus structure, and the relatively tilted (more than 90 degrees) orientation of N17 is noted in previous two dimensional solid state NMR spectra by Michaelek *et al.* [315]. In our work, we could note similar structures for membrane-associated N17. Moreover, we also delineate pathways for membrane association, which is led by the charged/polar N-terminal residues, and final membrane insertion and restructuring by the hydrophobic groups in the C-terminus.

These simulations also serve as a validation of the use of our CG model (ProMPT) towards peptide-membrane simulations; and membrane-assisted structural transitions.

7.4.4 Multi Peptide (N17) simulations in presence of a curved membrane

In the next phase, we extended our scope to study peptide aggregation in presence of a curved membranes. We started from peptides solvated in a grid-like fashion away from the membrane, with minimum peptide-membrane separation greater than 6 nm. The peptides could interact with each other, and the membrane in an unbiased manner. Three independent replicas were simulated for this biomolecular system for 300-350 ns. While, the number of absorbed peptides were different across the simulations, the observed biophysical mechanism were similar. We have added the VMD snapshot for replica 2 (Fig. A.1) and 3 (Fig. A.2) in Appendix A. The following analysis in this chapter is for the first replica simulation of N17 peptide in presence of

curved membrane.

Fig. 7.14-7.15 shows a gradual progression of the peptide from its solvated state to membrane-bound state, with the relevant lipid-peptide contacts (a) and the corresponding VMD snapshots (b). The peptides are absorbed primarily on top of the curved region as aggregates, and with some aggregates in solution. In another replica simulation, one of the solvated aggregate was in contact with the planar membrane, but most of the peptides still interacted exclusively with the curved region of the membrane. Here, we need to note that the absorbed peptides primarily reside on the curved region of the membrane, while 64 $\%$ of the available to interact membrane surface area is planar.

This shows the curvature sensing ability of N17 peptide previously discovered in the AFM study with nanospheres that enforced membrane curvature [60]. The initial interactions between the peptide and the membrane is driven through the polar and charged groups on the peptide (Fig. 7.14), with the peptides forming an aggregate in solution. The peptide aggregate consists of a hydrophobic core driven by the hydrophobic groups and the amphipathic face of N17. Over time (Fig. 7.14: 70-80 ns and 150-16 ns; Fig. 7.15: 200-210 ns and 330-340 ns, we note an increase in absorption with the larger PHE groups partitioning onto the membrane and interacting with the acyl core. Peptide-peptide interaction here can be between one aggregate solvated in solution, and another tethered to the membrane, or through peptides/aggregates laterally diffusing on top of the membrane. We noted similar trends for Aβ 16-22 peptide in Chapter 3.

To understand the structural features of the peptides absorbed on the membrane, in contrast to the aggregate in solution, we computed the distribution of each backbone dihedral angle (Fig. 7.16). Here, an angle of 50.4 degrees marks perfect α helix (Chapter 6). Overall, we note helical structures both in the solvated peptide aggregate and membrane-absorbed aggregate,

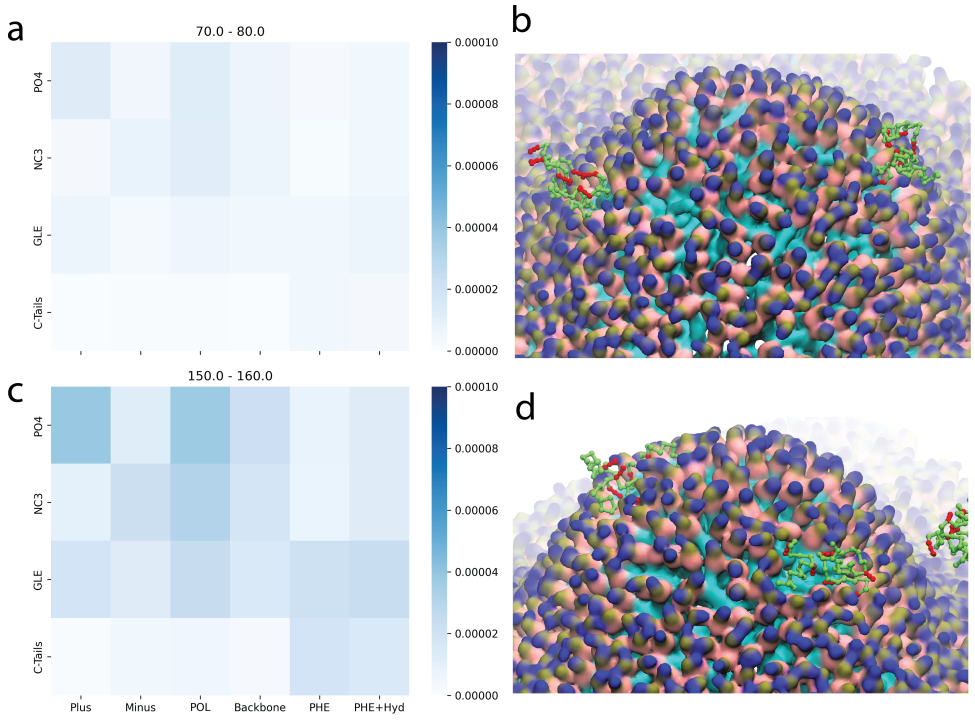

Figure 7.14: Characterization of membrane-peptide interactions at 70-80 ns (a/b) and 150-160 ns (c/d) : a,c-Contact map between peptide and membrane (NC3: Choline; PO4: Phosphate; GLE: Glycerol-Ester; Plus: Positive charged amino acid sidechain; Minus: Negative charged amino acid sidechain; POL: Polarizable amino-acid sidechains; Backbone: Peptide backbone; PHE: Phenylalanine sidechains; Hyd: Other hydrophobic group sidechains) c,d-A representative snapshot of peptide interacting with the membrane. Coloring scheme: blue/ochre/pink: membrane headgroup; cyan:acyl tail; green: peptide backbone; red: Phenylalanine

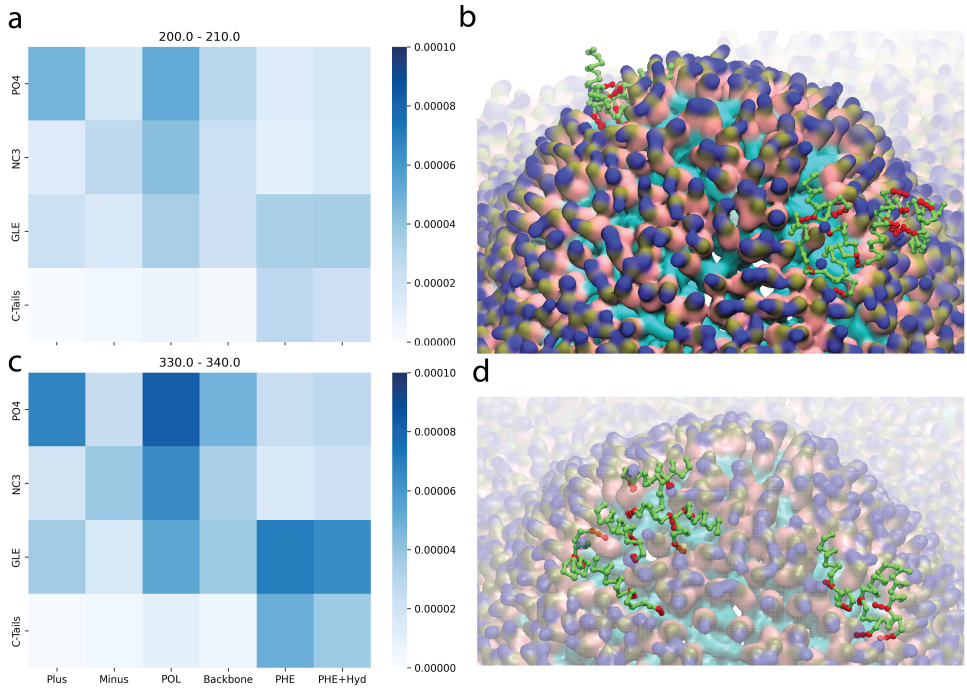

Figure 7.15: Characterization of membrane-peptide interactions at 200-210 ns (a/b) and 330-340 ns (c/d): a,b-Contact map between peptide and membrane (NC3: Choline; PO4: Phosphate; GLE: Glycerol-Ester; Plus: Positive charged amino acid sidechain; Minus: Negative charged amino acid sidechain; POL: Polarizable amino-acid sidechains; Backbone: Peptide backbone; PHE: Phenylalanine sidechains; Hyd: Other hydrophobic group sidechains) c,d-A representative snapshot of peptide interacting with the membrane. Coloring scheme: blue/ochre/pink: membrane headgroup; cyan:acyl tail; green: peptide backbone; red: Phenylalanine

with more disordered structures in solvated aggregates. Previous experiments have outlined the presence of helix-rich structures of N17 in their aggregate form. N17 forms amphipathic helical structures, with the hydrophobic face contributing to the hydrophobic core of the aggregate, and the hydrophilic side facing the solvent. This forces the initiation of peptide-membrane and aggregate-aggregate interaction through the hydrophilic residues.

On the membrane, although the helices are more structured, irregularities can be marked near A10-S13 region. We had previously noted the role of this fragment with peptide absorption onto the membrane. This can point to a somewhat bent helix being preferred on membranes. Early computational investigations by Kelly *et al.* with simulated tempering had observed such a population of bent double helix for N17 [317].

Fig. 7.17 shows the relative insertion of the peptides into the membrane, with distance calculated from a membrane surface constructed with phosphate headgroup. Only peptides with at-least one contact ($\leq$ 7 $\mathring{A}$) with the phosphate group are considered here. We observe the highest membrane insertion for PHE sidechains, followed by other hydrophobic groups. The least partitioned groups are the positive and negative charged sidechains. The polar backbone and sidechains of THR and SER are sandwiched between these extremes.

Fig. 7.18 shows the lipid density (Number of lipid headgroups per unit surface area) at 330 ns for a - simulation with N17 peptides; b - control curved membrane simulation without any peptides. The lipid density is lower in the curved membrane which results in peptides preferentially interacting with the curved region to engage with the membrane defects. Fig. 7.19 is the box-plot of area-per-lipid calculated at each time interval. We also mark the slow growth in these defects, with larger gaps at increased times when lateral absorption and aggregation of peptides is complete. This can also be noted in the long tailed distribution of the area-per-lipid

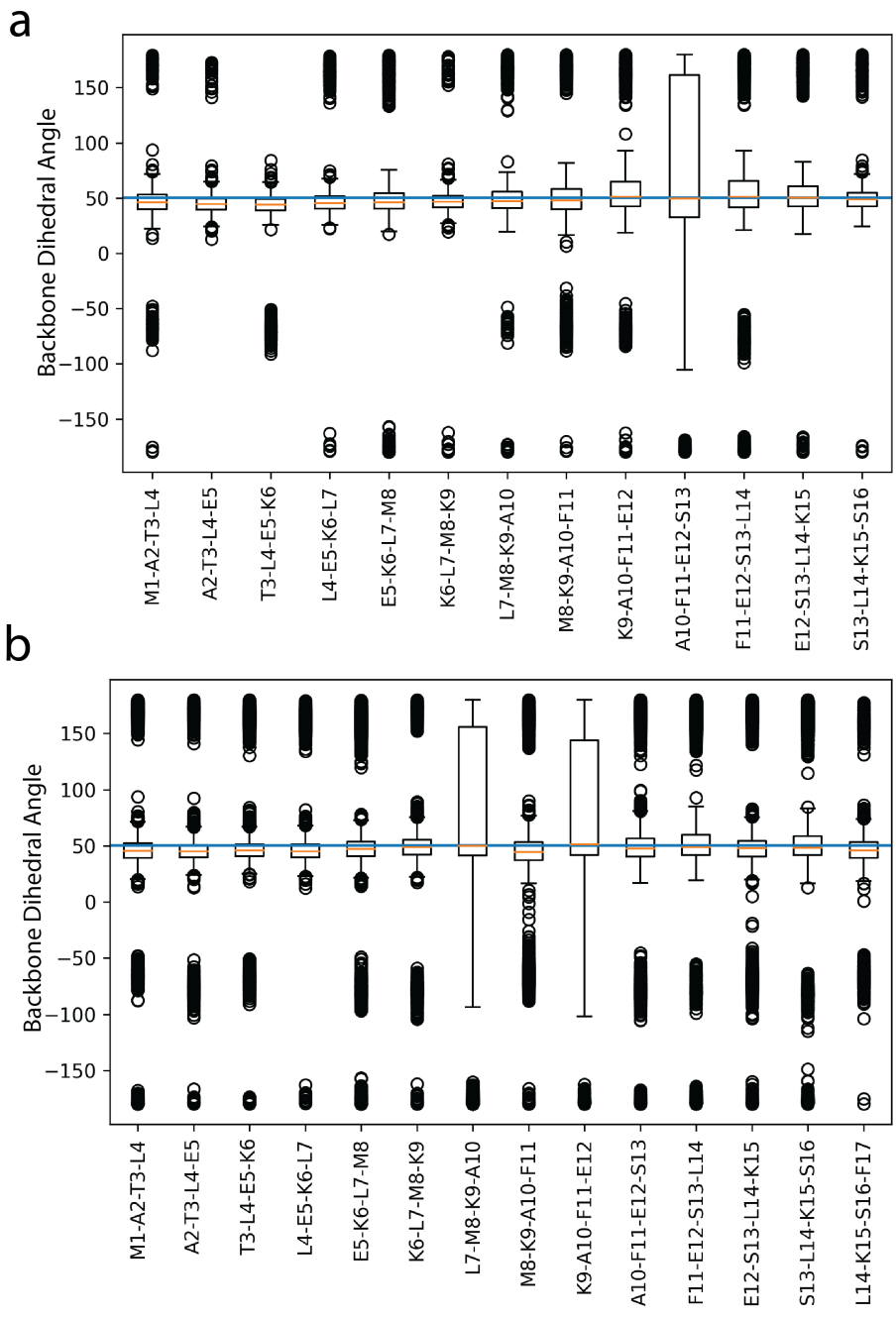

Figure 7.16: Boxplot marking distribution of peptide backbone dihedrals a-Peptides absorbed onto the membrane; b-Peptides as aggregates in solution

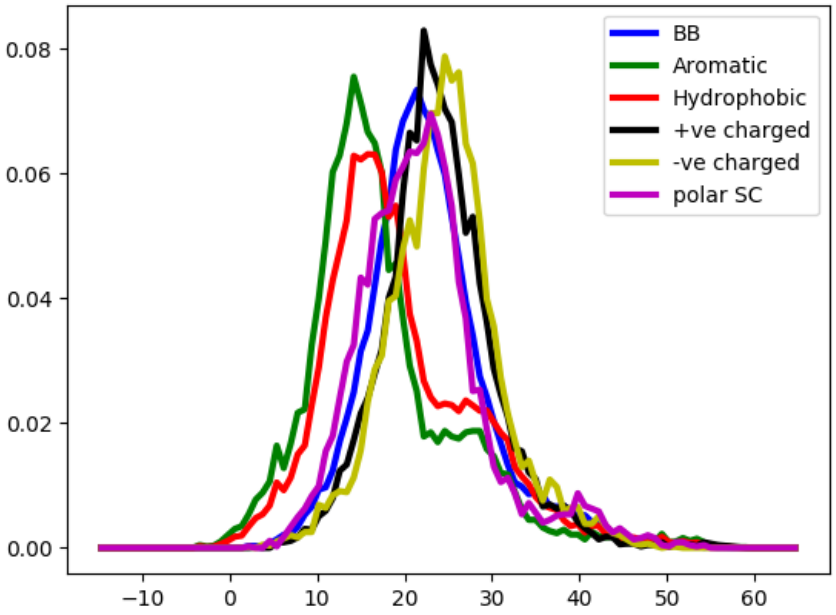

Figure 7.17: Relative distance of different groups from the membrane surface (determined by PO4). BB: Peptide Backbone; SC: Side-chain

between the control simulation and the simulation with peptide present (Fig. 7.20).

Finally, to further elucidate the role of bulky hydrophobic groups that we characterized as an essential part in membrane partitioning, we utilized a simulation of the F11L/F17L. A VMD snapshot of the membrane with the peptide aggregate is provided for 750 ns (Fig. 7.21). In our simulation, we did not observe a preference towards the curved membrane region here, with one peptide absorbed on the curved region compared to 6 on the planar region (Fig. 7.22). Instead, the peptide/peptide aggregate interacted both on the planar and the curved membrane. The peptides here, also could either partition onto the membrane as a single peptide or as a solvated aggregate. The helicity of the peptides are not affected (Fig. 7.23).

This could be because of the larger size of PHE groups, which makes them difficult to partition into membranes with smaller defects. Therefore, PHE prefers larger defects in the

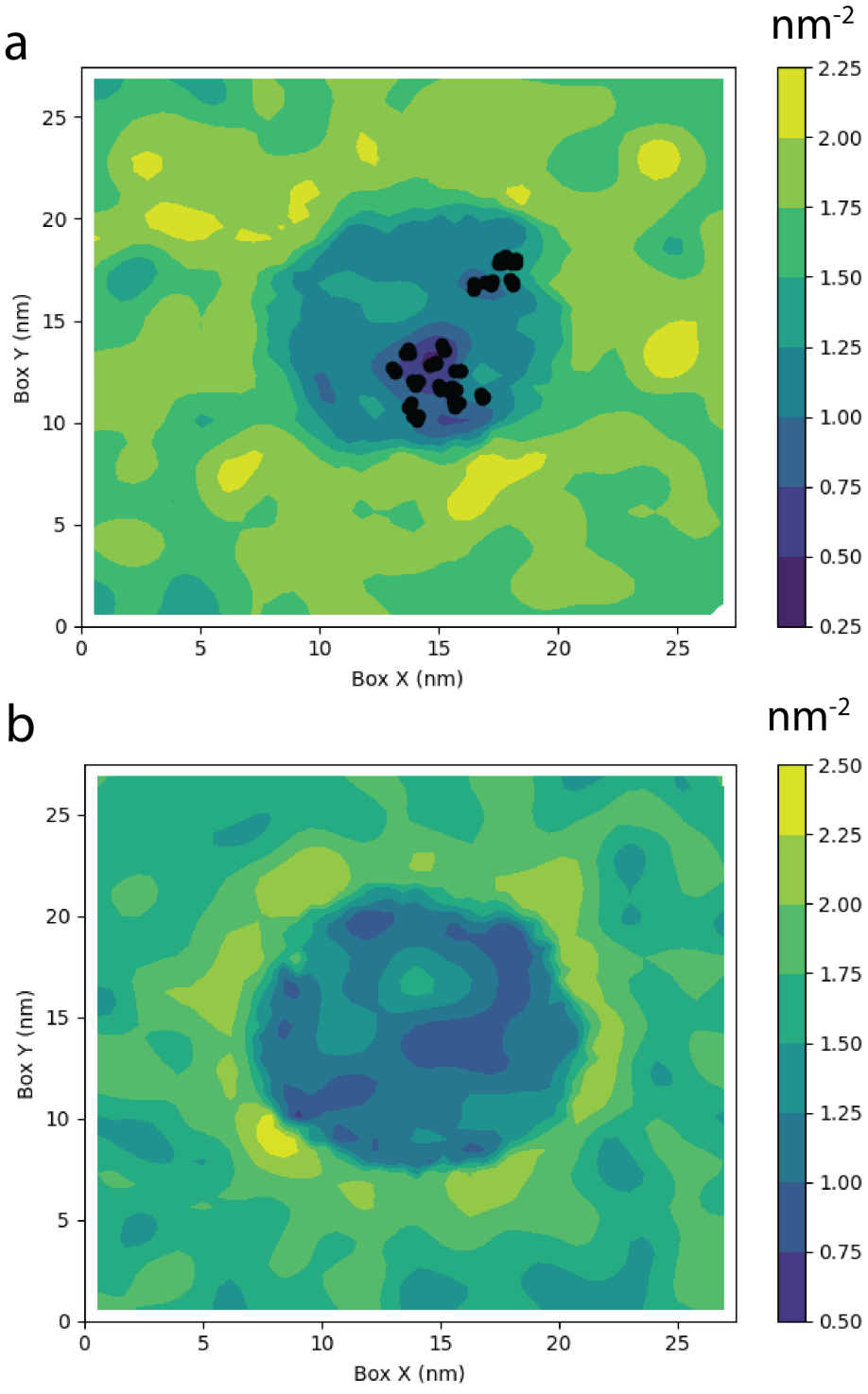

Figure 7.18: Snapshot of lipid density at 350 ns. a - Simulation with absorbed peptide. One PHE sidechain per peptide marked with black. b - Control simulation without peptides.

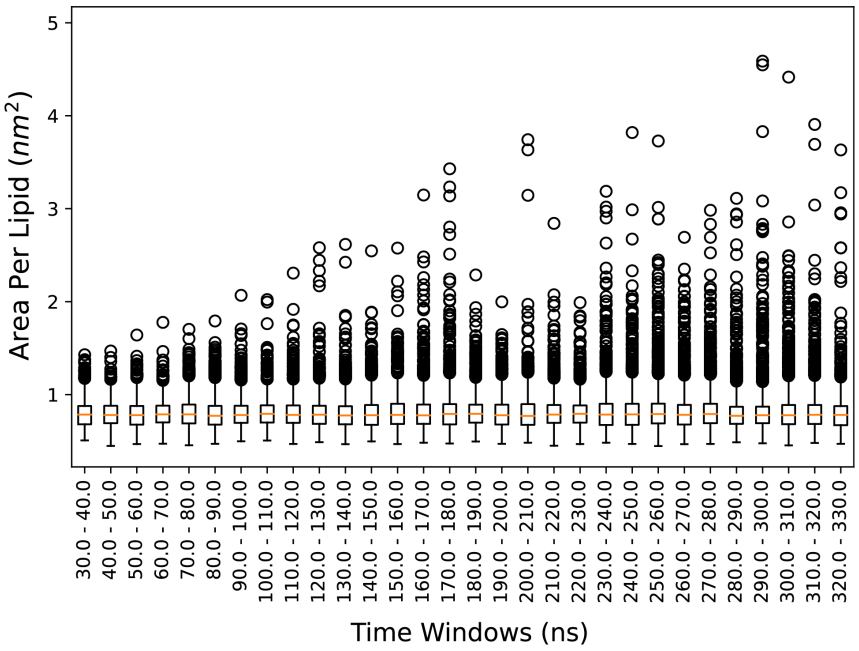

Figure 7.19: Boxplot of area per lipid over simulation time for the simulation with peptides on a curved membrane.

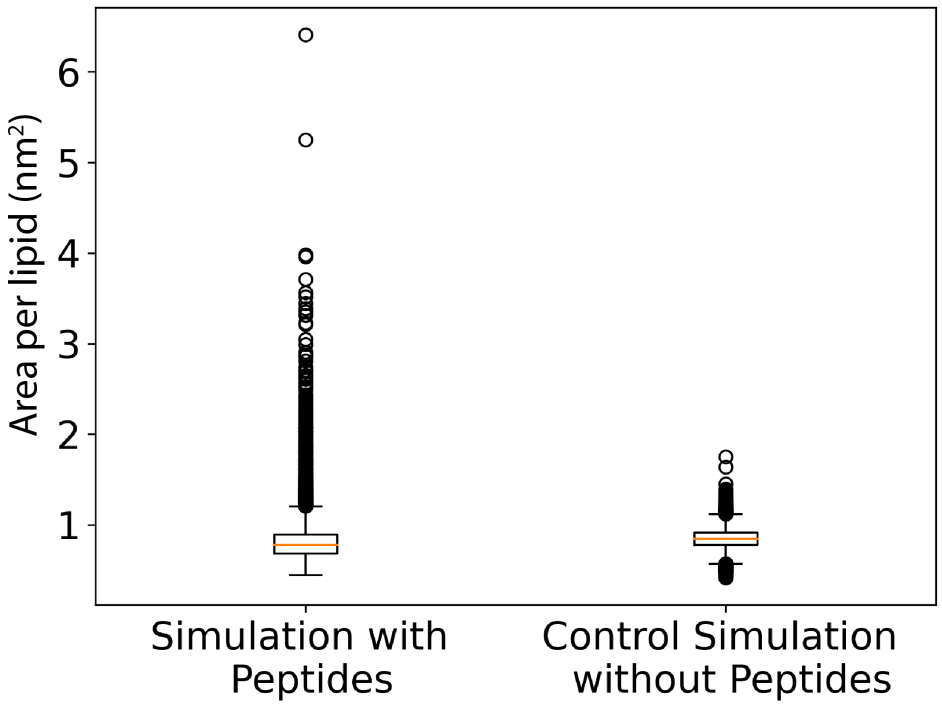

Figure 7.20: Comparison of area-per-lipid between the simulation with peptide and the control simulation without peptides.

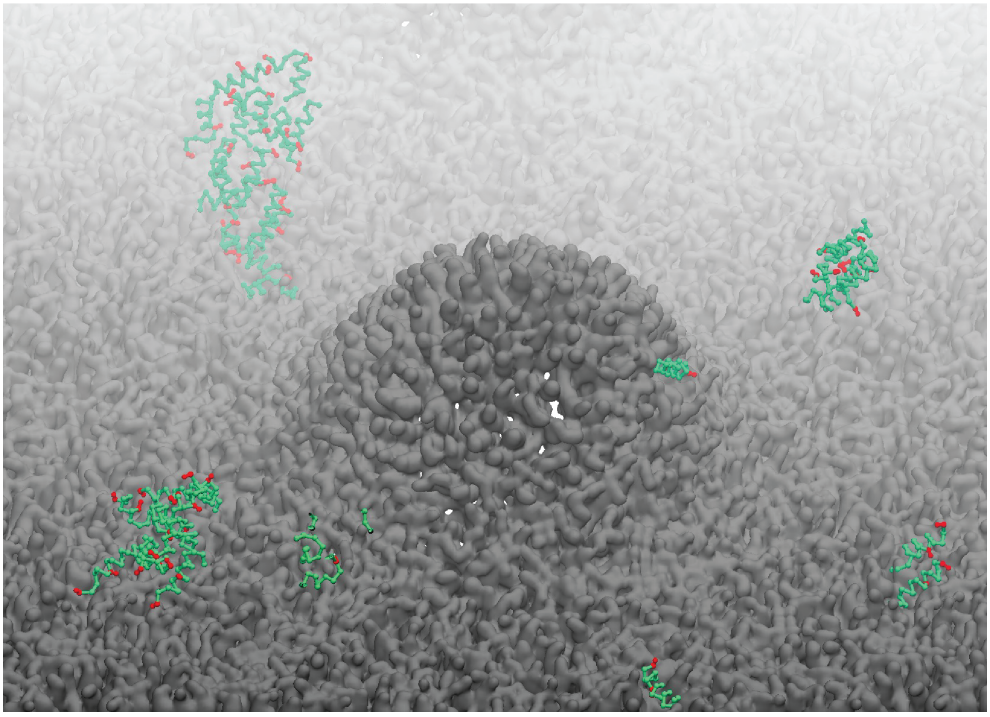

Figure 7.21: VMD snapshot at 750 ns for a F11L/F17L simulation. Color scheme - peptide backbone: Green; Leucines: Red; Membrane in surface representation: Black

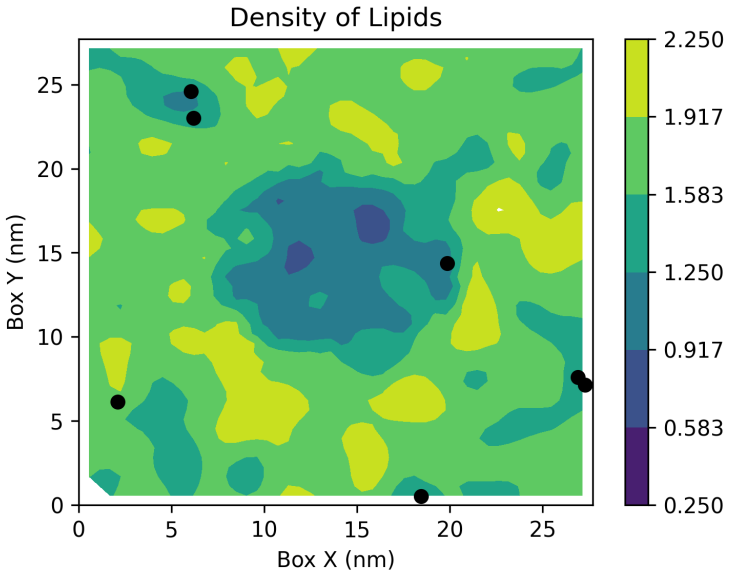

Figure 7.22: Density of lipid groups with center-of-mass of absorbed peptides marked in black.

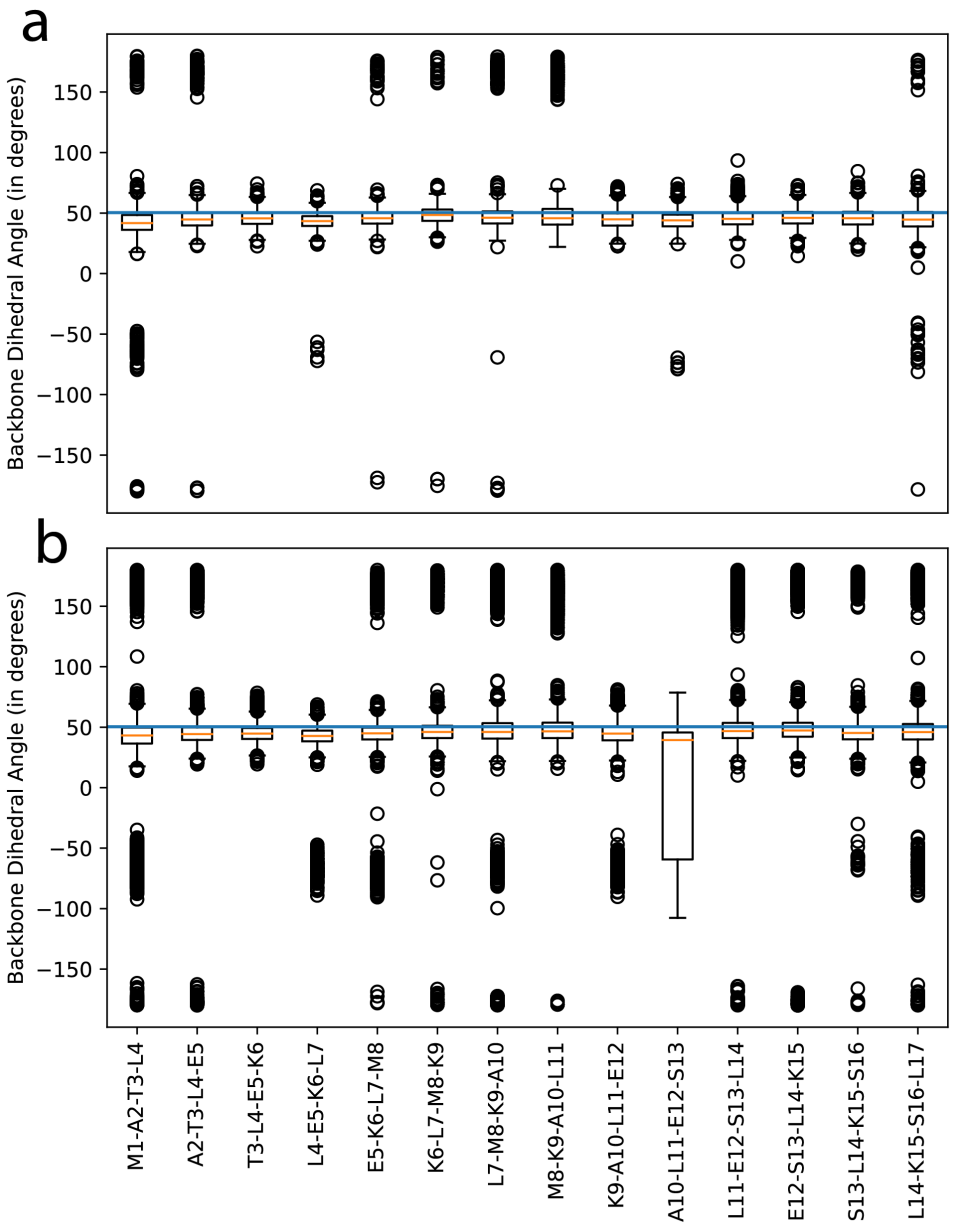

Figure 7.23: Boxplot of backbone dihedrals for a)peptides in contact with the membrane b) peptides not in contact with the membrane. The black horizontal line at 50.4 degrees marks backbone dihedral to match alpha helix structure.

curved region. On the other hand, the smaller LEU sidechain could partition into both the larger and smaller defects without any issues. This establishes the role of bulky groups such as PHE in curvature sensing of huntingtin protein. Two modes of amphipathic helices sensing curvature has been outlined in literature for smaller peptides — imbalance between polar/hydrophobic groups such as for α-synuclein and presence of bulky groups in the hydrophobic face for peptides such as ALPS and ArfGAP1 [168,328]. In this work, we position N17 into the second bin. This presence of PHE groups at the later part of N17 domain drives the curvature sensing nature of the peptide.

7.5 Conclusion

In this chapter, we focused on a slightly complex peptide sequence — Huntingtin Exon 1 (htt) with the updated CG model (ProMPT - that we developed in the previous chapter) to elucidate the role of flanking sequence and membrane curvature. Huntington's disease belongs to the larger class of diseases formed due to mutations in the poly-Q tract. The Huntingtin protein's N-terminal domain is characterized by an initial 17 residue N17 domain followed by a poly-Q tract. While previous research have presented a consensus overview of the structure and dynamics of N17 fragment, a similar overview of ther poly-Q tract is missing. There have been competing evidences of secondary structure in the poly-Q tract. Moreover, since this is a fast aggregating peptide, it is difficult to effectively discern between individual peptides and peptides in their aggregate form through structural studies. Here, we explored the conformations of N17-Qn (where n $\in \{7, 15, 35, 45\}$), both as a single peptide and peptides in their aggregate form. For the single peptide, we noted a trend towards more globular (decreasing asphericity) peptide structures with increasing lengths of polyglutamine tract. We also observed a significant

specificity towards Q-Q interactions also observed in previous experimental reports. Finally, we also found a gradual change from helical-rich structures to increasing sheet-like features with increase in number of trailing GLN residues. For the case of peptide aggregates, we found that the GLN-repeats drove peptide aggregation. The aggregate core was made with the GLN-GLN interactions that disrupted the number of N17-N17 contacts, with N17 situated more at the periphery of the aggregate. In another thread here, we also studied the structural features of the POPC membrane and N17 peptide, that drives the curvature sensing of the Huntingtin protein. Previous research on peptide-membrane interactions has been limited both from a experimental and computational perspective. While some structural features of N17-membrane interaction are known, detailed biophysical reports of peptide-membrane interactions are still missing. In this chapter we aim to understand these biomolecular details of peptide-membrane interaction and the molecular mechanism of curvature sensing. Our simulations suggest a gradual progression from a dominantly unstructured peptide in solution, to a more structured α-helical patch on the membrane. In presence of a curved membrane, we observed that N17 peptides preferentially interact with the curved region. While the polar and charged groups drive the initial membrane-peptide interaction, the membrane curvature sensing is dominantly controlled by the bulky hydrophobic groups such as PHE. Replacing these bulky residues from the sequence leads to loss of this curvature sensing ability of N17.

Chapter 8: Thesis Summary

The objective of this thesis was to understand various aspects of protein aggregation associated with neurodegenerative diseases. In particular, I explored how physiological biomolecules such as the presence of a membrane or solvated small bio-molecules and genetic mutations can affect/alter aggregation behavior. As case studies, I focused on Amyloid-beta (Aβ, associated with the Alzheimer's disease) and Huntingtin (htt, associated with the Huntington's disease). In this work, I explored various modalities on how protein aggregation is impacted by external stimulus such as membrane headgroup charge, applied surface-tension, hyperglycemic conditions, membrane curvature and added trailing GLN repeats.

We approached this research through the lens of computer simulations. Molecular dynamics is a class of computational methods that use statistical physics based underpinnings to present a description of a biomolecular system. Here, we represent a molecule as a collection of sites, that interact among themselves through a defined potential function. This generates a trajectory of moving sites, which map to the motions in that molecule. Molecular dynamics simulations aim to complement biomolecular experiments by addressing certain limitations that experiments pose such as structural transience and competing explanations of experimental observations. Among classical molecular dynamics simulation methods, all-atom simulations are the most commonplace. Here, all the atoms (or all heavy atoms) in a biomolecular system are represented

as interaction sites. But, this method can be inaccessible for studying larger spatiotemporal systems. To address this, we have followed a more reduced resolution approach of spatial coarse-graining. Here, we spatially group atoms into larger interaction site, and define interaction potential between them. This allows for a significant reduction in computational effort, allowing studies of complex and large systems.

In this thesis I modified and applied the coarse-grained model that was developed in our lab — Water Explicit Polarizable Coarse-Grained Model (WEPCGM) to study the effects of membrane headgroup charge, applied surface-tension and hyperglycemia on Aβ 16-22 aggregation. In chapter 2, I introduced the WEPCGM and detailed its properties. The coarse-grained model features explicit representation of the environment, that allows for studying environment assisted structural changes. The presence of explicit solvent also aids in creating a transferable forcefield without a need for reparameterization in order to study another biomolecular system. The current version of this forcefield consists of two segments — Water Explicit Polarizable MEMbrane Model (WEPMEM) and Water Explicit Polarizable PROtein Model (WEPPRO) [80–82]. In the next chapters (Chapter 3, 4 and 5) we applied WEPPROM in conjunction with WEPMEM towards presenting a biophysical picture of Aβ 16-22 aggregation.

In chapter 3, we used WEPCGM to to investigate the effect of lipid headgroup charge – zwitterionic (1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine: POPC) and anionic (1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine: POPS) – on Aβ 16–22 peptide aggregation. Our coarse grained molecular dynamics simulations provided a mechanistic explanation to a broad spectrum of experimental results [25, 26, 159, 164, 165]. We characterized multiple pathways for peptide absorption into membranes composed of POPC and POPS. Both lipid molecules have distinct effects on aggregation patterns of absorbed peptides. While rapid cumulative aggregation (ordered

+ disordered) was observed in zwitterionic PC bilayer, the emergence of ordered beta sheets and by extension, fibrillation was faster in presence of anionic POPS lipids. The results are in agreement with previous experimental studies [25, 137, 139–142, 164, 165] that had observed faster growth of amyloid fibrils in presence of anionic lipids. The discrepancy in the cumulative aggregation rates is a consequence of faster lateral diffusion of POPC lipid molecules compared to POPS. On the other hand, increased beta sheet content in POPS membranes is due to the differences in membrane compressibility. Higher membrane compressibility of POPC membrane compared to POPS, results in a relatively higher peptide insertion into the bilayer. This distorts the geometry of individual peptide molecules which hinders their participation in beta sheet formation. Some of the morphological aspects of membrane assisted Aβ aggregation reported in this study such as relatively higher membrane insertion of F19 compared to F20, have been supported by previous experimental evidences [159]. We also revealed the propensity of initial oligomeric deposits to act as nucleation seeds to enhance further fibrillation. Considering the presence of multiple, ordered aggregates in POPS due to slow cumulative aggregation and peptide aggregates operating as nucleation seeds, POPS membranes can have an increased progressive peptide aggregation rates. This work unravels how lipid headgroup driven biochemical interactions in homogeneous model membranes shape peptide absorption and aggregation.

In chapter 4, we leveraged molecular simulations to address and understand relationships between curved cellular membranes and aggregation of a model template peptide Aβ 16-22. Experimental results have suggested that both membranes and peptide can influence each other's structure and dynamics. Membrane curvature can promote extensive fibrillation, with faster formation of ordered structures in presence of small unilamellar vesicles, compared to larger ones. On the other hand, peptide aggregates can alter membrane's structure and packing. In

this work we investigated these intertwined effects with coarse-grained molecular simulations. Our results agree with previous observations of curvature-driven peptide aggregation into ordered structures, and suggest possible biophysical mechanisms for it [27,28]. Membranes with increased curvatures lead to increased peptide absorption, due to more exposed hydrophobic defects. The absorbed peptides can then laterally diffuse around, and interact through the peptide backbone to form ordered fibrillar structures. The presence of such peptide aggregates also affects the lipid membrane's local structure. The lipid groups closer to the aggregates are more ordered at the crowded interfacial region due to interfacial presence of peptides; and less ordered deep inside the membrane core. These observations align with previous experiments and molecular simulations that also highlighted membrane disruption by peptides [179, 180]. Our results supported the previously suggested "carpeting model" of membrane disruption, by increasing membrane fluidity inside the membrane core [166, 181]. These locally close lipids also have a broad distribution in $P\vec{N}$ vector tilt from the membrane normal, suggesting the heterogeneous nature of peptide-lipid interactions. Such local variations in $P\vec{N}$ vector tilt can manifest in changes to the membrane's electrostatic potential, ion distributions and alter electrostatics-assisted membrane signalling [182, 183]. This study unravels the effects of membrane curvature in aiding peptide aggregation, and the effects of peptide aggregation in reshaping membrane's local attributes.

In chapter 5, we explored the mechanisms through which solvated glucose molecules affect Aβ 16-22 aggregation in a concentration dependant manner. Previous reports have drawn correlations between pathogenesis of Alzheimer's disease and hyperglycemic conditions associated with type 2 diabetes. Experimental research has proposed chemical crosslinking, in presence of glucose molecules as responsible for increased toxicity. But the time-scale separation between increased aggregation and formation of chemical cross-links suggests the presence of an alternate

thermodynamic mechanism, which we explored in this chapter. We applied coarse-grained molecular dynamics simulations to study the how glucose molecules can drive Aβ 16-22 aggregation. We noted that that increasing the concentration of solvated glucose molecules can result in faster aggregation of Aβ 16-22, without any appreciable change in β-sheet content. This change in aggregation rates can be explained in terms of relative orientations of interfacial glucose molecules. The glucose molecules at the peptide-water interface feature a preferential orientation, resulting in loss of rotational entropy in a concentration dependent manner. This can assist in faster aggregation of the peptide, to reduce the cumulative availability of solvent accessible surface-area.

In chapter6, we revisited our coarse-grained model again and outlined its limitations. Firstly, WEPPROM does not have representation for all the essential amino-acids. Secondly, it has not been studied and validated for tertiary folding protein structures. To address these factors, we developed a transferable CG forcefield (Protein Model with Polarizability and Transferability - ProMPT) with an explicit representation of the environment for accurate simulations with proteins. The forcefield consists of a set of pseudo-atoms representing different chemical groups that can be joined/associated together to create different biomolecular systems. This preserves the transferability of the forcefield to multiple environments and simulation conditions. We added electronic polarization that can respond to environmental heterogeneity/fluctuations and couple it to protein's structural transitions. The non-bonded interactions are parametrized with physics-based features such as solvation, and partitioning free energies determined by thermodynamic calculations and matched with experiments and/or atomistic simulations. The bonded potentials are inferred from corresponding distributions in non-redundant protein structure databases. We presented validations of the CG model with simulations of well-studied aqueous protein systems

with specific protein fold types- TRP-cage, TrpZip4, Villin, WW-domain and β-α-β. We also explored the applications of the forcefield to study aqueous aggregation of Aβ 16-22 peptides and dimerization of glycophorin A (GPA) in micellar environments.

With this model we primarily address the shortcomings of WEPPROM. ProMPT now has access to all the amino acid representation. It has updated bonded and non-bonded interaction parameters with which we validated the conformational landscape of several small proteins and aggregates.

In chapter 7, I focused on a more complex peptide sequence — Huntingtin protein's N-terminal domain (htt) with the updated CG model (ProMPT) to elucidate the role of varying lengths of trailing sequence and membrane curvature. Huntington's disease belongs to the larger class of diseases formed due to mutations in the poly-Q tract. The htt is characterized by an initial 17 residue N17 domain followed by a poly-Q tract. While previous research have presented a consensus overview of the structure and dynamics of N17 fragment, a similar overview of ther poly-Q tract is missing [315]. There have been competing evidences of secondary structure in the poly-Q tract [304, 318]. Moreover, since this is a fast aggregating peptide, it is difficult to effectively discern between individual peptides and peptides in their aggregate form through structural studies. Here, we aim to explore the conformations of N17-Qn (where n $\in$ 7, 15, 35, 45), both as a single peptide and peptides in their aggregate form.

For the single peptide, we could mark a trend towards more globular (decreasing asphericity) peptide structures with increasing lengths of polyglutamine tract. We also observed a significant specificity towards Q-Q interactions also observed in previous experimental reports. Finally, we also found a gradual change from helical-rich structures to the growth of sheet-like features with increase in number of trailing GLN residues. For the the case of peptide aggregates,

we could establish that the GLN-repeats drove peptide aggregation. With longer poly-Q tract, the balance between hydrophobic interactions and polar interactions changes with GLN-GLN contacts creating a polar core of the aggregate.

In another thread here, we also studied the structural features of the POPC membrane and N17 peptide, that drives the curvature sensing of the Huntingtin protein. Previous research on peptide-membrane interactions has been limited both from a experimental and computational perspective. While some structural features of N17-membrane interaction are known, detailed biophysical reports of peptide-membrane interactions are still missing [60, 315]. In this chapter we aim to understand these biomolecular details of peptide-membrane interaction and the molecular mechanism of curvature sensing.

Our simulations propose a gradual progression from a dominantly unstructured peptide (with some helicity) in solution, to a more structured α-helical patch on the membrane. In presence of a curved membrane, we observed that N17 peptides preferentially interact with the curved region. Here, we noted that while the polar and charged groups drive the initial membrane-peptide interaction, the membrane curvature sensing is dominantly controlled by the bulky hydrophobic groups such as PHE. Omission of these bulky residues from the sequence leads to loss of this curvature sensing ability.

In summary, this thesis examined aggregation of neurodegenerative peptides associated with Alzheimer's disease (Aβ) and Huntington's (htt) disease and how this self-assembly is impacted by external factors. These factors can change both the structure and kinetics of peptide aggregation. In the case of both Aβ 16-22 and N17, the initial aggregation was driven by the hydrophobic core of the peptide. But the presence of a membrane initiated a competition between peptide-peptide and peptide-membrane interactions. In the case of Aβ 16-22 this membrane

interaction resulted in fibrillar structures, while in the case of N17 the peptide conformed into structured amphipathic interfacial helices. This suggests the ability of the membrane-water interface to render particular secondary structures and tertiary arrangements to these intrinsically disordered fragments.

In addition, we also discovered the importance of bulky hydrophobic groups in defining overall organization of the aggregate. For Aβ 16-22, these groups were essential for engaging with the membrane core, thereby allowing peptides to initiate contacts through the peptide backbone. On the other hand, for htt, these bulky groups were instrumental in orienting the peptides appropriately on membrane surface and for curvature sensing. These groups can be essential targets for rational design of therapeutics. For both N17 and Aβ 16-22, the aggregate core is primarily driven by the hydrophobic groups. In contrast, the presence of poly-Q segment initiated another set of strong competing interactions through the trailing GLN. These led to these GLN-repeats replacing the hydrophobic groups in forming the primary aggregate core at longer lengths of poly-Q tract, similar to a reverse micelle structure. This increased GLN repeats formed highly stable β-sheet rich structures, contrary to previously unstructured, and non-specific hydrophobic core.

Beyond specific interactions with the membrane, we also curated the effects of non-specific interactions that solvated small molecules have on aggregation rates. Here, we observed a concentration dependent effect driven by a specific way in which glucose molecules partition onto the aggregate surface.

Another contribution of this thesis is the development of a polarizable coarse-grained model for simulation of proteins in explicit environment. The key aspect of this forcefield is that it is transferable across biomolecular systems and does not require fitting to explicit biomolecular systems. This allows to decouple and study the effects of the environment on

biophysical processes. As such, it can be a adopted by the structural and molecular biophysics community towards large spatiotemporal studies of biomolecular systems of increased complexity.

8.1 Future Work

In this thesis we explored the effects of model membranes made with a single type of lipids (either 100% POPC or 100% POPS) to study peptide aggregation. To present a somewhat realistic scenario, the next objective can be to extend this study to binary mixture of lipid groups at varying ratios. This would present an *in-depth* overview of gradual introduction of membrane charge on aggregation behavior.

Beyond charge on lipid molecules, solvated divalent cations have been shown to affect peptide aggregation directly. Studies with surface plasmon resonance (SPR) and atomic force microscopy (AFM) have suggested the role of calcium ions in bridging Glutamate (residue 22) and the phosphate of lipids, anchoring the peptide to membrane-surface and driving aggregation [329]. Also, the local presence of calcium has been known to induce mesoscopic changes in bilayer organization and geometry. The binding of calcium ion with the model membranes such as vesicles has been investigated primarily with thermodynamic, microscopic, and spectroscopic techniques [330–335]. Several reports suggest lateral reorganization of lipid mixtures as a response to the presence of calcium ions. Doosti et al. used the localized application of calcium ions on vesicles enriched with negatively charged lipids to generate tubular protrusions [333]. Recently Graber et al. showed the development of negative curvature as a consequence of anionic lipid–phosphatidylinositol-4,5-bisphosphate (PI(4, 5)P2) and phosphatidylserine (PS) clustering in the presence of calcium ions [334]. Therefore, presence of calcium can result in local enrichment and

curvature changes, which in-turn can alter membrane's interaction pattern with peptides. These physico-chemical changes of lipid molecules in response to calcium ions can therefore lead to differential peptide aggregation properties.

In addition, investigations into a more realistic ternary — membrane, Aβ 16-22 peptide and calcium ions system is limited due to associated computational complexities. CG-MD can be an essential tool to investigate the complex effects of membrane assisted peptide aggregation in presence of calcium ions. We have already developed a coarse-grained model for calcium ions to work with WEPMEM [174], that can capture calcium driven demixing of binary lipid mixture (POPC and POPS). We plan to apply these coarse-grained models towards elucidating the microscopic interactions that shape aggregation in such complex systems.

The studies of peptide self-assembly, in this thesis, have always considered crowding biomolecules in isolation. The specific and non-specific interactions common in crowded systems such as an extra-cellular matrix (ECM) would be a more reasonable description of the fibrillation process. Hence, there is a great deal of interest in studying the aggregation process in with a realistic model for that environment. Recent reports have suggested of a particular biomolecule — chitosan (CHT), that can act as a mimetic for the extra-cellular matrix [336, 337]. Chitosan is a linear co-polymer composed of β-(1, 4) linked N-acetylglucosamine (GlcAc) units and glucosamine (GlcN) units [337].

A complete mechanistic picture of CHT self-assembly and CHT-Aβ interactions is missing. Molecular simulations can be an useful approach to connect microscopic level details to macroscopic statistical properties. Recent efforts in this direction has yielded several insights into atomic details of CHT self-assembly. But the associated length scales have made it difficult to study CHT self-assembly processes with atomistic molecular dynamics. Our lab has recently published

a coarse-grained model for CHT molecule that can capture both atomic micro-interactions and macroscopic mechanical properties of CHT hydrogels [338]. This positions us to study CHT-Aβ interactions by decoupling individual effects such as specificity of monomeric interactions and effects of polymeric structures.

For Huntington's disease, we explored the curvature sensing nature of Huntington's disease and the impact different lengths of poly-Q fragment has on peptide aggregation. Beyond that, we plan to study how poly-Q repeats could affect membrane interactions and membrane assisted aggregation. Moreover, we plan to increase the peptide length to study proline rich domains that have been known to have specific impacts on inhibiting peptide aggregation. Future plans also involve studying similar other amphipathic helices to provide an in-depth perspective of curvature sensing.

The coarse-grained models developed in our lab is always under constant updates. For the protein model, we are planning appropriate modifications to side-chain size, can achieve adequate protein-core packing required for correct tertiary structures. Along other directions, the model is currently being used to understand ion-assisted folding and unfolding of proteins due to explicit representation of the environment and transferability of interaction potential. The membrane model is also currently being expanded to include several other lipid and sterol varieties. Moreover our lab is also working on creating a representation for non-traditional solvents such as ionic liquids, as models for the environment.

Appendix A: Results From Other Replica Simulations of N17 with Curved Membrane

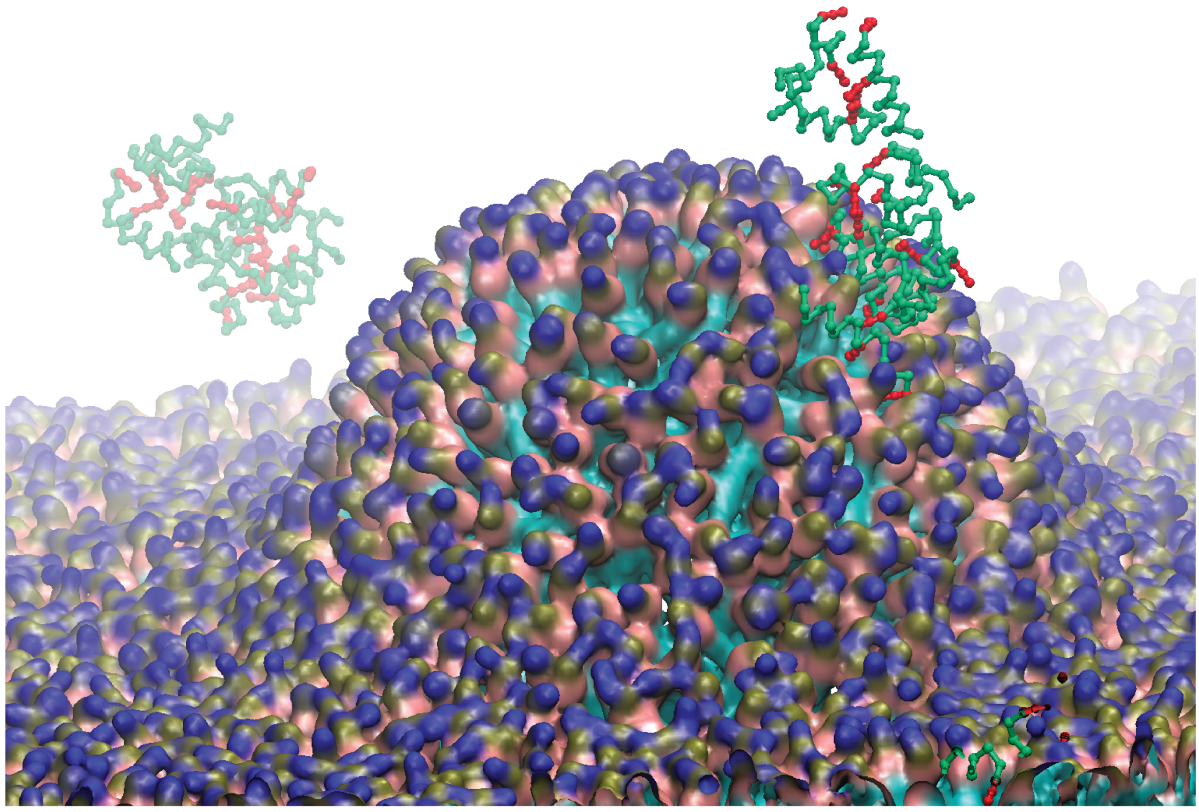

Figure A.1: A representative snapshot of peptide interacting with the membrane for replica-simulation 1. Coloring scheme: blue/ochre/pink: membrane headgroup; cyan:acyl tail; green: peptide backbone; red: Phenylalanine

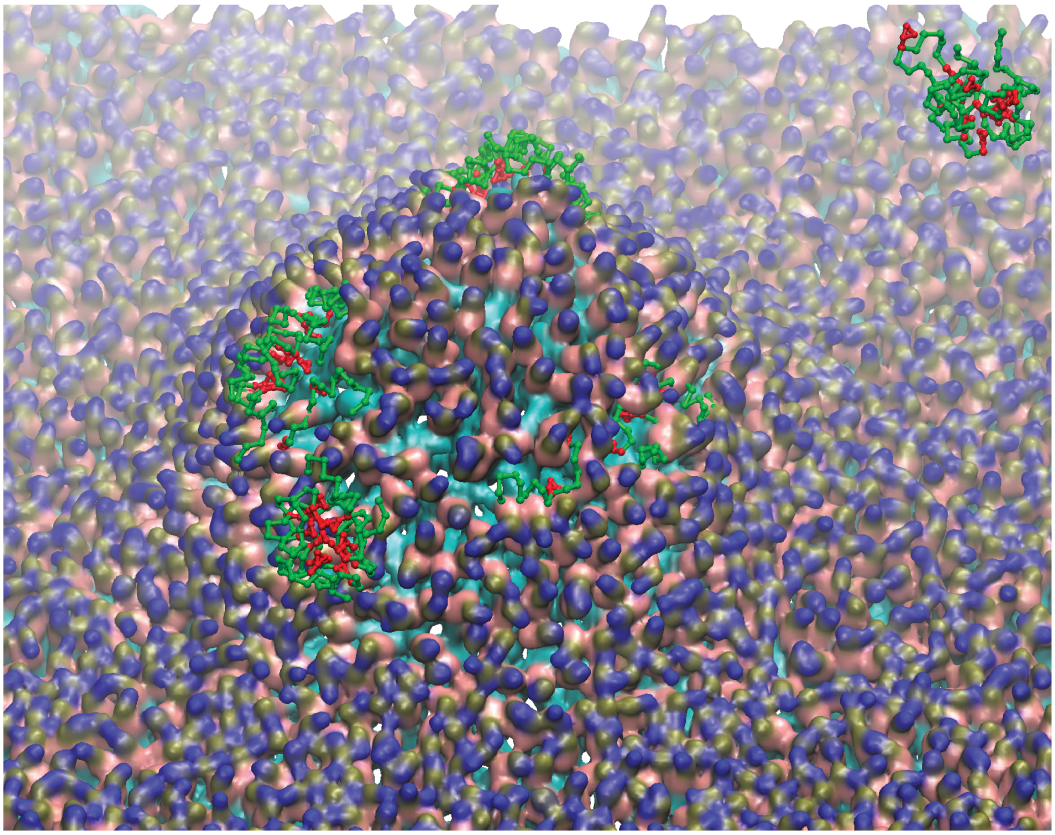

Figure A.2: A representative snapshot of peptide interacting with the membrane for replica-simulation 2. Coloring scheme: blue/ochre/pink: membrane headgroup; cyan:acyl tail; green: peptide backbone; red: Phenylalanine

Bibliography

[1] Christopher A Ross and Michelle A Poirier. Protein aggregation and neurodegenerative disease. *Nat. Med.*, 10(S7):S10–S17, 2004.

[2] Meenakshi Verma, Abhishek Vats, and Vibha Taneja. Toxic species in amyloid disorders: Oligomers or mature fibrils. *Ann Indian Acad Neurol*, 18(2):138–45, 2015.

[3] Fernanda G De Felice and Sérgio T Ferreira. Beta-amyloid production, aggregation, and clearance as targets for therapy in Alzheimer's disease. *Cell. Mol. Neurobiol.*, 22(5-6):545–63, 2002.

[4] William L Klein, Grant A Krafft, and Caleb E Finch. Targeting small Aβ oligomers: the solution to an Alzheimer's disease conundrum? *Trends Neurosci*, 24(4):219–224, 2001.

[5] Dominic M. Walsh and Dennis J. Selkoe. Aβ Oligomers — a decade of discovery. *J. Neurochem.*, 101(5):1172–1184, 2007.

[6] Hai Lin, Rajinder Bhatia, and Ratneswar Lal. Amyloid β protein forms ion channels: implications for Alzheimer's disease pathophysiology. *FASEB J.*, 15(13):2433–2444, 2001.

[7] Sara M. Butterfield and Hilal A. Lashuel. Amyloidogenic Protein-Membrane Interactions: Mechanistic Insight from Model Systems. *Angew. Chem.*, 49(33):5628–5654, 2010.

[8] N Arispe, H B Pollard, E Rojas, Rushana Azimova, Douglas Ng, Blas Frangione, Bruce Kagan, Jorge Ghiso, and Ratnesh Lal. Giant multilevel cation channels formed by Alzheimer disease amyloid beta-protein [A beta P-(1-40)] in bilayer membranes. *Proc. Natl. Acad. Sci. U.S.A.*, 90(22):10573–7, 1993.

[9] T. R. Serio, A G Cashikar, A S Kowal, G J Sawicki, J J Moslehi, L Serpell, M F Arnsdorf, and S L Lindquist. Nucleated Conformational Conversion and the Replication of Conformational Information by a Prion Determinant. *Science*, 289(5483):1317–1321, 2000.

[10] Hyunbum Jang, Fernando Teran Arce, Srinivasan Ramachandran, Ricardo Capone, Rushana Azimova, Bruce L Kagan, Ruth Nussinov, and Ratnesh Lal. Truncated beta-amyloid peptide channels provide an alternative mechanism for Alzheimer's Disease and Down syndrome. *Proc. Natl. Acad. Sci. U.S.A.*, 107(14):6538–43, 2010.

[11] Anne M. Brown and David R. Bevan. Molecular Dynamics Simulations of Amyloid β -Peptide (1-42): Tetramer Formation and Membrane Interactions. *Biophys. J.*, 111(5):937–949, 2016.

[12] Jo V. Rushworth and Nigel M. Hooper. Lipid Rafts: Linking Alzheimer's Amyloid- β Production, Aggregation, and Toxicity at Neuronal Membranes. *Int J Alzheimers Dis*, 2011:1–14, 2011.

[13] Charles H. Davis and Max L. Berkowitz. Interaction Between Amyloid-β (1–42) Peptide and Phospholipid Bilayers: A Molecular Dynamics Study. *Biophys. J.*, 96(3):785–797, 2009.

[14] Giovanni Bussi and Alessandro Laio. Using metadynamics to explore complex free-energy landscapes. *Nat. Rev. Phys.*, 2(4):200–212, April 2020.

[15] A. Laio and M. Parrinello. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.*, 99(20):12562–12566, October 2002.

[16] Alessandro Barducci, Massimiliano Bonomi, and Michele Parrinello. Metadynamics. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, 1(5):826–843, September 2011.

[17] Yuji Sugita and Yuko Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.*, 314(1-2):141–151, November 1999.

[18] Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.*, 116(14):7898–7936, July 2016.

[19] Sanja Pöyry and Ilpo Vattulainen. Role of charged lipids in membrane structures — Insight given by simulations. *Biochim. Biophys. Acta, Biomembr.*, 1858(10):2322–2333, 2016.

[20] Jason G. Kay and Sergio Grinstein. Phosphatidylserine-Mediated Cellular Signaling. In Daniel G.S. Capelluto, editor, *Lipid-mediated Protein Signaling*, volume 991, pages 177–193. Springer Netherlands, Dordrecht, 2013.

[21] Akiko Shiratsuchi, Masato Umeda, Yoshiki Ohba, and Yoshinobu Nakanishi. Recognition of Phosphatidylserine on the Surface of Apoptotic Spermatogenic Cells and Subsequent Phagocytosis by Sertoli Cells of the Rat. *Journal of Biological Chemistry*, 272(4):2354–2358, 1997.

[22] M Paul Murphy and Harry III LeVine. Alzheimer's disease and the amyloid-beta peptide. *J. Alzheimers Dis.*, 19(1):311–323, 2010.

[23] Anna Itkin, Vincent Dupres, Yves F Dufrêne, Burkhard Bechinger, Jean-Marie Ruysschaert, and Vincent Raussens. Calcium ions promote formation of amyloid β-peptide (1-40) oligomers causally implicated in neuronal toxicity of Alzheimer's disease. *PLoS ONE*, 6(3):e18250, 2011.

[24] Angelo Demuro, Ian Parker, and Grace E. Stutzmann. Calcium Signaling and Amyloid Toxicity in Alzheimer Disease. *J. Biol. Chem.*, 285(17):12463–12468, 2010.

[25] Nagarajan Sureshbabu, R. Kirubagaran, H. Thangarajah, E. J. Padma Malar, and R. Jayakumar. Lipid-Induced Conformational Transition of Amyloid β Peptide Fragments. *J. Mol. Neurosci.*, 41(3):368–382, 2010.

[26] Francis Hane, Elizabeth Drolle, Ravi Gaikwad, Erin Faught, and Zoya Leonenko. Amyloid-β Aggregation on Model Lipid Membranes: An Atomic Force Microscopy Study. *J. Alzheimers Dis.*, 26(3):485–494, 2011.

[27] Mayu S. Terakawa, Yuxi Lin, Misaki Kinoshita, Shingo Kanemura, Dai Itoh, Toshihiko Sugiki, Masaki Okumura, and Young-Ho Lee. Impact of membrane curvature on amyloid aggregation. *Biochim. Biophys. Acta, Biomembr.*, 1860(9):1741–1764, 2018.

[28] Yuuki Sugiura, Keisuke Ikeda, and Minoru Nakano. High Membrane Curvature Enhances Binding, Conformational Changes, and Fibrillation of Amyloid-β on Lipid Bilayer Surfaces. *Langmuir*, 31(42):11549–11557, 2015.

[29] Kyle L. Morris and Louise C. Serpell. X-Ray Fibre Diffraction Studies of Amyloid Fibrils. In Einar M. Sigurdsson, Miguel Calero, and María Gasset, editors, *Amyloid Proteins: Methods and Protocols*, Methods in Molecular Biology, pages 121–135. Humana Press, Totowa, NJ, 2012.

[30] Margaret Sunde, Louise C Serpell, Mark Bartlam, Paul E Fraser, Mark B Pepys, and Colin C.F Blake. Common core structure of amyloid fibrils by synchrotron X-ray diffraction. *J. Mol. Biol.*, 273(3):729–739, 1997.

[31] Holger A Scheidt, Isabel Morgado, Sven Rothemund, and Daniel Huster. Dynamics of amyloid β fibrils revealed by solid-state NMR. *J. Biol. Chem.*, 287(3):2017–2021, 2012.

[32] C Hilbich, B Kisters-Woike, J Reed, C L Masters, and K Beyreuther. Substitutions of hydrophobic amino acids reduce the amyloidogenicity of Alzheimer's disease beta A4 peptides. *J. Mol. Biol.*, 228(2):460–473, 1992.

[33] H Inouye, P E Fraser, and D A Kirschner. Structure of beta-crystallite assemblies formed by Alzheimer beta-amyloid protein analogues: analysis by x-ray diffraction. *Biophys. J.*, 64(2):502–519, 1993.

[34] D A Kirschner, H Inouye, L K Duffy, A Sinclair, M Lind, and D J Selkoe. Synthetic peptide homologous to beta protein from Alzheimer disease forms amyloid-like fibrils in vitro. *Proc. Natl. Acad. Sci. U.S.A.*, 84(19):6953–6957, 1987.

[35] S J Wood, R Wetzel, J D Martin, and M R Hurle. Prolines and amyloidogenicity in fragments of the Alzheimer's peptide beta/A4. *Biochemistry*, 34(3):724–730, 1995.

[36] K. Aurelia Ball, Aaron H. Phillips, Paul S. Nerenberg, Nicolas L. Fawzi, David E. Wemmer, and Teresa Head-Gordon. Homogeneous and Heterogeneous Tertiary Structure Ensembles of Amyloid-β Peptides. *Biochemistry*, 50(35):7612–7628, 2011.

[37] J J Balbach, Y Ishii, O N Antzutkin, R D Leapman, N W Rizzo, F Dyda, J Reed, and R Tycko. Amyloid fibril formation by A beta 16-22, a seven-residue fragment of the Alzheimer's beta-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry*, 39(45):13748–13759, 2000.

[38] Anil K. Mehta, Kun Lu, W. Seth Childers, Yan Liang, Steven N. Dublin, Jijun Dong, James P. Snyder, Sai Venkatesh Pingali, Pappannan Thiyagarajan, and David G. Lynn. Facial Symmetry in Protein Self-Assembly. *J. Am. Chem. Soc.*, 130(30):9829–9835, 2008.

[39] Kun Lu, Jaby Jacob, P. Thiyagarajan, V. P. Conticello, and D. G. Lynn. Exploiting Amyloid Fibril Lamination for Nanotube Self-Assembly. *J. Am. Chem. Soc.*, 2003.

[40] Kai Tao, Jiqian Wang, Peng Zhou, Chengdong Wang, Hai Xu, Xiubo Zhao, and Jian R. Lu. Self-Assembly of Short Aβ(16–22) Peptides: Effect of Terminal Capping and the Role of Electrostatic Interaction. *Langmuir*, 27(6):2723–2730, 2011.

[41] Li Na Zhao, See-Wing Chiu, Jérôme Benoit, Lock Yue Chew, and Yuguang Mu. Amyloid β Peptides Aggregation in a Mixed Membrane Bilayer: A Molecular Dynamics Study. *J. Phys. Chem. B*, 115(42):12247–12256, 2011.

[42] Anusha Jayaraman and Christian J. Pike. Alzheimer's Disease and Type 2 Diabetes: Multiple Mechanisms Contribute to Interactions. *Curr Diab Rep*, 14(4):476, April 2014.

[43] Wenqing Xia, Shaohua Wang, Zilin Sun, Feng Bai, Yi Zhou, Yue Yang, Pin Wang, Yan Huang, and Yang Yuan. Altered baseline brain activity in type 2 diabetes: A resting-state fMRI study. *Psychoneuroendocrinology*, 38(11):2493–2501, November 2013.

[44] Chris Moran, Thanh G. Phan, Jian Chen, Leigh Blizzard, Richard Beare, Alison Venn, Gerald Münch, Amanda G. Wood, Josephine Forbes, Timothy M. Greenaway, Susan Pearson, and Velandai Srikanth. Brain Atrophy in Type 2 Diabetes. *Diabetes Care*, 36(12):4036–4042, December 2013.

[45] Sanne M. Manschot, Augustina M.A. Brands, Jeroen van der Grond, Roy P.C. Kessels, Ale Algra, L. Jaap Kappelle, Geert Jan Biessels, and on behalf of the Utrecht Diabetic Encephalopathy Study Group. Brain Magnetic Resonance Imaging Correlates of Impaired Cognition in Patients With Type 2 Diabetes. *Diabetes*, 55(4):1106–1113, April 2006.

[46] Nobuyuki Sasaki, Ryo Fukatsu, Kayo Tsuzuki, Yorihide Hayashi, Taku Yoshida, Nobuhiro Fujii, Takao Koike, Ikuro Wakayama, Richard Yanagihara, Ralph Garruto, Naoji Amano, and Zenji Makita. Advanced Glycation End Products in Alzheimer's Disease and Other Neurodegenerative Diseases. *The American Journal of Pathology*, 153(4):1149–1155, October 1998.

[47] Yanyan Kong, Fushuai Wang, Jiao Wang, Cuiping Liu, Yinping Zhou, Zhengqin Xu, Chencheng Zhang, Bomin Sun, and Yihui Guan. Pathological Mechanisms Linking Diabetes Mellitus and Alzheimer's Disease: the Receptor for Advanced Glycation End Products (RAGE). *Front. Aging Neurosci.*, 12:217, July 2020.

[48] Shun-Yao Ko, Hshin-An Ko, Kuo-Hsiung Chu, Tzong-Ming Shieh, Tzong-Cherng Chi, Hong-I Chen, Weng-Cheng Chang, and Shu-Shing Chang. The Possible Mechanism of Advanced Glycation End Products (AGEs) for Alzheimer's Disease. *PLoS ONE*, 10(11):e0143345, November 2015.

[49] Niraja Kedia, Michael Almisry, and Jan Bieschke. Glucose directs amyloid-beta into membrane-active oligomers. *Phys. Chem. Chem. Phys.*, 19(27):18036–18046, 2017.

[50] Joana M. Gil and Ana Cristina Rego. Mechanisms of neurodegeneration in Huntington's disease. *Eur. J. Neurosci.*, 27(11):2803–2820, 2008.

[51] Kimberly B Kegel-Gleason. Huntingtin interactions with membrane phospholipids: strategic targets for therapeutic intervention? *J. Huntington's Dis.*, 2(3):239–50, 2013.

[52] Gillian Bates. Huntingtin aggregation and toxicity in Huntington's disease. *Lancet*, 361(9369):1642–1644, 2003.

[53] Henry L. Paulson, Vikram G. Shakkottai, H. Brent Clark, and Harry T. Orr. Polyglutamine spinocerebellar ataxias — from genes to potential treatments. *Nat Rev Neurosci*, 18(10):613–626, October 2017.

[54] Craig S. McIntosh, Dunhui Li, Steve D. Wilton, and May T. Aung-Htut. Polyglutamine Ataxias: Our Current Molecular Understanding and What the Future Holds for Antisense Therapies. *Biomedicines*, 9(11):1499, October 2021.

[55] Maciej Figiel, Wojciech J. Szlachcic, Pawel M. Switonski, Agnieszka Gabka, and Wlodzimierz J. Krzyzosiak. Mouse Models of Polyglutamine Diseases: Review and Data Table. Part I. *Mol Neurobiol*, 46(2):393–429, October 2012.

[56] Kathleen A. Burke, Jordan Godbey, and Justin Legleiter. Assessing mutant huntingtin fragment and polyglutamine aggregation by atomic force microscopy. *Methods*, 53(3):275–284, 2011.

[57] Justin Legleiter, Gregor P. Lotz, Jason Miller, Jan Ko, Cheping Ng, Geneva L. Williams, Steve Finkbeiner, Paul H. Patterson, and Paul J. Muchowski. Monoclonal Antibodies Recognize Distinct Conformational Epitopes Formed by Polyglutamine in a Mutant Huntingtin Fragment. *J. Biol. Chem.*, 284(32):21647–21658, 2009.

[58] Randy Singh Atwal, Jianrun Xia, Deborah Pinchev, Jillian Taylor, Richard M. Epand, and Ray Truant. Huntingtin has a membrane association signal that can modulate huntingtin aggregation, nuclear entry and toxicity. *Hum. Mol. Genet.*, 16(21):2600–2615, 2007.

[59] Kathleen A. Burke, Kaitlin M. Hensal, C. Samuel Umbaugh, Maxmore Chaibva, and Justin Legleiter. Huntingtin disrupts lipid bilayers in a polyQ-length dependent manner. *Biochim Biophys Acta Biomembr*, 1828(8):1953–1961, 2013.

[60] Maxmore Chaibva, Kathleen A. Burke, and Justin Legleiter. Curvature Enhances Binding and Aggregation of Huntingtin at Lipid Membranes. *Biochemistry*, 53(14):2355–2365, 2014.

[61] Anthony R. Braun, Michael M. Lacy, Vanessa C. Ducas, Elizabeth Rhoades, and Jonathan N. Sachs. Alpha-Synuclein's uniquely long amphipathic helix enhances its membrane binding and remodeling capacity. *The Journal of membrane biology*, 250(2):183–193, 2017.

[62] Elka R. Georgieva, Shifeng Xiao, Peter P. Borbat, Jack H. Freed, and David Eliezer. Tau Binds to Lipid Membrane Surfaces via Short Amphipathic Helices Located in Its Microtubule-Binding Repeats. *Biophysical Journal*, 107(6):1441–1452, 2014.

[63] Manuel Giménez-Andrés, Alenka Čopič, and Bruno Antonny. The Many Faces of Amphipathic Helices. *Biomolecules*, 8(3), 2018.

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

[65] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant Kalé, and Klaus Schulten. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[66] Ilian T. Todorov, William Smith, Kostya Trachenko, and Martin T. Dove. DL_POLY_3: new dimensions in molecular dynamics simulations via massive parallelism. *Journal of Materials Chemistry*, 16(20):1911, 2006.

[67] James A. Maier, Carmenza Martinez, Koushik Kasavajhala, Lauren Wickstrom, Kevin E. Hauser, and Carlos Simmerling. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.*, 11(8):3696–3713, 2015.

[68] Callum J. Dickson, Benjamin D. Madej, Åge A. Skjevik, Robin M. Betz, Knut Teigen, Ian R. Gould, and Ross C. Walker. Lipid14: The Amber Lipid Force Field. *J. Chem. Theory Comput.*, 10(2):865–879, 2014.

[69] Robert B. Best, Xiao Zhu, Jihyun Shim, Pedro E. M. Lopes, Jeetain Mittal, Michael Feig, and Alexander D. MacKerell. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backboneϕ, ψ and Side-Chain χ_1 and χ_2 Dihedral Angles. *J. Chem. Theory Comput.*, 8(9):3257–3273, 2012.

[70] R. W. Pastor and A. D. MacKerell. Development of the CHARMM Force Field for Lipids. *J. Phys. Chem. Lett.*, 2(13):1526–1532, 2011.

[71] Jeffery B. Klauda, Viviana Monje, Taehoon Kim, and Wonpil Im. Improving the CHARMM Force Field for Polyunsaturated Fatty Acid Chains. *J. Phys. Chem. B*, 116(31):9424–9431, 2012.

[72] Chris Oostenbrink, Alessandra Villa, Alan E. Mark, and Wilfred F. Van Gunsteren. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.*, 25(13):1656–1676, 2004.

[73] See-Wing Chiu, Sagar A. Pandit, H. L. Scott, and Eric Jakobsson. An Improved United Atom Force Field for Simulation of Mixed Lipid Bilayers. *J. Phys. Chem. B*, 113(9):2748–2763, 2009.

[74] William L. Jorgensen and Julian Tirado-Rives. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.*, 110(6):1657–1666, 1988.

[75] David S. Maxwell William L. Jorgensen and Julian Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.*, pages 11225–11236, 1996.

[76] Mookyung Cheon, Iksoo Chang, and Carol K Hall. Extending the PRIME model for protein aggregation to all 20 amino acids. *Proteins*, 78(14):2950–2960, 2010.

[77] Aram Davtyan, Nicholas P. Schafer, Weihua Zheng, Cecilia Clementi, Peter G. Wolynes, and Garegin A. Papoian. AWSEM-MD: Protein Structure Prediction Using Coarse-grained Physical Potentials and Bioinformatically Based Local Structure Biasing. *J Phys Chem B*, 116(29):8494, 2012.

[78] Fabio Sterpone, Simone Melchionna, Pierre Tuffery, Samuela Pasquali, Normand Mousseau, Tristan Cragnolini, Yassmine Chebaro, Jean-Francois St-Pierre, Maria Kalimeri, Alessandro Barducci, Yoann Laurin, Alex Tek, Marc Baaden, Phuong Hoang Nguyen, and Philippe Derreumaux. The OPEP protein model: from single molecules, amyloid formation, crowding and hydrodynamics to DNA/RNA systems. *Chem. Soc. Rev.*, 43(13):4871–4893, 2014.

[79] Siewert J. Marrink, H. Jelger Risselada, Serge Yefimov, D. Peter Tieleman, and Alex H. de Vries. The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J. Phys. Chem. B*, 2007.

[80] Sai J. Ganesan and S. Matysiak. Role of Backbone Dipole Interactions in the Formation of Secondary and Supersecondary Structures of Proteins. *J. Chem. Theory Comput.*, 10(6):2569–2576, June 2014.

[81] Sai J. Ganesan and Silvina Matysiak. Interplay between the hydrophobic effect and dipole interactions in peptide aggregation at interfaces. *Phys. Chem. Chem. Phys.*, 18(4):2449–2458, 2016.

[82] Abhilash Sahoo, Hongcheng Xu, and Silvina Matysiak. Pathways of amyloid-beta absorption and aggregation in a membranous environment. *Phys. Chem. Chem. Phys.*, 21(16):8559–8568, April 2019. Publisher: The Royal Society of Chemistry.

[83] Semen O. Yesylevskyy, Lars V. Schäfer, Durba Sengupta, and Siewert J. Marrink. Polarizable Water Model for the Coarse-Grained MARTINI Force Field. *PLoS Comput Biol*, 6(6):e1000810, June 2010.

[84] Sai J. Ganesan, Hongcheng Xu, and Silvina Matysiak. Influence of Monovalent Cation Size on Nanodomain Formation in Anionic–Zwitterionic Mixed Bilayers. *J. Phys. Chem. B*, 121(4):787–799, February 2017.

[85] Sai J. Ganesan, Hongcheng Xu, and Silvina Matysiak. Effect of lipid head group interactions on membrane properties and membrane-induced cationic β-hairpin folding. *Phys. Chem. Chem. Phys.*, 18(27):17836–17850, 2016.

[86] David R. Howlett, Kevin H. Jennings, David C. Lee, Michael S.G. Clark, Frank Brown, Ronald Wetzel, Stephen J. Wood, Patrick Camilleri, and Gareth W. Roberts. Aggregation State and Neurotoxic Properties of Alzheimer Beta-Amyloid Peptide. *Neurodegeneration*, 4(1):23–32, 1995.

[87] G Brent Irvine, Omar M El-Agnaf, Ganesh M Shankar, and Dominic M Walsh. Protein aggregation in the brain: the molecular basis for Alzheimer's and Parkinson's diseases. *Mol. Med.*, 14(7-8):451–64, 2008.

[88] Montserrat Arrasate and Steven Finkbeiner. Protein aggregates in Huntington's disease. *Exp Neurol.*, 238(1):1–11, 2012.

[89] George G. Glenner and Caine W. Wong. Alzheimer's disease and Down's syndrome: Sharing of a unique cerebrovascular amyloid fibril protein. *Biochem. Biophys. Res. Commun.*, 122(3):1131–1135, 1984.

[90] George G. Glenner and Caine W. Wong. Alzheimer's disease: Initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.*, 120(3):885–890, 1984.

[91] C L Masters, G Simms, N A Weinman, G Multhaup, B L McDonald, and K Beyreuther. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc. Natl. Acad. Sci. U.S.A.*, 82(12):4245–9, 1985.

[92] Carlo Ballatore, Virginia M.-Y. Lee, and John Q. Trojanowski. Tau-mediated neurodegeneration in Alzheimer's disease and related disorders. *Nat. Rev. Neurosci.*, 8(9):663–672, 2007.

[93] H. Braak and E. Braak. Evolution of neuronal changes in the course of Alzheimer's disease. pages 127–140. Springer, Vienna, 1998.

[94] 2018 Alzheimer's disease facts and figures. *Alzheimers Dement*, 14(3):367–429, 2018.

[95] B. Yankner, L. Dawes, S Fisher, L Villa-Komaroff, M. Oster-Granite, and R. Neve. Neurotoxicity of a fragment of the amyloid precursor associated with Alzheimer's disease. *Science*, 245(4916):417–420, 1989.

[96] Dennis J. Selkoe. The molecular pathology of Alzheimer's disease. *Neuron*, 6(4):487–498, 1991.

[97] M. Goedert and M. G. Spillantini. A Century of Alzheimer's Disease. *Science*, 314(5800):777–781, 2006.

[98] Dennis J. Selkoe. Alzheimer's Disease: Genes, Proteins, and Therapy. *Physiol. Rev.*, 81(2):741–766, 2001.

[99] R. Tanzi, J. Gusella, P. Watkins, G. Bruns, P St George-Hyslop, M. Van Keuren, D Patterson, S Pagan, D. Kurnit, R. Neve, and al. Et. Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science*, 235(4791):880–884, 1987.

[100] Dominic M. Walsh and Dennis J. Selkoe. Deciphering the Molecular Basis of Memory Failure in Alzheimer's Disease. *Neuron*, 44(1):181–193, 2004.

[101] Charles C. Glabe. Amyloid accumulation and pathogensis of alzheimer's disease: Significance of monomeric, oligomeric and fibrillar aβ. In *Amyloid Accumulation and Pathogensis of Alzheimer's Disease: Significance of Monomeric, Oligomeric and Fibrillar Aβ*, pages 167–177. Springer US, 2005.

[102] Jessica Meinhardt, Carsten Sachse, Peter Hortschansky, Nikolaus Grigorieff, and Marcus Fändrich. Aβ(1-40) Fibril Polymorphism Implies Diverse Interaction Patterns in Amyloid Fibrils. *J. Mol. Biol.*, 386(3):869–877, 2009.

[103] Julian D. Gillmore, Philip N. Hawkins, and Mark B. Pepys. Amyloidosis: A review of recent diagnostic and therapeutic developments. *Br. J. Haematol.*, 99(2):245–256, 1997.

[104] E H Koo, P T Lansbury, and J W Kelly. Amyloid diseases: abnormal protein aggregation in neurodegeneration. *Proc. Natl. Acad. Sci. U.S.A.*, 96(18):9989–9990, 1999.

[105] Roma N Rambaran and Louise C Serpell. Amyloid fibrils: abnormal protein assembly. *Prion*, 2(3):112–117, 2008.

[106] Kyle L. Morris and Louise C. Serpell. X-Ray Fibre Diffraction Studies of Amyloid Fibrils. In *Methods in molecular biology (Clifton, N.J.)*, volume 849, pages 121–135. 2012.

[107] Anders Olofsson, Malin Lindhagen-Persson, Monika Vestling, A. Elisabeth Sauer-Eriksson, and Anders Öhman. Quenched hydrogen/deuterium exchange NMR characterization of amyloid-β peptide aggregates formed in the presence of Cu^{2+} or Zn^{2+} . *FEBS J.*, 276(15):4051–4060, 2009.

[108] Valerie L Anderson and Watt W Webb. Transmission electron microscopy characterization of fluorescently labelled amyloid β 1-40 and α-synuclein aggregates. *BMC Biotechnol*, 11:125, 2011.

[109] Carsten Sachse, Marcus Fändrich, and Nikolaus Grigorieff. Paired beta-sheet structure of an Abeta(1-40) amyloid fibril revealed by electron microscopy. *Proc. Natl. Acad. Sci. U.S.A.*, 105(21):7462–7466, 2008.

[110] Lothar Gremer, Daniel Schölzel, Carla Schenk, Elke Reinartz, Jörg Labahn, Raimond B. G. Ravelli, Markus Tusche, Carmen Lopez-Iglesias, Wolfgang Hoyer, Henrike Heise, Dieter Willbold, and Gunnar F. Schröder. Fibril structure of amyloid-β(1–42) by cryo–electron microscopy. *Science*, 358(6359):116–119, 2017.

[111] Elizabeth Drolle, Alexander Negoda, Keely Hammond, Evgeny Pavlov, and Zoya Leonenko. Changes in lipid membranes may trigger amyloid toxicity in Alzheimer's disease. *PloS ONE*, 12(8):e0182194, 2017.

[112] Mun'delanji Vestergaard, Tsutomu Hamada, and Masahiro Takagi. Using model membranes for the study of amyloid beta:lipid interactions and neurotoxicity. *Biotechnol. Bioeng.*, 99(4):753–763, 2008.

[113] Justin A. Lemkul and David R. Bevan. Lipid composition influences the release of Alzheimer's amyloid β-peptide from membranes. *Protein Sci.*, 20(9):1530–1545, 2011.

[114] Isabel Morgado and Megan Garvey. Lipids in Amyloid-β Processing, Aggregation, and Toxicity. pages 67–94. Springer, Cham, 2015.

[115] Laura Dominguez, Leigh Foster, John E. Straub, and D. Thirumalai. Impact of membrane lipid composition on the structure and stability of the transmembrane domain of amyloid precursor protein. *Proc. Natl. Acad. Sci. U.S.A.*, 113(36):E5281–E5287, 2016.

[116] Katsumi Matsuzaki. Physicochemical interactions of amyloid β-peptide with lipid bilayers. *Biochim. Biophys. Acta, Biomembr.*, 1768(8):1935–1942, 2007.

[117] Wenxin Yu, Kun Zou, Jian Sheng Gong, Mihee Ko, Katsuhiko Yanagisawa, and Makoto Michikawa. Oligomerization of amyloid β-protein occurs during the isolation of lipid rafts. *J. Neurosci. Res.*, 80(1):114–119, 2005.

[118] Dominic M. Walsh, Aleksey Lomakin, George B. Benedek, Margaret M. Condron, and David B. Teplow. Amyloid β-protein fibrillogenesis: Detection of a protofibrillar intermediate. *J. Biol. Chem.*, 272(35):22364–22372, 1997.

[119] Rakez Kayed, Elizabeth Head, Jennifer L. Thompson, Theresa M. McIntire, Saskia C. Milton, Carl W. Cotman, and Charles G. Glabel. Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis. *Science*, 300(5618):486–489, 2003.

[120] Payel Das, Silvina Matysiak, and Jeetain Mittal. Looking at the Disordered Proteins through the Computational Microscope. *ACS Cent Sci*, 4(5):534–542, 2018.

[121] Rongliang Wu, Jun Liu, Xinlong Qiu, and Manli Deng. Molecular dynamics simulation of the nanofibrils formed by amyloid-based peptide amphiphiles. *Mol Simul*, 43(13-16):1227–1239, 2017.

[122] B. Urbanc, L. Cruz, F. Ding, D. Sammond, S. Khare, S.V. Buldyrev, H.E. Stanley, and N.V. Dokholyan. Molecular Dynamics Simulation of Amyloid β Dimer Formation. *Biophys. J.*, 87(4):2310–2321, 2004.

[123] Matthew Biancalana and Shohei Koide. Molecular mechanism of Thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta, Proteins Proteomics*, 1804(7):1405–1412, 2010.

[124] Nikolaos G. Sgourakis, Myrna Merced-Serrano, Christos Boutsidis, Petros Drineas, Zheming Du, Chunyu Wang, and Angel E. Garcia. Atomic-Level Characterization of the Ensemble of the Aβ(1–42) Monomer in Water Using Unbiased Molecular Dynamics Simulations and Spectral Algorithms. *J. Mol. Biol.*, 405(2):570–583, 2011.

[125] Nikolaos G. Sgourakis, Yilin Yan, Scott A. McCallum, Chunyu Wang, and Angel E. Garcia. The Alzheimer's Peptides Aβ40 and 42 Adopt Distinct Conformations in Water: A Combined MD / NMR Study. *J. Mol. Biol.*, 368(5):1448–1457, 2007.

[126] Naohiro Nishikawa, Phuong H. Nguyen, Philippe Derreumaux, and Yuko Okamoto. Replica-exchange molecular dynamics simulation for understanding the initial process of amyloid peptide aggregation. *Mol Simul*, 41(10-12):1041–1044, 2015.

[127] Dmitri K Klimov and D Thirumalai. Dissecting the assembly of Abeta16-22 amyloid peptides into antiparallel beta sheets. *Structure*, 11(3):295–307, 2003.

[128] Sébastien Santini, Normand Mousseau, and Philippe Derreumaux. In Silico Assembly of Alzheimer's Aβ 16-22 Peptide into β-Sheets. *J. Am. Chem. Soc.*, 126(37):11509–11516, 2004.

[129] Sebastian Kmiecik, Dominik Gront, Michal Kolinski, Lukasz Wieteska, Aleksandra Elzbieta Dawid, and Andrzej Kolinski. Coarse-Grained Protein Models and Their Applications. *Chem. Rev.*, 116(14):7898–7936, 2016.

[130] Sanne Abeln, Michele Vendruscolo, Christopher M. Dobson, and Daan Frenkel. A Simple Lattice Model That Captures Protein Folding, Aggregation and Amyloid Formation. *PLoS ONE*, 9(1):e85185, 2014.

[131] Ran Friedman, Riccardo Pellarin, and Amedeo Caflisch. Amyloid Aggregation on Lipid Bilayers and Its Impact on Membrane Permeability. *J. Mol. Biol.*, 387(2):407–415, 2009.

[132] Alex Morriss-Andrews, Frank L. H. Brown, and Joan-Emma Shea. A Coarse-Grained Model for Peptide Aggregation on a Membrane Surface. *J. Phys. Chem. B*, 118(28):8420–8432, 2014.

[133] Mara Chiricotto, Thanh Thuy Tran, Phuong H. Nguyen, Simone Melchionna, Fabio Sterpone, and Philippe Derreumaux. Coarse-grained and All-atom Simulations towards the Early and Late Steps of Amyloid Fibril Formation. *Isr. J. Chem.*, 57(7-8):564–573, 2017.

[134] Yassmine Chebaro, Samuela Pasquali, and Philippe Derreumaux. The Coarse-Grained OPEP Force Field for Non-Amyloid and Amyloid Proteins. *J. Phys. Chem. B*, 116(30):8741–8752, 2012.

[135] Mookyung Cheon, Mooseok Kang, and Iksoo Chang. Polymorphism of fibrillar structures depending on the size of assembled Aβ17-42 peptides. *Sci. Rep.*, 6(1):38196, 2016.

[136] Weihua Zheng, Min-Yeh Tsai, Mingchen Chen, and Peter G Wolynes. Exploring the aggregation free energy landscape of the amyloid-β protein (1-40). *Proc. Natl. Acad. Sci. U.S.A.*, 113(42):11835–11840, 2016.

[137] Canay Ege, Jaroslaw Majewski, Guohui Wu, Kristian Kjaer, and Ka Yee C. Lee. Templating Effect of Lipid Membranes on Alzheimer's Amyloid Beta Peptide. *ChemPhysChem*, 6(2):226–229, 2005.

[138] R. O. Calderon, B. Attema, and G. H. DeVries. Lipid Composition of Neuronal Cell Bodies and Neurites from Cultured Dorsal Root Ganglia. *J. Neurochem.*, 64(1):424–429, 2002.

[139] Xiang Yu, Qiuming Wang, Qingfen Pan, Feimeng Zhou, and Jie Zheng. Molecular interactions of Alzheimer amyloid-β oligomers with neutral and negatively charged lipid bilayers. *Phys. Chem. Chem. Phys*, 15(23):8878, 2013.

[140] JoAnne McLaurin and Avijit Chakrabartty. Characterization of the Interactions of Alzheimer beta-Amyloid Peptides with Phospholipid Membranes. *Eur. J. Biochem.*, 245(2):355–363, 1997.

[141] Evelyne Terzi, Guenter Hoelzemann, and Joachim Seelig. Alzheimer .beta.-Amyloid Peptide 25-35: Electrostatic Interactions with Phospholipid Membranes. *Biochemistry*, 33(23):7434–7441, 1994.

[142] Evelyne Terzi, Günter Hölzemann, and Joachim Seelig. Interaction of Alzheimer β-Amyloid Peptide(140) with Lipid Membranes†. *Biochemistry*, 1997.

[143] Claudio Canale, Silvia Seghezza, Silvia Vilasi, Rita Carrotta, Donatella Bulone, Alberto Diaspro, Pier Luigi San Biagio, and Silvia Dante. Different effects of Alzheimer's peptide Aβ(1–40) oligomers and fibrils on supported lipid membranes. *Biophys. Chem.*, 182:23–29, 2013.

[144] Marc-Antoine Sani, John D. Gehman, and Frances Separovic. Lipid matrix plays a role in Abeta fibril kinetics and morphology. *FEBS Lett.*, 585(5):749–754, 2011.

[145] Tong-Lay Lau, John D. Gehman, John D. Wade, Keyla Perez, Colin L. Masters, Kevin J. Barnham, and Frances Separovic. Membrane interactions and the effect of metal ions of the amyloidogenic fragment Aβ(25-35) in comparison to Aβ(1–42). *Biochim Biophys Acta Biomembr*, 1768(10):2400–2408, 2007.

[146] Yan Lu, Philippe Derreumaux, Zhi Guo, Normand Mousseau, and Guanghong Wei. Thermodynamics and dynamics of amyloid peptide oligomerization are sequence dependent. *Proteins*, 75(4):954–963, 2009.

[147] Sander Pronk, Szilárd Páll, Roland Schulz, Per Larsson, Pär Bjelkmar, Rossen Apostolov, Michael R. Shirts, Jeremy C. Smith, Peter M. Kasson, David van der Spoel, Berk Hess, and Erik Lindahl. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 2013.

[148] Shūichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52(2):255–268, 1984.

[149] Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31(3):1695–1697, 1985.

[150] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52(12):7182–7190, 1981.

[151] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An nlogn method for ewald sums in large systems. *J. Chem. Phys.*, 98(12):10089–10092, 1993.

[152] Ulrich Essmann, Lalith Perera, Max L. Berkowitz, Tom Darden, Hsing Lee, and Lee G. Pedersen. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103(19):8577–8593, 1995.

[153] Nabin Kandel, Tianyu Zheng, Qun Huo, and Suren A. Tatulian. Membrane Binding and Pore Formation by a Cytotoxic Fragment of Amyloid β Peptide. *J. Phys. Chem. B*, 121(45):10293–10305, 2017.

[154] Siewert J Marrink and Alan E Mark. Molecular Dynamics Simulation of the Formation, Structure, and Dynamics of Small Phospholipid Vesicles. *J. Chem. Phys.*, 2003.

[155] Chun-Min Lin, Chun-Shian Li, Yu-Jane Sheng, David T Wu, and Heng-Kwong Tsao. Size-Dependent Properties of Small Unilamellar Vesicles Formed by Model Lipids. *Langmuir*, 28:689–700, 2012.

[156] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD – Visual Molecular Dynamics. *J Mol Graph*, 14:33–38, 1996.

[157] Fernando E. Herrera and Sergio Pantano. Structure and dynamics of nano-sized raft-like domains on the plasma membrane. *J. Chem. Phys.*, 136(1):015103, 2012.

[158] Jianjun Pan, Xiaolin Cheng, Luca Monticelli, Frederick A. Heberle, Norbert Kučerka, D. Peter Tieleman, and John Katsaras. The molecular structure of a phosphatidylserine bilayer determined by scattering and molecular dynamics simulations. *Soft Matter*, 10(21):3716, 2014.

[159] Yi Hu, Pascal Kienlen-Campard, Tzu-Chun Tang, Florian Perrin, Rémi Opsomer, Marie Decock, Xiaoshu Pan, Jean-Noel Octave, Stefan N. Constantinescu, and Steven O. Smith. β-Sheet Structure within the Extracellular Domain of C99 Regulates Amyloidogenic Processing. *Sci. Rep.*, 7(1):17159, 2017.

[160] Neil R. Anthony, Anil K. Mehta, David G. Lynn, and Keith M. Berland. Mapping amyloid-β(16-22) nucleation pathways using fluorescence lifetime imaging microscopy. *Soft Matter*, 10(23):4162–4172, 2014.

[161] Sandra Chimon, Medhat A Shaibat, Christopher R Jones, Diana C Calero, Buzulagu Aizezi, and Yoshitaka Ishii. Evidence of fibril-like β-sheet structures in a neurotoxic amyloid intermediate of Alzheimer's β-amyloid. *Nat. Struct. Mol. Biol.*, 14(12):1157–1164, 2007.

[162] Nicolae-Viorel Buchete and Gerhard Hummer. Structure and dynamics of parallel beta-sheets, hydrophobic core, and loops in Alzheimer's A beta fibrils. *Biophys. J.*, 92(9):3032–3039, 2007.

[163] Jie Li, Ruiwu Liu, Kit S Lam, Lee-Way Jin, and Yong Duan. Alzheimer's disease drug candidates stabilize A-β protein native structure by interacting with the hydrophobic core. *Biophys. J.*, 100(4):1076–1082, 2011.

[164] David J. Lindberg, Emelie Wesén, Johan Björkeroth, Sandra Rocha, and Elin K. Esbjörner. Lipid membranes catalyse the fibril formation of the amyloid-β (1–42) peptide through lipid-fibril interactions that reinforce secondary pathways. *Biochim Biophys Acta Biomembr*, 1859(10):1921–1929, 2017.

[165] John J Kremer and Regina M Murphy. Kinetics of adsorption of beta-amyloid peptide Abeta(1-40) to lipid bilayers. *J. Biochem. Biophys. Methods*, 57(2):159–169, 2003.

[166] Abhilash Sahoo and Silvina Matysiak. Computational insights into lipid assisted peptide misfolding and aggregation in neurodegeneration. *Phys. Chem. Chem. Phys.*, 21(41):22679–22694, October 2019.

[167] Karine Berthelot, Christophe Cullin, and Sophie Lecomte. What does make an amyloid toxic: Morphology, structure or interaction with membrane? *Biochimie*, 95(1):12–19, 2013.

[168] Iwona M. Pranke, Vincent Morello, Joëlle Bigay, Kimberley Gibson, Jean-Marc Verbavatz, Bruno Antonny, and Catherine L. Jackson. α-Synuclein and ALPS motifs are membrane curvature sensors whose contrasting chemistry mediates selective vesicle binding. *J. Cell Biol.*, 194(1):89, July 2011. Publisher: The Rockefeller University Press.

[169] Maxmore Chaibva, Kathleen A. Burke, and Justin Legleiter. Curvature enhances binding and aggregation of huntingtin at lipid membranes. *Biochemistry*, 53(14):2355–2365, April 2014.

[170] G. C. A. da Hora, N. L. Archilha, J. L. S. Lopes, D. M. Müller, K. Coutinho, R. Itri, and T. A. Soares. Membrane negative curvature induced by a hybrid peptide from pediocin PA-1 and plantaricin 149 as revealed by atomistic molecular dynamics simulations. *Soft Matter*, 12(43):8884–8898, November 2016.

[171] Denys E. S. Santos, Laércio Pol-Fachin, Roberto D. Lins, and Thereza A. Soares. Polymyxin Binding to the Bacterial Outer Membrane Reveals Cation Displacement and Increasing Membrane Curvature in Susceptible but Not in Resistant Lipopolysaccharide Chemotypes. *J. Chem. Inf. Model*, 57(9):2181–2193, September 2017.

[172] Federico Elías-Wolff, Martin Lindén, Alexander P. Lyubartsev, and Erik G. Brandt. Curvature sensing by cardiolipin in simulated buckled membranes. *Soft Matter*, 15(4):792–802, 2019.

[173] Jordi Gómez-Llobregat, Federico Elías-Wolff, and Martin Lindén. Anisotropic Membrane Curvature Sensing by Amphipathic Peptides. *Biophys. J.*, 110(1):197–204, January 2016.

[174] Abhilash Sahoo and Silvina Matysiak. Microscopic picture of calcium-assisted lipid demixing and membrane remodeling using multiscale simulations. *J. Phys. Chem. B*, 124(34):7327–7335, August 2020.

[175] Keon A. Reid, Caitlin M. Davis, R. Brian Dyer, and James T. Kindt. Binding, folding and insertion of a beta-hairpin peptide at a lipid bilayer surface: Influence of electrostatics and lipid tail packing. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1860(3):792–800, 2018.

[176] Semen O. Yesylevskyy, Timothée Rivel, and Christophe Ramseyer. The influence of curvature on the properties of the plasma membrane. Insights from atomistic molecular dynamics simulations. *Sci. Rep*, 7(1):16078, December 2017.

[177] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Páll, Jeremy C. Smith, Berk Hess, and Erik Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, September 2015.

[178] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.*, 81(8):3684–3690, October 1984.

[179] Michele F.M. Sciacca, Carmelo Tempra, Federica Scollo, Danilo Milardi, and Carmelo La Rosa. Amyloid growth and membrane damage: Current themes and emerging perspectives from theory and experiments on Aβ and hIAPP. *Biochim Biophys Acta Biomembr*, 1860(9):1625–1638, September 2018.

[180] Justin A. Lemkul and David R. Bevan. Perturbation of membranes by the amyloid β-peptide – a molecular dynamics study. *FEBS J.*, 276(11):3060–3075, 2009.

[181] Thomas L. Williams and Louise C. Serpell. Membrane and surface interactions of Alzheimer's Aβ peptide – insights into the mechanism of cytotoxicity. *FEBS J.*, 278(20):3905–3917, 2011.

[182] Xubo Lin, Vinay Nair, Yong Zhou, and Alemayehu A. Gorfe. Membrane potential and dynamics in a ternary lipid mixture: insights from molecular dynamics simulations. *Phys. Chem. Chem. Phys.*, 20(23):15841–15851, 2018.

[183] Carl-Johan Högberg and Alexander P. Lyubartsev. Effect of Local Anesthetic Lidocaine on Electrostatic Properties of a Lipid Bilayer. *Biophys. J.*, 94(2):525–531, January 2008.

[184] Guo-fang Chen, Ting-hai Xu, Yan Yan, Yu-ren Zhou, Yi Jiang, Karsten Melcher, and H Eric Xu. Amyloid beta: structure, biology and structure-based therapeutic development. *Acta Pharmacol Sin*, 38(9):1205–1235, September 2017.

[185] Chetan N. Patel, Schroeder M. Noble, Gresham T. Weatherly, Ashutosh Tripathy, Donald J. Winzor, and Gary J. Pielak. Effects of molecular crowding by saccharides on α-chymotrypsin dimerization. *Protein Sci.*, 11(5):997–1003, May 2002.

[186] Sumra Shahid, Ikramul Hasan, Faizan Ahmad, Md. Imtaiyaz Hassan, and Asimul Islam. Carbohydrate-Based Macromolecular Crowding-Induced Stabilization of Proteins: Towards Understanding the Significance of the Size of the Crowder. *Biomolecules*, 9(9):477, September 2019.

[187] Lihui Bai, Xiaocui Guo, Xue Zhang, Wenting Yu, and Dayong Yang. Saccharides Create a Crowding Environment for Gene Expression in Cell-Free Systems. *Langmuir*, 35(17):5931–5936, April 2019.

[188] Irina Kuznetsova, Konstantin Turoverov, and Vladimir Uversky. What Macromolecular Crowding Can Do to a Protein. *IJMS*, 15(12):23090–23140, December 2014.

[189] David C. Latshaw, Mookyung Cheon, and Carol K. Hall. Effects of Macromolecular Crowding on Amyloid Beta (16–22) Aggregation Using Coarse-Grained Simulations. *J. Phys. Chem. B*, 118(47):13513–13526, November 2014.

[190] F. Musiani and A. Giorgetti. Protein Aggregation and Molecular Crowding. In *International Review of Cell and Molecular Biology*, volume 329, pages 49–77. Elsevier, 2017.

[191] Larissa A. Munishkina, Atta Ahmad, Anthony L. Fink, and Vladimir N. Uversky. Guiding Protein Aggregation with Macromolecular Crowding. *Biochemistry*, 47(34):8993–9006, August 2008.

[192] Sneha Menon and Neelanjana Sengupta. Influence of Hyperglycemic Conditions on Self-Association of the Alzheimer's Amyloid β (Aβ $_{1–42}$) Peptide. *ACS Omega*, 2(5):2134–2147, May 2017.

[193] Cesar A. López, Andrzej J. Rzepiela, Alex H. de Vries, Lubbert Dijkhuizen, Philippe H. Hünenberger, and Siewert J. Marrink. Martini Coarse-Grained Force Field: Extension to Carbohydrates. *J. Chem. Theory Comput.*, 5(12):3195–3210, December 2009.

[194] Philipp S. Schmalhorst, Felix Deluweit, Roger Scherrers, Carl-Philipp Heisenberg, and Mateusz Sikora. Overcoming the Limitations of the MARTINI Force Field in Simulations of Polysaccharides. *J. Chem. Theory Comput.*, 13(10):5039–5053, October 2017.

[195] Moritoki Egi, Rinaldo Bellomo, Edward Stachowski, Craig J. French, Graeme K. Hart, Colin Hegarty, and Michael Bailey. Blood glucose concentration and outcome of critical illness: The impact of diabetes*:. *Critical Care Medicine*, 36(8):2249–2255, August 2008.

[196] A. Lerbret, F. Affouard, P. Bordat, A. Hédoux, Y. Guinet, and M. Descamps. Molecular dynamics simulations of lysozyme in water/sugar solutions. *Chemical Physics*, 345(2-3):267–274, April 2008.

[197] Na Zhang, Fu-Feng Liu, Xiao-Yan Dong, and Yan Sun. Molecular Insight into the Counteraction of Trehalose on Urea-Induced Protein Denaturation Using Molecular Dynamics Simulation. *J. Phys. Chem. B*, 116(24):7040–7047, June 2012.

[198] A. Lerbret, P. Bordat, F. Affouard, A. Hédoux, Y. Guinet, and M. Descamps. How Do Trehalose, Maltose, and Sucrose Influence Some Structural and Dynamical Properties of Lysozyme? Insight from Molecular Dynamics Simulations. *J. Phys. Chem. B*, 111(31):9410–9420, August 2007.

[199] Fu-Feng Liu, Luo Ji, Xiao-Yan Dong, and Yan Sun. Molecular Insight into the Inhibition Effect of Trehalose on the Nucleation and Elongation of Amyloid β-Peptide Oligomers. *J. Phys. Chem. B*, 113(32):11320–11329, August 2009.

[200] C. Molteni and M. Parrinello. Glucose in Aqueous Solution by First Principles Molecular Dynamics. *J. Am. Chem. Soc.*, 120(9):2168–2171, March 1998.

[201] Nicholas F. Dupuis, Erik D. Holmstrom, and David J. Nesbitt. Molecular-crowding effects on single-molecule RNA folding/unfolding thermodynamics and kinetics. *Proc. Natl. Acad. Sci. U.S.A.*, 111(23):8464–8469, June 2014.

[202] Michael Senske, Lisa Törk, Benjamin Born, Martina Havenith, Christian Herrmann, and Simon Ebbinghaus. Protein Stabilization by Macromolecular Crowding through Enthalpy Rather Than Entropy. *J. Am. Chem. Soc.*, 136(25):9036–9041, June 2014.

[203] Kim A. Sharp. Analysis of the size dependence of macromolecular crowding shows that smaller is better. *Proc. Natl. Acad. Sci. U.S.A.*, 112(26):7990–7995, June 2015.

[204] Mohona Sarkar, Conggang Li, and Gary J. Pielak. Soft interactions and crowding. *Biophys Rev*, 5(2):187–194, June 2013.

[205] William M. Aumiller, Bradley W. Davis, Emmanuel Hatzakis, and Christine D. Keating. Interactions of Macromolecular Crowding Agents and Cosolutes with Small-Molecule Substrates: Effect on Horseradish Peroxidase Activity with Two Different Substrates. *J. Phys. Chem. B*, 118(36):10624–10632, September 2014.

[206] A D McLachlan. Protein Structure and Function. *Annu. Rev. Phys. Chem.*, 23(1):165–192, October 1972.

[207] Joseph A. Marsh and Sarah A. Teichmann. Structure, Dynamics, Assembly, and Evolution of Protein Complexes. *Annu. Rev. Biochem.*, 84(1):551–575, June 2015.

[208] Freddie R Salsbury Jr. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Curr Opin Pharmacol*, 10(6):738–744, December 2010.

[209] M. Karplus and J. Kuriyan. Molecular dynamics and protein function. *Proc. Natl. Acad. Sci. U.S.A.*, 102(19):6679–6685, May 2005.

[210] Ron O. Dror, Robert M. Dirks, J.P. Grossman, Huafeng Xu, and David E. Shaw. Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annu. Rev. Biophys.*, 41(1):429–452, June 2012.

[211] Paul Maragakis, Kresten Lindorff-Larsen, Michael P. Eastwood, Ron O. Dror, John L. Klepeis, Isaiah T. Arkin, Morten. Jensen, Huafeng Xu, Nikola Trbovic, Richard A. Friesner, Arthur G. Palmer, and David E. Shaw. Microsecond Molecular Dynamics Simulation Shows Effect of Slow Loop Dynamics on Backbone Amide Order Parameters of Proteins. *J. Phys. Chem. B*, 112(19):6155–6158, May 2008.

[212] Peter L. Freddolino, Feng Liu, Martin Gruebele, and Klaus Schulten. Ten-Microsecond Molecular Dynamics Simulation of a Fast-Folding WW Domain. *Biophys. J.*, 94(10):L75–L77, May 2008.

[213] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw. How Fast-Folding Proteins Fold. *Science*, 334(6055):517–520, October 2011.

[214] Kresten Lindorff-Larsen, Paul Maragakis, Stefano Piana, and David E. Shaw. Picosecond to Millisecond Structural Dynamics in Human Ubiquitin. *J. Phys. Chem. B*, 120(33):8313–8320, August 2016.

[215] José Nelson Onuchic, Zaida Luthey-Schulten, and Peter G. Wolynes. THEORY OF PROTEIN FOLDING: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.*, 48(1):545–600, October 1997.

[216] Ronald Hills and Charles Brooks. Insights from Coarse-Grained Gō Models for Protein Folding and Dynamics. *Int. J. Mol. Sci.*, 10(3):889–905, March 2009.

[217] Cecilia Clementi. Coarse-grained models of protein folding: toy models or predictive tools? *Curr. Opin. Struct. Biol.*, 18(1):10–15, February 2008.

[218] Ram Samudrala and John Moult. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction 1 1Edited by F. Cohen. *J. Mol. Biol.*, 275(5):895–916, February 1998.

[219] Parimal Kar, Srinivasa Murthy Gopal, Yi-Ming Cheng, Alexander Predeus, and Michael Feig. PRIMO: A Transferable Coarse-Grained Force Field for Proteins. *J. Chem. Theory Comput.*, 9(8):3769–3788, August 2013.

[220] Aram Davtyan, Nicholas P. Schafer, Weihua Zheng, Cecilia Clementi, Peter G. Wolynes, and Garegin A. Papoian. AWSEM-MD: Protein Structure Prediction Using Coarse-Grained Physical Potentials and Bioinformatically Based Local Structure Biasing. *J. Phys. Chem. B*, 116(29):8494–8503, July 2012.

[221] Yassmine Chebaro, Samuela Pasquali, and Philippe Derreumaux. The Coarse-Grained OPEP Force Field for Non-Amyloid and Amyloid Proteins. *J. Phys. Chem. B*, 116(30):8741–8752, August 2012.

[222] Karolina Zieba, Magdalena Ślusarz, Rafał Slusarz, Adam Liwo, Cezary Czaplewski, and Adam K. Sieradzan. Extension of the UNRES Coarse-Grained Force Field to Membrane Proteins in the Lipid Bilayer. *J. Phys. Chem. B*, 123(37):7829–7839, September 2019.

[223] Henry Chan, Mathew J. Cherukara, Badri Narayanan, Troy D. Loeffler, Chris Benmore, Stephen K. Gray, and Subramanian K. R. S. Sankaranarayanan. Machine learning coarse grained models for water. *Nat. Commun.*, 10(1):379, December 2019.

[224] Jiang Wang, Simon Olsson, Christoph Wehmeyer, Adrià Pérez, Nicholas E. Charron, Gianni de Fabritiis, Frank Noé, and Cecilia Clementi. Machine Learning of Coarse-Grained Molecular Dynamics Force Fields. *ACS Cent. Sci.*, 5(5):755–767, May 2019.

[225] Cecilia Clementi, Hugh Nymeyer, and José Nelson Onuchic. Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins. *J. Mol. Biol.*, 298(5):937–953, May 2000.

[226] Scott Brown, Nicolas J. Fawzi, and Teresa Head-Gordon. Coarse-grained sequences for protein folding and design. *Proc. Natl. Acad. Sci. U.S.A.*, 100(19):10712–10717, September 2003.

[227] Alex Morriss-Andrews, Frank L. H. Brown, and Joan-Emma Shea. A Coarse-Grained Model for Peptide Aggregation on a Membrane Surface. *J. Phys. Chem. B*, 118(28):8420–8432, July 2014.

[228] Payel Das, Silvina Matysiak, and Cecilia Clementi. Balancing energy and entropy: A minimalist model for the characterization of protein folding landscapes. *Proc. Natl. Acad. Sci. U.S.A.*, 102(29):10141–10146, July 2005.

[229] Hao Wu, Yamini Dalal, and Garegin A. Papoian. Binding Dynamics of Disordered Linker Histone H1 with a Nucleosomal Particle. *J. Mol. Biol.*, 433(6):166881, March 2021.

[230] David S. Yang, Arash Saeedi, Aram Davtyan, Mohsen Fathi, Michael B. Sherman, Mohammad S. Safari, Alena Klindziuk, Michelle C. Barton, Navin Varadarajan, Anatoly B. Kolomeisky, and Peter G. Vekilov. Mesoscopic protein-rich clusters host the nucleation of mutant p53 amyloid fibrils. *Proc. Natl. Acad. Sci. U.S.A.*, 118(10):e2015618118, March 2021.

[231] Scott P. O. Danielsen, James McCarty, Joan-Emma Shea, Kris T. Delaney, and Glenn H. Fredrickson. Molecular design of self-coacervation phenomena in block polyampholytes. *Proc. Natl. Acad. Sci. U.S.A.*, 116(17):8224–8232, April 2019.

[232] Bogdan Barz and Brigita Urbanc. Minimal Model of Self-Assembly: Emergence of Diversity and Complexity. *J. Phys. Chem. B*, 118(14):3761–3770, April 2014.

[233] Stefan Auer, Filip Meersman, Christopher M. Dobson, and Michele Vendruscolo. A Generic Mechanism of Emergence of Amyloid Protofilaments from Disordered Oligomeric Aggregates. *PLoS Comput. Biol.*, 4(11):e1000222, November 2008.

[234] K. Yue, K. M. Fiebig, P. D. Thomas, H. S. Chan, E. I. Shakhnovich, and K. A. Dill. A test of lattice protein folding algorithms. *Proc. Natl. Acad. Sci. U.S.A.*, 92(1):325–329, January 1995.

[235] Jian Zhou, Ian F. Thorpe, Sergey Izvekov, and Gregory A. Voth. Coarse-Grained Peptide Modeling Using a Systematic Multiscale Approach. *Biophys. J.*, 92(12):4289–4303, June 2007.

[236] S. Y. Mashayak, Mara N. Jochum, Konstantin Koschke, N. R. Aluru, Victor Rühle, and Christoph Junghans. Relative Entropy and Optimization-Driven Coarse-Graining Methods in VOTCA. *PLOS ONE*, 10(7):e0131754, 2015. Publisher: Public Library of Science.

[237] Sergei Izvekov and Gregory A. Voth. A Multiscale Coarse-Graining Method for Biomolecular Systems. *J. Phys. Chem. B*, 109(7):2469–2473, February 2005.

[238] S.M. Vaiana, M. Manno, A. Emanuele, M.B. Palma-Vittorelli, and M.U. Palma. The role of solvent in protein folding and in aggregation. *J. Biol. Phys.*, 27(2/3):133–145, 2001.

[239] Hiraku Oshima and Masahiro Kinoshita. Essential roles of protein-solvent many-body correlation in solvent-entropy effect on protein folding and denaturation: Comparison between hard-sphere solvent and water. *J. Chem. Phys.*, 142(14):145103, April 2015.

[240] C. Nick Pace, Saul Treviño, Erode Prabhakaran, and J. Martin Scholtz. Protein structure, stability and solubility in water and other solvents. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 359(1448):1225–1235, August 2004.

[241] D. K. Klimov and D. Thirumalai. Mechanisms and kinetics of beta -hairpin formation. *Proc. Natl. Acad. Sci. U.S.A.*, 97(6):2544–2549, March 2000.

[242] D. K. Klimov and D. Thirumalai. Linking rates of folding in lattice models of proteins with underlying thermodynamic characteristics. *J. Chem. Phys.*, 109(10):4119–4125, September 1998.

[243] Yuzo Ueda, Hiroshi Taketomi, and Nobuhiro G? Studies on protein folding, unfolding, and fluctuations by computer simulation. II. A. Three-dimensional lattice model of lysozyme. *Biopolymers*, 17(6):1531–1548, June 1978.

[244] Paulo C. T. Souza, Riccardo Alessandri, Jonathan Barnoud, Sebastian Thallmair, Ignacio Faustino, Fabian Grünewald, Ilias Patmanidis, Haleh Abdizadeh, Bart M. H. Bruininks, Tsjerk A. Wassenaar, Peter C. Kroon, Josef Melcr, Vincent Nieto, Valentina Corradi, Hanif M. Khan, Jan Domański, Matti Javanainen, Hector Martinez-Seara, Nathalie Reuter, Robert B. Best, Ilpo Vattulainen, Luca Monticelli, Xavier Periole, D. Peter Tieleman, Alex H. de Vries, and Siewert J. Marrink. Martini 3: a general purpose force field for coarse-grained molecular dynamics. *Nat. Methods*, 18(4):382–388, April 2021.

[245] Paulo C. T. Souza, Sebastian Thallmair, Paolo Conflitti, Carlos Ramírez-Palacios, Riccardo Alessandri, Stefano Raniolo, Vittorio Limongelli, and Siewert J. Marrink. Protein–ligand binding with the coarse-grained Martini model. *Nat Commun*, 11(1):3714, December 2020.

[246] Paulo C. T. Souza, Vittorio Limongelli, Sangwook Wu, Siewert J. Marrink, and Luca Monticelli. Perspectives on High-Throughput Ligand/Protein Docking With Martini MD Simulations. *Front. Mol. Biosci.*, 8:657222, March 2021.

[247] Cristian Gobbo, Isabelle Beurroies, David de Ridder, Rienk Eelkema, Siewert J. Marrink, Steven De Feyter, Jan H. van Esch, and Alex H. de Vries. MARTINI Model for Physisorption of Organic Molecules on Graphite. *J. Phys. Chem. C*, 117(30):15623–15631, August 2013.

[248] C. Arnarez, S. J. Marrink, and X. Periole. Molecular mechanism of cardiolipin-mediated assembly of respiratory chain supercomplexes. *Chem. Sci.*, 7(7):4435–4443, 2016.

[249] Fude Sun, Carsten F.E. Schroer, Lida Xu, Huiwei Yin, Siewert J. Marrink, and Shi-Zhong Luo. Molecular Dynamics of the Association of L-Selectin and FERM Regulated by PIP2. *Biophys. J.*, 114(8):1858–1868, April 2018.

[250] Tsjerk A. Wassenaar, Kristyna Pluhackova, Anastassiia Moussatova, Durba Sengupta, Siewert J. Marrink, D. Peter Tieleman, and Rainer A. Böckmann. High-Throughput Simulations of Dimer and Trimer Assembly of Membrane Proteins. The DAFT Approach. *J. Chem. Theory Comput.*, 11(5):2278–2291, May 2015.

[251] John Karanicolas and Charles L. Brooks. The origins of asymmetry in the folding transition states of protein L and protein G. *Protein Science*, 11(10):2351–2361, April 2009.

[252] Adolfo B. Poma, Marek Cieplak, and Panagiotis E. Theodorakis. Combining the MARTINI and Structure-Based Coarse-Grained Approaches for the Molecular Dynamics Studies of Conformational Transitions in Proteins. *J. Chem. Theory Comput.*, 13(3):1366–1374, March 2017.

[253] Md. Iqbal Mahmood, Adolfo B. Poma, and Kei-ichi Okazaki. Optimizing Gō-MARTINI Coarse-Grained Model for F-BAR Protein on Lipid Membrane. *Front. Mol. Biosci.*, 8:619381, February 2021.

[254] Abhilash Sahoo and Silvina Matysiak. Effects of applied surface-tension on membrane-assisted aβ aggregation. *Phys. Chem. Chem. Phys.*, 23(36):20627–20633, 2021.

[255] Luca Monticelli, Senthil K. Kandasamy, Xavier Periole, Ronald G. Larson, D. Peter Tieleman, and Siewert-Jan Marrink. The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J. Chem. Theory Comput.*, 4(5):819–834, May 2008.

[256] Jing Huang, Sarah Rauscher, Grzegorz Nawrocki, Ting Ran, Michael Feig, Bert L de Groot, Helmut Grubmüller, and Alexander D MacKerell. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods*, 14(1):71–73, January 2017.

[257] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, 79(2):926–935, 1983.

[258] Sunhwan Jo, Taehoon Kim, Vidyashankara G. Iyer, and Wonpil Im. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.*, 29(11):1859–1865, August 2008.

[259] J. P. Gallivan and D. A. Dougherty. Cation-pi interactions in structural biology. *Proc. Natl. Acad. Sci. U.S.A.*, 96(17):9459–9464, August 1999.

[260] Neal J. Zondlo. Aromatic–Proline Interactions: Electronically Tunable CH/π Interactions. *Acc. Chem. Res.*, 46(4):1039–1049, April 2013.

[261] Emily G Baker, Christopher Williams, Kieran L Hudson, Gail J Bartlett, Jack W Heal, Kathryn L Porter Goff, Richard B Sessions, Matthew P Crump, and Derek N Woolfson. Engineering protein stability with atomic precision in a monomeric miniprotein. *Nat Chem Biol*, 13(7):764–770, July 2017.

[262] Motohiro Nishio, Yoji Umezawa, Jacques Fantini, Manfred S. Weiss, and Pinak Chakrabarti. Ch–π hydrogen bonds in biological macromolecules. *Phys. Chem. Chem. Phys.*, 16(25):12648–12683, 2014.

[263] Jing Huang, Pedro E. M. Lopes, Benoît Roux, and Alexander D. MacKerell. Recent Advances in Polarizable Force Fields for Macromolecules: Microsecond Simulations of Proteins Using the Classical Drude Oscillator Model. *J. Phys. Chem. Lett.*, 5(18):3144–3150, September 2014.

[264] Haibo Yu, Troy W. Whitfield, Edward Harder, Guillaume Lamoureux, Igor Vorobyov, Victor M. Anisimov, Alexander D. MacKerell, and Benoît Roux. Simulating Monovalent and Divalent Ions in Aqueous Solution Using a Drude Polarizable Force Field. *J. Chem. Theory Comput.*, 6(3):774–786, March 2010.

[265] Eng-Hui Yap, Nicolas Lux Fawzi, and Teresa Head-Gordon. A coarse-grained α-carbon protein model with anisotropic hydrogen-bonding. *Proteins*, 70(3):626–638, February 2008.

[266] Harald A. Posch, William G. Hoover, and Franz J. Vesely. Canonical dynamics of the Nos$\tilde{e}$ oscillator: Stability, order, and chaos. *Phys. Rev. A*, 33(6):4253–4265, June 1986.

[267] Lingle Wang, Richard A Friesner, and BJ Berne. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (rest2). *The Journal of Physical Chemistry B*, 115(30):9431–9438, 2011.

[268] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.

[269] Michael R Shirts and John D Chodera. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.*, 129(12):124105, 2008.

[270] Dietmar Paschek, Sascha Hempel, and Angel E García. Computing the stability diagram of the trp-cage miniprotein. *Proc. Natl. Acad. Sci. U.S.A.*, 105(46):17754–17759, 2008.

[271] Jia Lin Huang, Michael E Noss, Karson M Schmidt, Leigh Murray, and Michelle R Bunagan. The effect of neat ionic liquid on the folding of short peptides. *Chemical Communications*, 47(28):8007–8009, 2011.

[272] Ryan Day, Dietmar Paschek, and Angel E Garcia. Microsecond simulations of the folding/unfolding thermodynamics of the trp-cage miniprotein. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1889–1899, 2010.

[273] Andrea G Cochran, Nicholas J Skelton, and Melissa A Starovasnik. Tryptophan zippers: Stable, monomeric β-hairpins. *Proc. Natl. Acad. Sci. U.S.A.*, 98(10):5578–5583, 2001.

[274] Deguo Du, Yongjin Zhu, Cheng-Yen Huang, and Feng Gai. Understanding the key factors that control the rate of β-hairpin folding. *Proc. Natl. Acad. Sci. U.S.A.*, 101(45):15915–15920, 2004.

[275] Aimee Byrne, D Victoria Williams, Bipasha Barua, Stephen J Hagen, Brandon L Kier, and Niels H Andersen. Folding dynamics and pathways of the trp-cage miniproteins. *Biochemistry*, 53(38):6011–6021, 2014.

[276] Dietmar Paschek, Hugh Nymeyer, and Angel E García. Replica exchange simulation of reversible folding/unfolding of the trp-cage miniprotein in explicit solvent: On the structure and possible role of internal water. *J. Mol. Biol.*, 157(3):524–533, 2007.

[277] Lucas N. R. Wafer, Werner W. Streicher, and George I. Makhatadze. Thermodynamics of the Trp-cage Miniprotein Unfolding in Urea. *Proteins*, 78(6):1376–1381, May 2010.

[278] Beatrice N. Markiewicz, Lijiang Yang, Robert M. Culik, Yi Qin Gao, and Feng Gai. How Quickly Can a β-Hairpin Fold from Its Transition State? *J. Phys. Chem. B*, 118(12):3317–3325, March 2014.

[279] Andreas Reiner, Peter Henklein, and Thomas Kiefhaber. An unlocking/relocking barrier in conformational fluctuations of villin headpiece subdomain. *Proc. Natl. Acad. Sci. U.S.A.*, 107(11):4955–4960, 2010.

[280] Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules*, 22(12):2577–2637, 1983.

[281] Yiming Wang, Samuel J. Bunce, Sheena E. Radford, Andrew J. Wilson, Stefan Auer, and Carol K. Hall. Thermodynamic phase diagram of amyloid-β (16–22) peptide. *Proc. Natl. Acad. Sci. U.S.A.*, 116(6):2091–2096, February 2019.

[282] Dongdong Lin, Yin Luo, Shan Wu, Qianqian Ma, Guanghong Wei, and Xinju Yang. Investigation of the Aggregation Process of Amyloid-β-(16-22) Peptides and the Dissolution of Intermediate Aggregates. *Langmuir*, 30(11):3170–3175, March 2014.

[283] Suman Samantray, Feng Yin, Batuhan Kav, and Birgit Strodel. Different Force Fields Give Rise to Different Amyloid Aggregation Pathways in Molecular Dynamics Simulations. *J. Chem. Inf. Model.*, 60(12):6462–6475, December 2020.

[284] John J Balbach, Yoshitaka Ishii, Oleg N Antzutkin, Richard D Leapman, Nancy W Rizzo, Fred Dyda, Jennifer Reed, and Robert Tycko. Amyloid fibril formation by aβ16-22, a seven-residue fragment of the alzheimer's β-amyloid peptide, and structural characterization by solid state nmr. *Biochemistry*, 39(45):13748–13759, 2000.

[285] John J. Balbach, Yoshitaka Ishii, Oleg N. Antzutkin, Richard D. Leapman, Nancy W. Rizzo, Fred Dyda, Jennifer Reed, and Robert Tycko. Amyloid Fibril Formation by Aβ $_{16\text{-}22}$, a Seven-Residue Fragment of the Alzheimer's β-Amyloid Peptide, and Structural Characterization by Solid State NMR †. *Biochemistry*, 39(45):13748–13759, November 2000.

[286] Phuong H Nguyen, Mai Suan Li, and Philippe Derreumaux. Effects of all-atom force fields on amyloid oligomerization: Replica exchange molecular dynamics simulations of the aβ 16–22 dimer and trimer. *Phys. Chem. Chem. Phys.*, 13(20):9778–9788, 2011.

[287] Kevin R MacKenzie, James H Prestegard, and Donald M Engelman. A transmembrane helix dimer: structure and implications. *Science*, 276(5309):131–133, 1997.

[288] Madhusoodanan Mottamal, Jinming Zhang, and Themis Lazaridis. Energetics of the native and non-native states of the glycophorin transmembrane helix dimer. *PROTEINS: Structure, Function, and Bioinformatics*, 62(4):996–1009, 2006.

[289] Emi Psachoulia, Philip W Fowler, Peter J Bond, and Mark SP Sansom. Helix- helix interactions in membrane proteins: Coarse-grained simulations of glycophorin a helix dimerization. *Biochemistry*, 47(40):10503–10512, 2008.

[290] P. McColgan and S. J. Tabrizi. Huntington's disease: a clinical review. *Eur J Neurol*, 25(1):24–34, January 2018.

[291] Christopher A Ross and Sarah J Tabrizi. Huntington's disease: from molecular pathogenesis to clinical treatment. *The Lancet Neurology*, 10(1):83–98, January 2011.

[292] Gillian Bates. Huntingtin aggregation and toxicity in Huntington's disease. *The Lancet*, 361(9369):1642–1644, May 2003.

[293] Francis O Walker. Huntington's disease. *The Lancet*, 369(9557):218–228, January 2007.

[294] Matthew P. Parsons and Lynn A. Raymond. Huntington Disease. In *Neurobiology of Brain Disorders*, pages 303–320. Elsevier, 2015.

[295] Hueng-Chuen Fan, Li-Ing Ho, Ching-Shiang Chi, Shyi-Jou Chen, Giia-Sheun Peng, Tzu-Min Chan, Shinn-Zong Lin, and Horng-Jyh Harn. Polyglutamine (PolyQ) Diseases: Genetics to Treatments. *Cell Transplant*, 23(4-5):441–458, May 2014.

[296] J. Shao and M. I. Diamond. Polyglutamine diseases: emerging concepts in pathogenesis and therapy. *Human Molecular Genetics*, 16(R2):R115–R123, July 2007.

[297] Sunjay Barton, Ron Jacak, Sagar D. Khare, Feng Ding, and Nikolay V. Dokholyan. The Length Dependence of the PolyQ-mediated Protein Aggregation. *Journal of Biological Chemistry*, 282(35):25487–25492, August 2007.

[298] Tetyana Yushchenko, Elke Deuerling, and Karin Hauser. Insights into the Aggregation Mechanism of PolyQ Proteins with Different Glutamine Repeat Lengths. *Biophysical Journal*, 114(8):1847–1857, April 2018.

[299] Keizo Sugaya, Shiro Matsubara, Yasuhiro Kagamihara, Akihiro Kawata, and Hideaki Hayashi. Polyglutamine Expansion Mutation Yields a Pathological Epitope Linked to Nucleation of Protein Aggregate: Determinant of Huntington's Disease Onset. *PLoS ONE*, 2(7):e635, July 2007.

[300] Andreas Vitalis, Xiaoling Wang, and Rohit V. Pappu. Quantitative Characterization of Intrinsic Disorder in Polyglutamine: Insights from Analysis Based on Polymer Theories. *Biophysical Journal*, 93(6):1923–1937, September 2007.

[301] Scott L. Crick, Murali Jayaraman, Carl Frieden, Ronald Wetzel, and Rohit V. Pappu. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form collapsed structures in aqueous solutions. *Proc. Natl. Acad. Sci. U.S.A.*, 103(45):16764–16769, November 2006.

[302] Natasha H. Rhys, Alan K. Soper, and Lorna Dougan. Hydrophilic Association in a Dilute Glutamine Solution Persists Independent of Increasing Temperature. *J. Phys. Chem. B*, 119(51):15644–15651, December 2015.

[303] Jeremy T Lant, Rashmi Kiri, Martin L Duennwald, and Patrick O'Donoghue. Formation and persistence of polyglutamine aggregates in mistranslating cells. *Nucleic Acids Research*, 49(20):11883–11899, November 2021.

[304] Hongsuk Kang, Francisco X. Vázquez, Leili Zhang, Payel Das, Leticia Toledo-Sherman, Binquan Luan, Michael Levitt, and Ruhong Zhou. Emerging β-Sheet Rich Conformations in Supercompact Huntingtin Exon-1 Mutant Structures. *J. Am. Chem. Soc.*, 139(26):8820–8827, July 2017.

[305] Nitin K. Pandey, J. Mario Isas, Anoop Rawat, Rachel V. Lee, Jennifer Langen, Priyatama Pandey, and Ralf Langen. The 17-residue-long N terminus in huntingtin controls stepwise aggregation in solution and on membranes via different mechanisms. *Journal of Biological Chemistry*, 293(7):2597–2605, February 2018.

[306] Maria Lucia Pigazzini, Mandy Lawrenz, Anca Margineanu, Gabriele S. Kaminski Schierle, and Janine Kirstein. An Expanded Polyproline Domain Maintains Mutant Huntingtin Soluble in vivo and During Aging. *Front. Mol. Neurosci.*, 14:721749, October 2021.

[307] James R. Arndt, Maxmore Chaibva, and Justin Legleiter. The emerging role of the first 17 amino acids of huntingtin in Huntington's disease. *Biomolecular Concepts*, 6(1):33–46, March 2015.

[308] Guillaume Drin and Bruno Antonny. Amphipathic helices and membrane curvature. *FEBS Letters*, 584(9):1840–1847, May 2010.

[309] Nikos S Hatzakis, Vikram K Bhatia, Jannik Larsen, Kenneth L Madsen, Pierre-Yves Bolinger, Andreas H Kunding, John Castillo, Ulrik Gether, Per Hedegård, and Dimitrios Stamou. How curved membranes recruit amphipathic helices and protein anchoring motifs. *Nat Chem Biol*, 5(11):835–841, November 2009.

[310] Anna Giangaspero, Luca Sandri, and Alessandro Tossi. Amphipathic α helical antimicrobial peptides.: A systematic study of the effects of structural and physical properties on biological activity. *European Journal of Biochemistry*, 268(21):5589–5600, November 2001.

[311] Ashwani K Thakur, Murali Jayaraman, Rakesh Mishra, Monika Thakur, Veronique M Chellgren, In-Ja L Byeon, Dalaver H Anjum, Ravindra Kodali, Trevor P Creamer, James F Conway, Angela M Gronenborn, and Ronald Wetzel. Polyglutamine disruption of the huntingtin exon 1 N terminus triggers a complex aggregation mechanism. *Nat Struct Mol Biol*, 16(4):380–389, April 2009.

[312] Cody L. Hoop, Hsiang-Kai Lin, Karunakar Kar, Gábor Magyarfalvi, Jonathan M. Lamley, Jennifer C. Boatz, Abhishek Mandal, Józef R. Lewandowski, Ronald Wetzel, and Patrick C. A. van der Wel. Huntingtin exon 1 fibrils feature an interdigitated β-hairpin–based polyglutamine core. *Proc. Natl. Acad. Sci. U.S.A.*, 113(6):1546–1551, February 2016.

[313] Valentina Fodale, Natalie C. Kegulian, Margherita Verani, Cristina Cariulo, Lucia Azzollini, Lara Petricca, Manuel Daldin, Roberto Boggio, Alessandro Padova, Rainer Kuhn, Robert Pacifici, Douglas Macdonald, Ryan C. Schoenfeld, Hyunsun Park, J. Mario Isas, Ralf Langen, Andreas Weiss, and Andrea Caricasole. Polyglutamine- and Temperature-Dependent Conformational Rigidity in Mutant Huntingtin Revealed

by Immunoassays and Circular Dichroism Spectroscopy. *PLoS ONE*, 9(12):e112262, December 2014.

[314] Yoshitaka Nagai, Takashi Inui, H Akiko Popiel, Nobuhiro Fujikake, Kazuhiro Hasegawa, Yoshihiro Urade, Yuji Goto, Hironobu Naiki, and Tatsushi Toda. A toxic monomeric conformer of the polyglutamine protein. *Nat Struct Mol Biol*, 14(4):332–340, April 2007.

[315] Matthias Michalek, Evgeniy S. Salnikov, and Burkhard Bechinger. Structure and Topology of the Huntingtin 1–17 Membrane Anchor by a Combined Solution and Solid-State NMR Approach. *Biophysical Journal*, 105(3):699–710, 2013.

[316] Meixin Tao, Nitin K. Pandey, Ryan Barnes, Songi Han, and Ralf Langen. Structure of Membrane-Bound Huntingtin Exon 1 Reveals Membrane Interaction and Aggregation Mechanisms. *Structure*, 27(10):1570–1580.e4, October 2019.

[317] Nicholas W. Kelley, Xuhui Huang, Stephen Tam, Christoph Spiess, Judith Frydman, and Vijay S. Pande. The Predicted Structure of the Headpiece of the Huntingtin Protein and Its Implications on Huntingtin Aggregation. *Journal of Molecular Biology*, 388(5):919–927, May 2009.

[318] Yanting Wang and Gregory A. Voth. Molecular Dynamics Simulations of Polyglutamine Aggregation Using Solvent-Free Multiscale Coarse-Grained Models. *J. Phys. Chem. B*, 114(26):8735–8743, July 2010.

[319] Maciej Długosz and Joanna Trylska. Secondary Structures of Native and Pathogenic Huntingtin N-Terminal Fragments. *J. Phys. Chem. B*, 115(40):11597–11608, October 2011.

[320] Vinal V. Lakhani, Feng Ding, and Nikolay V. Dokholyan. Polyglutamine Induced Misfolding of Huntingtin Exon1 is Modulated by the Flanking Sequences. *PLoS Comput Biol*, 6(4):e1000772, April 2010.

[321] Leili Zhang, Hongsuk Kang, Francisco X. Vázquez, Leticia Toledo-Sherman, Binquan Luan, and Ruhong Zhou. Molecular Mechanism of Stabilizing the Helical Structure of Huntingtin N17 in a Micellar Environment. *The Journal of Physical Chemistry B*, 121(18):4713–4721, 2017. Publisher: American Chemical Society.

[322] Sébastien Côté, Vincent Binette, Evgeniy S. Salnikov, Burkhard Bechinger, and Normand Mousseau. Probing the Huntingtin 1-17 Membrane Anchor on a Phospholipid Bilayer by Using All-Atom Simulations. *Biophysical Journal*, 108(5):1187–1198, 2015.

[323] Kevin J. Boyd and Eric R. May. BUMPy: A model-independent tool for constructing lipid bilayers of varying curvature and composition. *Journal of chemical theory and computation*, 14(12):6642–6652, 2018.

[324] Mingchen Chen and Peter G. Wolynes. Aggregation landscapes of Huntingtin exon 1 protein fragments and the critical repeat length for the onset of Huntington's disease. *Proc. Natl. Acad. Sci. U.S.A.*, 114(17):4406–4411, April 2017.

[325] Sébastien Buchoux. FATSLiM: a fast and robust software to analyze MD simulations of membranes. *Bioinformatics*, 33(1):133–134, January 2017.

[326] Annika Urbanek, Matija Popovic, Anna Morató, Alejandro Estaña, Carlos A. Elena-Real, Pablo Mier, Aurélie Fournet, Frédéric Allemand, Stephane Delbecq, Miguel A. Andrade-Navarro, Juan Cortés, Nathalie Sibille, and Pau Bernadó. Flanking Regions Determine the Structure of the Poly-Glutamine in Huntingtin through Mechanisms Common among Glutamine-Rich Human Proteins. *Structure*, 28(7):733–746.e5, July 2020.

[327] E.P Melo, M.R Aires-Barros, and J.M.S Cabral. Reverse micelles and protein biotechnology. In *Biotechnology Annual Review*, volume 7, pages 87–129. Elsevier, 2001.

[328] Joëlle Bigay, Jean-François Casella, Guillaume Drin, Bruno Mesmin, and Bruno Antonny. ArfGAP1 responds to membrane curvature through the folding of a lipid packing sensor motif. *EMBO J*, 24(13):2244–2253, July 2005.

[329] Xinyao Yi, Yi Zhang, Ming Gong, Xiang Yu, Narek Darabedian, Jie Zheng, and Feimeng Zhou. Ca2+ Interacts with Glu-22 of Abeta(1–42) and Phospholipid Bilayers to Accelerate the Abeta(1–42) Aggregation Below the Critical Micelle Concentration. *Biochemistry*, 54(41):6323–6332, 2015. Publisher: American Chemical Society.

[330] Marina E. Monteiro, Maria J. Sarmento, and Fábio Fernandes. Role of calcium in membrane interactions by PI(4,5)P2-binding proteins, 2014.

[331] M. Mirza, Y. Guo, K. Arnold, C. J. Van Oss, and S. Ohki. Hydrophobizing effect of cations on acidic phospholipid membranes. *J Dispers Sci Technol*, 19(6-7):951–962, 1998.

[332] H Hauser and R M Dawson. The binding of calcium at lipid-water interfaces. *Eur. J. Biochem.*, 1(1):61–619, 1967.

[333] Baharan Ali Doosti, Weria Pezeshkian, Dennis S. Bruhn, John H. Ipsen, Himanshu Khandelia, Gavin D.M. Jeffries, and Tatsiana Lobovkina. Membrane Tubulation in Lipid Vesicles Triggered by the Local Application of Calcium Ions. *Langmuir*, 33(41):11010–11017, 2017.

[334] Z. T. Graber, Z. Shi, and T. Baumgart. Cations induce shape remodeling of negatively charged phospholipid membranes. *Phys. Chem. Chem. Phys.*, 19(23):15285–15295, 2017.

[335] Mijo Simunovic, Ka Yee C. Lee, and Patricia Bassereau. Celebrating Soft Matter's 10th anniversary: screening of the calcium-induced spontaneous curvature of lipid membranes. *Soft Matter*, 11(25):5030–5036, 2015.

[336] Ewa Stodolak-Zych, Piotr Jeleń, Ewa Dzierzkowska, Małgorzata Krok-Borkowicz, Łukasz Zych, Maciej Boguń, Alicja Rapacz-Kmita, and Beata Kolesińska. Modification of chitosan fibers with short peptides as a model of synthetic extracellular matrix. *Journal of Molecular Structure*, 1211:128061, 2020.

[337] A. Muxika, A. Etxabide, J. Uranga, P. Guerrero, and K. de la Caba. Chitosan as a bioactive polymer: Processing, properties and applications. *International Journal of Biological Macromolecules*, 105:1358–1368, 2017.

[338] Hongcheng Xu and Silvina Matysiak. Effect of pH on chitosan hydrogel polymer network structure. *Chemical Communications*, 53(53):7373–7376, 2017. Publisher: The Royal Society of Chemistry.