Elucidating the impact of glycolipid-containing microenvironments on amyloid structure and aggregation

[redacted]

Abstract. We aim to investigate the structural and biophysical properties of amyloids in relevant biological microenvironments using all-atom simulations. Our specific interests are in determining how lipid membrane composition and the presence of glycosaminoglycans impact amyloid folding and aggregation. This project will delineate the basic biological understanding of amyloid biophysics and how other biomolecules influence the aggregation pathway and contribute to disease pathways.

Research Objectives. Amyloids are insoluble protein aggregates that have functional roles in some biological systems but are also hallmark features of neurodegenerative diseases such as Alzheimer's and Parkinson's. These structures form through the aggregation of intrinsically disordered protein (IDP) monomers, where central hydrophobic regions promote β-sheet formation, and intermolecular hydrogen bonding drives formation of oligomers and fibrils with the potential to be cytotoxic. Although all amyloids share the ability to form fibrillar structures, their cytotoxicity varies widely. For example, β-endorphin (βE) forms functional amyloids involved in paracrine signaling and exhibits no known cytotoxic effects, while amyloid-β (Aβ) is strongly associated with Alzheimer's disease pathogenesis. Other amyloidogenic proteins such as Islet amyloid polypeptide (IAPP) and α-synuclein (αS), also serve physiological signaling roles but are implicated in Type II diabetes and Parkinson's disease, respectively. This variability raises a fundamental question: what factors determine whether the level of amyloid cytotoxicity?

Our overall goal is to uncover the structural, environment, and biophysical characteristics that distinguish cytotoxic from functional amyloids. A growing body of evidence suggests that the surrounding microenvironment, including the presence and composition of lipids and glycosaminoglycans (GAGs), plays a key role in modulating amyloid aggregation and cytotoxicity. To probe these influences, we will employ molecular dynamics (MD) simulations to resolve atomic-level interactions and conformational landscapes of the amyloids during aggregation in complex environments. These simulations will help elucidate the molecular mechanisms underlying amyloid functionality versus cytotoxicity.

Lipids. The lipid-chaperone hypothesis proposes that free lipids in aqueous environments can interact with amyloidogenic proteins to form lipid-protein complexes that promote membrane insertion, causing cellular damage and contributing to disease pathology.[1] In our prior work, we used MD simulations to investigate the oligomerization pathway and membrane interactions of $A\beta_{42}$ tetramers.[2,3] In the presence of model membranes containing phosphatidylcholine (POPC) and cholesterol-raft domains, tetramers underwent secondary structure rearrangements, adopting conformations resembling intermediates in fibril formation pathways and showing agreements with experimental data. We have since expanded this work to examine larger oligomeric forms of $A\beta_{42}$, such as hexamers, octamers, and decamers, finding that POPC alters traditional oligomerization mechanisms. In these simulations, residues that typically mediate oligomer packing in lipid-free controls become more solvent-exposed in the presence of POPC (Figure 1). This rearrangement may promote $A\beta_{42}$ aggregation by increasing formation of hydrophobic β-strand motifs, which can serve as nucleation scaffolds for fibril formation in disease states.

Building on our preliminary findings that lipids alter the aggregation pathway of $A\beta_{42}$, we will next evaluate how the lipid microenvironment influences aggregation of full-length βE_{31}, $IAPP_{37}$, and αS_{140}. We will run initial control systems containing only the amyloid of interest to obtain insights into the lipid-free oligomerization pathway for hexameric, octomeric, decameric systems. These simulations will be conducted using the CHARMM36m[4] force field. Secondary structure formation will be monitored

(1) Sciacca, M. F.; Lolicato, F.; Tempra, C.; Scollo, F.; Sahoo, B. R.; Watson, M. D.; García-Viñuales, S.; Milardi, D.; Raudino, A.; Lee, J. C.; Ramamoorthy, A.; La Rosa, C. Lipid-Chaperone Hypothesis: A Common Molecular Mechanism of Membrane Disruption by Intrinsically Disordered Proteins. *ACS Chem. Neurosci.* **2020**, *11* (24), 4336–4350. https://doi.org/10.1021/acschemneuro.0c00588.
(2) Brown, A. M.; Bevan, D. R. Molecular Dynamics Simulations of Amyloid β-Peptide (1-42): Tetramer Formation and Membrane Interactions. *Biophys. J.* **2016**, *111* (5), 937–949. https://doi.org/10.1016/j.bpj.2016.08.001.
(3) Brown, A. M.; Bevan, D. R. Influence of Sequence and Lipid Type on Membrane Perturbation by Human and Rat Amyloid β-Peptide (1-42). *Arch. Biochem. Biophys.* **2017**, *614*, 1–13. https://doi.org/10.1016/j.abb.2016.11.006.

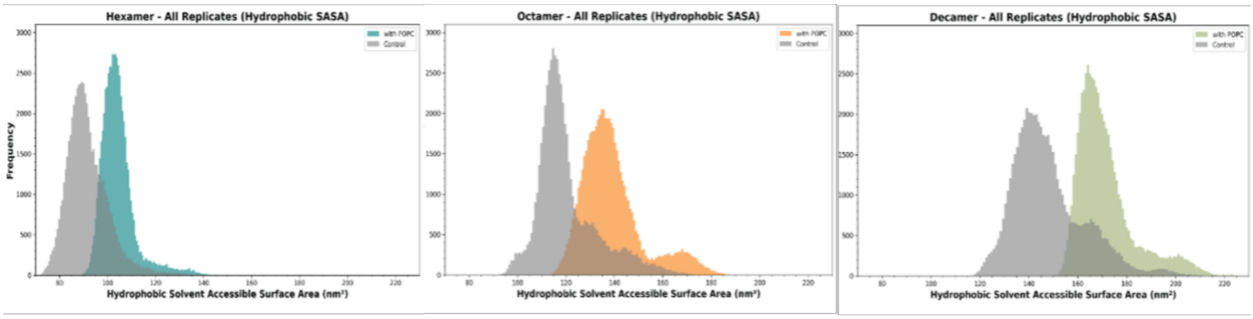

Figure 1. Hydrophobic solvent accessible surface area is increased in Aβ$_{42}$:POPC oligomers compared to Aβ$_{42}$ controls. Hydrophobic solvent accessible surface area distributions for the Aβ$_{42}$ controls (gray) and Aβ$_{42}$:POPC oligomers (blue, yellow, green).

throughout the simulations using the dictionary of protein secondary structure[5,6] to track β-strand formation and deformation. Radius of gyration and eccentricity will be utilized to determine the level of monomer compaction and estimate the extent of oligomerization. RMSD-based clustering will be utilized to identify representative oligomer structures from each system for future simulations involving lipids.

Glycosaminoglycans. GAGS are long, negatively charged sugar chains found within the extracellular matrix and on cell surfaces. GAGs have been shown to accelerate amyloid fibril formation by engaging in electrostatic interactions with positively charged residues on amyloidogenic proteins.[7,8] Notably, our group demonstrated that such charged residue interactions influence membrane binding and integrity.[2] Experimental studies have shown that cells deficient in heparan sulfate (HS), a common GAG, exhibit resistance to Aβ-induced toxicity[9,10], supporting the hypothesis that GAGs can protect against cytotoxicity by promoting conversion of toxic oligomers into inert amyloid fibrils.[11] HS is typically present as a component of transmembrane proteoglycans known as syndecans (Figure 2), which aid in the internalization of amyloids into cells.[12,13] Despite, the importance of GAGs in modulating amyloid behavior, the structural and functional role of HS, when linked to syndecans and in the context of amyloid aggregation, remain poorly understood. In this work, we aim to determine how HS influences both amyloid aggregation and cellular internalization when modeling in its native microenvironment.

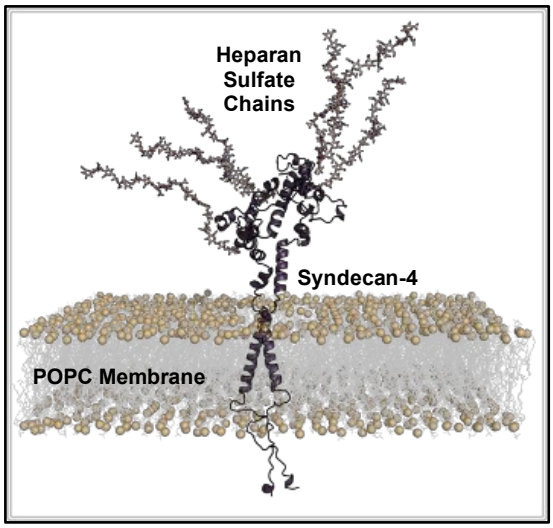

Figure 2. Snapshot of the Syndecan-4 dimer transmembrane glycoprotein system construction post-equilibration.

Currently, we have setup systems containing HS bound to a dimer of syndecan-4 within a POPC membrane using GROMACS and the CHARMM36m force field (Figure 2). This system will serve as a control for comparison against systems containing Aβ$_{42}$, βE$_{31}$, IAPP$_{37}$, and αS$_{140}$. The control simulation will provide insight into the binding interactions between syndecan and its HS chains and their dynamics, helping to infer how these structures recruit and bind amyloids. Non-covalent interactions between the HS and syndecan will be tracked across the simulations and solvent-accessible surface area will be utilized to identify regions of HS that are most likely to bind amyloids. An evaluation of the membrane within this control environment will also be conducted using deuterium order parameters to measure the

(4) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71–73. https://doi.org/10.1038/nmeth.4067.
(5) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers* **1983**, *22* (12), 2577–2637. https://doi.org/10.1002/bip.360221211.
(6) Touw, W. G.; Baakman, C.; Black, J.; te Beek, T. A. H.; Krieger, E.; Joosten, R. P.; Vriend, G. A Series of PDB-Related Databanks for Everyday Needs. *Nucleic Acids Res.* **2015**, *43* (Database issue), D364-368. https://doi.org/10.1093/nar/gku1028.
(7) Iannuzzi, C.; Irace, G.; Sirangelo, I. The Effect of Glycosaminoglycans (GAGs) on Amyloid Aggregation and Toxicity. *Molecules* **2015**, *20* (2), 2510–2528. https://doi.org/10.3390/molecules20022510.
(8) Nguyen, K.; Rabenstein, D. L. Interaction of the Heparin-Binding Consensus Sequence of β-Amyloid Peptides with Heparin and Heparin-Derived Oligosaccharides. *J. Phys. Chem. B* **2016**, *120* (9), 2187–2197. https://doi.org/10.1021/acs.jpcb.5b12235.
(9) Sandwall, E.; O'Callaghan, P.; Zhang, X.; Lindahl, U.; Lannfelt, L.; Li, J.-P. Heparan Sulfate Mediates Amyloid-Beta Internalization and Cytotoxicity. *Glycobiology* **2010**, *20* (5), 533–541. https://doi.org/10.1093/glycob/cwp205.

order or disorder of surrounding lipids; area per lipid to determine how tightly packed the lipid molecules are within the membrane, and bilayer thickness to determine the structural integrate of the membrane.

Estimate of Compute, Storage, and Other Resources. All oligomer simulations conducted thus far have been performed on GPU-accelerated GROMACS.[14] Typically, each system is simulated for 2-μs (2000 ns), with three replicates per system. For $A\beta_{42}$ oligomers, we achieve simulation speeds of approximately 20-30 ns/day on a single NVIDIA V100 GPU, although the precise rate varies with oligomer size and system complexity. Using a conservative estimate of 20 ns/day, the GPU hours required for one 2-μs replicate are:

$$2000\ \text{ns} \times \frac{1\ \text{day}}{20\ \text{ns}} \times \frac{24\ \text{hr}}{\text{day}} = 2{,}400\ \text{GPU hr}$$

Thus, for three oligomer constructs with three replicates per amyloid across three amyloid systems (βE_{31}, $IAPP_{37}$, and αS_{140}), we estimate a total of 64,800 GPU hours for the oligomer lipid control systems.

For the GAG project, control simulations of HS-bound syndecan embedded in a membrane are also conducted using GROMACS at comparable speeds (20 ns/day). Each system will be simulated for 1-μs (1000 ns), with three replicates requiring:

$$1000\ \text{ns} \times \frac{1\ \text{day}}{20\ \text{ns}} \times \frac{24\ \text{hr}}{\text{day}} = \frac{1{,}200\ \text{GPU hr}}{\text{replicate}} \times 3\ \text{replicates} = 3{,}600\ \text{GPU hr}$$

Total GPU time requested for the full set of proposed control simulations is approximately **68,400 GPU-hours**, but to accommodate any variability in system size we round to **72,000 GPU-hours** for all 30 simulations. Post-simulation analyses will be conducted using CPU-based workflows with moderate CPU demands (estimated at ~200 CPU-hours per system). This can be accommodated by standard CPU partitions on the selected resource. SDSC Expanse is our preferred compute resource due to its availability of GPUs and proven performance with similar workflows. However, PSC Bridges-2 is also a viable alternative. Based on prior runs, a 2-μs simulation of an oligomer system produces around 100 GB of data per replicate. If all systems are run simultaneously, we will require **4 TB of disk space during active simulations**. Long-term storage will be managed through our internal lab infrastructure.

Support Needs. Our team has extensive experience using GROMACS for molecular dynamics simulations and analysis. The PI has extensive experience working with amyloid and membrane systems and protocols for these simulations already exist within the group and have been utilized on other HPC resources. As such, we do not anticipate needing extensive user support but would welcome consultation if issues arise related to storage, or performance optimization.

Team Preparedness. [redacted]

Our team is prepared to begin the proposed work immediately. We have experience setting up and running MD simulations on HPC resources, and protocols for the oligomer systems are already established and ready to be employed. The HS-syndecan systems have been constructed and are ready for production simulations. Our lab also has infrastructure in place, including a dedicated NAS, for long-term storage and management of project data.

(10) Zhou, X.; Wang, Y.; Zheng, W.; Deng, G.; Wang, F.; Jin, L. Characterizing Heparin Tetrasaccharides Binding to Amyloid-Beta Peptide. *Front. Mol. Biosci.* **2022**, *9*. https://doi.org/10.3389/fmolb.2022.824146.
(11) Di Domizio, J.; Zhang, R.; Stagg, L. J.; Gagea, M.; Zhuo, M.; Ladbury, J. E.; Cao, W. Binding with Nucleic Acids or Glycosaminoglycans Converts Soluble Protein Oligomers to Amyloid. *J. Biol. Chem.* **2012**, *287* (1), 736–747. https://doi.org/10.1074/jbc.M111.238477.
(12) Letoha, T.; Hudák, A.; Kusz, E.; Pettkó-Szandtner, A.; Domonkos, I.; Jósvay, K.; Hofmann-Apitius, M.; Szilák, L. Contribution of Syndecans to Cellular Internalization and Fibrillation of Amyloid-β(1–42). *Sci. Rep.* **2019**, *9* (1), 1393. https://doi.org/10.1038/s41598-018-37476-9.
(13) Hudák, A.; Kusz, E.; Domonkos, I.; Jósvay, K.; Kodamullil, A. T.; Szilák, L.; Hofmann-Apitius, M.; Letoha, T. Contribution of Syndecans to Cellular Uptake and Fibrillation of α-Synuclein and Tau. *Sci. Rep.* **2019**, *9* (1), 16543. https://doi.org/10.1038/s41598-019-53038-z.
(14) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. https://doi.org/10.1016/j.softx.2015.06.001.