Abstract

Title of dissertation: Study of Membrane Binding Proteins and Related Signaling Molecules

Robert James Allsopp, Doctor of Philosophy, 2023

Dissertation directed by: Professor, Associate Chair and Graduate Program Director (ChBE) co-director of the Biophysics Program, Jeffery B. Klauda

University of Maryland, Chemical Engineering

The membrane contact site theory is a critical theory to understanding lipid transport. The Osh protein is a yeast lipid transport protein theorized to form membrane contact sites. We investigated the contact site theory by identifying a second binding domain and studying the Osh Amphipathic Lipid Packing Sensor (ALPS) to explain better why each protein might target different organelles. The α_6- α_7 domain appears more charged and prefers lipids with oppositely charged inositol sugars, making it ideal for binding to the Trans Golgi Network (TGN) and the plasma membrane. The ALPS peptide is another dedicated binding domain bound in several membrane types with varied Phosphatidylcholines (PC) tails to vary the lipid packing. If the force field was valid, the results indicate that Osh4 ALPS prefers the loose packing of POPC, and Osh5 ALPS prefers the tighter packing of DMPC. More input from the wet lab is needed before researchers can make predictions from the force field. Another vital area of research is antimicrobial peptides (AMPs) that disrupt the membrane. Part of the dissertation focused on determining the dual placement of the AMPs on the surface and inserted into the membrane. For the first time, the membrane properties of bilayers with AMPs were studied, using the combination of all-atom simulation

informed by x-ray scattering. The surface tension was a critical parameter that enabled us to compare the simulation to the wet lab results and became vital in allowing the peptide to be inserted into the membrane and remain stable. The 5-HT_{3A} project simulated predicted structures of toxins with computational tools. Our work simulated these toxins for the first time, and we observed the unbiased binding of σ-GVIIIA conotoxin to the allosteric binding pocket. In the first trajectories, the ion channel pore remained closed, similar enough to the native apo crystal structure that water could form a partially water-filled channel for a few microseconds. In one example, the 5-HT_{3A} had serotonin in all of the binding pockets for close to 1 μs. The long simulation of the conotoxin showed that the extracellular domain (ECD) was deformed by more than a nanometer compared to a control. This deformation was the first indication that such a conformation is possible and might be related to the presence of the toxin. Finally, traumatic brain injury was studied by identifying new molecules that activate fibroblast growth factor (FGF) and toll-like receptor (TLR) proteins. The focus on FGF resulted in identifying a critical conformational change and potential new binding sites (previously unknown) that activate FGF without activating damaging inflammatory TLR responses.

STUDY OF MEMBRANE BINDING PROTEINS AND RELATED SIGNALING MOLECULES

by

Robert J. Allsopp

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
2023

Advisory Committee:
Prof. Jeffery B Klauda, Chair/Advisor
Prof. Amy Karlsson
Prof. Hannah Zierden
Prof. Pratyush Tiwary
Dr. Michael Kio

Personal Thoughts

Many readers in the molecular dynamics field likely already understand the appeal of computational biology and the satisfaction of writing custom computer codes to investigate unique questions. I want to write a few sentences to the broader audience to explain my motivation to use computers to research biological systems and share some thoughts regarding the process of selecting specific research projects. First, why biology and why computer coding? Before starting my Ph.D. studies, I knew developing computer skills would be useful. However, I feared that significant additional computer training would be needed to conduct effective research. I learned from several books that computer coding was not particularly difficult and that even inexperienced users could use computers and make new improvements. Although challenges were inevitable, there would always be ways to work around problems. While progressing through my Ph.D. studies, I realized I preferred wrestling with computer codes to calculating complex computations. My motivation to study biology is that this area is at the forefront of identifying new treatments that significantly impact people's lives. Biological proteins are the most advanced on the scale of nanostructures. Imagine trying to make a cell from scratch using artificial materials. Material scientists are still pushing the limits of making new catalysts, but biology has a huge head start since it uses highly advanced materials that have evolved over millions of years. Finally, selecting new research topics is a critical milestone for any researcher. When selecting a research topic, one must consider factors such as their interest in and knowledge level of the topic, available resources, expected benefits, and time to complete. I greatly benefited from listening to relevant seminars, podcasts, and talking with experts. I also spoke with classmates and family members about what topics they found interesting. The 5-HT_{3A} and TLR were some of my first projects, where I wanted to study cone snail venom deactivating a receptor and how the immune system detects polysaccharides. The most important thing to do is read as much as possible, figure out what past researchers thought, and determine if you agree or disagree. I wanted to pick a side in an argument, find evidence, and be as knowledgeable as possible. When starting a new project, it is also vital to identify available structures or sequences to predict how much computer time would be needed to achieve specific outcomes. Although the

polysaccharide project was unsuccessful, I became interested in the fibroblast growth factor's role in preventing traumatic brain injury. In high school, I became interested in the brain when I researched Alzheimer's disease. High school sports teams required me to take tests to identify future TBI, which reinforced my interest in the brain. The fact that I developed an interest in the brain in high school made the FGF project particularly exciting to work on.

Dedication:

To my parents and family for their love and support.

Acknowledgments

This dissertation would not have been possible without the help of a great many people. Interacting with many talented peers and mentors in Dr. Klauda's lab enriched my skills, broadened my knowledge, and fanned the flames of my motivation. These are people who I admire and who helped me improve far beyond anything I could have done alone. I benefitted from hearing positive experiences from my brother, Matthew, who volunteered in Dr. Klauda's lab. I sought Dr. Klauda as a mentor because of his passionate and engaging teaching style. Dr. Klauda also dedicates considerable time to teaching and monitoring his students' progression. I drew inspiration from Dr. Klauda's belief that UMD students are well-prepared to be top-notch in their fields and excel in graduate school if desired. I also want to thank Dr. Karlsson for her mentorship during my time in the Chemical Graduate Association and her involvement in the group meetings on the Osh project. Her insights, along with Dinara Konakbayeva, enriched my understanding of the complexities of wet lab experiments while obtaining results for a shared project. Dinara and other lab mates were beneficial in their instructional roles in the wet lab and supportive when discussing my research. Finally, I would like to thank my parents for their unconditional love, encouragement, and unwavering support. This research is supported by the NSF grant MCB-1951425. Simulations ran on the Extreme Science and Engineering Discovery Environment (XSEDE) now Access supercomputers by grant number MCB-100139.

Table of Contents

Abbreviation List

NMR - nuclear magnetic resonance

HMMM - Highly Mobile Membrane Mimetic

AMP - Antimicrobial Peptide

MAP - Membrane active peptide

NAMD - Nanoscale molecular dynamics

PMF - Potential of mean force

EDP - Electron Density Profile

DCLE - 1,1-dichloroethane

MD - molecular dynamics

Osh4 - oxysterol homology protein 4

WLBU2 - antimicrobial peptide

X-ray FF - X-ray form factor

CHARMM - Chemistry at HARvard Macromolecular Mechanics

5-HT-5 - hydroxytrypatmine

CHARMM-GUI - CHARMM graphical user interface

NTER - N terminal capping residue

CTER - C terminal capping residue

NPT - Constant Number Pressure Temperature

NPAT - Constant Number Pressure Area Temperature

NPgT - Constant Number Pressure Surface Tension Temperature

PIP2 - POPI24 named Phosphatidylinositol 4,5-bisphosphate

POPC - 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPS - 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine

POPG - 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphatidylglycerol

POPE - 1-palmitoyl-2-arachidonoyl-sn-glycero-3-phosphoethanolamine

POPA - 1-palmitoyl-2-oleoyl phosphatidic acid

DMPC - 1,2-dimyristoyl-sn-glycero-3-phosphocholine

DMPS - 1,2-Dimyristoyl-*sn*-glycero-3-PS

DOPC - 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine

SDPC - 1-stearoyl-2-docosahexaenoylphosphatidylcholine

TOCL2 - tetraoleoyl cardiolipin

nAChRs - nicotinic acetylcholine receptor

MCS - membrane contact sites

Chapter 1: Overview

1. Introduction

Proteins are ubiquitous in living organisms; all living cells require these to regulate potent molecules, detect foreign pathogens, and process environmental stimuli. Proteins are polymers made up of 20 building blocks called amino acids. The protein's functionality depends on the 3D structure, which can change due to solvent, solute, and ligand binding effects. The standard functions discussed in this dissertation include lipid transport[1], cellular protective signal transduction[2], ion channel allostery[3-6], and pattern recognition.[7-9] Numerous cutting-edge experimental techniques have been used to study proteins include nuclear magnetic resonance (NMR)[10], Fluorescence Resonance Spectroscopy[11], Crystallography[12], 2D IR[13, 14] and many other techniques aimed at quantifying spatio-temporal changes in protein structure, thermodynamic interactions, and other biophysical properties. Despite significant efforts in developing ever increasingly sophisticated techniques, there are limitations in understanding the transition pathways[15] of these fundamental biological processes. Molecular dynamics (MD) is one of a few other techniques that show promise for answering these questions about mechanisms and pathways.[15] While thermodynamic constants are already a point of active scientific research, MD simulations can quantify free energies and other properties that compare with other techniques such as isothermal titration calorimetry (ITC) and Langmuir Adsorption Isotherms.[11, 16] These experiments excel at calculating the thermodynamics of molecules binding to the membranes or binding to the active sites of proteins. Force field optimization has recently improved the representation of any given protein structure, including reproducing the correct percentages of secondary structure and other metrics, such as conformational dynamics.[17-19] The membrane force field parameters have been modeled based on surface area per lipid, density profiles with neutron and X-ray diffraction experiments, NMR deuterium order parameter (S_{CD}) including the proper splitting of the aliphatic carbon adjacent to the carbonyl, area compressibility modulus and frequency dependence of ^{13}C NMR relaxation rates, and water distribution profiles in low hydration bilayers.[20, 21]

MD simulations were recognized to be so crucial to chemistry that the Nobel Prize for Chemistry was awarded in 2013 to Martin Karplus, Michael Levitt, and Arieh Warshall for their pioneering work of MD methodology for biomolecular systems. In short, classical MD uses Newton's equation of motion to predict molecular movement based on 3D coordinates of interacting systems and requires a force field trained to reproduce biophysical properties. By solving Newton's laws of motion, the algorithm can predict how the system evolves and represent the transitions between states. The two main factors include the force field and the system's sampling. Supercomputers are the best way to run long simulations, while enhanced sampling techniques can surpass the time scale limitations and measure the thermodynamics of various transitions.[22, 23]

1.1 Membranes

Membranes are essential to the cell because they protect DNA and maintain homeostasis by controlling pH and separating intracellular and extracellular ions. An insightful question to ask is how membranes differ from all other biomaterials. A significant difference is that lipids lack covalent bonds linking monomers to form larger structures.[24, 25] In membranes, the individual lipids are held together entirely by intermolecular forces primarily caused when a bilayer is created to shield the hydrophobic core. The phenomena consist of many forces and not merely one specific intermolecular force. Lipids have other interactions, including hydrogen bonds breaking and re-forming. No single energy minimum defines the membrane, but instead, many acceptable configurations exist separated by $< k_B T$ (boltzmann constant multiplied by temperature), and many membrane proteins also have high residual entropy.[24-26] The membranes themselves have been studied to reveal a range of necessary time scales.[27, 28]

1.1.1 Lipid Packing and Membrane Shape

Valuable metrics to measure the reconfigurability of the membrane include the area per lipid and order parameters that are sensitive to the atomic level interactions and help correct the force field.[29] The area per lipid is an important physical property of cell membranes. Area per lipid is usually a bulk property along with the area compressibility modulus related by surface area.[30] Through bulk properties like area

per lipid and compressibility modulus, it is possible to compare experimentally measured properties to the simulations.[30] This is an important parameter because it helps probe the state of the membrane. For example, adding a peptide to the membrane could cause changes in the lipid packing by adjusting the surface tension.[31] Membranes can have irregular shapes and come within nanometers of other organelles, contrary to textbook pictures.[32] The primary examples of re-configurations include interactions with simple perturbations such as the interactions with ions[33], diffusion, and the binding of proteins to the cell surface.[34] For example, cholesterol and other lipids play essential roles in the lipid environment that which impacts neural ion channels.[35, 36] Antimicrobial peptides are known to disrupt the membrane properties[31], and similarly, specific ions induce PS lipids to form microdomains.[33]

1.1.2 Regulation of Membrane Composition for Functionality

The molecular composition of the membrane controls the level of lipid packing.[37] Different membrane organelles are known to have different lipid compositions.[37] The Osh4 protein is a lipid transport protein. Many details of lipid transport are not understood, such as how individual mechanisms contribute to the overall observed kinetics. For example, one poorly understood mechanism is the membrane contact site which requires no ATP. The membrane contact site hypothesis states that two membrane organelles are bridged together[32] to enable lipid transport. Osh is a family of several proteins, each targeting a different membrane organelles.[38] It was hypothesized in this work that there are binding domains on the Osh protein that account for the different binding affinities for different membrane compositions. The Osh proteins are an excellent example of how proteins regulate both the lipid and membrane composition, impacting the protein's function.[38-40] Endocytosis is an example of a complex process that some viruses use to enter the cell. Endocytosis was recently highlighted due to the impact of COVID-19, which uses this mechanism after binding to ACE2.[41] This process features the bi-directionality of lipids influencing the structure and dynamics of proteins, and proteins influencing conformations and organization of lipids.[24] During this process, the membrane remains primarily intact except for the presence of viral proteins and DNA passing through the center and entering into the cell. Covid-19 and endocytosis demonstrate the importance of the reconfigurability of the cell membrane. Another process that is bi-directional is the formation of membrane

contact sites.[32] Membrane contact sites form when two organelle membranes come within 2-3 nm of each other, and proteins hold them together. Different Osh proteins have a preference to bind to different organelle membranes.[38] The function of Osh proteins is to transport and sort lipids to enrich specific organelles, particularly lipids with notable roles and functions unique to that organelle.[42] The membrane composition signals the correct proteins to bind to the proper locations on the membrane. The proteins further sort and adjust the membrane composition to maintain homeostasis as lipids degrade and the cell progresses through stages of growth.[42]

The membrane is also an essential consideration for drug treatments.[43] This is because drugs frequently target membrane proteins. Membrane proteins are receptors of a significant category of drugs that bind and interact on the surfaces of membranes, which can then activate different signaling pathways.[43] Good examples of signaling proteins in this dissertation include the fibroblast growth factor (FGF)[44] and 5-HT_{3A} ion channel.[45] FGF folds in the aqueous phase and binds to the membrane, while TLR and 5-HT_{3A} have separate ECD, TM, and intracellular domains. The membrane is essential to accurately simulate the TM regions of 5-HT_{3A} while the ECD can, to a first approximation, be modeled in the aqueous phase.

1.1.3 Summary of Projects and Motivating Papers

During my Ph.D., I extended Dr. Monje-Galvan's work on the Osh4 project.[34, 46-48] We identified a second dedicated binding domain[49], and conducted extensive experiments on Osh4 ALPS to measure the effect of changing lipid packing on binding free energy. The first new project was the Conotoxin binding to 5-HT_{3A} and was primarily motivated by the discovery of the conotoxin[50] and the past work that Dr. Guros had done with 5-HT_{3A}.[6] I wanted to further validate the binding site with thermodynamics and use that to better understand if allostery occurs to help improve the force field. The growth factor project was motivated by the intriguing work from the 90s showing the remarkable neuroprotective benefits to rabbits and gerbils that survived occluded arteries for several minutes after having been administered F-1,6-BP.[51, 52]

1.1.4 Osh4 α_6-α_7 and ALPS Peptide

The Osh4 protein of yeast has a significant function of moving a PI4P lipid from the trans-Golgi network to the endoplasmic reticulum with the aid of ergosterol.[40, 53] This transport is believed to occur when the protein forms a MCS between these organelles.[32] Membrane contact sites are physical contacts where the membrane's lipids are a few nanometers apart, and a protein tethers them together.[54] To be consistent with the membrane contact sites theory, there needs to be two different binding domains that have affinities for other organelles. These contact sites and lipid transport proteins enable the organization of the cell organelles that work without ATP and knockouts of more than 3 Osh proteins, causing severe changes in membrane composition, but only fatal to the cell when all Osh proteins get knocked out.[1, 55, 56] My work focuses on two domains of Osh4 (the α_6-α_7 helical domain and the ALPS domain) to determine if these regions play a role in forming and maintaining contact sites (Fig. 1.1).[49] The ALPS peptide is a lipid packing sensor believed to sense the membrane curvature and lipid packing to help direct the Osh4 protein to the proper locations.[57-59] This process is accomplished by filling in the natural packing defects in the membranes.[34, 46, 47] Membranes with more curvature are more likely to have packing defects.[34, 46, 47] The HMMM simulations benefited from acceleration in the binding of the peptide because otherwise, it could take as long as a few microseconds for specific dense membranes.[46, 47]

This study aims to gain additional information and insights into the penetration depth and binding free energy of the peptide in particular membranes. Specifically, this study focuses on the following five questions: 1) For Osh4 ALPS, how does varying the membrane properties such as area/lipid impact the binding orientation and strength? 2) Does the Osh4 α_6-α_7 play a role in forming contacts on the cell surface? 3) What is the role of PIP2 in relation to this α_6-α_7 region, and could it increase the residence time of

binding? 4) Which amino acids are essential for binding? 5) How does improving the representation of the electrostatic interactions impact the membrane interactions?

Experimental mutational studies on charged residues of the Osh4 α_6-α_7 motif demonstrate a reduction in the lipid transport rate.[40, 60] Since this motif is essential for function and based on its structure, we hypothesized that it was necessary for binding to the membrane. However, there were no computational studies of the work to understand this domain's role. The significance of the work is to provide initial insight into how Osh4 can form MCS by focusing at the atomic level.[42, 54, 61] By identifying the various domains of the protein that include the physical membrane contact sites.[54, 61] The longer-term goals are to develop simulations illustrating the full extent of protein conformational changes needed to create the membrane contact site. Showing the functional protein will give a more detailed picture of the protein membrane contact sites.[61]

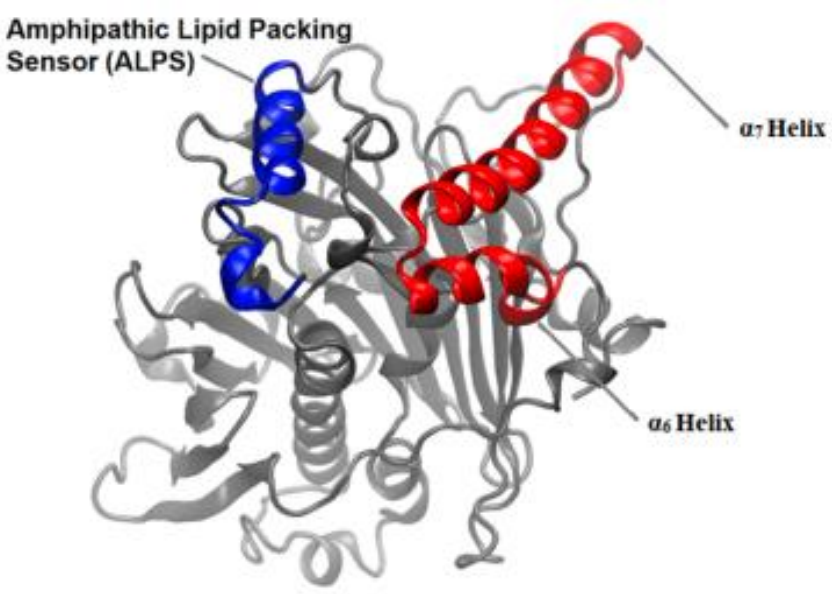


Figure 1.1) Structure of the Osh4 protein with the two membrane binding domains of the Osh4 for this objective colored in red and blue.

1.1.5 WLBU2

Antimicrobial peptides (AMPs) are important alternatives to traditional antibiotics. These peptides also have antiviral properties that could help the immune system respond to viruses like COVID-19. Additionally, drug resistance is exacerbated by the COVID-19 pandemic and the recent wars that increase hospitalization and transmission and lead to the overuse of antibiotics. These current events have increased

the number of people resting in hospital beds, expanding the number of patients coming in contact with drugs that already develop quick resistance. AMPs are unique because many target the membrane by disrupting the function by forming holes, dissolving, and potentially interfering with other membrane proteins. Dr. S. T.-Nagle's lab at Carnegie Mellon University experimentally studied WLBU2 (**RR**WV **RR**V**R** **R**WV**R** **R**VV**R** VV**RR** WV**RR**) to probe its structure in the membrane with X-ray Scattering and Neutron Reflectometry. The research aims to characterize a specific synthetic AMP to better understand how it disrupts the membrane.[62] Future researchers will build models of the peptide to study how to effectively kill bacteria.[62] Further optimization can help create new, more active, or selective therapeutic drugs to replace those becoming less effective due to increasing bacterial drug resistance,[62, 63] and the lentivirus lytic peptide was used as a template. This peptide is a naturally occurring[64] sequence with some membrane and antimicrobial properties used by the virus to defend against bacteria. Mutations to the peptide limited the composition to only arginine, valine, or tryptophan.[62] Then, there were variations in different parameters, including the length and presence of tryptophan residues. The result was that the WLBU2[62], a medium (24 residue peptide), had the best activity for the shortest length.[62] From these data, it is possible to develop simulations guided by the wet lab experiments.[65] This information from the wet lab is vital to match the membrane properties and the position of the peptide.[65]

1.2 Ligand Gated Ion Channels

Ligand gated ion channels are another important fundamental molecular machine critical for animals to grow and respond to stimuli from the outside environment.[66] These ion channels are typically used to trigger depolarization, a critical step in neural synapses.[67] These ion channels are found in many tissues in addition to the brain, including muscle,[68] heart,[69] and digestive tract.[70] High concentrations of GABAR and glutamate receptors in the brain[71] signify the importance of the corresponding neurotransmitters. GABA is an inhibitor of neural synapses, and glutamate is a stimulator of synapses. Other notable neurotransmitters are dopamine and serotonin, which are associated with focus and happiness. Another neurotransmitter is acetylcholine, which activates the nicotinic acetylcholine receptors

in muscular tissues, including the heart. The $KCNH_1$ ion channel plays a vital role in the heart[69] and neuronal excitability. The GI tract[72] produces 95% of all serotonin, where 5-HT_{3A} receptors play an essential role in signaling processes, including nausea[73] and some interactions with sleep and depression, possibly linked to bipolar depression.[74] The main point is that ion channels play essential roles in the brain and in linking the brain to other parts of the body, such as the muscles and digestive tract, and the connections between these tissues can potentially contribute to diseases.

1.2.1 5-HT_{3A}

Serotonin is an essential neurotransmitter for the brain, mainly produced in the intestines and digestive tract.[72] Serotonin is most well-known for its role in depression through 5-HT_{2A}.[74] The 5-HT_{3A} ion channel is unique because it is part of the cystine loop family of ligand-gated ion channels, including nicotinic acetylcholine receptor, GABAR, and glutamate receptors.[75] Each receptor appears self-contained, contrasting the 5-$HT_{2A,}$ a G-coupled protein, meaning the pore can couple to several different receptors that are initially separate proteins.[76] The ECD mainly contains an allosteric site covered by a b-sandwich loop with considerable flexibility. GABA, glutamate, serotonin, dopamine, nicotine, or acetylcholine bind to the allosteric binding site and activate allosteric transitions that open the transmembrane region and allow the passage of ions through the membrane. Past MD simulations have focused on improving the serotonin parameters from the automatically predicted values to ones that better match the free energy of solvation via FEP.[6] Despite testing a range of serotonin concentrations, the past simulations did not fully activate the 5-$HT_{3A,}$ and the pore remained physically closed. Closer inspection did reveal pre-activation of the 5-HT_{3A} caused by changes to the pore lining that essentially opened 4/5 transmembrane (TM) helixes.[6] This transition took approximately one μs, with four total allosteric pockets bound with serotonin, and is called the pre-activated state because the serotonin had interacted with the protein for considerable periods and altered the TM pores.[6]

1.2.2 Conotoxins Inhibit Receptors, including 5-HT_{3A}.

Several well-known animal species, such as spiders, snakes, jellyfish, sting rays, killer bees, and fish, have developed offensive and defensive venoms. One animal, the cone snail, is known to have one of the most potent venoms in the world.[77] This venom is used to block critical components on nerve and heart cell membranes to prevent the prey from escaping before the snail begins to feed.[78]

Some of these peptide venoms have therapeutic effects by targeting pain in new ways that are less addictive.[79] Peptide venoms have been studied to find new treatments and better understand the envenomation process. By understanding how the toxins work, treatments should be able to be developed to respond to the different venomous bites. With the advent of 3D peptide structural prediction, it is increasingly possible to analyze and use MD simulations to probe structures that represent the realistic 3D geometries of actual toxins.[80, 81] This new study aims to simulate conotoxins and understand what these peptides can do differently from small molecule drugs.

The conotoxin venom is ranked in the top ten most deadly venoms globally and includes over 8,134 known conopeptides in the ConoServer database.[82] The snails use these peptides to kill prey by disrupting the nervous system to stop respiration. Previously, an experimental study investigated potential protein receptors for the toxin peptide.[50] This study tested the activity of various protein complexes and determined that the venom had high specificity for the 5-HT_{3A} protein complex. However, more extensive testing is needed to determine if there are other interactions. Additionally, inhibition of other cystine loop superfamily members, including GABAR could lead to seizures.[50] Although this peptide has therapeutic properties, it is only examined for its neurological properties in a small set of receptors. Details about the binding affinities show that this is competitive with other small molecule drugs.[50] 5-HT_{3A} is in the same family of the cystine loop superfamily[3] that includes several Nicotinic Acetylcholine receptors implicated in addictive behaviors[83], and some conotoxins bind to those receptors.[84] There are many possibilities, such as further engineering these already existing peptides or, *de novo,* engineering new peptides that can effectively block the signaling regions without activating them. Past computational work in Dr. Klauda's lab exclusively focused on the 5-HT_{3A} and serotonin. However, my current work focused on the toxin interactions and the ion flux through the 5-HT_{3A} ion channel.

The peptide studied in this dissertation is the σ-GVIIIA conotoxin, one of the first toxins found to inhibit serotonin receptors.[50] Characteristics that stood out about the conotoxin were the unusual post-translational modifications, including a brominated tryptophan and a hydroxylated proline. It only deactivated the 5-HT_{3A} receptor and not any other receptors, such as nicotinic acetylcholine or even chimeras of the two.[50] The toxin was found to be competitive and to displace other inhibitors such as ondansetron and zacopride. It also had a competitive $K_{i,}$ meaning that it was physically effective at inhibiting the channel in addition to having a high affinity. Radiolabeled zacopride measured the K_i of the conotoxin by determining how the changes in toxin concentration impacted the fraction of bound zacopride, thus calculating the IC50 of the toxin.[50] Researchers estimate an actual thermodynamic constant by predicting the reaction mechanism. The Ki is equivalent to a simulated PMF.

The search for activating and deactivating allosteric conformational changes was common to the FGF and the 5-HT_{3A}. Past work from Dr. Klauda's lab[6] hinted that allosteric movements were possible with the current force field. At the same time, the toxin might inhibit the ion channel by causing a deactivating allosteric movement. By comparing the dynamics of activated and deactivated 5-HT_{3A}, it was possible to test for deactivating activity. The objectives for this work are: 1) Generating dynamics of the open unbiased 5-HT_{3A} pore to compare with the toxin pose, and 2) Observing the toxin binding in an unbiased manner to the allosteric binding site as theorized by the unbiased toxin binding objective. The first objective will use the newer available crystal structures to build upon Dr. Guros's work[6], simulate the open pose with improved equilibration protocol, and quantify the conductance of the first-ever unbiased material transport through the membrane through a correctly opened pore. If the second objective is accomplished, it will displace the serotonin and could alter the allosteric conformation of the ECD.

1.2.3 Traumatic Brain Injury-Glutamate Receptor Connection

Fibroblast growth factor (FGF) and toll-like receptors (TLR) are usually unrelated, with the former being associated with immunity and the latter related to cellular division. They are both activated by highly charged glycosaminoglycans, such as heparin and heparan sulfate, that have a higher production rate during hypoxic stress.[85, 86] FGF is known to regulate cellular migration, division, and re-growth. TLR is expressed

on sentinel immune cells and detects foreign chemicals mainly from bacterial cell walls and membranes, but also bacterial RNA.[87-89] Hypoxic traumatic brain injury (TBI) conditions cause heparin-like glycosaminoglycan production as a response. [85, 86, 90] Knockouts of growth factor activators exacerbate the injury[44], while knockouts of TLR are beneficial during TBI, although there are proposed processes to sensitize the immune system to various stresses.[91, 92] These sensitization schemes seem impractical because they might require daily injections. It is unknown if it is better to continually introduce artificial sensitization of the TLRs and how that impacts the immune system efficiency during actual bacterial infections.

Several decades of TBI research yielded two potential treatment molecules identified as fructose-1,6-diphosphate and enoxaparin.[51, 52, 93-96] There are many hypothesized mechanisms and considerable efforts to explore a variety of possible pathways, even though there is no universally accepted mechanism. Some approaches include improving blood flow,[94] ATP production,[51, 52, 95] and intracellular [Ca^{2+}]. [94, 96] The work assumes that the catastrophic rise in intracellular [Ca^{2+}] causes cell death. Glutamate neurotransmitters can over-activate glutamate receptors, causing excitotoxic neural depolarization with corresponding increases in intracellular [Ca^{2+}].[94] This hypothesis does not explain how F-1,6-BP also protects non-neurological cells, which may not have glutamate receptors. There could be other crucial mechanisms everyone takes for granted. Both F-1,6-BP and enoxaparin stabilized the intracellular Ca^{2+}.[2, 94] If scientists prove this theory to be correct, then it could motivate the exploration of treatments for mechanically induced TBI as was done with F-1,6-BP and DMSO[97], because DMSO is known to suppress excitotoxic effects of several ion transport receptors. DMSO is not ideal because it permeabilizes the skin.

The current treatments for TBI are unacceptable because of unreasonable concentrations or unacceptable anti-thrombotic effects. It is essential to mention that FGF receptors exist in the brain and have crucial roles in the learning process.[98, 99] Studies have proven that the administration of F-1,6-BP and enoxaparin reduced the intracellular Ca^{2+} levels corresponding to better neural protection.[93-95] The original studies indicated that the sugars restored the metabolism, but further analysis revealed that phospholipase C, a step in Erk 1/2, is essential to Ca^{2+} stabilization. This process could be related to IP3/IP4 molecules

interacting with Ca^{2+} ion channels.[2, 96] Multiple studies are showing how F1,6BP effectively protects neural cells.[2, 51, 52, 100, 101] The concentration of F1,6BP needed can be relatively high to produce results. Several grams (as low as 3.5mM) of F-1,6-BP are required, and the half-life of sugars is well known to be short[102], making compliance with this treatment difficult.[2, 95, 103, 104] While many aspects of the neuroprotection are known, more questions still need to be addressed, such as the following. How F-1,6-BP activates the phospholipase C enzyme? Does this happen directly, or does F-1,6-BP trigger an intermediate Erk step via IP3 Ca^{2+} channels? How much amplification of the signal is involved?

Identifying these connections is an essential step in developing treatments for TBI. As mentioned above, another approach is to use enoxaparin to protect the brain.[93] Enoxaparin and heparin will prevent blood clotting, which would make medical procedures overly complicated if first responders had to diagnose the level of bleeding before treating head injuries. Additionally, this approach would entirely exclude the treatment of hemorrhagic shock and hypoxia caused by blood loss. Further, these longer chains activate side reactions such as the innate immune system through TLR4.[105] In conclusion, short carbohydrate molecules can pass through the blood-brain barrier and activate neurologically protective pathways[106, 107] while preventing unwanted side reactions will require further studies.

The fibroblast growth factor is studied to identify alternative initiators of the Erk 1/2 pathway that have notable differences from heparin. A unique feature of the FGF is the well-known primary binding site that is known to prevent degradation caused by proteolytic and thermal denaturation. This site is, however, not essential to any one process.[108] The research tested two hypotheses. **Hypothesis 1.** *Multi-site polysaccharide (PS) stabilization*. PSs assist in the proper folding of FGF with a variety of binding motifs— **Hypothesis 2.** A clash between the bFGF and the receptor is responsible for differentiating the activated and deactivated bFGF protein. The bFGF point mutations were overlayed on top of the wild-type bFGF to compare the past crystal structure sugar interactions with the thermally unstable binding sites. Better fine-tuning the glycosaminoglycan processes will enable the better treatment of diseases by working with cellular defense mechanisms rather than against them.

1.3 Overview

The second chapter details the molecular dynamics methods and theoretical support. The second chapter includes the microstates, force fields, and essential equations for different systems—other concepts like calculating the PMF, cutoffs, and periodic boundary conditions. The third chapter details how the Osh α_6-α_7 forms membrane contact sites between specific membranes. Several hypotheses exist, but the most important was that this was a dedicated membrane binding domain and that certain critical amino acids could be mutated to disrupt the capacity to bind. Another hypothesis is that the presence of the PIP2 lipids alters the binding free energy, but more sampling is required to answer this adequately. The fourth chapter focuses on the ALPS peptide, another membrane-binding peptide that might play a role in differentiating between membrane compositions to help direct certain Osh proteins to specific locations. This chapter focuses on the area per lipid and the binding free energy associated with each trial. The top binding membrane in the Osh4 was different from the Osh5 membrane. The binding could indicate a preference for one membrane over the other. The fifth chapter is a study on antimicrobial peptides that disrupt the membrane. The chapter is supported with experimental results and details the results of a collaboration that yielded simulations that minimized the error between the X-ray form factor and incorporated neutron reflectometry data. Chapter six describes a docking study on a cone snail toxin and reveals how the toxin was bound unbiasedly and docked. The docked simulation ran for 15 μs and perturbed the ion channel structure in ways that could disrupt the ion channel from opening.

Additionally, the pore opened from an Apo crystal structure, indicating that the serotonin force field worked and that the toxins had artificial off-target binding to the outer shell of the protein but did not seem to disrupt the mechanism significantly. Interestingly, the majority of serotonin contacts close to the pocket were not actually productive interactions in the pocket. Chapter seven focused on treating hypoxic traumatic brain injury with different compounds and expanding the understanding of where the compounds bind to the protein. Traditional compounds were studied, along with less common compounds, to show that alternative binding sites exist where the peptide can bind and stabilize the protein folding.

Chapter 2: Methods

2.1 MD and Statistical Thermodynamics

2.1.1 Microstates Lead to Macro Phenomena

Physicists led the way in establishing a well-known connection between the microscopic states and macroscopic phenomena with the photoelectric effect. There are ground states and excited states of all atoms. Analogous comparisons have grown and motivated the field of statistical mechanics over several decades. Perhaps the simplest example is the excited states of electrons, where the transition from an excited state to a lower energy statereleases photons with different wavelengths, such as those used to light neon signs. The phase space is related to the probability density of finding the system in a particular configuration.[109] Each particle has three momentum variables and three position variables.[109] The position variables are used to calculate potential energy and the momentum variables are used to calculate kinetic energy.[109] In the biophysics of proteins, microstates focus on a given protein's various 3D coordinates (potential energy) and momentum (kinetic energy).[109] There can be ground state conformations and excited conformations of the protein. A distribution of states is known as an ensemble. A handful of states will be much more dominant than others in the ensemble and will be kept at a high probability for some proteins. The dominance of certain states explains why transition states can be challenging to observe. They can be relatively rare even if essential to the system, thus slow kineticscontrol the timescales of transitions between stable states. An excellent example of a crucial excited state is the opening and closing of ion channels in neurons. Without the proper ligands, such as glutamate, serotonin, dopamine, or GABA, the ion channels will remain in the ground state with the ion pore closed. However, in the presence of the proper ligands, the ion channel enters an excited state where it is more likely to open the pore than close it. The excited state allows Ca^{2+}

ions to enter the neurons via diffusion and initiate depolarization, a critical step in processing information in the brain. Activated and deactivated forms of a protein can connect microscopic states to a macroscopic disease. For example, one microstate corresponds to a disease and others correspond to a normal functioning protein.

2.1.2 Force Fields

The force field contains parameters that describe the interactions between any two atoms in the system. Some of the first simulations that used the Lennard-Jones (LJ) equations were performed in the 1950s by Rosenbluth et al. [110] to look at dispersive interactions of Nobel gasses. Today, the LJ equations continue to represent dispersion-repulsion interactions and have proven useful to study biomolecules. Below is the force field used by CHARMM, which contains the LJ and various other terms [111]

$$V(R) = \sum_{bonds} K_b(b - b_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} K_\chi(1 + \cos(n\chi - \delta)) + \sum_{UB} (S - S_0)^2 + \sum_{impropers} K_{imp}(\phi - \phi_0)^2 + \sum_{Lennard\ Jones} \varepsilon\left[\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{12} - \left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{6}\right] + \sum_{Electrostatic} \frac{q_i q_j}{\varepsilon_1 r_{ij}} \quad (eq\ 1.00)$$

The Lorentz-Berthelot rules generate predicted values of σ_{ij} and ϵ_{ij}.[112]

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \quad (eq\ 1.01)$$

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \quad (eq\ 1.02)$$

$$R_{min_{ij}} = \sigma_{ij} * 2^{\frac{1}{6}} \quad (eq\ 1.03)$$

In the above equation, I) is the total energy detected by the Chemistry at Harvard Molecular Mechanics (CHARMM) force field, and it incorporates the bonds, angles, dihedrals, and other parameters necessary for the topology of the molecule. At the same time, the last two parameters are the LJ and coulombic interactions.

CHARMM only programs 124 of the 7626 sets of LJ parameters (as of 2018)[113] as atom-specific sets of LJ parameters used to describe solute-water interactions that directly affect hydration. The Lorentz-Berthelot combinations estimate LJ parameters using solute-solute atom pair estimates. Corrections to the force field include defining more and more pair-specific sets rather than using the Lorentz-Berthelot combinations, and this is particularly useful because it targets and corrects specific interactions without the need to completely re-adjust other force field parameters. The Non-Bonded Fix (NBFIX) corrections to the force field edit the $R_{min_{ij}}$values as was done with the CUFIX parameters. The '12' '6' combination was chosen because the $\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{12}$ can be easily calculated as the square of $\left(\frac{R_{min_{ij}}}{r_{ij}}\right)^{6}$ which is computationally efficient. Additionally, the 12th power approximates the Pauli repulsion, or repulsion in general, reasonably well.

2.1.3 NVE Hamiltonian Equations of Motion

The simplest way to define differences between simulation trajectories is to start by defining the Hamiltonian of a system as the sum of the potential and kinetic energies. It is worth noting that all MD simulations are deterministic instead of non-deterministic Monte Carlo simulations. MD is deterministic because it creates an ensemble of states that match the phase space of the ensemble, while Monte Carlo may not and may predict non-physical configuration. The Hamiltonian is the combination of the coordinates (potential energy) and the velocities (kinetic energies) of the system written as:

$$H(q_1, \dots, q_n; p_1, \dots, p_n) = \frac{1}{2}\sum_{n=1}^{n} \frac{p_i^2}{m_i} + U(q_1, \dots, q_m) \qquad (eq\ 2.00)$$

where m_i Is the mass, q_i Is the coordinate and p_i is the momenta of the particle i.

$$q = \frac{\partial \dot{H}}{\partial p_\iota} = \frac{p_\iota}{m_\iota} \qquad (eq\ 2.01)$$

$$\dot{p_\iota} = -\frac{\partial \dot{H}}{\partial p_\iota} = \frac{\partial U}{\partial q_\iota} = F_\iota \qquad (eq\ 2.02)$$

The F_i is the force on any given particle in the system *i*. Solving these equations for all particles leads to the trajectory of all particles in the system.

$$Isolated\ System\text{:}\ \frac{\partial H}{\partial t} = 0\ all\ microstates\ are\ visited\ equally\ (eq\ 2.03)$$

$$Heat\ Bath\text{:}\ \frac{\partial H}{\partial t} \neq 0\ energy\ exchanges, canonical\ ensemble\ p(q,p)\ (eq\ 2.04)$$

$$Probability\ of\ Visiting\ a\ Given\ Phase\ Space\text{:}\ P(q,p) \sim e^{-\frac{H(q,p)}{k_B T}}\ (eq\ 2.05)$$

Canonical ensembles hold a fixed number of particles.

2.1.4 Modeling Membrane Systems For NPAT, NPγT

For many membrane systems in this dissertation, the thermodynamic properties of the system's pressure are sometimes different between the lateral direction and the normal direction. These ensembles are NPAT and NPγT with constant lateral area and surface tension, respectively.[114] The lateral directions are x or y, and the normal direction is z. The first law of thermodynamics can be extended with the second law to yield to the fundamental equation to explain the different ensembles:

$$dE = TdS - P_n dV + \gamma dA + \sum_{i=1}^{C} \mu_i dN_i \qquad (eq\ 2.06)$$

E is the internal energy, *T* is the temperature, *S* is the entropy, and *V* is the volume. In the equation above, variable γ is the interfacial surface tension, N_i the number of particles of component *i*, C is

the number of components, and μ_i is the chemical potential. The surface tension[114] is given generally by:

$$\overline{P_t} = \frac{1}{2h_z}\int_{-\frac{1}{2h_z}}^{+\frac{1}{2h_z}} dz[P_{xx}(z) + P_{yy}(z)] \qquad (eq\ 2.07)$$

$$P_n = P_{zz} \qquad (eq\ 2.08)$$

$$\gamma = h_z\left(P_n - \overline{P_t}\right) \qquad (eq\ 2.09)$$

where P_n is the normal pressure, and the $\overline{P_t}$ is the tangential pressure. A planar surface's normal pressure is uniform, but the tangential pressure depends on position. Far from the interface, the normal and tangential pressures are roughly equal but become large and negative at the interface. This results from how surface tension is defined and how surface tension is observed at the interface of two fluids. We define the simulation box dimensions as h_x, h_y, h_z and are helpful to define the PdV work and the surface tension. The Gibbs phase rule calculates the number of degrees of freedom which are intensive variables needed to be specified to characterize an equilibrium system. The number of degrees of freedom (F) $= c - p + 2$ where *c* is the number of components, *p* is the number of phases, and the degrees of freedom can is specified by intensive properties such as *T, P, A*, and γ, while N is an extensive property.[114] There are also generalizations of the Gibbs phase rule that apply to interfaces that enable the simulation of *N, P*, γ*,T* systems because a two-component two-phase system only allows the definition of two intensive variables as is done in standard (*N, P, T*). The four most used ensembles in this paper are discussed below.

NVT: Hamiltonian, Phase Space Probability, Free Energy of Ensemble, Partition Function

A few additional properties of the previously defined Hamiltonian equations of motion include:

$$H(q_1, \dots, q_n; p_1, \dots, p_n) = \frac{1}{2}\sum_{n=1}^{n}\frac{p_i^2}{m_i} + U(q_1, \dots, q_m) \qquad (eq\ 2.10)$$

The Hamiltonian equations remain the same because there is no PV work to incorporate into the system.[114] Because of this, the Hamiltonian of the system remains the same for NVE and NVT systems. Along with the other properties such as the phase space and other parameters.[114]

$$NVT, \qquad F = E - TS \qquad (eq\ 2.11)$$

The expression for the Helmholtz free energy, *F*, is defined above. The expression for the partition function of the ensemble is defined below:

$$\Delta \sim \int dh_x dh_y dh_z d^N \boldsymbol{r} d^N p d^{Nr} \boldsymbol{q} d^{Nr} \boldsymbol{L} \times e^{-\beta H} \qquad (eq\ 2.12)$$

Temperature is held constant with a thermostat (details provided below).

NPT: Hamiltonian, Phase Space Probability, Free Energy of Ensemble, partition function

The expression for the Hamiltonian of the NPT ensemble is defined below:

$$H(q_1, \dots, q_n; p_1, \dots, p_n) = \frac{1}{2}\sum_{n=1}^{n} \frac{p_i^2}{m_i} + U(q_1, \dots, q_m) + P_{ext} h_x h_y h_z \ (eq\ 2.13)$$

A significant change is a PV work term to the system. Other variations of these systems will also have similar PV terms, but each corresponds to different free energies.

Because the NPT ensemble allows for volume changes in the system, the $\underline{P}\mathrm{d}(V)$ term is non-negligible even with fixed pressure, and the energy of the system is then[114]:

$$NPT, \qquad F = E - TS + PV = H - TS = Gibbs\ Free\ Energy \ (eq\ 2.14)$$

The partition function of the ensemble is defined below:

$$\Delta \sim \int dh_x dh_y dh_z d^N \boldsymbol{r} d^N p d^{Nr} \boldsymbol{q} d^{Nr} \boldsymbol{L} \times e^{-\beta H} \qquad (eq\ 2.15)$$

The pressure is held constant with a barostat (Langevin for NAMD[115], MTK[116] for Anton2, and Parrinello-Rahman for Gromacs[117, 118])

NPAT: Hamiltonian, Phase Space Probability, Free Energy of Ensemble, partition function

The Hamiltonian of the NPAT ensemble is defined below:

$$H(q_1, \dots, q_n; p_1, \dots, p_n) = \frac{1}{2}\sum_{n=1}^{n} \frac{p_i^2}{m_i} + U(q_1, \dots, q_m) + P_{ext}Ah_z \quad (eq\ 2.16)$$

The NP_nAT ensemble is slightly different from the NPT ensemble, changing to P_n instead of P because of the constant area requirement. [114]

$$NP_nAT,\ F_{pn} = E - TS + P_nAh_z \quad (eq\ 2.17)$$

The NP_nAT system barely changes from Gibbs free energy except that there is a fixed lateral area, so those do not contribute to the PV work needed in the simulation box to derive Gibbs's free energy. The partition function of the ensemble is defined below:

$$\Delta \sim \int dh_x dh_y dh_z d^N\boldsymbol{r} d^N p d^{Nr}\boldsymbol{q} d^{Nr}\boldsymbol{L} \times e^{-\beta H} \quad (eq 2.18)$$

NPγT: Hamiltonian, Phase Space Probability, Free Energy of Ensemble, partition function

The Hamiltonian of the NPγT ensemble is defined below:

$$H(q_1, \dots, q_n; p_1, \dots, p_n) = \frac{1}{2}\sum_{n=1}^{n} \frac{p_i^2}{m_i} + U(q_1, \dots, q_m) + P_{ext}h_xh_yh_z - \gamma_{ext}h_xh_y \ (eq\ 2.19)$$

Finally, the $NP\gamma T$ ensemble free energy accounts for an additional $-\gamma A$ term for the imposed surface tension.[114]

$$NP\gamma T, \qquad F_{pn} = E - TS + P_nV - \gamma A \quad (eq\ 2.20)$$

By changing A to γ the system is closer to an NPT system because the box is less constrained, but this also differs from the true NPT that generates Gibbs free energy.[114] This is

because the lateral area has the g-term that is pre-defined in the simulations. The partition function of the ensemble is defined below:

$$\Delta \sim \int dh_x dh_y dh_z d^N \boldsymbol{r} d^N p d^{Nr} \boldsymbol{q} d^{Nr} \boldsymbol{L} \times e^{-\beta H} \qquad (eq\ 2.21)$$

2.1.5 Integrators

Numerical integration is required to solve the complex systems of equations. If given initial positions, velocities, and a time step the ordinary differential equations can be solved using various algorithms, such as Verlet[119] and Leapfrog[120]. The Verlet algorithm uses two Taylor series expansions to derive the class:

$$q(t+\delta t) = 2q(t) - q(t-\delta t) + \delta^2 q(t) + O(\delta t^4) \qquad (eq\ 3.00)$$

The first-order central difference calculates the velocities,

$$q(t) = \frac{1}{2\delta t}[q(t+\delta t) \dot{-} q(t-\delta t)] + O(\delta t^2) \qquad (eq\ 3.01)$$

while the Leapfrog integrator differs by offsetting the velocity and position update timing by integrating velocities at time $t + \frac{1}{2\delta t}$ and positions at time $t + \delta t$.

$$q\left(t + \tfrac{1}{2}\delta t\right) = q(t \dot{-} \tfrac{1}{2}\delta t) + \frac{\delta t}{m} F(t) \qquad (eq\ 3.02)$$

$$q(t+\delta t) = q(t) + \delta t \dot{q}\left(t + \frac{1}{2}\delta t\right) \ (eq\ 3.03)$$

The velocity computed at time t by averaging the two half-time steps.

$$\dot{q}(t) = \tfrac{1}{2}[\dot{q}\left(t + \tfrac{1}{2}\dot{\delta} t\right) + q\left(t \dot{-} \tfrac{1}{2}\delta t\right) \qquad (eq\ 3.04)$$

2.1.6 Periodic Boundary Conditions

Simulating full-sized systems with Avogadro's number of molecules is not feasible, so periodic boundary conditions[121] must embed the system into an infinite number of unit cells on a continuous condensed state. Embedding the system prevents vacuum effects at the boundaries that

would otherwise exist without calculating an entire mole of atoms. The downside of this approach is that it is only possible to study one protein, ion channel, or one of the many different molecules placed in the primary unit cell. Limitations in today's computing make the upper limits of feasible systems around 500,000 atoms for a few μs. However, if one wanted to run several replicates, troubleshoot simulation parameters, or perform thermodynamic PMFs, the simulations would have to be around 50,000-100,000 atoms and run for shorter timescales (<0.5 μs).

2.1.7 Cutoff Methods

The nonbonded forces are usually the most time-consuming part of the MD calculations. Usually, these can be accelerated using a cutoff.[122, 123] For example, the van der Waals interactions decay quickly as distance increases; thus, a spherical cutoff r_c can be applied so than any interactions beyond a distance of r_c are ignored. Another approach is to use the Lennard Jones-PME, which is an approach that avoids the use of a cutoff. The Lennard Jones-PME works not only for electrostatics but also for Lennard Jones.[21, 124] However, since parameters require re-optimization with LJ-PME, the lipid force field still requires more extensions to varied lipids and was not used in this work.

2.1.8 Thermostat

In MD, the system's temperature is maintained constant in most ensembles. Coupling the system to a thermostat with a fixed temperature achieves the goal of a constant temperature. Several algorithms employ this, including the Anderson thermostat[125], Berendsen thermostat[126], Nosé-Hoover thermostat[127-129], and Langevin (stochastic) thermostat.[109] NAMD uses the Langevin thermostat, while the work with Anton2 involved the Nosé-Hoover thermostat. The weakness of the Nosé-Hoover is that this thermostat directly couples the baths, so there can be oscillations between them, and the system typically needs to be close to equilibrium already. The weakness of

the Langevin thermostat is that it randomly applies force terms to the equations of motion to account for frictional and random forces that combine to represent the ensemble. The limitation is that this thermostat can have artificial and incorrect system dynamics if the friction or frequency are set too high.

2.1.9 Barostats

There must also be ways to control the pressure of fixed-pressure systems. The main barostats used in the simulations were the Langevin barostat for NAMD and the Martyna, Tobias, and Klein (MTK) barostat for Anton2. Although each makes different assumptions, these are reasonable approximations of the ensembles. The Langevin barostat is essentially modeled based on the Langevin thermostat. It is similar in some ways to the MTTK, which is an extension of the MTK method except for the friction damping and the stochastic force. Both barostats reproduce identical "correct" ensembles and are adequate. The Langevin barostate converges faster and oscillates less due to the stochastic collisions and oscillations.[130] The MTK applied energy corrections to equations similar to the Anderson and Hoover equations to satisfy the incompressibility conditions and other issues. The heat bath more directly couples to the simulation box, so there can be oscillations between the system and the bath.[109]

2.1.10 Enhanced Sampling PMF

The basic theory is that the free energy along a pathway depends entirely on the potential energy that is increased or released. Typically, one dimensional (1D) PMFs are sufficient. 1D PMFs heavily utilize restrained configurations of the target molecule. The restraints limit the molecule of interest from unfolding or re-orienting too much unless that re-orientation is the specific dimensions being probed and varied in the PMF.

The sum of the potential energy contributions defines the partition coefficient. The partition function captures changes in the potential energy as a peptide is unbound from a given surface.

$$\boldsymbol{Z_{NVT}} = \int \boldsymbol{e}^{-\frac{\boldsymbol{U(X_0,\ldots,X_{3N})}}{\boldsymbol{k_BT}}} \boldsymbol{dx_0 \ldots dx_{3N}} \qquad (eq\ 4.00)$$

A simple relation converts the partition function into free energy. The partition function would incorporate those effects if the ensemble were NPT or NPAT. Depending on restraints to area or surface tension, the corresponding free energy will change from Helmholtz free energy to some variant with enthalpic contributions.

$$\boldsymbol{A_{NVT}} = -\boldsymbol{k_BTln(Z_{NVT})} \qquad (eq\ 4.01)$$

The Helmholtz free energy then defines the probability of the system. The x' is typically used to denote the extracted degree of freedom. Essentially, one variable is adjusted along the reaction coordinate and allowed to vary, which is the focus of the PMF.

$$\boldsymbol{P(x')} \sim \boldsymbol{e}^{-\frac{\boldsymbol{A(x')}}{\boldsymbol{k_BT}}} \qquad (eq\ 4.02)$$

The WHAM method is implemented by running the MD simulations and calculating the probability of the systems. The raw MD data calculates a probability distribution, from which free energy is calculated along the profile. A challenge with unbiased sampling is that the simulation gets stuck in the lowest energy wells and requires umbrella sampling to apply external forces to sample other parts of the system. Otherwise, there will not be enough barrier crossings.

$$A(x) = -k_BTln\big(P'(x)\big) - U'(x) + F \qquad (eq\ 4.03)$$

U'(x) is the umbrella biasing potential introduced into the system. The challenge is to run separate simulations, each with different offsets in placing the external biases. How much force that is applied needs to be incorporated into the calculations. WHAM determines how the other simulations can be best combined since the probability distribution profile is continuous.

$$P(x) = \frac{\sum_{i=1}^{N_{sims}} n_i(x)}{\sum_{i=1}^{N_{sims}} N_i \exp\left(\frac{\left[F_i - U_{bias,i}(x)\right]}{k_B T}\right)} \qquad (eq\ 4.04)$$

$$F_i = -k_B T ln(\sum_{X_{bins}} P(x) \exp\left(-\frac{U_{bias,i}(x)}{k_B T}\right)) \qquad (eq\ 4.05)$$

The value N_{sims} represents the number of simulations. The parameter $n_i(x)$ stands for the number of counts in histogram bin associated with x. The parameter $U_{bias,i}, F$ represents the biasing potential and free energy shift from simulation i. The function $P(x)$ represents the best estimate of unbiased probability distribution. The parameter F_i and $P(x)$ are unknowns that need to be calculated. The solution involves solving by iteration to self-consistency.

2.2 Specific Computational Methods:

Standard methods involved CHARMM Force Field, NAMD, and GROMACS for calculating the molecular dynamics trajectories. This software involved using different force fields for proteins,[19, 131] lipids,[18, 132, 133] sugars,[134] water,[135, 136] and ligands.[6, 137, 138] Several systems probed the aggregation of proteins using a CUFIX force field, including peptides and carbohydrate sugars. The CHARMM-GUI website generated system inputs. CHARMM-GUI processes the input PDB and generates standardized equilibration and input files compatible with the CHARMM force field for various biological molecules. Excluding the scaled1-4 term prevented overestimating the atoms closest to any given amino acid. Still, the calculation of the long-range electrostatics involved using a particle mesh Ewald (PME)[139] with a force-based switching function set to switch at 10 Å and to cutoff at 12 Å.[140] It is too costly to compute all the effects, especially for short-range Lennard-Jones. The most primitive approach is to cut off the forces beyond a specific range, but causes discontinuity at the cutoff radius. The switching method gradually reduces the potential beyond a specific cutoff to represent interactions more accurately without creating discontinuities. A switch distance function differentiated the van der Waals from the columbic interactions. The pair list distance screens out only pairs of atoms within a cutoff distance. This prevents the computer from cycling through every possible pair. This list is then updated every ten steps. The PME interpolates using a spline order six and a

PMEGridSpacing of 1.0. Standard simulations set the time step to 2 fs/step, and HMR accelerated systems set the time step to 4fs/step. The α6-α7 used the charged NTER and CTER terminal patch to the protein because there was no noticeable problem. The Osh4 ALPS used a CT2 and an NNEU terminal patch to match the experimental conditions specifically and prevent aggregation. Other projects used a variety of patches, but most were neutral. In the 5-HT_{3A} work, the termini were too far apart for problems to occur. Software like Pymol, Chimera, and VMD manipulated the PDB structures. Finally, structure prediction software such as Robetta, AlphaFold2, and Modeler predicted small peptides between 26-41 amino acids.

NAMD software[22, 48, 141, 142] simulated most of the systems in this work, while Anton2 processed the 5-HT_{3A} work, and Gromacs simulated the REMD and some control trials. The general approach for all studies was to start with unbiased binding of ligands to membranes, receptors, and proteins. Place the peptide, protein, or ion channel in the aqueous phase and use NVT or NPT for membranes.

The membranes with the Osh were initially setup using HMMM builder from the *Membrane Builder* of CHARMM-GUI with various orientations of the peptide and the TIP3P water model. This involves replacing the lipid tails with DCLE solvent instead of the actual lipid tails to enhance the diffusion of the lipids by an order of magnitude without affecting the mechanism of attachment.[34, 145] Neutralizing potassium ions were added to the simulation mixtures. These were then equilibrated using a standard 6-step equilibration process before production runs began to obtain equilibrated structures ranging from several μs to 200-300 ns. The HMMM algorithm required setting values for the lipid area scaling factor and terminal acyl carbon position. A value of around 1.3 defined the lipid area scaling factor, although it varied across all the projects. A value of 6 defined the terminal acyl carbon. The HMMM ran using NPAT ensemble. All non-Anton2 simulations used the SHAKE algorithm to constrain the hydrogen atoms.[135] We simulated everything with the CHARMM36 (C36) lipid force field and C36m additive all-atom protein force field because it is the most advanced membrane force field.[18, 131, 132] The full-length simulations are essentially the same as the HMMM except that the full-length simulations delete all of the DCLE and replace the DCLE with the actual atoms of the lipid carbon tails. The full-length simulation uses NPT instead of NPAT because NPT eliminates the artificial restraint on the bilayer lipid packing.

The conversion from HMMM to standard full-length membranes involved taking the input system's final ".pdb", ".psf", and ".str" files. The CHARMM-GUI web server accepted these files to convert the system from HMMM to full length. The CHARMM-GUI server essentially removes the DCLE and extends the lipid tails. The tails then need to be equilibrated. The conversion approach helps preserve the varied peptide orientations and structures on the membrane head and the packing of the different membranes. Non-membrane setups require an initial input ".pdb" from RCSB or AlphaFold that can equilibrate in an aqueous solution. All of these systems ran in NPT but without a semi-isotropic condition. The solution also had ions to neutralize the system and simulate physiological conditions. RCSB complexes were sound starting points for the aFGF and the TLR dimer systems. The ECDs of the TLR dimers could be aligned over the full AlphaFold structures to generate full structures.

To estimate the free energy of binding, an artificial spring was added into the system on a selection of atoms with a specified spring constant.[150] The positions of the particle center of mass are recorded at a high frequency and compared against the equilibrium spring position to calculate the force applied to the selection. The WHAM method is required to combine the different windows with different spring equilibrium positions to provide the combined energy surface.[150] For membrane PMFs, the C2 lipid atom is restrained with strong 10 kcal/mol restraints. The protein was forced into the correct shape and orientation by using RMSD collective variable orientation to prevent it from unfolding, and this used a harmonic spring constant of 20 kcal/mol. A spring constant of 5-10 kcal/(Å^2 * mol) was typically used to pull the peptide off the membrane.

Chapter 3: Osh α_6-α_7

The contents of this chapter are modified from the following publication with minor adjustments to the results to showcase free energy calculations to some membranes:

- Reprinted (adapted) with permission from Allsopp, R. J.; Klauda, J. B. Impact of PIP2 Lipids, Force Field Parameters, and Mutational Analysis on the Binding of the Osh4's α6–α7 Domain. *The Journal of Physical Chemistry B* **2021**, *125* (20), 5296-5308. DOI: 10.1021/acs.jpcb.0c10393. Copyright {2021} American Chemical Society

3.1 Introduction

Since all cells have lipid membranes, it is crucial to understand the role of lipid transport proteins in maintaining membrane composition. Model eukaryotic cells such as yeast help to study lipid transport, including transporting lipids between and within different cellular membranes.[151] Since lipid transport proteins are in control of modulating the lipid composition of membranes, they also regulate many cellular processes, including vesicular trafficking, signal transduction, and lipid metabolism.[1, 151-153] Lipid trafficking is controlled by several different transporters. In addition, lipid transport between unconnected organelles occurs via vesicular transport machinery such as the mitochondria and peroxisomes. Also known is that non-vesicular lipid transport plays a vital role in total lipid transport.[151, 154, 155] Finally, there is evidence that intracellular lipid transport enhances at membrane contact sites, which are small gaps of 10-20 nm that form between organelles.[56, 151, 156, 157]

Intracellular membrane composition is partially controlled in yeast by a family of seven oxysterol-binding homology (Osh) proteins with overlapping roles.[40] Although the Osh family of proteins is specific to yeast, mammals have homologous proteins.[40] A key feature is that each of the proteins has an affinity for specific membrane compositions and, thus, organelles.[38] Osh4 localizes to the endoplasmic reticulum[53] and Golgi complex.[40] At the same time, the plasma

membrane has an affinity for Osh2, Osh6, and Osh7.[38] The Osh family consists of seven total proteins, with the first three (Osh1-Osh3) being significantly longer in sequence (>1100 residues) compared to the Osh4-7 proteins, which are smaller (~430 residues). Osh1-3 binding sites include the phosphatidylinositol (PI) binding pleckstrin homology (PH) domains[54] with Osh1's PH domain selective for lipids in the Golgi.[158] All Osh proteins have an oxysterol binding domain, which also has an affinity for lipids with two phosphates, such as the one studied here: 1-palmitoyl-2-oleoyl-*sn*-glycero-phosphoinositide 4,5-bisphosphate (PIP2).[38] This domain holds lipids and sterols as they shuttle between organelles during the lipid transfer process.

This work focuses on one of the lipid transport proteins in yeast, i.e., Osh4. This protein is composed of multiple domains that are involved in the membrane binding mechanism and protein function. The oxysterol binding protein related domain (ORD, grey in Figure 3.1) is responsible for the uptake and placement of the transported lipids. All Osh proteins have this conserved attribute.[38] Another common domain is the amphipathic lipid packing sensor (ALPS, blue in Figure 3.1),[40] necessary for binding to membranes and sensing their curvature or lipid packing due to its amphipathic structure.[58]

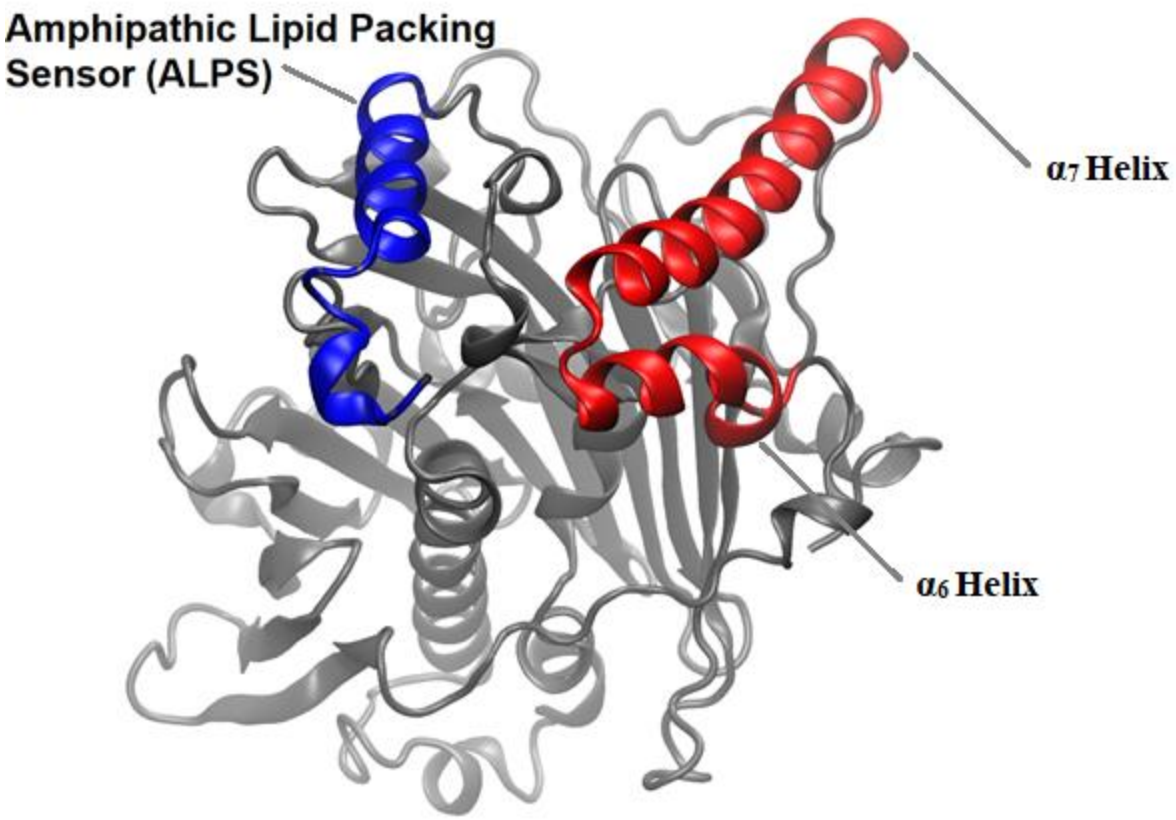


Figure 3.1. This image shows the ALPS (blue) and α_6-α_7 helix (red) with the entire Osh4 protein (silver).

The binding region in question for our work ranges from residues E312 on α_6 to I355 on α_7 (red in Figure 3.1). It is unknown if the α_6-α_7 domain independently interacts with the membrane, but evidence shows these residues play a crucial role in the protein's function. The objective of this research was to investigate these interactions further. First, from the location of this domain in the crystal structure, the α_6-α_7 domain is far away from the lipid transport section of the protein, so that loss of function is not necessarily related to a loss of ability to extract PI4P.[40] Experiments showed that K336 was essential,[40, 60] because the K336 mutation to non-polar alanine significantly changed the binding affinity and the labeled cholesterol transference of Osh4. Moreover, the E312K mutant resulted in a loss of function.[159] In the original article that published Osh4's

structure, sulfate ions bound to K336 and potentially mimic the role of membrane phospholipid phosphates.[40] It is also noted that the region surrounding E312K lies within a domain that exhibits features of a Pleckstrin homology (PH) domain, which is a known domain that binds to PIP lipids.[159]

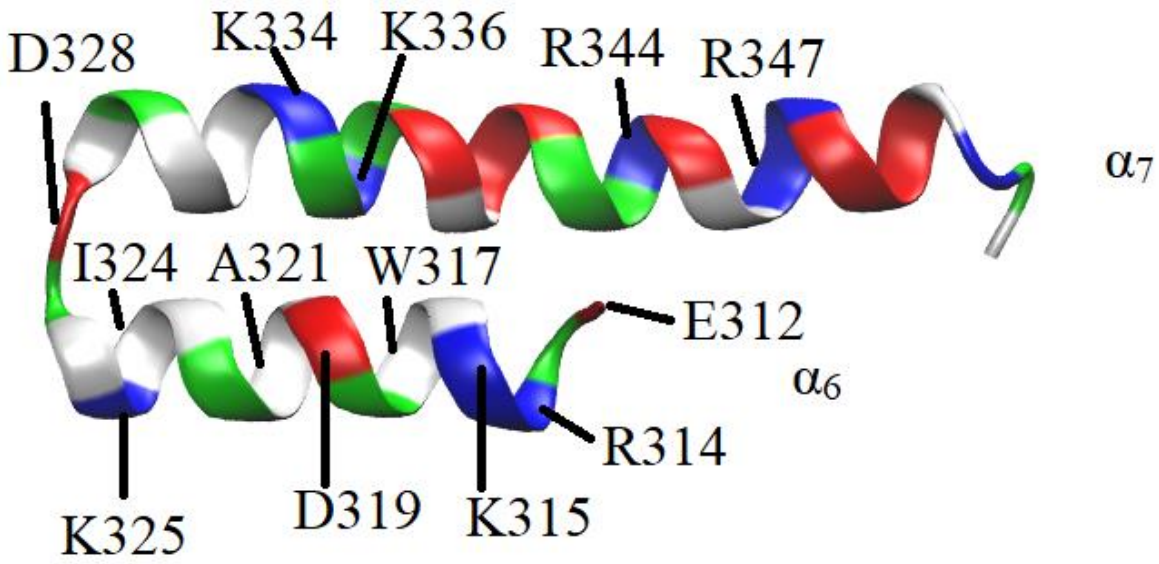


Figure 3.2. The α_6-α_7 peptide is oriented so that the facing residues, K334, K315, D319, and E312, tend to have less interaction with the membrane while the inward residues, K325, R314, and D328 interact nearest to the membrane. Negatively charged amino acids were colored red, positively charged amino acids were colored blue, white is hydrophobic, and green is polar. The helices also have non-polar residues on the opposite side of the polar face. For the α_7 helix, the primary amino acid residue interacting with the membrane is K336. However, additional residues such as R347 and R344 hover above the membrane but still interact with the membrane, consistent with the simulations.

A limited scope of research focused on the α_6-α_7 helices (Figure 3.2), while other domains on Osh4 interact with the membrane before considering the more complicated full protein. This research tests whether this domain can bind to model membranes. We investigated the effect of membranes having a mixture of 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) and 1-

palmitoyl-2-oleoyl-glycero-3-phosphidylserine (POPS) lipids and the presence of a few PIP2 lipids. Osh4 extracts PI4P from membranes functionally, but this capability was not the focus because the α_6-α_7 helical domain is away from the PI4P binding pocket.[160] It is known that PIP2 lipids are not needed for binding.[54, 159] However, PIP2 is speculated to assist the critical process by increasing the dwell time of the peptide on the liposome surface,[54, 159] which is why this lipid was included in our study. Some amino acids that are hydrophobic and likely attracted to the membrane have I324, A321, and W317. Simulations will be used to determine which residues of the peptide drive the binding process and the orientation of the bound peptide before and after binding.

Four *hypotheses* that were probed in this work include: *1.* Certain residues play a vital role in the binding process, *2*. The mechanism of peptide binding is related to the angle of the helical approach, 3. Peptide-bound orientation differs in membranes with and without PIP2, and *4*. Salt-bridge interactions between the two helices stabilize the membrane binding structure.

This report is organized into a Methods Section, then Results and Discussion, and finally, the Conclusion. The Methods Section explains how the simulations were set up and the details needed to reproduce the work. Specifically, this section describes two lipid compositions, initial builds for equilibration, simulation conditions, and how different analyses were performed. The Results and Discussion Section discusses PIP2 and non-PIP2 binding parameters, including critical mechanisms, force field dependence, structure of peptide and placement in the membrane, the residue-membrane interactions meaningful to binding, PIP2-specific interactions, and mutational analysis to understand the binding process.

3.2 Methods

MD simulations were used to study the Osh4 α_6 and α_7 helices interacting with POPC, POPS, and PIP2 membranes. The MD simulations were divided into two sets, one with PIP2 lipids and the other without. The simulations without the PIP2 lipids were composed of 40 POPC and 10 POPS in bilayers of 50 lipids per leaflet (80% POPC and 20% POPS). The simulations with PIP2 lipids were composed of 40 POPC, 10 POPS, and 3 PIP2 in bilayers of 53 lipids per leaflet, where the lipids were evenly distributed between the two leaflets (75.4% POPC, 18.9% POPS, and 5.7% PIP2). Although the simulation box was relatively small, it was sufficient for this research since there were no interactions between the primary space peptide and its image.

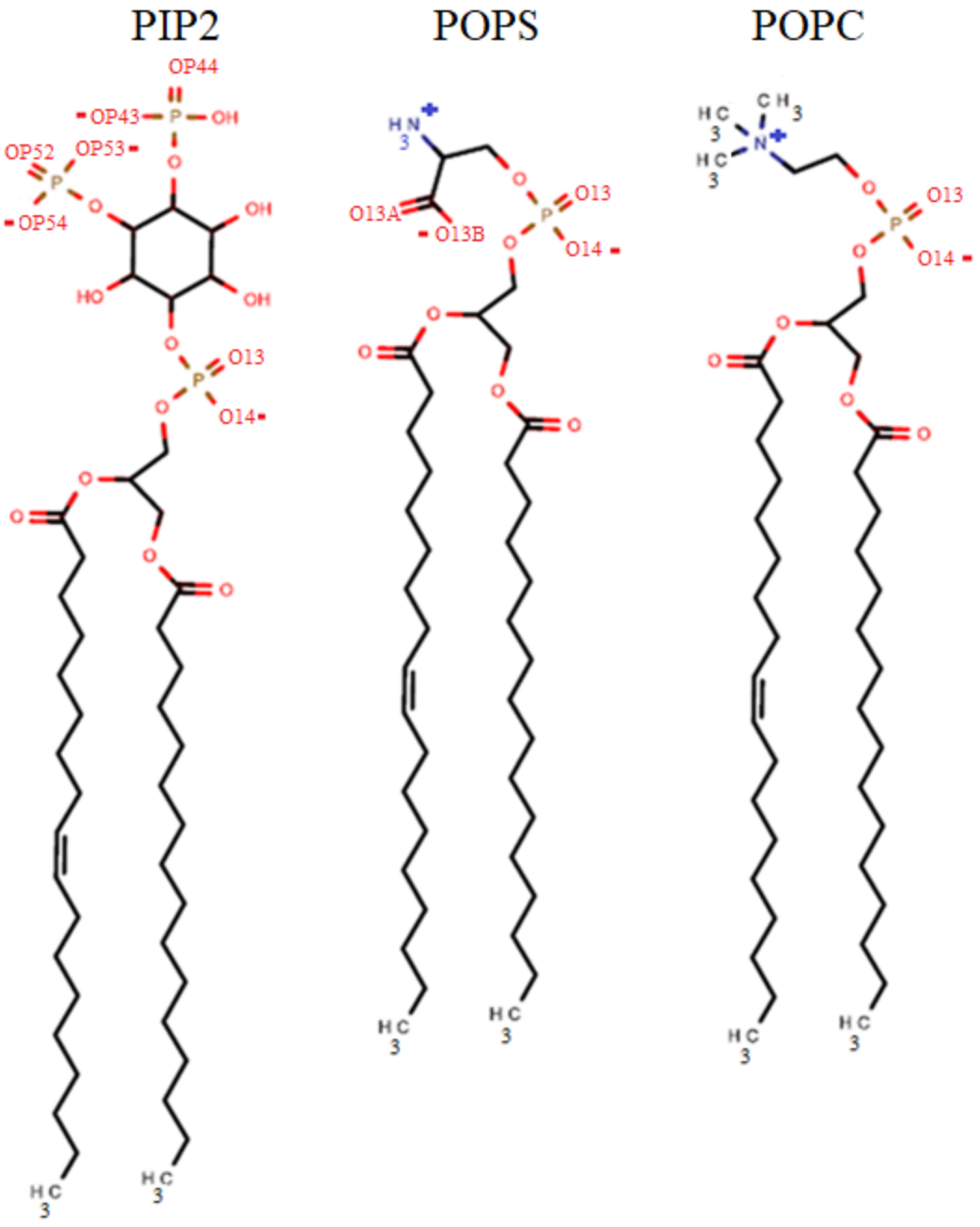


Figure 3.3: This is an image of the three lipid types that were used in the simulations 1-palmitoyl-2-oleoyl-*sn*-glycero-phosphoinositide 4,5-bisphosphate (PIP2), 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine (POPS), and 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC). The images were created using Marvin to draw the chemical structures.[161]

The membrane models were built with the HMMM builder feature within the *Membrane Builder* of CHARMM-GUI (www.charmm-gui.org)[48, 144-149, 162] with either 83 or 120 molecules of water per lipid for either the horizontal or vertical orientations using the TIP3P model[143] for water. Neutralizing potassium ions were added to the mixture. The HMMM builder feature accelerates the simulation with a simplified approach by modeling only the individual lipid head groups and replacing the tails with DCLE solvent without the extended lipid tails, increasing lipid diffusion.[145] It is known that the HMMM model accelerates the diffusion of the lipids by an order of magnitude without affecting the attachment mechanism, thereby saving computational resources without losing binding details.[34, 145] Each of the models was equilibrated using the standard six-step CHARMM-GUI[144, 147] protocol for 225 ps. The peptide coordinates that range from E312 to I355 were extracted from the full protein (PDBID: 1zhy[40]) to simulate the α_6-α_7 helix. The terminal group was patched with the charged NTER and CTER patch without acetylation or methylation. The NTER and CTER capping method was used based on a review of the initial results of simulations that did not identify any adverse interactions of significance with these charged terminal residues. Next, the peptide was inserted into the aqueous phase with different orientations and distances away from the membrane in CHARMM-GUI without forcing major peptide-membrane interaction. The system was equilibrated for several short 125-225 ps segments with gradually reduced restraints before it was prepared to run with Nanoscale Molecular Dynamics (NAMD) [142, 163] by completing the six standard CHARMM-GUI equilibration runs that also used NAMD. The process involved minimizing the system and then simulating it with decreasing levels of restraints. The lipid area scaling factor was set to 1.3 with a terminal acyl carbon number of 6.[145] These settings help accelerate the binding process by exposing the lipid tails to the solvent and allowing for increased packing defects in the membrane.[145] The

equilibration and production runs ran with a 2 fs time step. Other settings for the equilibration include a restart frequency of 500 steps, a trajectory, and an extended system output frequency of 1,000 steps. The frequency for all was changed to 5,000 steps for the production runs. All of these initial equilibration and production runs used NAMD on a constant Number, Pressure, Area, and Temperature (NPAT) ensemble for approximately 150-200 ns of total simulation time.[142, 163] All simulations used the SHAKE algorithm to constrain hydrogen atoms.[135] Each membrane system has ten replicates that have five unique initial peptide orientations (horizontal (H) and vertical (V)), with details in Table A1. For these five initial placements, two replicas are simulated (10 total) to provide a better sampling of the binding to model membranes. For all simulations, the CHARMM36 (C36) lipid force field[132] was used since it has the most updated parameters for the PIP lipids.[164] The C36m additive all-atom protein force field was used[18, 131] along with the TIP3P water model.[18, 143] Finally, the force field was changed after 150 ns of full-lipid simulation to the CUFIX parameters,[113] which was run for an additional 150 ns with an adjustment in the strength of ionic interactions so that the results of this work will be comparable with other ongoing research.

Van der Waals and electrostatic forces were computed using a Lennard-Jones force-switching function over 10 to 12 Å.[140] The temperature was kept constant at 303.15 K to ensure a fluid phase membrane using the Nosé-Hoover thermostat and Langevin dynamics.[129] All simulations were run with periodic boundary conditions (PBC) that were used to evaluate long-range electrostatic interactions using the Particle Mesh Ewald (PME).[139]

Four of the HMMM simulations from each membrane model were chosen for conversion to full-length simulations via CHARMM-GUI *HMMM* to full-lipid converter and run again with NAMD for 150 additional ns. For the all-atom systems, the temperature was held constant like the HMMM simulations, but the box size varied semi-isotropically (X=Y but not Z) while keeping the

pressure at a value of 1.01325 bar using a Langevin piston.[22, 115, 116] This allowed for the membrane to more closely resemble natural membrane lipid packing. The non-PIP2 simulations were chosen based on differences in orientation from the deepest residue graphs, and the PIP2 simulations were selected to depict the full range of interactions between the PIP2 lipids and the peptide.

For the analysis of peptide binding to the membranes, the minimum penetration distance graphs were generated by measuring the vertical positions of all the peptide residues in VMD[141] around the binding region and the average phosphate position over every frame. This method was used for the minimum position graphs and the angle graphs. The figures were rendered with VMD Tachyon internal memory processing.[165] It was determined that the binding was impacted by two different angles, the angle of approach (θ) and the angle of rotation (ϕ) (Figure 3.4). The angle of approach (θ) measures the angle between the peptide and the membrane. The rolling angle is measured between residues I324 and N330, while the angle of approach is calculated separately for each helix between residues K315 to L326 on the α_6 and residues N330 to K353 on the α_7.

A steep angle of approach (θ) means that the trailing ends of the helices are further away from the membrane. The angle of rotation (ϕ) determines whether the α_6 or the α_7 is closest to the membrane. A flat angle of rotation means that the α_6 and α_7 helices are at the same distance from the peptide, while a 90° rotation results in the α_7 above the α_6 helix. This orientation results in the α_6 helix being closer to the membrane and is referred to as the peptide approaching the membrane on the side. The angle was calculated using the VMD orient package by defining the angles with several residues, selecting only the z component, and subtracting 90°.

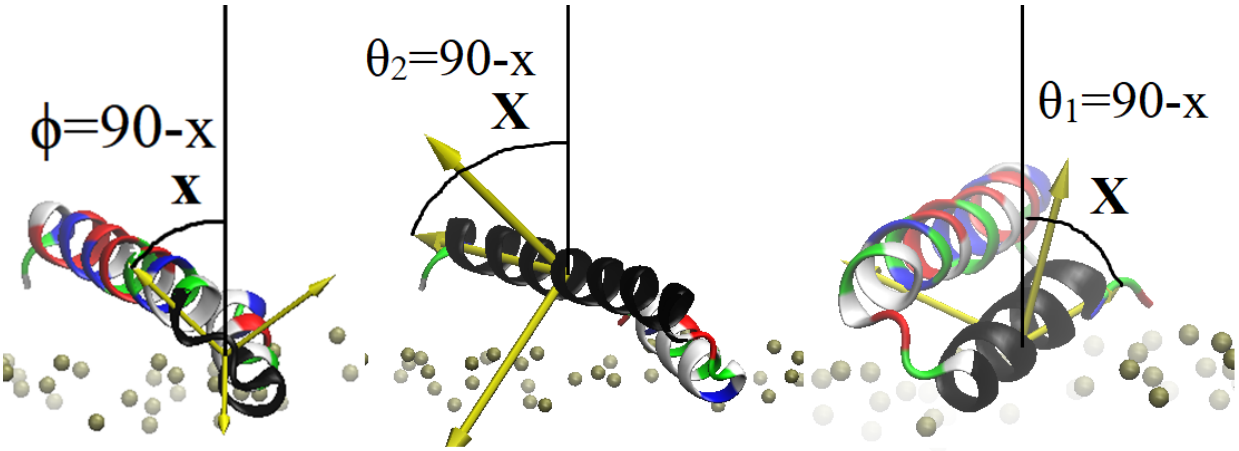


Figure 3.4. Left shows the ϕ angle, defined by the vector from residues I324 to N330 (black), and its angle with the bilayer normal, which measures the angle of rotation. The middle shows angle θ_2, defined by the vector of residues N330 to K353 (black) with the bilayer normal and measures the angle of approach of the α_7 helix. Θ_1 is defined similarly except with residues K315 to L326 on the α_6 helix. The color scheme for the peptide is cationic (blue), anionic (red), non-polar (white), and polar (green).

The interaction energy (calculated with the CHARMM program) per residue is the sum of the van der Waals and the electrostatic contributions. The hydrogen bonds are based on a 3 Å cutoff and a 20º angle cutoff using the VMD Hbonds plugin. The distance used to calculate the hydrogen bond cutoff is defined as the distance from the heavy atom of the hydrogen bond acceptor to the hydrogen on the hydrogen bond donor. VMD defines the internal angle as an angle between the donor and acceptor atoms. The salt bridges between the peptide and the membrane were measured with CHARMM, and the CHARMM program counts more salt bridges than hydrogen bonds because there is no angle cut-off. The salt bridges were calculated using an hbonds function in CHARMM, but only salt bridge interactions with no hydrogen bonds were examined. The distance cutoff was set to 3Å for these peptide-membrane salt bridges, differing from the peptide-peptide salt bridge cutoff of 4 Å. There was no angle cutoff set, and the atoms that were being examined include resonance, and the PIP-specific atoms include OP43, OP44, OP52, OP53, OP54, along with O13, O14 which are the two resonance pairs shared among all of the lipids and lastly

O13A and O13B that are on the POPS lipid. There were two positively charged atoms on the lipids, including POPS and POPC named N.

An essential aspect of the PIP2 bound simulations is that the interaction energy was lower in the trials where the peptide bound directly on top of the PIP2 lipids. The peptide bound on top of the PIP2 lipids in trials V2a PIP2 and V2b PIP2, indicating that there is an energetic penalty if the peptide binds on top of the PIP2 lipids because the peptide is not as deeply bound. When the membrane is too highly enriched in PIP2 lipids, the peptide does not bind as deeply, while when there is just the right amount, the peptide binds with PIP2 on the sides, giving it a deeper placement than it would have without the PIP2 lipids.

3.3 Results and Discussion

This study evaluates the role of PIP2 lipids on the binding mechanism of the α_6- α_7 peptide and its fully bound state to the membrane, focusing on the PIP2-specific interactions, peptide structure, and residues critical to the binding. The crucial residues that help orient the peptide before binding appear to be determined by a few charged lysine amino acids that are attracted to the membrane phosphates. Several parameters of the peptide binding are explored, such as the peptide orientation during and after binding, the deepest residue center of mass (COM), the role the salt bridge plays in stabilizing the peptide, PIP2 sensitive conformational change, and the influence of binding with an updated force field. The binding interactions and energy were also evaluated to determine where the PIP2 lipids interact with the peptide. Replicas for these simulations are noted based on their initial configuration, i.e., H=horizontal and V=vertical, relative to the membrane surface. Five orientations were used (H1, H2, H3, V1, and V2), and duplicates of these were also simulated so that H1a and H1b have the same starting point but have varied trajectories.

3.3.1 Peptide Binding to non-PIP2 Lipid Membranes

Peptide Binding Mechanism to HMMM

The following section was included to gain insights into the binding mechanism, which addresses *hypothesis 1* (the key binding residues) and *hypothesis 2* (the mechanism of peptide binding). These show the binding process and mechanism. Two different binding approaches were observed during the HMMM phase of the simulations. The first approach (50% of the time) was with the peptide on its side as it approached the membrane (Figure 3.5A); Figures A1-A2 show the position of the instantaneous deepest residue COM over time and were used to visualize the binding process. In the second approach, the peptide remains flat with both helices at the same level as the peptide approaches the membrane (Figure 3.5B). A critical residue that was identified, K325, is positioned on the side, which allowed the peptide to interact with the lipid head groups as it comes closer to the membrane, especially during the side approach (Figure 3.6). K325 and R314 are the only charged residues on the otherwise relatively uncharged α_6 helix. As the non-polar residues lodge deep in the membrane, K325 is brought to a similar level, as seen in Figure 3.6.

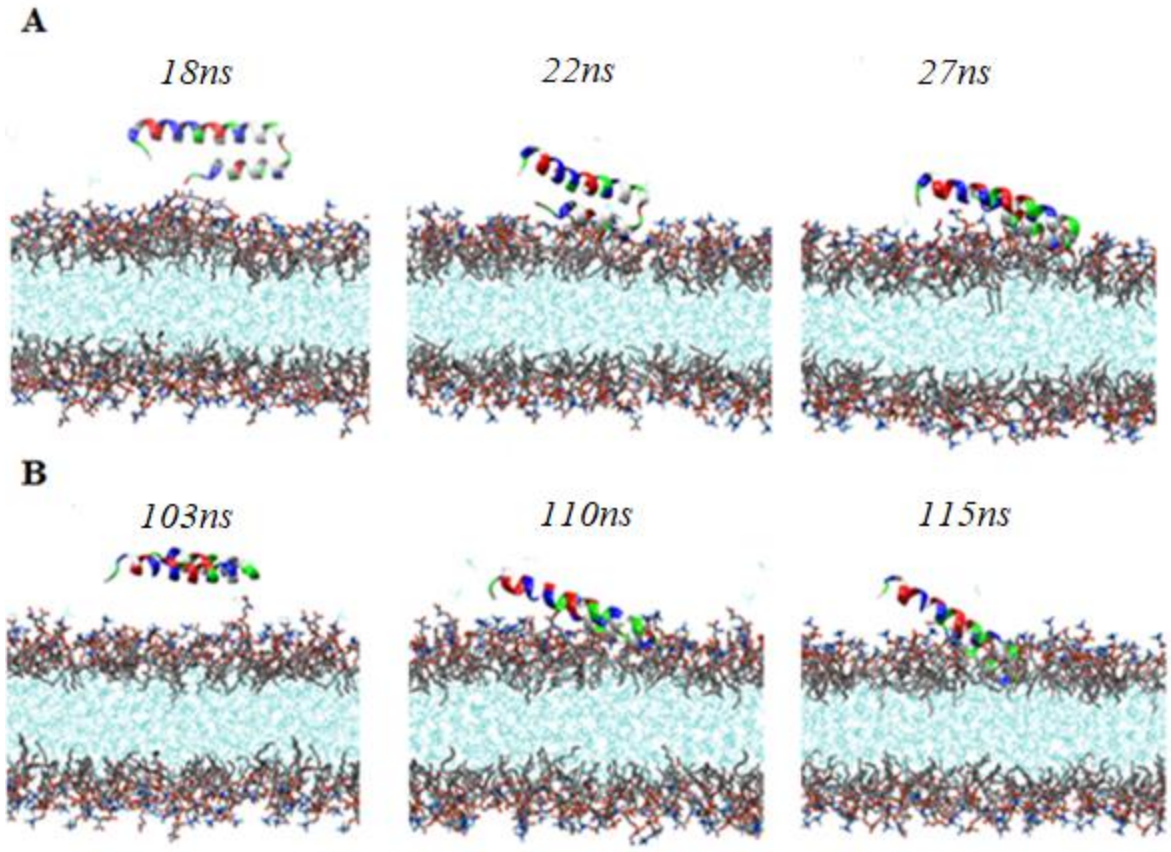


Figure 3.5: The binding approach for the rolling motion around the horizontal axis where the peptide transitions from a 90° angle of rotation ϕ (the longer α_7 helix is on top of the α_6 helix) to 0° (the two helices at the same level) (**A**) peptide rolls sideways in H1a, (**B**) peptide approaches flat and rotates to a less extreme ϕ angle in H3a. The hydrophobic DCLE core is shown as cyan with short-tailed lipids in silver (carbon), red (oxygen), gold (phosphorous), and blue (nitrogen).

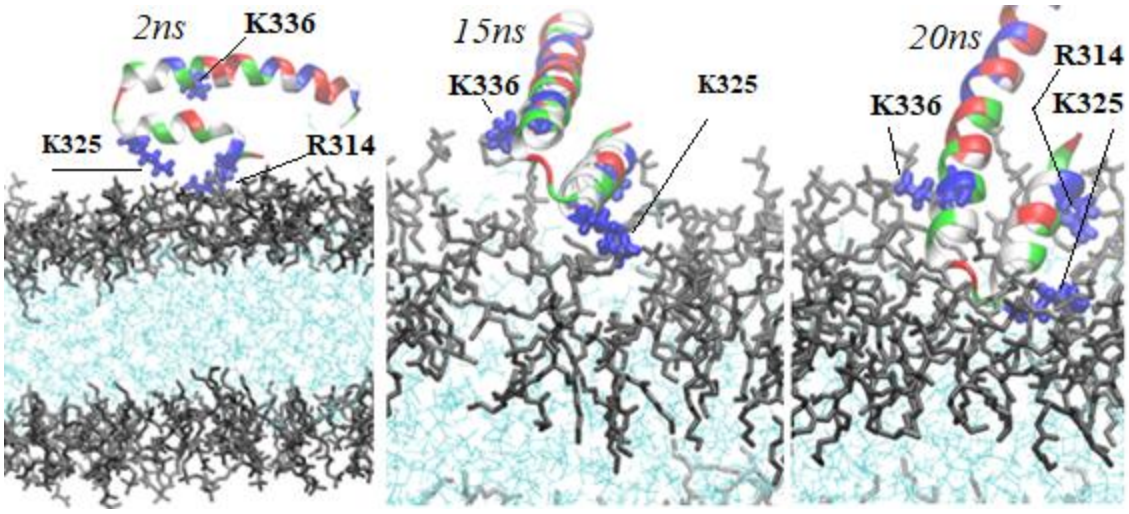


Figure 3.6: H1a shows the three key residues, R314, K325 on the α_6 helix, and K336 on the α_7 helix as the peptide approaches to bind onto a membrane in H1a. The hydrophobic DCLE core is shown as cyan, with the short-tailed lipids in silver.

Due to equilibration and eventual binding process, the peptide's structure altered from the original crystallographic structure, as is expected with MD simulations in a varied environment. The RMSD of the peptide was calculated to illustrate changes to the peptide structure over time (Figure A3-A4). Generally, the RMSD is 3-4 Å with the exception of the V2a replica due to deformation of the α_7 helix. This indicates that nearly all runs maintained most of the starting crystal structure, as this RMSD range is typical for MD simulations.

Based on the simulations, three positively charged residues (R314, K325, and K336) drive the binding process, narrowing *hypothesis 1* to these three residues. These simulations further revealed that the passage of the peptide through the main phosphate layer was initiated by the lysine residues that make first contact due to electrostatic attraction to the phosphates.

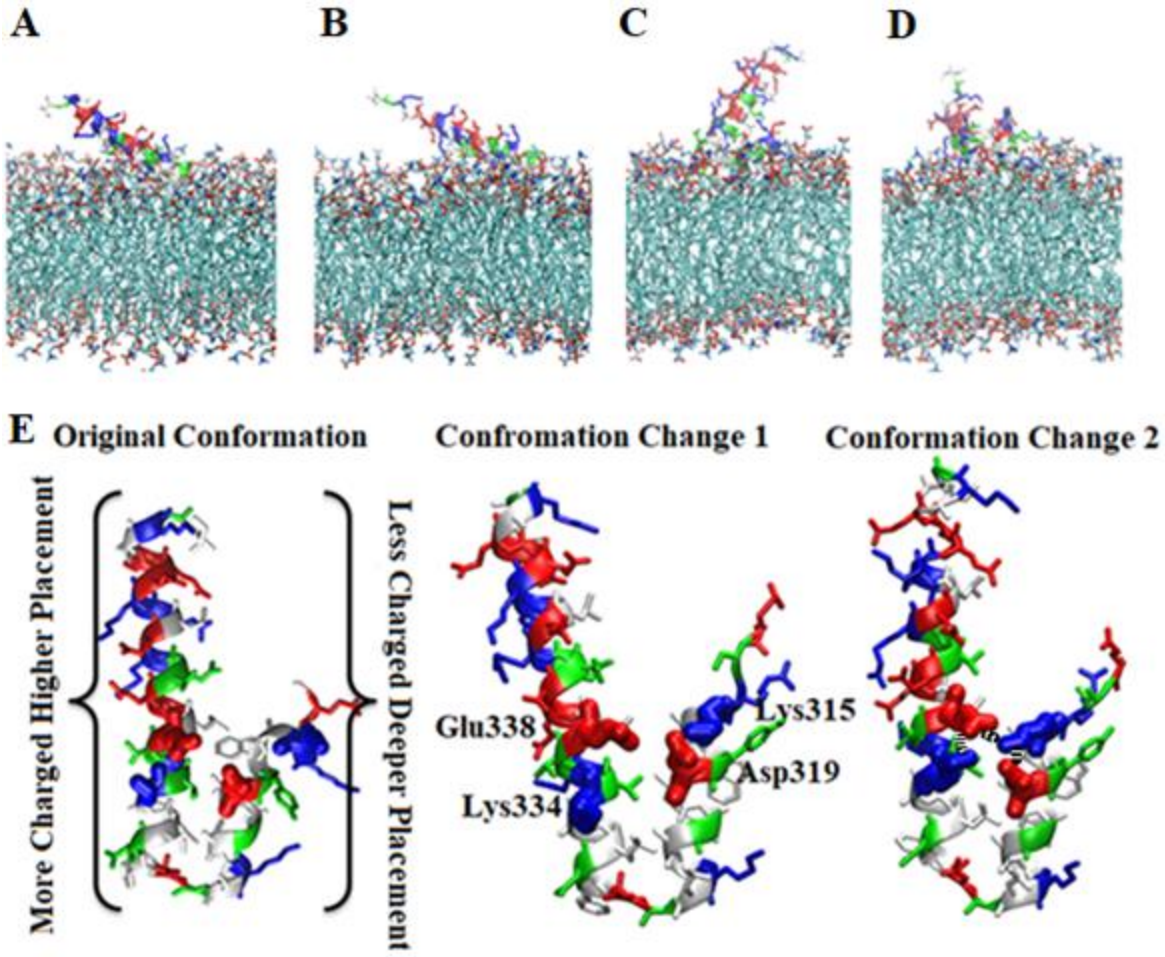
Regarding *hypothesis 2*, the orientation of the peptide is measured to assess the impact on the binding process. The rolling motion of the peptide is measured by the angle of rotation ϕ in Figure A5-A6, but nothing noteworthy was identified from this parameter other than the wide

variation during binding. The angle of approach θ was measured individually for each helix during and after binding, which revealed that some simulations had high and low angles of approach after binding. Specifically, simulations H2a and H2b had the lowest angles of approach (Figure A7-A8) after binding. All of the measured angles can be seen in Figures A5-A10. Although the binding mechanism varied with these angles, there was a precise single mechanism of binding. Thus, our tested h*ypothesis 2* was shown to be invalid. Figures A11-A12 show the final bound states for the non-PIP2 HMMM simulations.

Structure of Peptide and its Placement in the Full-Membrane

The full-length lipid simulations were used to refine the HMMM bound state to better understand this peptide's membrane-binding states, with four of the non-PIP2 simulations (H1a, H2a, H1b, and H2b) chosen. The decision to use these four simulations was based on differences in the HMMM results in the angles of approach and bound states. The RMSD of the peptide was calculated for all the full-length simulations (Figures A13-A14) to analyze how the relative protein structure varied over time. The final bound states of the two full-length simulations are shown in Figure 3.7, along with an image of the peptide with intra-peptide salt bridges.

In Figure 3.7A-D, there are several end snapshots taken from the simulations. The two different orientations refer to the high and low angle of approach. These images indicate that there are wide variations in the angle of approach θ_2 during the simulations and that there are different states accordingly. The high angle state (H1a and H2a) had an average θ_2 angle of 29.0° and the others had a low θ_2 angle of average 14.4°. Although the results support *hypothesis 3* because the orientation of the peptide can vary based on the angle of approach, additional study of the link to the PIP2 is needed and included later in this paper.


A
B
C
D
E
Original Conformation
Confromation Change 1
Conformation Change 2
More Charged Higher Placement
Less Charged Deeper Placement
Glu338
Lys334
Lys315
Asp319

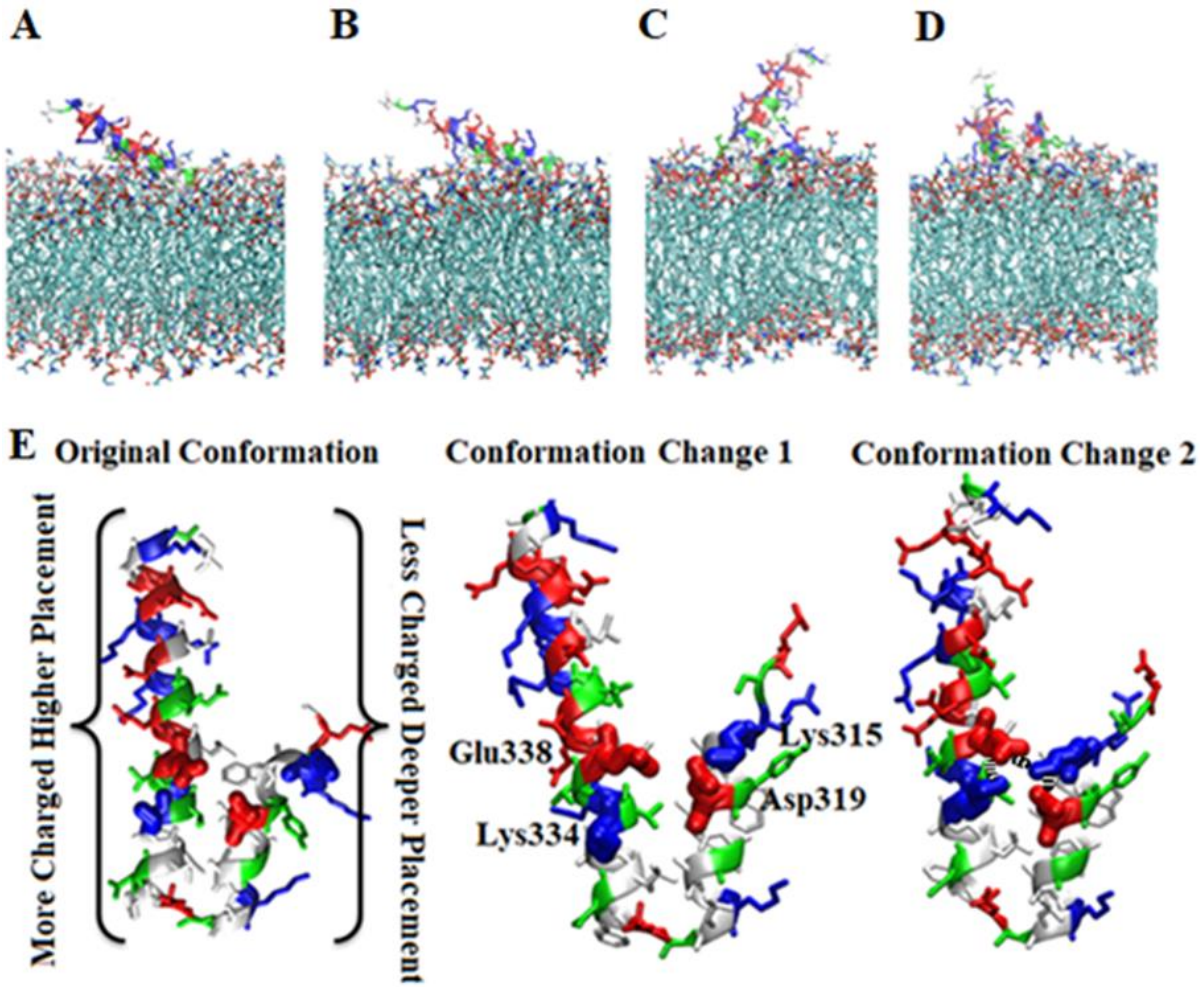


Figure 3.7: Images from full-length simulations **A**. is the final bound state of H1a **B**. is the final bound state of H2a, **C**. is nearly the final bound state of H1b, **D**. is nearly the final bound state of H2b. **E.** is an image showing the structure of the peptide from H2a, with four electrostatic residues labeled to show that an intra-peptide salt bridge exists, which helps stabilize this critical region of the peptide. There was one simulation where the distances were close enough to be classified as salt bridges, H1a.

While *hypothesis 1* asserts that specific residues play a crucial role, other low-energy residues that support the binding are less important due to how they are placed in the membrane. The per-residue interaction energies (Figure A15), the hydrogen bonds (Figure A16), and the salt bridges (Figure A17) are shown in the supporting materials. R314, K325, and K336 have elevated values for all three parameters and are thought to be crucial to the binding process by helping to

bring the hydrophobic portion into the hydrocarbon region. On the other hand, R344 and R347 appear to not be essential because of their placement, meaning they do not reliably contact the membrane surface, as opposed to R314, K325, and K336, which are believed to help orient the peptide. The other residue that appears to be non-essential is K315 since it only has a high value of interaction in simulation H2a. Upon further examination, the increase in the interaction energy is believed to have resulted from a POPS lipid hydrogen bonding to the nitrogen backbone. Since there are only a few salt bridges with the charged portion of the peptide, there are many hydrogen bonds.

Regarding *hypothesis 4*, the simulations were examined to correlate the salt bridge interactions with the stability of the membrane binding structure. As mentioned above, there are a series of interactions that range from direct peptide-peptide salt bridges (H1a) to no peptide-peptide interaction but allow for interaction with lipids on the sides. The most observed conformation was the one in the middle of the bottom panel of Figure 3.7, which had a slight rearrangement from the original structure. The conformation on the right was observed only once, and that occurred in the H1a full-length simulation but was sustained for about 120 ns. This is the first spontaneous transition from a state of no interaction during the end of the HMMM to a much closer conformation with a salt bridge between K315-E338, and finally, to have it naturally break up. This confirms the validity of *hypothesis 4.* However, these peptide-peptide interactions are not essential for maintaining the structure of the peptide since there were no drastic changes after the conformational change was broken up. The peptide-to-peptide salt bridges may promote folding in the aqueous phase by forming an attraction between the α_6-α_7 helices to create a side-by-side motif. The four amino acids involved in these potential salt bridges are marked explicitly in Figure 3.7, i.e., K315, D319, K334, and E338. The distances between these amino acids versus time are

shown for all the non-PIP2 simulations in Figures A19-A21. The ones that are believed to be close enough to be salt bridges have distances below 4 Å and are less frequently observed for K315-E338, which bridges between the α_6-α_7 helices. There is no indication of water-mediated interactions, and the majority of the time, there are either direct salt bridges or no interactions at all. The most stable simulation of all the non-PIP2 full-length simulations was H1a which had bridges during full-length simulation less than 4 Å between D319-K315, 52% of the time, between K315-E338, 35% of the time, and between E338-K334, 16% of the time. All of the salt bridge occurrences for all runs can be found in Tables A3 and A4, based on residues close enough to qualify as salt bridges.

In terms of peptide placement, Figures A22-A25 show the relative position of each residue. In general, the α_6 helix fully contacts, and most of it penetrates the membrane, while only a portion of the α_7 helix is in complete contact with the membrane, and most of it lies above the membrane. An exciting aspect of the peptide is that R347 and R344 approach close enough to hover around and interact with the lipid head groups. The placement graphs were also helpful for observing differences in the position of K315 to explain why some simulations had POPS lipids around K315 but did not interact with them. Specifically, H1a and H2b had lipids around K315 that did not interact with K315. The reason why the POPS lipids did not interact could be explained by the higher placement above the membrane, as shown in the placement graphs. It is intriguing that this residue that may specifically interact with POPS, can vary significantly between simulations.

3.3.2 Peptide Binding to PIP2 Lipid Membranes

Peptide Binding Mechanism to HMMM

In general, the peptide had the same tendency to orient to one of two main binding approaches. The significant difference is that PIP2 lipids can interact electrostatically with residues

on both helices before the peptide fully binds to the membrane. As seen in Figure 8A, the residues interact with the PIP2 lipids first before interacting with the other lipids. There are several different figures that feature the bound states of the peptide, including Figures A26-A28.

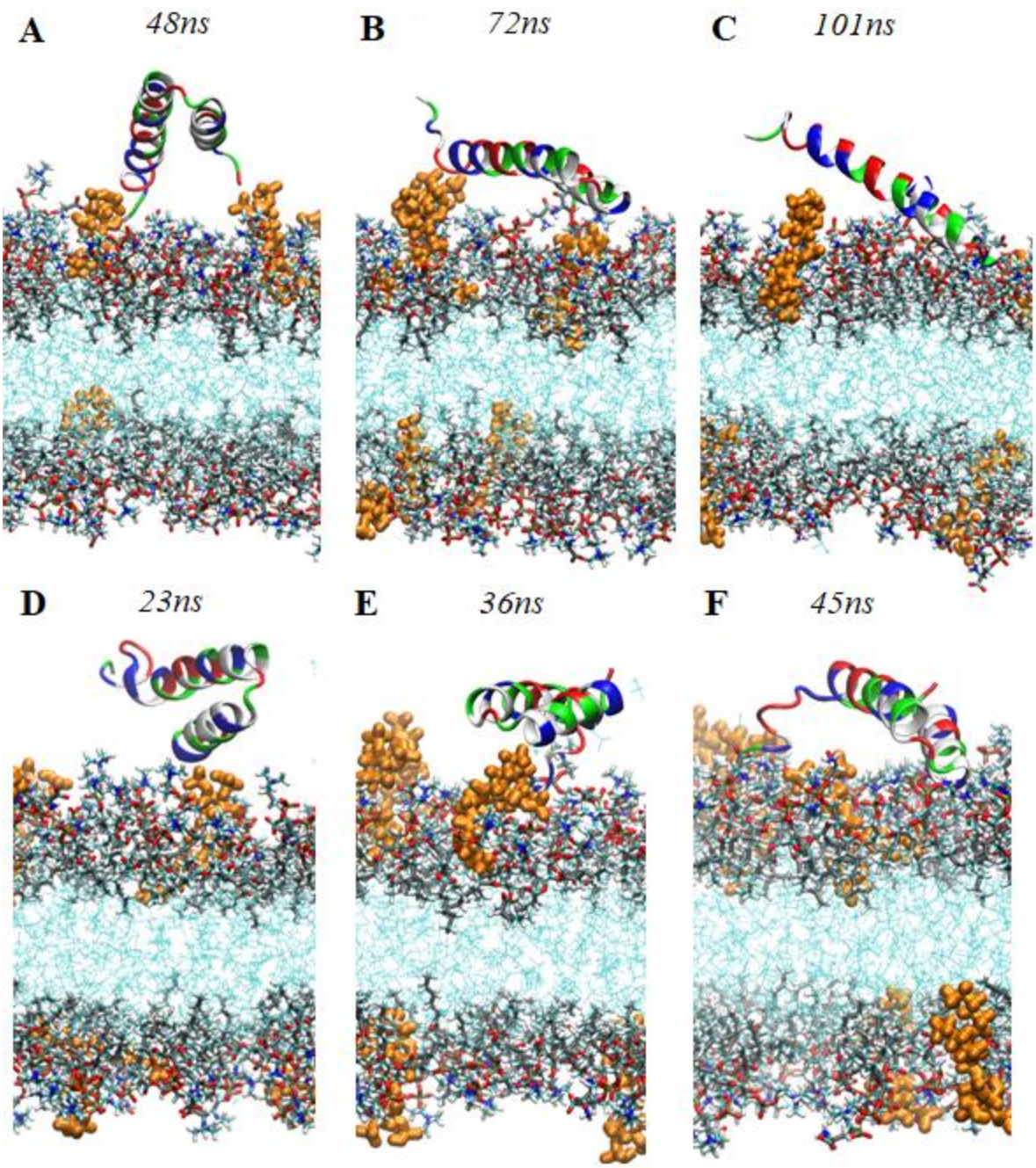


Figure 3.8: In these images, the darker-colored components represent the lipid heads; the light blue is DCLE, and the orange is PIP lipids. The sequence of binding: (**A-C**), simulation H2a displays the result of PIP2 lipids recruiting the α_7 helix before binding and with positively charged residues on the α_7 and α_6 helices. (**D-F**), simulation V1a displays a PIP2 lipid on the α_6 helix that was stabilizing K325 and is transitioning to the α_7 helix.

Although the PIP2 lipid trials aimed to sample as many different orientations as possible that had peptide-PIP2 lipid interaction, only four stable states were observed with interactions close enough for PIP2 lipid interaction. These observed states with one replica each were run for an additional 150 ns as full-tailed lipids and are shown in Figures A29 and A30.

Structure of Peptide and its Placement in Full-Membrane

The placement of the peptide in the membrane is much different in membranes with PIP2 lipids, as these interact with the terminal ends of the peptide (Figure 8A-C). The interactions with the terminal ends of the peptide are illustrated in Figure A31, where the deepest residues are around I324-K325, and the higher residue numbering steadily increases in distance from the membrane. Not all simulations have PIP2 lipid interactions. Five of the simulations have interaction on both helices, H1a PIP2, V1a PIP2, V2a PIP2, V1b PIP2, and V2b PIP2. One of these structure, V1a PIP2, unfolded at the point of the possible stabilizing salt bridge interactions. The placement of the peptide in the membrane can be described by the hydrophobic residues on the α_6 helix that imbed deep into the membrane with the assistance of essential electrostatic residues (see the mutation section below) that guide the process. The placement of the peptide can also be described by the many positively charged residues that are placed above the hydrophobic/hydrophilic interface and have greater selectivity for interacting with the POPS and PIP2 head group.

The α_7 helix differs from α_6 because the position is higher, and some amino acid residues point down and lie above the membrane (R344 and R347). During the initial binding phase, K336, K325, and R314 are essential to guide the peptide into place and remain significant contributors after binding to keep the peptide on the membrane. Also, residues K315 and K334 orient upward and tilt to interact with the PIP2 and POPS lipids. In general, the α_6 helix was deeper than the α_7 helix, and the connection region of the peptide between the α_6 to α_7 helix is the deepest. The

placement of these residues can be visualized in Figure A31-A34. Interestingly, residue K315 showed some variation in its placement. Different views and snapshots of the final states of the four full-length simulations are shown in Figures A29, A30, and A35. The RMSD for HMMM is shown in Figures A49 and A50.

In terms of *hypothesis 3,* the peptide-bound orientations were measured again to see if the PIP2 lipids had any impact on the orientation or, more specifically, the angle of approach. The angle of approach does appear to increase in the trials that interacted with the PIP2 lipids. This result is likely caused by PIP2 lipids that have phosphate groups elevated above the main membrane. The phosphates can interact with the peptide, bringing R344 and R347 higher up on the α_7 side. Initial interactions with R344 and R347 could alter the angle of approach of the bound peptide and the orientation. The validity of *hypothesis* 3 could not be proven as there was no difference in orientation between peptide bound on PIP2 and non-PIP2 membranes.

PIP2 Specific Interactions and Importance to Salt Bridges within Peptide

While testing *hypothesis 4* to determine the impact of PIP2 interactions, it was discovered that the PIP2 can compete for the same residues that create salt bridges between the two helices. An exciting discovery of the PIP2 simulations was the frequency and strength of the interactions residue K315 developed between PIP2 and POPS compared to other residues. However, the frequency of this interaction was not found to be significant ($p=0.19$) based on a T-test assuming unequal variance and a hypothesized mean difference of 0 for K315 interacting with anionic lipids. This higher p-value is believed to be due to the heavy dependence on lipid sampling of the K315 and the fact that the PIP2 lipids usually interact with R344 and R347. These were observed to occur in a peptide state similar to that of the middle of the bottom panel of Figure 3.7, with the side chains tilted even more out to the sides. More sampling may be needed to fully explore lipid

interaction with the peptide due to slow PIP2 and POPS diffusion, but this was beyond the general scope of the project.

Of the full-length lipid simulations, H1a PIP2 had many interactions closer than 4 Å over the entire full-length lipid simulation. Specifically, 43% of the time between residues K315-D319 and 33% of the time between K334-E338. The HMMM salt bridge data can be seen in Table A5, and the full-length data can be seen in Table A4. The position versus time for all amino acids in the simulations can be seen in Figures A36-A38. Figures A39 and A40 feature further conformational changes and K315 interacting with the POPS or PIP2 lipids. Figure A41 shows the specific conformational change on residue K315, which only interacts with the PIP2 lipid. Figure A40 shows simulation V1b before the conformational change, and Figure A40 shows it after a conformational change with lipids to stabilize it. Finally, Figures A42 and A43 show the conformational change on residue K334, which was much less frequently observed and poorly sampled.

3.3.3 Residues Critical for Binding

While *hypothesis 1* asserts that specific residues play a crucial role in binding, differences are caused by parameters such as the interaction energy, hydrogen bonding, and salt bridges between the peptide and membrane. Some of these parameters are investigated in this section. The energy per residue (Figure A44) indicates that the residues with the lowest and most attractive energy are at the ends of the peptide and that the residues in the center, such as Y318-T342, have higher energy. Figure A45 shows an unexpectedly high level of hydrogen bonding among R344 and R347 with the membrane because these residues only hover above the membrane surface and are not tightly bound to the membrane. Additionally, there are similarly high levels of salt bridges

for the residues R344 and R347 in Figure A46. The high number of salt bridges indicates that these residues are important but not as essential as the other residues.

Additionally, there were low levels of hydrogen bonds in the center of the peptide (V320-E340) and higher levels of hydrogen bonds among the remaining residues. K325 and K336 are an interesting exception to the observed trends of low hydrogen bonding in the center. These charged lysine residues are well-oriented to interact with the membrane. The salt bridges between the membrane and the peptide indicate that the most frequent bridges occur with R314. Figures A47 and A48 for the PIP2 simulations record the position of the deepest residue versus time.

3.3.4 Effect of Mutating the Key Binding Residues

Mutations were performed to verify the effect of the critical residues identified while researching hypothesis 1. Two of the key binding residues (K325 and K336) were mutated to confirm their role in the binding process. In the wild-type simulations, these residues were pulled down underneath the membrane to the phosphate region, pulling some of the hydrophobic residues with them. Simulations with the K325D/K336D mutation caused the peptide not to bind to any extent onto the membrane, thus proving their key role in the binding process. The effects of the mutations appear in the supporting materials, and the image of the final state of the mutated peptide is in Figure A51.

3.3.5 Effect of Changing Force Field on Amino Acid-Lipid Interactions

The force field was changed by adjusting pair-specific corrections to non-bonded interactions between the amine group of positively-charged amino acids and lipids (phosphate or carboxylate) known as CUFIX.[113] This change resulted in a drastic reduction in intermolecular interactions based on simulations that started from the end of the full-length simulations and ran for an additional 150 ns with CUFIX. The side-by-side comparison for hydrogen bonds can be

found in Figures A16 and A45, while the salt bridges are shown in Figures A17 and A46. Another factor is the interaction energy per residue, which is higher because of weaker electrostatic interactions. The CUFIX force field provides more of a balance between electrostatic and hydrophobic interactions.

Table 3.1. Demonstrates that there is a significant decrease in the total interaction energy when switching to the CUFIX force field. The most stable binding was H1a PIP2 and V1b PIP2, where the PIP2 lipids are located on either side; the most stable of the non-PIP2 simulations is H2a, which has a comparable free energy.

	Interaction Energy (kcal/mol)	**ΔG (kcal/mol)**
H1a	-287.0 ± 38.8	-19.2 ± 0.03
H2a	-358.7 ± 44.7	-34.0 ± 0.03
H1b	-351.7 ± 50.3	-25.2 ± 0.03
H2b	-288.8 ± 46.3	-23.5 ± 0.03
H1a PIP2	-425.0 ± 33.0	-36.2 ± 0.03
V1b PIP2	-405.4 ± 51.0	-31.3 ± 0.03
V2a PIP2	-331.7 ± 33.1	
V2b PIP2	-374.1 ± 53.1	
H1a CUFIX	-155.3 ± 36.4	
H2a CUFIX	-221.4 ± 40.9	
H1b CUFIX	-210.3 ± 36.2	
H2b CUFIX	-151.6 ± 39.3	
H1a PIP2 CUFIX	-206.8 ± 56.4	
V2a PIP2 CUFIX	-171.7 ± 33.9	
V1b PIP2 CUFIX	-210.1 ± 32.8	
V2b PIP2 CUFIX	-225.6 ± 40.9	

The ΔG values, when analyzed with a Boltzmann average, indicate that the top non-PIP2 simulation and the top PIP2 simulations were roughly equal. There were two separate PIP PMF calculations, but that is not enough sampling to compare with the non-PIP. The possible explanation is that many more viable non-PIP2 simulations were sampled, so a lower energy configuration was found while the PIP2 simulations were not sampled as extensively. Since the less sampled PIP2 simulations were roughly equal to the higher sampled non-PIP2 simulation, it

makes sense that the PIP2 lipids are assisting in the binding pose. Overall, these results indicate that the PIP2 is much closer in free energy compared to the non-PIP2 than would be expected from the total interaction energy alone.

An essential aspect of the PIP2 bound simulations is that the interaction energy was lower in the trials where the peptide bound directly on top of the PIP2 lipids. The interaction with the PIP2 lipids occurred in trials V2a PIP2 and V2b PIP2, indicating that there is an energetic penalty if the peptide binds on top of the PIP2 lipids because the peptide is not as deeply bound. It is important to note here that V2b PIP2 CUFIX was not on top of lipids. The lipids moved out from under it. When the membrane is too highly enriched in PIP2 lipids, the peptide does not bind as deeply, however, at a more physiological concentration of PIP2, the peptide binds deeper with PIP2 on the sides.

The results of the total binding energy indicate that the standard force field has an average value of -352.8±49.7 kcal/mol, and the CUFIX has a value of -194.1±29.8 kcal/mol, and no correlation between the angle of approach. This data shows that the PIP2 simulations clearly had lower energy than the non-PIP2 simulations. Specifically, the interaction energy for the non-PIP systems was -321.6 kcal/mol and -384.1 kcal/mol for those with PIP. However, with CUFIX, the difference is smaller at -184.7 kcal/mol for non-PIP2 and -203.6 kcal/mol for PIP. There was a weak statistical significance in the PIP2 vs. non-PIP2 membranes in the original force field ($p=0.068$) based on a T-test assuming unequal variance and a hypothesized difference of 0, but almost no statistical difference using the same T-test in the CUFIX ($p=0.42$). The standard force field predicted a larger difference, which indicates the PIP2 played a larger role and supports the overlapping roles of Osh proteins.[40] The PIP2 lipid membrane is capable of the lowest energy, but the peptide can achieve a similar binding energy on the non-PIP2 membrane.

On the other hand, there was evidence that certain amino acid residues were selective for PIP2 and POPS lipids and were out of reach of the main lipid POPC. These residues include K315, K334, R344, and R347. Of these, the sum of the interaction energy of residues R344 and R347 is statistically higher ($p=0.018$) using a T-test assuming unequal variance and a hypothesized difference of 0 under the standard force field conditions. The T-test had an even higher significance using the same statistical test under the CUFIX conditions ($p<10^{-5}$) with PIP2 lipids compared to non-PIP2 simulations with the same force field. The other two amino acids were not sampled as frequently since most of the PIP2 lipids were interacting in different areas of the peptide. Neither of these was statistically significant in the CUFIX force field, but they still selectively interacted with the POPS and PIP2 head group over the O13 and O14 atoms that are deeper in the membrane. It is interesting that the only exceptions to the high selectivity are the amino acids that were mutated and found to prevent any level of binding, which makes sense that these residue salts bridge with the deeper oxygen atoms. The connection between the amino acids, their placements, and the effect of mutations is further evidence to support *hypothesis 1*.

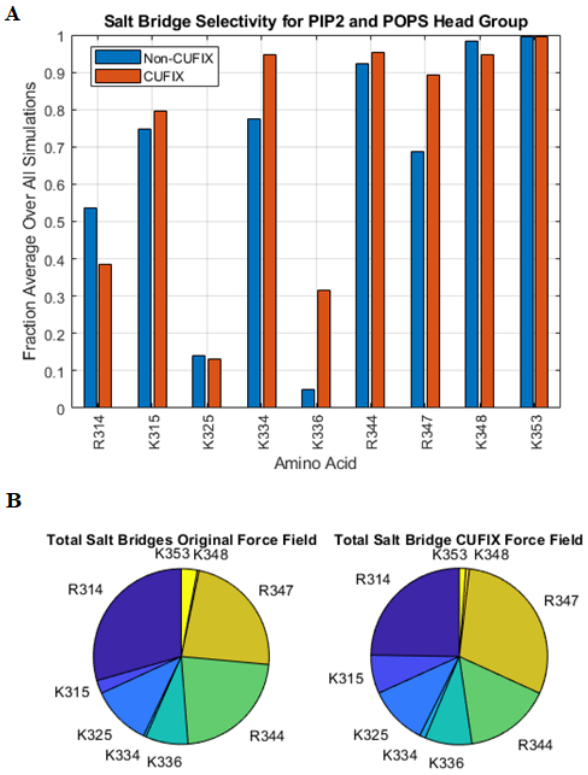


Figure 3.9) A. The selectivity of the different amino acids for the PIP2 and POPS head group, i.e., the fraction that is not interacting with the O13 and O14 (Figure 3.3) that are deeper in the bilayer. The results show that all of these amino acid residues are selective except for R314, K325, and K336, which are believed to play a role in guiding the insertion of the hydrophobic core of the peptide, which is consistent with the low selectivity, indicating more frequent salt bridges with the deeper oxygen atoms. **B.** Pie chart of total salt bridges. Note that while residues K348 and K353 have high selectivity, they also have fewer overall salt bridges compared to other cationic amino acids.

The relative selectivity of the different positively charged amino acid residues are shown in Figure 3.9 and demonstrates that all amino acids have selectivity above 0.5 for both the original force field and CUFIX force field except R314, K325, and K336. The exception of R314, K325, and K336 is interesting because these amino acids were identified as being critical to the binding process. Although residues K325 and K336 are the most selective for the O13 and O14 (deeper charged oxygens), residue R314 has a higher overall number of salt bridges (Figure 3.9B). The pie charts suggest that although the force field affects the total number of interactions, it does not change the overall fractions that each amino acid contributes.

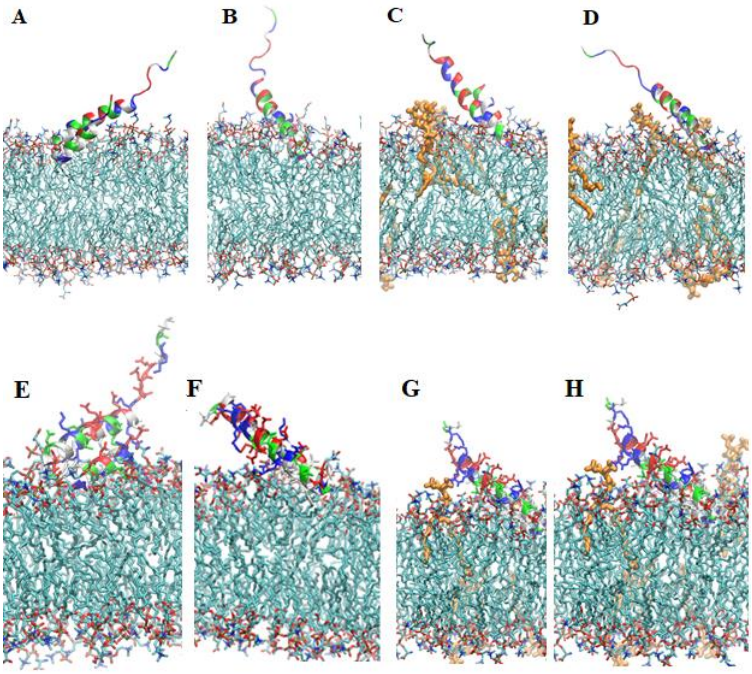


Figure 3.10) Images of the final bound states of the Osh α_6-α_7 helices **A** is H1a, **B** is H2a, **C** H1a PIP, and **D** is V2a PIP, **E** H1b, **F** H2b, **G** V1b, and **H** V2b.

Salt bridge interactions were found to promote the stability of the tertiary structure of Osh4's α_6-α_7 peptide. Figure 3.10 shows the final bound states of the peptide, where the PIP2 lipids interact with the peptide, and where the peptide contacts the membrane. There were salt bridge interactions between residues D319-K315, K315-E338, and E338-K334, as shown in Figures A19-A21 and A36-A38. Although the salt bridges were not always present, they were still observed to exist for some time and, therefore, play some role in the stability of the peptide. In the case of H1a, the CUFIX simulations relaxed the peptide-peptide interactions, increasing the number of conformations that residues K315 and K334 could adopt. For example, the peptide-peptide interactions existed in the original force field, but as soon as the force field was changed to CUFIX, the salt bridge spanning the two helices was broken. In other cases, CUFIX may have helped to speed up the diffusion of lipids that were already interacting with the peptide.

3.4 Conclusions

Of the four *hypotheses* that were probed in this work, there were two that were confirmed to be true. *Hypothesis 1.* Specific residues play a vital role in the binding process; this was established initially by observation, then verified by our peptide-membrane interaction analysis, and later by mutating two of the three proposed critical residues to find that the peptide had

severely reduced binding capacity. *Hypothesis 2.* The mechanism of peptide binding is related to the angle of the helical approach; this was proven to be invalid since there was a wide range of angles for both PIP2 and non-PIP2 trials. *Hypothesis 3.* Peptide-bound orientation differs in membranes with and without PIP2; there was data supporting the hypothesis that the orientation of the peptide was affected by the presence of PIP2 lipids because the angle of approach was altered. *Hypothesis 4.* Salt-bridge interactions between the two helices stabilize the membrane binding structure; this was found to be partially valid, at least in the original C36/C36m force field, and there are indications of the same for the CUFIX force field as well.

Understanding the way in which peptides bind is an essential first step in assessing their overall mechanism of action. Osh4's α_6-α_7 domain was able to bind to membranes with three critical amino acids, R314, K325, and K336, that yielded strong membrane binding. This binding caused insertion of much of the hydrophobic amino acids beneath the lipid head salt bridges. This location between the two helices stabilized the membrane binding sections K315-E338, as well as other salt bridges on the same helix K315-D319 and K334-E338. However, some of these salt bridge pairs destabilized in some simulations, but only one of the 20 total simulations showed a drastic unfolding.

From past experimental work, it is known that Osh4 has a preference for PIP2 lipids and that the protein will preferentially transport PI(4)P over other lipid types and the PI(4,5)P lipid accelerated transport of sterol.[160] Our simulation results demonstrate that signaling PIP2 lipids offers additional stability to the Osh α_6-α_7 domain. The concept that PIP2 lipids stabilize peptide binding supports the idea that PIP2 lipids signal lipid transport by stabilizing this binding domain. Specifically, our results indicate that K315, K334, R344, and R347 are critical residues for interactions with the PIP2 and POPS lipids but not essential for function. In addition, our double

mutation of the wildtype to K325D/K336D showed that the peptide had no binding with the removal of the positively charged lysine. Removing the lysine confirms that these amino acid residues are positioned to interact specifically with the deeper oxygen atoms that are shared among all the simulated lipids. There is experimental evidence indicating that PIP2 lipids speed up the lipid transfer process, and it was hypothesized that the increase in speed was caused by lengthening the dwell time on the liposome surface.[54, 159] However, our interaction energy results indicate that the peptide can be stable in membranes without PIP2 lipids and harness similar binding energies compared to PIP2 membranes. The ability of the peptide to bind to membranes with and without PIP2 is in line with the idea that the Osh proteins have overlapping roles since the α_6-α_7 peptide was capable of similar binding energy and structure with and without any PIP2 lipids. Specific details on the free energy of binding in varied lipid environments will require enhanced sampling and replicas due to slow lipid diffusion and were beyond the scope of this work.

The current work shows that in addition to the ALPS motif, the α_6-α_7 helices are another membrane binding motif of Osh4.[34, 47] This work will guide new simulations of the full-length Osh4 that membrane binding will undoubtedly be influenced by the b-crease,[46] but we will be probing how the α_6-α_7 helices simultaneously bind to membranes with the b-crease that might occur during binding events that consist of two opposing membranes. Future work will be done to simulate other parts of the Osh4 protein, such as the ALPS section to learn more about how it can sense lipid packing.

Acknowledgements:

All simulations ran on the Extreme Science and Engineering Discovery Environment (XSEDE) supercomputer by grant number MCB-100139. Data analysis was performed on the

University of Maryland's Deeptought2, which is maintained by the Division of Information Technology. This research is supported by the NSF grant MCB-1951425.

Supporting Information:

Appendix A has several different kinds of information, including data figures, images of simulations, and tables, and they are organized by non-PIP2 and full length first, PIP2 second, and the mutations last. There are two tables, one showing the distance between the main phosphate and the peptide center of mass for all the simulation orientations. The second table shows the details of the peptide binding, including the time of binding, minimum position or maximum penetration, and the center of mass relative to the bilayer phosphate. Then there are several data figures: A1-A2 shows the deepest amino acid residue COM over time, A3-A4 shows the RMSD of the simulations, and A5-A10 shows the different angles measured over time. Figures A11-A12 show the final bound states of the HMMM simulations, and A13-A14 shows the RMSD of the full-length simulation. Figures A15-A17 show the interaction energy per residue, hydrogen bonds per residue, and salt bridges per residue. Figure A18 shows the accessibility of residue K315 to hydrogen bonding with POPS lipids under the correct conditions, and figures A19-A21 show the distance between the different residues that form peptide-peptide salt bridges. Table A3-A4 shows the exact frequency of salt bridge formation as a fraction of 1, which is 100% of the time. Figures A22-A25 show the placement graphs that are useful for seeing the position of all the amino acids over time and show that there is some variation in the residue K315. Figure A26 shows snapshots of the peptide binding with PIP2 lipids around it, and A27 and A28 show the final bound HMMM states of the PIP2 trials. Figures A29 and A30 show the final bound states of the PIP2 simulations and Figures A31-A34 show the placement graphs of PIP2 simulations with an additional image of the final bound states in Figure A35. Figures A36-A38 show the salt bridges for the PIP2 simulations,

with Table A5 showing the salt bridge frequency as a fraction of 1. Figures A39-A43 show the different conformations that appear to be stabilized by PIP2 and POPS salt bridges interrupting the peptide-peptide salt bridges. Figures A44-A46 show the interaction energy per residue, hydrogen bonding per residue, and the salt bridges per residue. Figures A47-A48 show the position vs. time of the COM of the deepest amino acid residue, while Figures A49-A50 show the RMSD of all the PIP2 simulations. The final image, Figure A51, shows the mutated peptide never bound.

Chapter 4: ALPS Domain of Osh4 and Osh5

4.1 Introduction

The building blocks of cells are proteins and lipids. The Osh proteins are yeast lipid transporters at specific membrane organelle junctions. Living cells need to reorganize the lipids in their membranes to function correctly. Lipid packing sensors contribute to this process by targeting Osh lipid transporters to the specific membrane organelles in need of transport. Osh proteins fulfill the role of selective transport of hydrophobic oxysterols and phospholipids such as PIP2. Thermodynamics and diffusion drive the transport without the use of ATP. The lack of ATP makes this mechanism one of the most fundamental processes for maintaining membrane compositions. Each Osh protein targets a different pair of organelles to transport lipids efficiently. However, there is some overlap among the Osh proteins so that knockouts of multiple proteins are not fatal. The overlap is done, in part, by exploiting differences in the membrane packing.[166, 167] The Osh4 protein favors the endoplasmic reticulum (ER) and trans-Golgi network (TGN). The Osh5 protein favors the part of the ER that borders the plasma membrane. The Amphipathic Lipid Packing Sensor (ALPS) is a domain on various proteins, including the Osh4 protein from the yeast (Figure 4.1), similar to mammalian lipid transporters. Other motifs, like the Osh ALPS, assist in binding the proteins to the membrane, including the α_6-α_7 domain, the Phe-loop, the beta-crease, and the distal site.[46, 168]

Two main driving forces, kinetic and thermodynamic, could bind ALPS to specific locations. The thermodynamic explanation is that the peptide has a thermodynamic affinity for one membrane over another due to the lipid packing or curvature. The kinetic description is that each membrane is equally favored, but the kinetics dictate looser membrane packing needed for insertion. The thermodynamic explanation is more consistent because rates overall remain

constant, and kinetically relevant knockouts would seem to imbalance the mass transfer. Whatever the answer, the lipid transport is robust to any single protein failure. Statistically significant increases in the steady-state ergosterol levels resulted from the Osh5 and Osh6 knockouts.[169]

ALPS is critical to how the Osh4 protein senses and binds to the desired membrane organelle. The ALPS peptide's characteristics include amphipathic, low charge, and many serine and threonine amino acids.[58, 160, 170, 171] The goal of studying the Osh proteins is to understand the lipid transport function of these proteins that exchange lipids between organelles. However, there is limited understanding of how Osh proteins exchange lipids between organelles. Current wet lab techniques enable the quantification of the overall organelle lipid concentration and exchange rates. Still, there are limited details of the spatial distribution and how the concentrations change close to other organelles.

The Osh4 and Osh5 have a point mutation in the first 20 amino acids that could be studied to understand why Osh4 prefers to bind to the specific ER region nearest the TGN. The goal is to collaborate with wet lab experimentalists to understand the physical processes better and improve the force fields to more accurately model physiological properties.

Osh4 ALPS:	MSQYASSSSWTSFLKSIASF
Osh5 ALPS:	MSQ**H**ASSSSWTSFLKSI**S**SF**NGDLS**

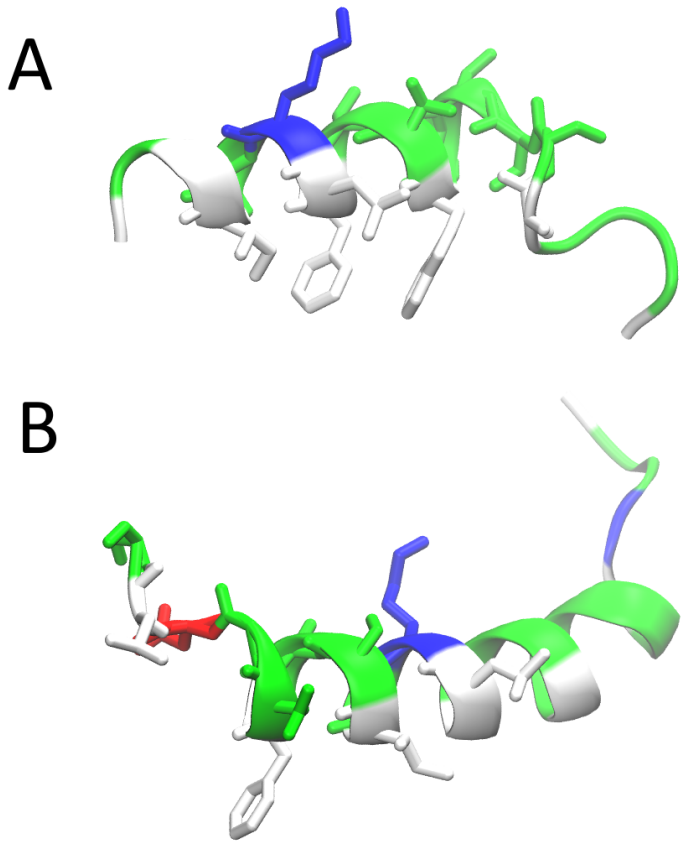


Figure 4.1) Panel A shows the Osh4 peptide structure and panel B shows the Osh5 peptide structure. The region of similarity is depicted with licorice sticks to show the amino acids.

This chapter examines the detailed membrane interactions between the Osh4 and Osh5 ALPS peptides. The chapter begins with HMMM simulations, which convert to full-length membranes. The time series demonstrates the transient fluctuations of the peptide on the membrane bilayer. Snapshots of the Osh5 reveal the different Osh peptides' varied structure and penetration depths. The protein folding landscape for Osh4 and Osh5 is also studied to show the effect of some mutations and different ion concentrations. Finally, a free energy analysis shows how differences in lipid packing affect the binding free energy.

4.2 Methods

The same equilibration process from the last section, Osh α_6- α_7 (Section 3.2), is used for these protein structures. The significant difference is in the membrane type studied 80:20mol% (SDPC-POPS, POPC-POPS, PhPC-POPS, DMPC-POPS). The chemical structures of these lipids are shown below in Figure 4.2.

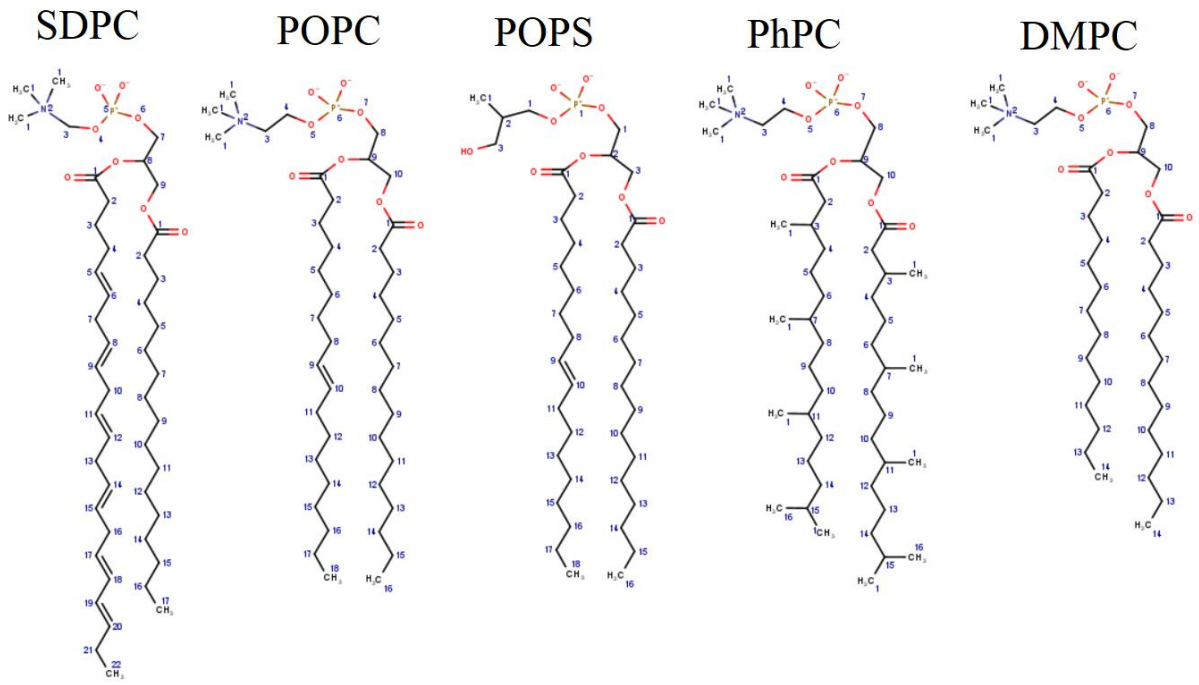


Figure 4.2) Shows the types of lipids studied in this chapter. Holding the lipid head constant was essential, we employed a fixed ratio of PC-PS, and only the lipid tails varied. The DMPC is the tightest packing lipid due to its shortened chain length with a packing roughly equal to POPE. SDPC and PhPC are the more densely packed relative to POPC, which is a good reference in the middle.

The highly mobile membrane mimetic (HMMM) builder feature within the *Membrane Builder* of CHARMM-GUI (www.charmm-gui.org) built the model membranes studied in this manuscript.[48, 144-149] These membrane models had around 50-60 molecules of water per lipid depending on the area per lipid and used the TIP3P model[143] for water and added neutralizing

potassium ions to the mixture. The HMMM builder feature accelerates the simulation with a simplified approach by modeling only the individual lipid head groups and replacing most lipid tails with dichloroethane (DCLE) solvent without the extended lipid tails, increasing lipid diffusion.[145] The HMMM model accelerates the diffusion of the lipids by an order of magnitude without affecting the binding mechanism, thereby saving computational resources without losing details of the binding process.[34, 145] Each model was equilibrated using the standard six-step CHARMM-GUI[144, 147] protocol for 225 ps. The peptide coordinates range from 1-20 for Osh4 and 1-25 for Osh5. The peptide structures studied in this work formed a subset extracted from the full protein (PDBID: 1zhy40 for Osh4 and AlphaFold prediction for Osh5 ALPS). All peptides employed the NNEU (neutral) and CT2 (acetylation) patch to prevent aggregation. The NNEU and CT2 capping methods are best compared to experiments that use the same patches, as was done with our work. Next, CHARMM-GUI defines an aqueous phase around the peptide in different initial system seeds, each with different orientations and distances from the membrane. CHARMM-GUI was used to set up the Nanoscale Molecular Dynamics (NAMD)[142, 163] simulations. The process involved minimizing the system and then simulating it with decreasing levels of restraints. The lipid area scaling factor was varied to achieve box sizes of 74 $Å^2$/lipid other than pHPC, which was not run in HMMM but instead as a full-length membrane from the beginning, which had an 89 $Å^2$/lipid. The system was defined with CHARMM-GUI using a terminal acyl carbon number of 6 for all simulations.[145] These settings help accelerate the binding process by exposing the lipid tails to the solvent and allowing for increased packing defects in the membrane.[145] The equilibration and production runs ran with a 2 fs time step. Other settings for the equilibration include a restart frequency of 500 steps, a trajectory, and an extended system output frequency of 1,000 steps. The NAMD program saved data with a frequency of 5,000 steps

for the production runs. These initial equilibration and production runs used NAMD on a constant Number, Pressure, Area, and Temperature (NPAT) ensemble for approximately 150-200 ns of total simulation time.[142, 163] All simulations used the SHAKE algorithm to constrain hydrogen atoms.[135] All of these simulations were extended longer using full-length membranes. The CHARMM36 (C36) lipid force field[132] was used for all simulations since it has the most updated parameters for the PIP lipids.[164] The C36m additive all atom protein force field[18, 131] and the TIP3P water model were used.[143]

A Lennard-Jones force-switching function over 10 to 12 Å defined the electrostatic and van der Waals forces.[140] The temperature was kept constant at 298 K to ensure a fluid phase membrane using the Langevin thermostat and Langevin dynamics.[129] Periodic boundary conditions (PBC) defined the boundaries of all simulations as the mirror image of the system from the opposite sides. The PBC evaluated long-range electrostatic interactions using the Particle Mesh Ewald (PME).[139]

The conversion to full-length required selecting four HMMM simulations chosen from each membrane model and input into the CHARMM-GUI *HMMM* to full-lipid converter. After conversion, the system inputs were simulated again with NAMD for an additional 150 ns. The temperature was constant for the all-atom systems as in the HMMM simulations. Still, the box size varied semi-isotropically (X=Y but not Z) while keeping the pressure at a value of 1.01325 bar using a Langevin piston.[22, 115, 116] This allowed the membrane to more closely resemble natural membrane lipid packing.

Calculating a potential of mean force (PMFs) required pulling the peptide off the membrane using an external biasing force. The external bias was applied sequentially, progressing over 1 Å increments to increase the separation distance. A repeated cycle of pulling the peptide by

1 Å and relaxing for 0.1 ns, continued until the peptide completely separated from the membrane. Harmonic restraints prevented the membrane from stretching or bending, creating a clean separation of the peptide from the membrane. There was also a RMSD restraint on the peptide to prevent unfolding and maintain a steady, controlled separation distance.

To analyze the peptide binding to the membranes, a record generated by visual molecular dynamics (VMD)[141] tracked the vertical positions of all the peptide residues around the binding region and the average phosphate position over every frame. The VMD data generated the minimum penetration over time graphs. Renderings of the membrane figures also employed VMD with Tachyon internal memory processing.[165] The interaction energy (calculated with the CHARMM program) per residue is the sum of the van der Waals and the electrostatic contributions.

2Struc was used to calculate the secondary structure of the peptides in the PMF. The 2Struc[172] webserver accepted an input PDB that could be uploaded to analyze the peptide secondary structure. The secondary structure was held constant throughout the PMF using an RMSD colvar. The constant RMSD colvar meant that the input PDB structure was equivalent to the structure in the PMF.

4.3 Results

All simulations of the ALPS peptides with membranes consisted of three replicas. The chapter focuses on the final full-length simulations with the lowest free energy for each membrane type shown in the main document (not Appendix B), along with the replica (r#). Images of the other models appear in Appendix B. There were a variety of carbon alpha penetration depths for HMMM of Osh4 in Figures B1 and B2 and the full length in Figures B3 and B4. The specific time series position of M1 is in Figures B5 and B6. The time series position of F13 appears in Figures

B7 and B8. The particular time series position of the L14 appears in Figures B9 and B10. Figure B11 shows the time series of π–π stacking between W10-F13 to demonstrate that this is favored. Snapshots of the π–π stacking are presented in Figure B12. An alternative free energy landscape from the one presented in the main text appears in Figure B13 with varied ion concentrations and the different mutations F13S and K15L. The PMF free energy profiles of the standard trials are shown in Figure B14 and for the mutants in Figure B15. The HMMM for the Osh5 carbon alpha penetration depth appears in Figure B16, and the full-length carbon alpha appears in B17. The M1 position of the Osh5 appears in Figure B18, the F13 position in Figure B19, and L14 in Figure B20. Finally, the alternative free energy landscape to the one presented in the main text appears in Figure B21. The Osh5 PMF profiles appear in Figure B22. The temperature replica exchange history for the system of interest is shown in Figure B23. The results suggest that there is good variability between the HMMM and full-length replicas. The aqueous ALPS peptide folding free energy was computed with replica exchange to compare with CD data. The PMFs were used to compare binding free energy in the simulations with the wet lab data.

4.3.1 Osh4 ALPS

Osh4 HMMM

The HMMM structures of the Osh4 ALPS peptide that bound to the membrane are similar to those shown in Figure 4.1. The first five amino acids typically unfold or have a kink or bend separating them from the rest of the α-helix. Figures 4.3 A and D show clearly that this section can refold as an α-helix with a bend, while Figures 4.3 B and C show the unstructured folds of this peptide.

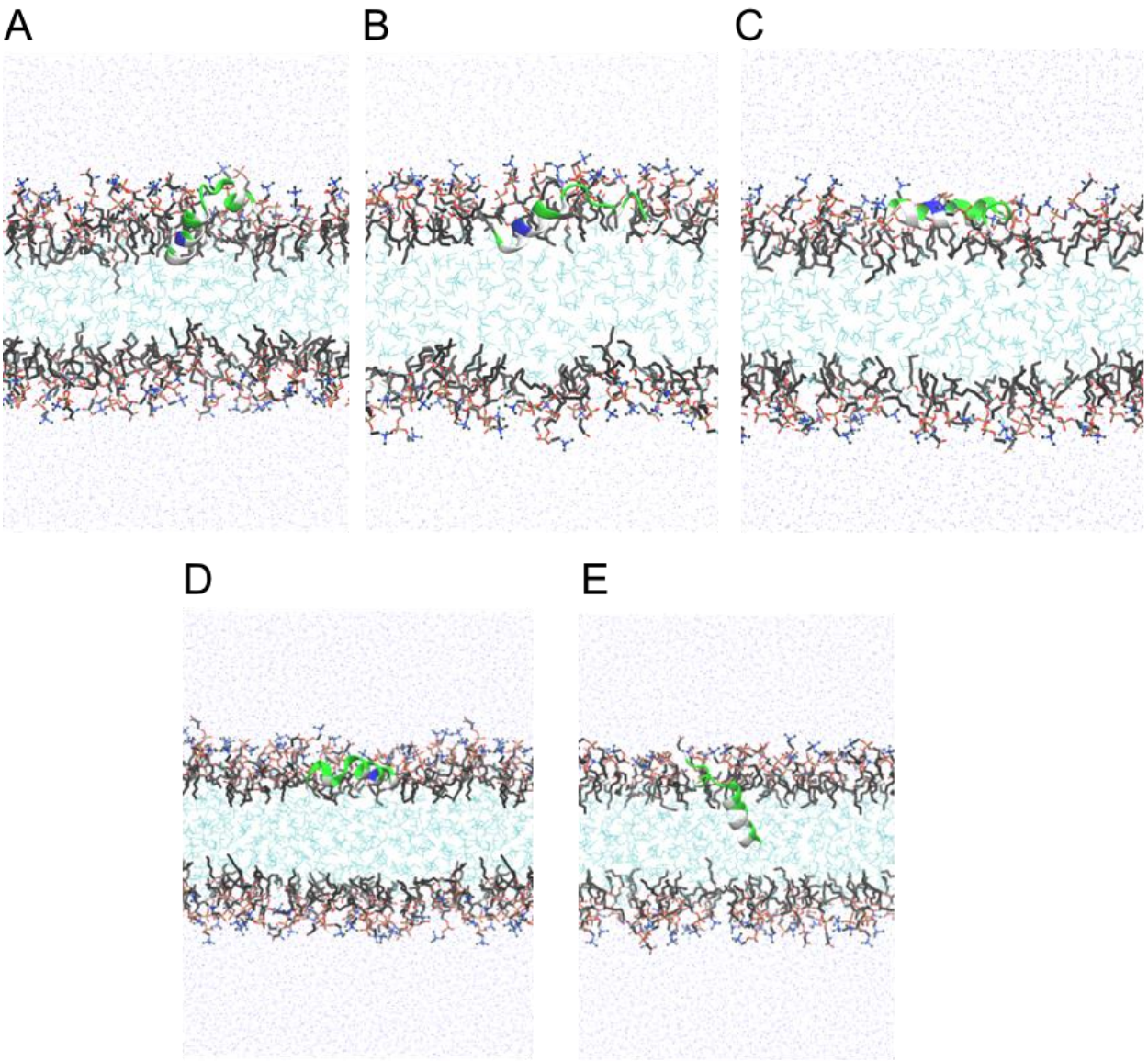


Figure 4.3) HMMM images of the peptide featuring the end snapshots of the HMMM equilibration. **A** is DMPC-POPS r1 (300 ns), **B** is SDPC-POPS r2 (200 ns), **C** is POPC-POPS r2 (150 ns), **D** is F13S PC-PS r2 (150 ns), **E** is K15L PC-PS r3 (150 ns)

The position versus time data shows the top "lowest free energy" peptides bound stochastically over 150-300 ns. The top simulations penetrated at least 5 Å beneath the phosphate

plane. However, only two of the five simulations approached an equilibrium position that maintained 5 Å below the phosphate. Panel C in Figure 4.4 (POPC 80% - POPS 20%) was interesting because it reached the -5 Å position and re-surfaced. The final placement before conversion was around 0 Å. All replicas for HMMM simulations appear in Figures B1-B2. As the simulations continued, the simulation in panel C of Figure 4.4 was re-arranged to become the lowest free energy simulation. Another different simulation is the K15L because it penetrates and inserts into the membrane as a transmembrane (TM) helix. The use of the HMMM protocol could have impacted the results. HMMM overcame the kinetic barriers while maintaining the peptides' final structure and placement.

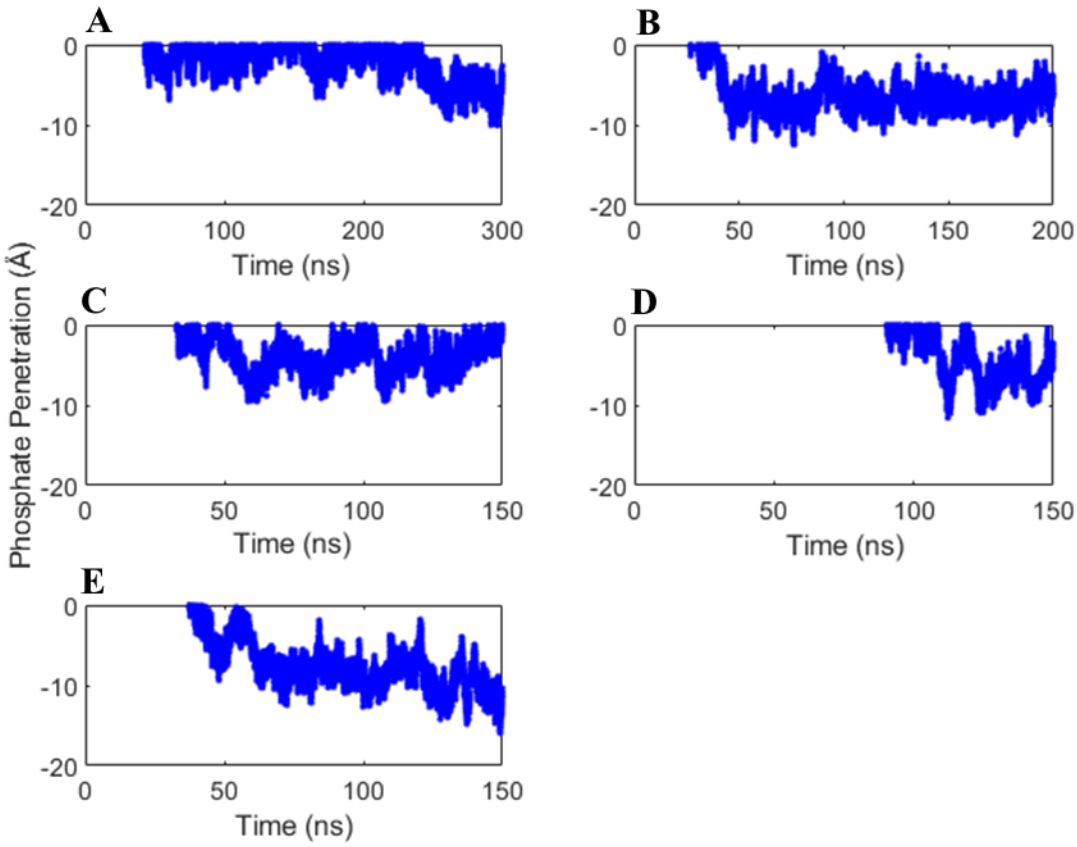


Figure 4.4) Shows the HMMM simulation's center of mass of the C_a of the peptide amino acids bound and can move deeper or re-surface. A) DMPC 80%-POPS 20% r1, B) SDPC 80% POPS 20% r2, C) POPC 80% POPS 20% r2, D) F13S POPC 80% POPS 20% r2, E) K15L POPC 80% POPS 20% r3. Y=0 is the phosphate plane defined as the average of the phosphorus atom. The simulation in panel E is a peptide fully inserted into the membrane. The K15L mutation caused the ALPS peptide to partition into the membrane core. The other peptides remained tethered to the surface due to the charge that anchored them to the head group.

Figure 4.4 is an excellent representation of the HMMM penetration beneath the membrane phosphate P over time. The pHPC model membrane was run as a full-length simulation because converting the HMMM pHPC models back to full-length models was impossible.

Osh4 Final Snapshots of the Full-Length Simulations

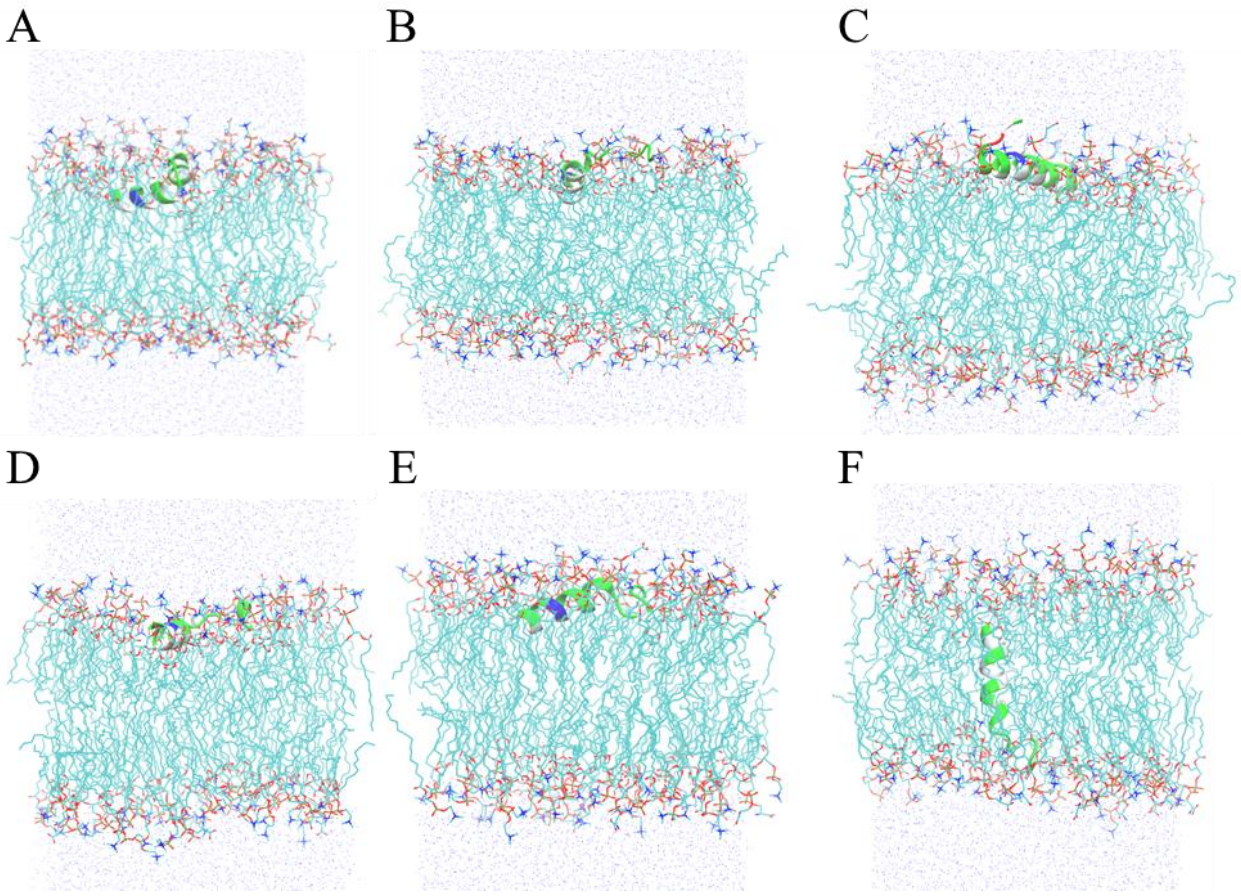


Figure 4.5) Shows the lowest free energy simulations from each membrane system. **A** is DMPC-POPS r1, **B** is PhPC-POPS r1, **C** is SDPC-POPS r2, **D** is POPC-POPS r2, **E** is F13S PC-PS r2, **F** is K15L PC-PS r3.

The final snapshots of the lowest free energy profiles for the all-atom runs are depicted above in Figure 4.5. The complete list of penetration distances appears in Figures B3-B4. Figures B5-B10 measure the same atoms as isotope labeling would for Phe, Met, and Leu. The snapshots shown above are those used to start the PMF profiles. Most of these peptides bind at the interface by burying the hydrophobic side in the lipid tails. It is also apparent from these graphs that some peptides penetrated to different levels beneath the phosphate layer, which suggests some peptides folded more than others. The one exception was the K15L simulation, which showed that the α-helix was fully inserted into the membrane and fully partitioned into the membrane as a

transmembrane helix. Figure 4.6 shows the penetration of the C_a of the peptide over time and indicates that the POPC-POPS mixture had the deepest penetration for the starting point of the PMF.

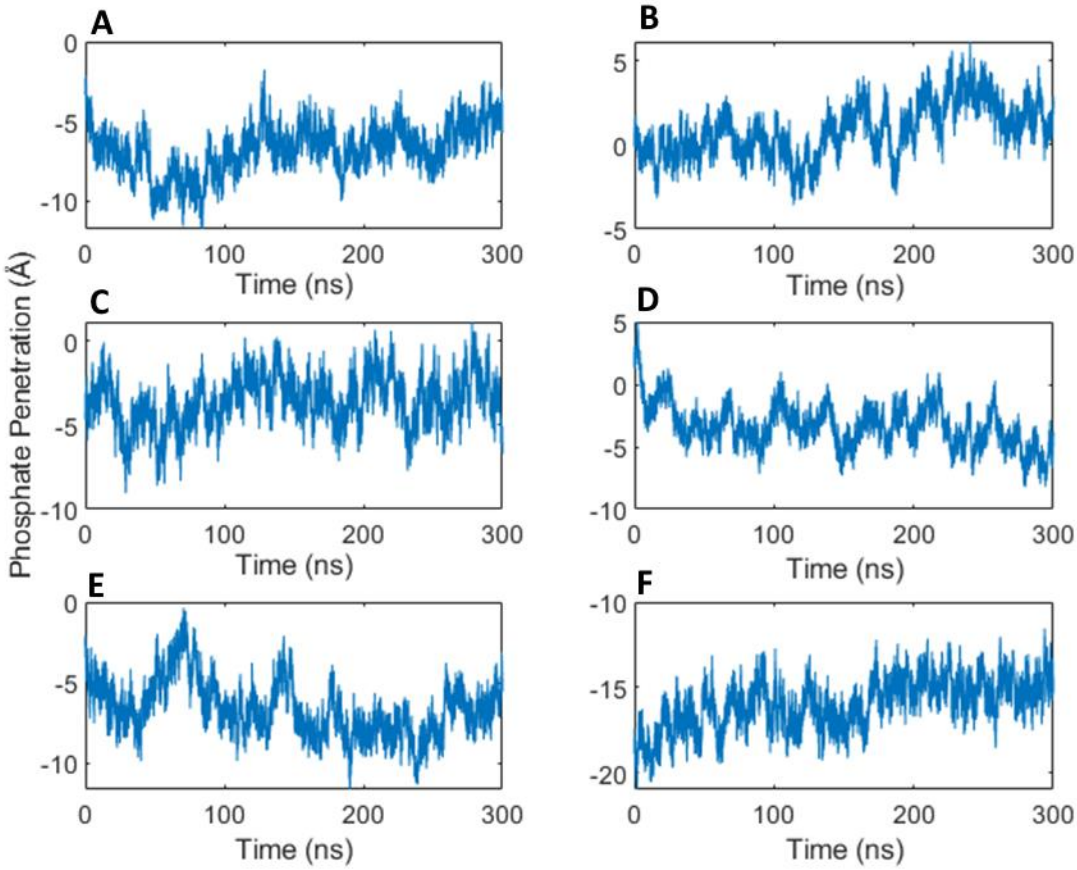


Figure 4.6). The center-of-mass C_a penetration distances beneath the membrane phosphate. **A** is DMPC-POPS r1, **B** is PhPC-POPS r1, **C** is SDPC-POPS r2, **D** is POPC-POPS r2, **E** is F13S PC-PS r2, **F** is K15L PC-PS r3.

The data from Figure 4.6 indicates that all top binding poses were stable around -5 Å other than pHPC and K15L. The K15L mutant essentially became a transmembrane helix. There was a lot of variation in the penetration depth among the simulations, but the lowest binding free energy simulations appear more consistent. Additional data includes the detailed breakdown for the Met 1, Phe 13, and Leu 14 amino acids shown in Figures B5-B10.

Table 4.1 displays two other parameters: helix and membrane area per lipid. Each parameter varied across the different simulations, although the penetration depths among the lowest free energy profiles were similar for the surface-bound poses.

Table 4.1). The binding free energy of the Osh4 ALPS peptide is bound to varied membranes. Those in red are the lowest free energy states per membrane system.

Membrane	Run	Osh4	ΔG (kcal/mol)	Boltzmann Average ΔG (kcal/mol)	Area Per Lipid (Å^2)	% Helix
80% DMPC / 20% POPS	1	WT	-29.9	-29.9	61.8+-0.3	75
	2	WT	-21.4		66.0+-0.3	55
	3	WT	-19.4		66.0+-0.3	45
80% DpHPC / 20% POPS	1	WT	-29.3	-29.2	80.0+-0.3	50
	2	WT	-19.2		75.9+-0.2	50
	3	WT	-28.1		79.1+-0.3	70
80% SDPC / 20% POPS	1	WT	-24.0	-26.9	71.6+-0.2	55
	2	WT	-26.9		69.6+-0.2	35
	3	WT	-23.1		72.5+-0.2	40
80% POPC / 20% POPS	1	WT	-22.8	-38.5	68. 6+-0.2	50
	2	WT	-38.5		65.4+-0.2	55
	3	WT	-28.7		66.6+-0.3	60
80% POPC / 20% POPS	1	F13S	-19.9	-26.1	68.0+-0.3	75
	2	F13S	-26.1		66.1+-0.3	75
	3	F13S	-21.7		65.1+-0.3	55
80% POPC / 20% POPS	1	K15L	-29.4	-38.6	63.6+-0.2	75
	2	K15L	-17.1		64.4+-0.2	50
	3	K15L	-38.6		65.2+-0.5	50

Osh4 ALPS Peptide Free Energy of Folding in Water

In Prof. Karlsson's lab, circular dichroism (CD) measurements showed that the peptide was unstructured in the aqueous phase. Ultimately, we are interested in comparing the free energy of ALPS binding with experimental fluorescence. This reaction consists of the ALPS peptide transitioning from the unfolded state in water to the α-helical state on the membrane surface. We must consider the free energy of unfolding because the unbinding path maintains the initial peptide structure. Figure 4.7 shows the reference structures used in the RMSD calculation for wild type,

F13S, and K15L. Replica exchange simulations address the unfolding and inform the comparison between ITC and simulated PMFs. Our restrained 1D umbrella sampling yielded PMFs that cannot simultaneously account for folding and binding. However, since the free energy is a pathway-independent state function, the sequence of events does not matter. Separate reactions of first unbinding with a fixed peptide structure and secondly unfolding pathways are equivalent to the realistic and likely simultaneous unbinding and unfolding.

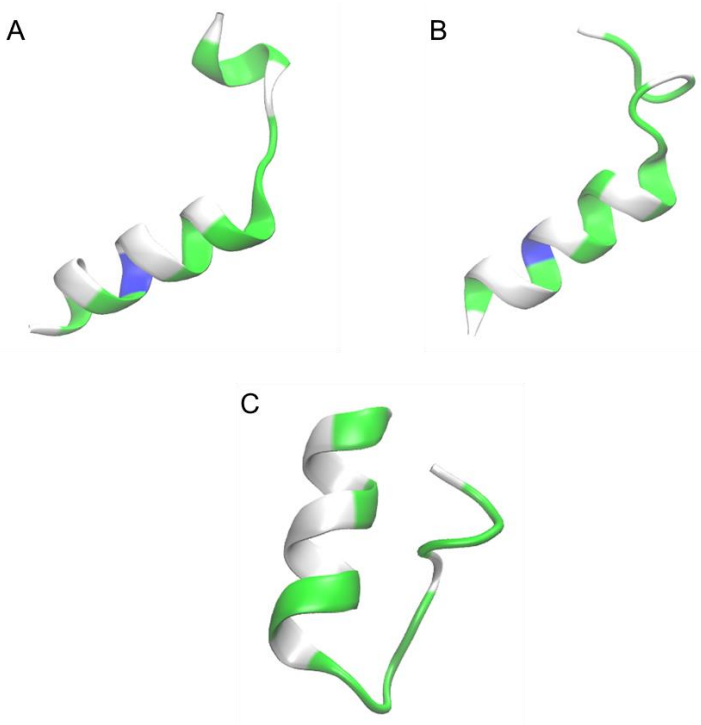


Figure 4.7) Shows the reference structures used for the RMSD calculations. The structure in panel A is from the snapshot before the PMF for the POPC-POPS 2 simulation, which had the lowest free energy of the Osh4 simulations. Panel B F13S and panel C K15L are the RMSD targets.

Figure 4.8 shows the free energy profiles and that the folded energy well was impacted by the presence of ions. The folded "well" at 50% helicity expands with only neutralizing ions. Interestingly, the point mutations F13S effectively eliminated the folded "well" in both ion conditions. The elimination of the folded "well" could be caused by disrupting the amphipathic surface and removing a bulky hydrophobic amino acid that was otherwise shielding the

hydrophobic surface. Additionally, W10 and F13 are neighboring amino acids on the amphipathic helix, contributing to stabilizing π–π stacking. The F13S free energy landscape might be a more appropriate result for the overall free energy landscape of Osh4 ALPS. Is the π–π stacking too strong? Figure B11 shows the center of mass distances between W10 and F13 over time from one of the simulations that only had neutralizing counter ions. There appears to be frequent stacking because this is a center of mass measurement. If instead, closest contacts were calculated, then the distances would be smaller. There are also examples of the peptide completely unfolded, with the π–π stacking and alpha helix around the hydrophobic amino acids remaining intact as shown in Figure B12. On the other hand, the Osh4 K15L mutation appears to have stabilized the peptide. An alternative energy landscape defined by the radius of gyration and the RMSD is in Figure B13.

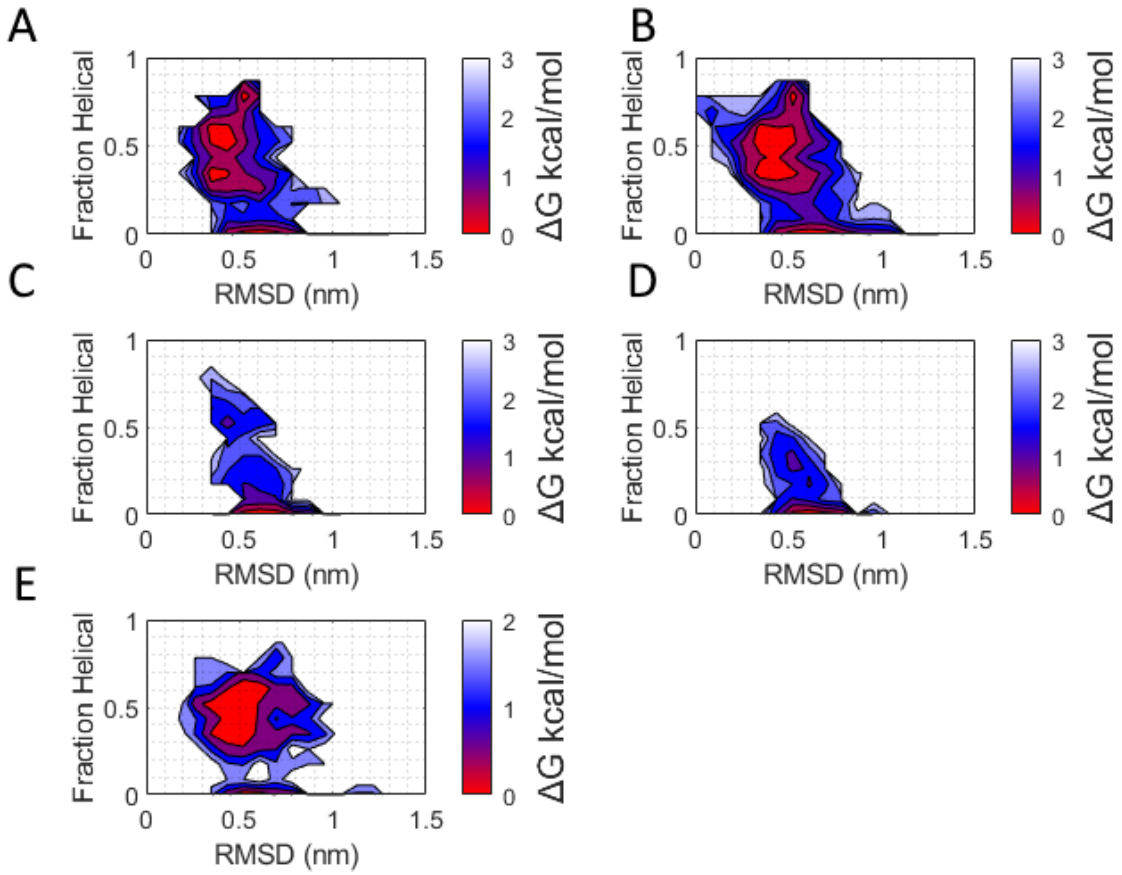


Figure 4.8) Panel A shows the protein folding heat map with 0.15 M KCl, which consisted of

2,200 frames. The first 80 ns of all simulations were discarded to ensure only equilibrated systems were analyzed. Panel B shows the same with only one neutralizing counterion, composed of 4,400 frames. A 1-2 kcal/mol energy barrier exists for the unfolded peptide to fold. Panels C and D show the F13S mutation with and without excess ions, each with 1,100 frames. Panel D is F13S with 0.15 M KCl. Panel E shows the K15L mutation with 0.15 M KCl.

Despite having a broader helical well than a non-helical well, the trial with only neutralizing counter ions further increased the population in the helical peptide states and a slightly lower free energy minimum by -0.1 kcal/mol. As for the F13S point mutation, there was a much higher 2 kcal/mol energy cost to fold the peptide. The π–π stacking between the W10 and F13 could cause this. Finally, the Osh4 K15L peptide was similar to the wild-type Osh4 in that the helical well was broader than the well with a 0 helicity. The two wells were approximately equal at the deepest point with -0.1 kcal/mol, favoring 0 helicity. The target structures used to calculate the RMSD came from the final membrane simulations (Figure 4.7) to be consistent with the Osh5 work that lacked crystal structures. Additionally, the centers of the RMSD wells are depicted in Figure 4.9.

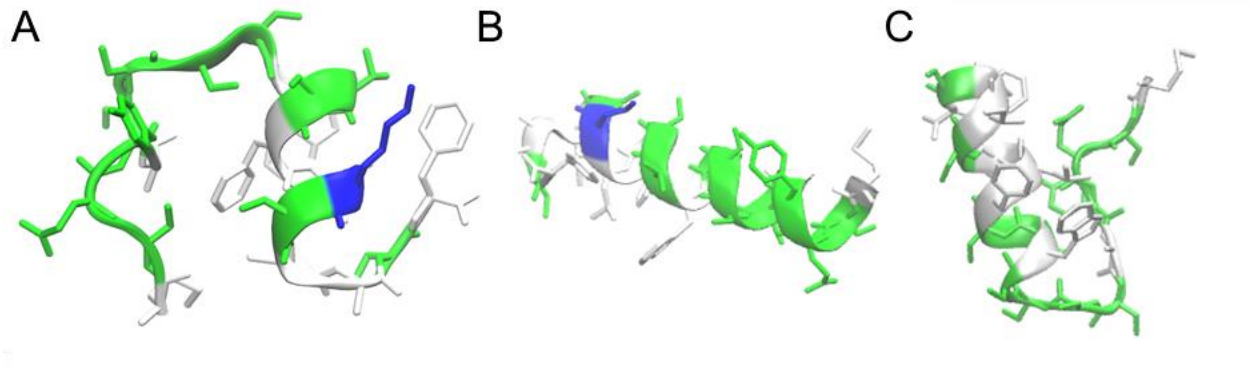


Figure 4.9) Shows the alternative energy wells as determined by the lowest free energy based on probability. Panel A shows Osh4 ALPS with a helicity of 0.37 and an RMSD of 0.37 nm. Panel B shows the Osh4 ALPS with only minimal counter ions and indicates greater helicity with a helicity of 0.8 and an RMSD of 0.5. Finally, the Osh4 K15L mutant appears in panel C with a helicity of 0.54 and an RMSD of 0.615.

Osh4 ALPS Free Energy of Binding to Membrane

Osh4 simulations quantify the thermodynamic dependence of differences in lipid packing. The PC:PS head group ratio was fixed, while the PC lipid tails were varied. As noted above, we designed the umbrella sampling to hold the peptide structure fixed. A harmonically enforced RMSD colvar maintained the initial peptide structure to keep the structure fixed even as it separated from the membrane. Several specific lipids were varied to achieve this result, including DMPC, DpHPC, SDPC, and POPC. The Boltzmann average helped identify each membrane type's lowest free energy simulation. The lowest free energy simulations were assumed to represent best the dominant system membrane-peptide interactions, such as penetration depth and membrane pinching. Although HMMM enhanced the sampling of ALPS binding to the membrane, more sampling is required to acquire a converged ensemble of peptide structure. However, we believe there is sufficient sampling to provide an approximation for the free energy minima for different lipid types. Figures B14 and B15 show the complete PMF profiles for the F13S and K15L mutants to demonstrate full convergence.

Table 4.1 shows that the weakest binding membrane was to the SDPC-POPS. The DMPC and DpHPC lipid components were near the median. The tightest binding affinity of Osh4 ALPS simulation rivals the ΔG of the tightest Osh4 α_6-α_7 (Table 3.1). The varying binding free energies indicate that the peptide favors one membrane over another without the variation in overall charge or introduction of PIP2 lipids. The simulations showed that the DMPC and DpHPC simulations were approximately equal. Although these two membranes differ significantly in lipid packing, they have roughly equal free energies. The F13S mutant showed a reduced attraction to the membrane by 12 kcal/mol from the wild-type Osh but also did not bend the membrane as noticeably as in the lowest free energy wild-type simulation. The K15L mutant resulted in a

transmembrane (TM) peptide. The change in function highlights the importance of the charged lysine amino acid to keep the peptide on the surface of the membrane.

ALPS Interaction Energy per Residue with Membrane

The ALPS-membrane interaction energy is a more detailed energy analysis between the peptide residues and the membrane (Figure 4.10). In addition to the serine and threonine amino acids, the tryptophan and phenylalanine amino acids also play essential roles in enthalpic binding. The W10 had the most vital interaction energy of all the bulky hydrophobic amino acids. F13 and F20 had very little interaction in comparison. SDPC or POPC membranes favor interactions with the phenylalanine amino acids.

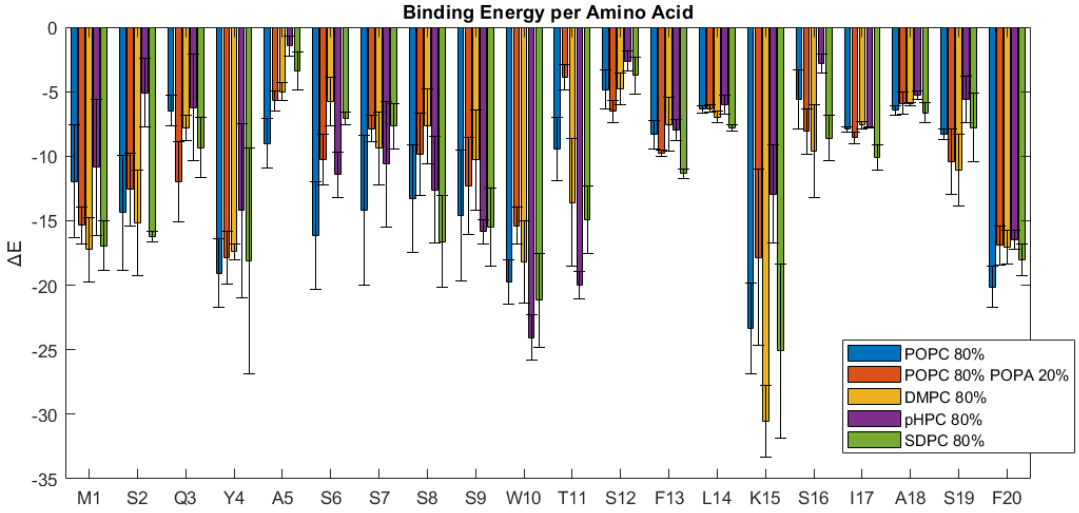


Figure 4.10) Averaged interaction energies between the Osh4 ALPS bound to various membrane types. The peptides were attached to 80% POPC and 20% POPS. The error bars represent the standard error of the three replicas. The error bars were calculated as the standard error of 3 trials.

The K15L and F13S point mutations were somewhat predictable, as shown in Figure 4.10. The K15L removed a dominant amino acid from the peptide. Notably, the F13S had a minimal reduction in the absolute value of free energy. However, the solution REMD indicates that these point mutations play significant roles in the structure of the peptide. The main point is that the

F13S mutation does not alter the interaction energy significantly, but it does impact the folding of the peptide, which, in the worst case, could cause little to no binding in the first place. The interaction energies appear in Figure 4.11. The K15L mutation appeared to improve the peptide's percent helical nature because the energy well's centers remained folded despite elevated RMSDs.

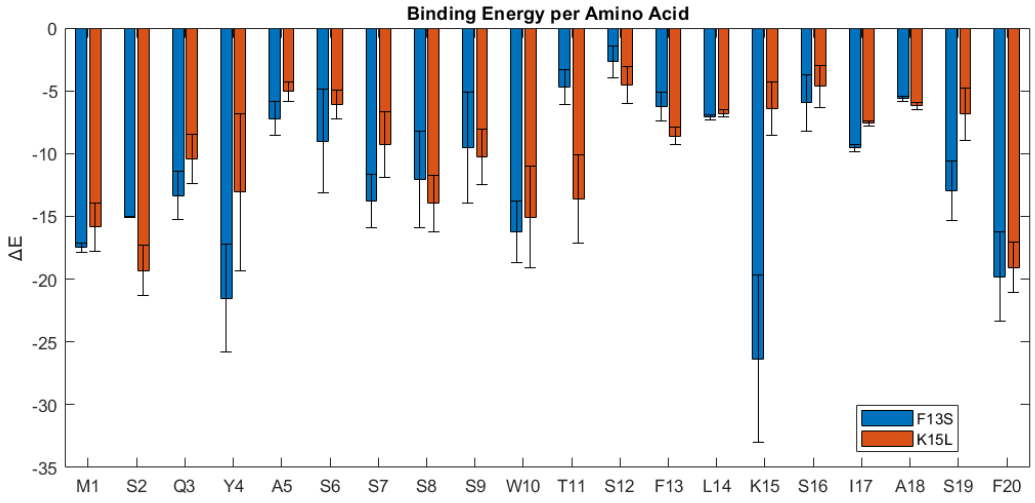


Figure 4.11) Averaged interaction energies for the F13S and K15L mutants bound to POPC 80% POPS 20%. The error bars represent the standard error of the three replicas. The error bars were calculated as the standard error of 3 trials.

4.3.2 Osh5 ALPS

Other proteins in the Osh family have ALPS peptides. The purpose was to examine how membrane binding from different family members is affected by the variation in the peptide sequence. Osh4 and Osh5 have a similar sequence for the ALPS with a single-point mutation of Y4H. Due to this mutation and the extended N-terminus tail, the Osh5 ALPS might target a different organelle from the Osh4 ALPS. The results below show Osh5 ALPS binding to model membranes and how it compares with Osh4 ALPS results.

Osh5 HMMM

The structure of the Osh5 peptide is almost entirely α-helical and tends to lack the bend in the first five amino acids. However, Osh5 ALPS is bound similarly to Osh4 ALPS. Structures shown in Figure 4.12A and B appear deeply penetrating, while panel C is slanted but approximately in the same position. The fully α-helical peptide penetrates only down to the 6th carbon position of the lipid and thus at the interface between the lipid tail and the DCLE solvent.

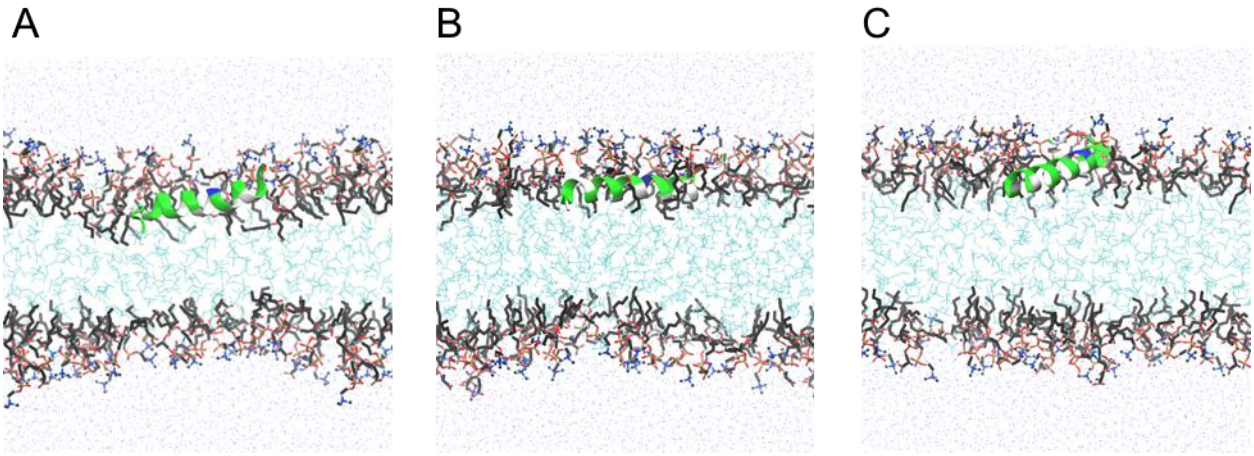


Figure 4.12) Final HMMM simulations of the Osh5 peptide. A) is the DMPC 80% - POPS 20% r1, B) is POPC 80%-POPS 20% r2, and C) is SDPC 80% – POPS 20% r2.

The distance versus time graph of the HMMM shows that the peptides bound in a stochastic manner as shown in Figure 4.13. All of the top Osh5 peptides approached the exact placement, which might be necessary if there is an optimal position, as shown in Figure 4.12. The complete set of HMMM binding sequences appears in Figures B1-B2.

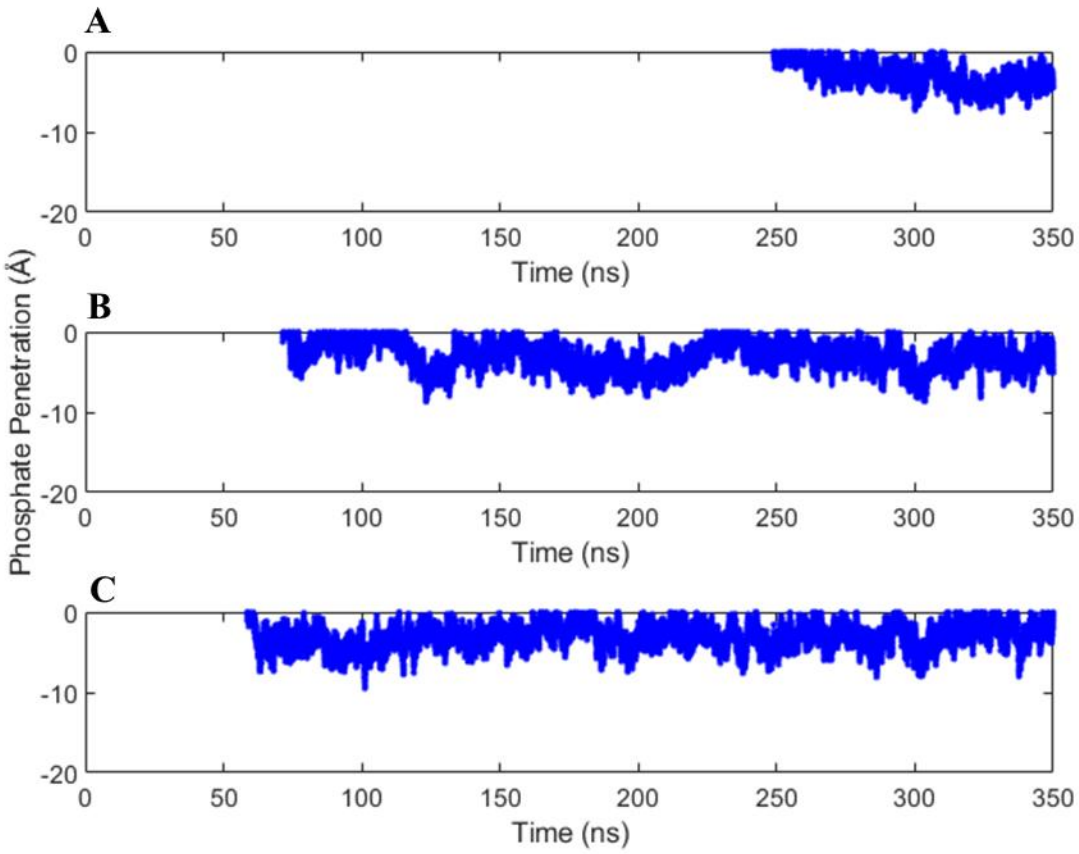


Figure 4.13). The center-of-mass C_a penetration distances beneath the membrane phosphate for Osh5 ALPS indicate that the lowest free energy simulations had similar depths of penetration around -2.5 Å. The depths slightly differ from the Osh4 ALPS, which appears to have penetrated closer to -5 Å. **A)** DMPC 80%-POPS 20% Osh5 r1 **B)** POPC 80% POPS 20% Osh5 r2 **C)** SDPC 80% POPS 20% Osh5 r2.

Osh5 Full Length Membrane

The final simulations show that the peptides studied in the DMPC membrane unfolded slightly more compared to the other peptides. All of the Osh5 ALPS peptides were flat except for that in DMPC. The slant in the peptide orientation might have resulted in a better interaction with the membrane, as shown in Figure 4.12A. The biggest surprise was that in the membranes composed of 80% DMPC- 20% POPS, the Osh5 histidine point mutation was closer to the surface and had little interaction with the membrane. The POPS could have caused looser placement

because the POPS protrudes above the POPC. This protrusion has a negative charge that attracted the histidine even though the histidine is a neutral amino acid in this model. Pyrole-like and imine nitrogen are more polar than a single OH group, making histidine more polar than threonine. However, this unfolding did not appear to occur in the other simulations. Overall, the bound Osh5 ALPS simulations appeared to have less structure variance as depicted in figure 4.14 and be more folded than Osh4. The difference in the extent of folding could be due to the initial structure being more helical.

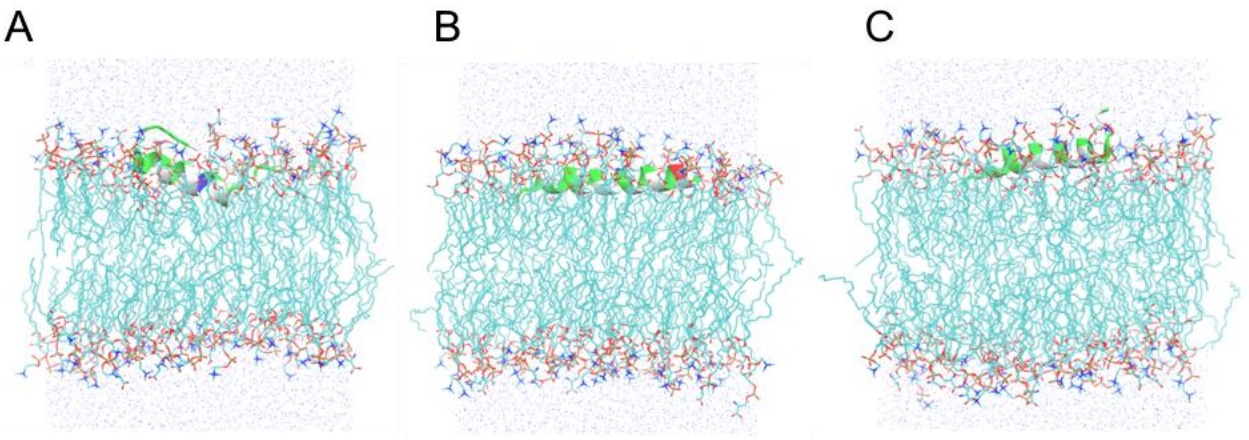


Figure 4.14) Shows the lowest binding simulations, which are A) DMPC r1, B) POPC r2, and C) SDPC r2.

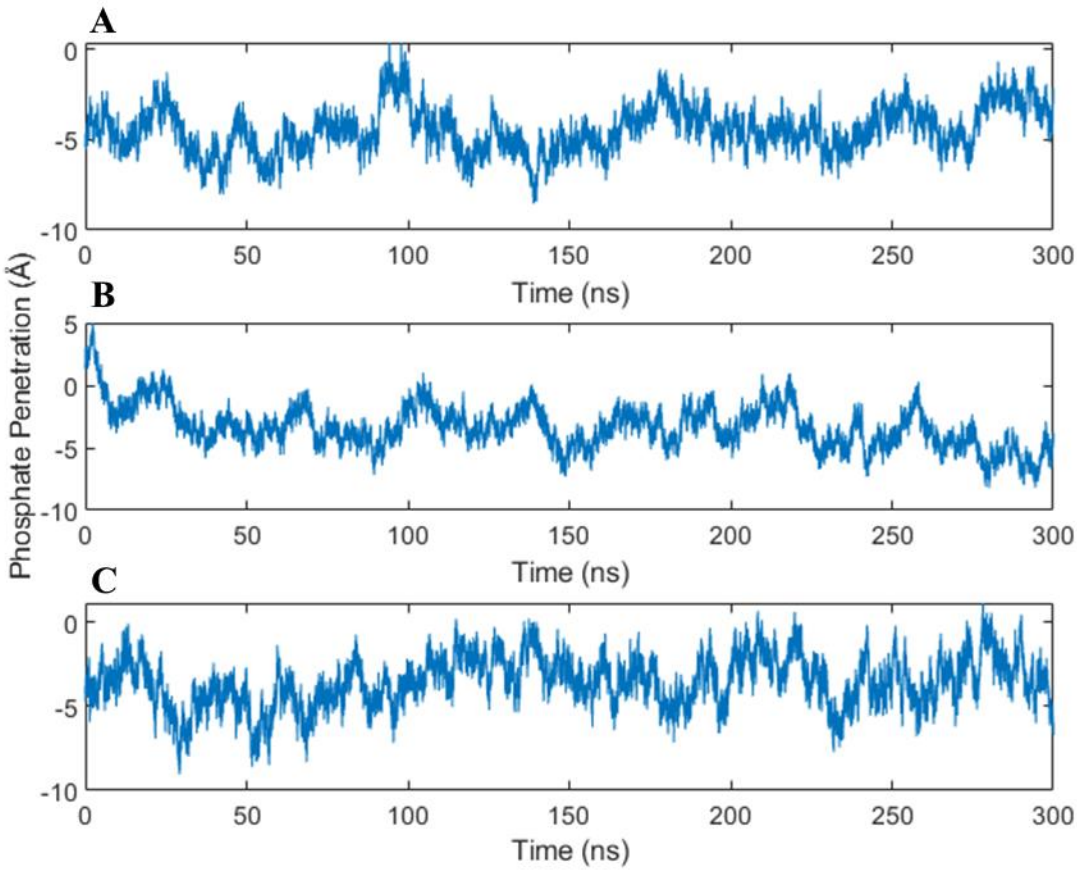


Figure 4.15) Shows the C_a penetration distances beneath the membrane phosphate. All of these lowest free energy simulations penetrated around -5 Å. **A)** DMPC r1 **B)** POPC r2 **C)** SDPC r2.

The distance versus time data of the full-length Osh5 peptides is shown in Figure 4.15 and illustrates some rearrangement from the HMMM simulation positions. The lowest free energy simulations of the Osh5 peptide positioned closer to -5 Å regardless of the lipid type studied. The complete set of position vs. time graphs appear in Figures B18-B20, including the Met1, Phe13, and Leu14 amino acids positions measured as if they were isotope labeled. Neutron experiments are analogous experiments that compare the positions of the simulated lipids with isotope-labeled lipids.

Osh5 ALPS Folding Energy in Water

The Osh5 ALPS was thoroughly studied with simulations, including T-REMD, to determine whether it was thermally unstable. The ALPS of Osh4 and Osh5 only differ by a single point mutation and several additional amino acids that could be helical. A good metric for folding the Osh5 peptide is the RMSD; the target for the RMSD was a highly helical Osh5 peptide from a previous simulation shown in Figure 4.16.

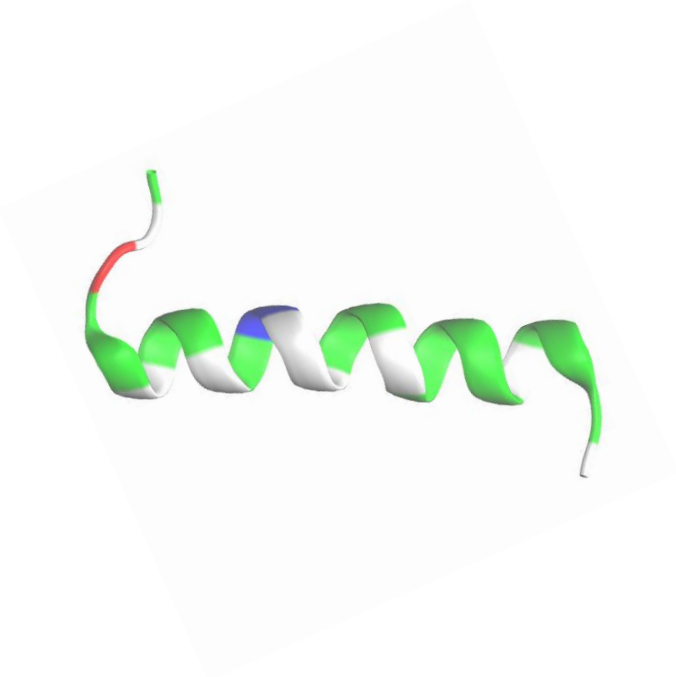


Figure 4.16) The Osh5 target came from SDPC 3 for Osh5. The structure was almost entirely helical in structure.

This Osh5 peptide was slightly longer but has an energy difference close to -0.5 kcal/mol because the folded "well" is deeper than the unfolded "well" Figure 4.17. Surprisingly, if helicity is the only parameter of interest, the true minima of the Osh5 is a fraction of the helicity value of 0.37. As with the Osh4 ALPS, two main wells in the energy landscape were observed, with one corresponding to the completely unfolded and another corresponding to a helicity of 0.4-0.5, as shown in Figure 4.17. The alternative free energy landscape appears in Figure A2, which defines

an energy landscape with the radius of gyration and the RMSD. The different variables lead to varying interpretations of the data. Based on the helical interpretation, there are two energy wells and a reasonable amount of folded peptide. At the same time, the RMSD and radius of gyration show one energy well and a higher free energy to attain a folded state. There is undoubtedly a difference in the interpretation of the free energy to access states closer to the folded state. The helical content does not suggest the peptide favors a completely folded alpha helix.

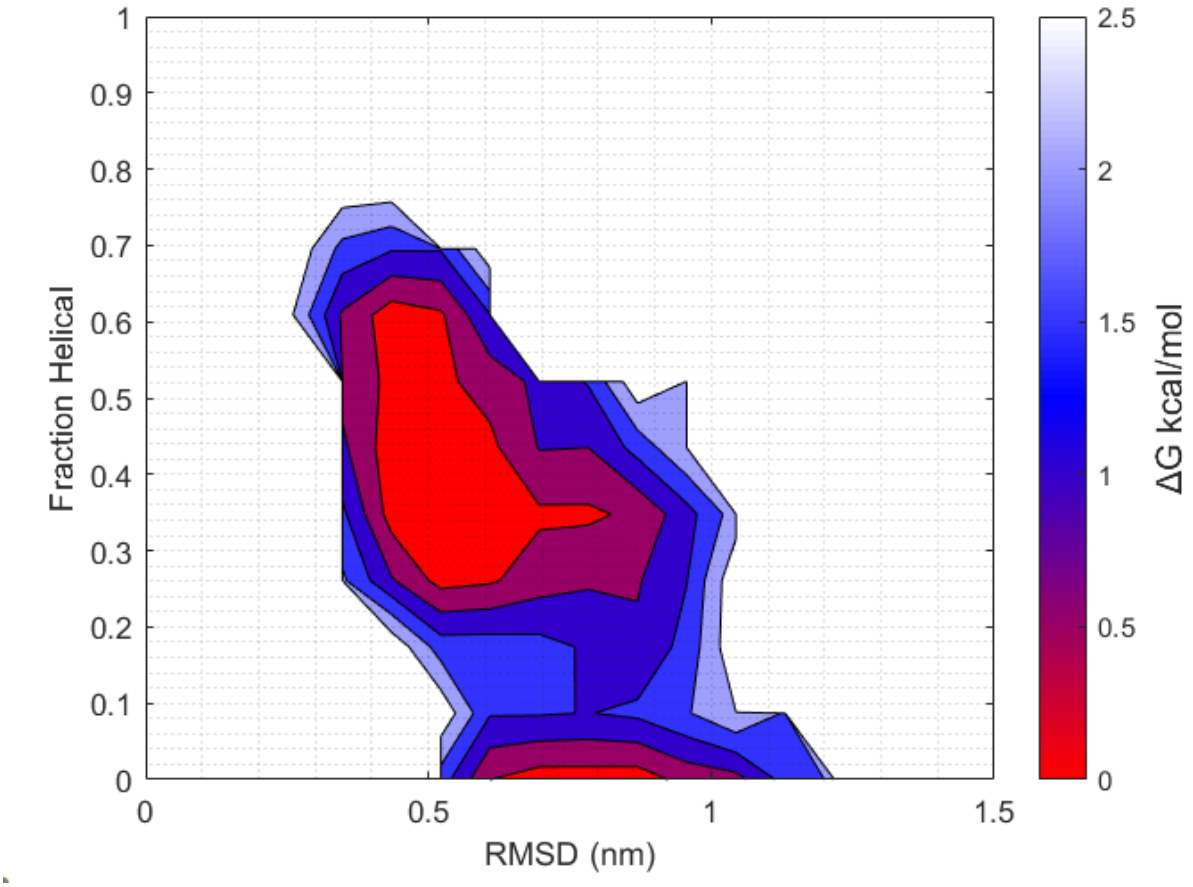


Figure 4.17) Energy of folding for the Osh5 peptide has a 1.25 kcal/mol energy barrier to fold and unfold closest to the target RMSD. However, if helicity is the only parameter, the peptide is favored to be helical with a lower energy of -0.23 kcal/mol at the lowest point of the folded energy well. The calculation for the above image included 2,200 frames and discarded the first 80 ns.

Notably, the Osh5 ALPS has an energy well minima of 0.37 helicity, corresponding to the structure shown in Figure 4.18.

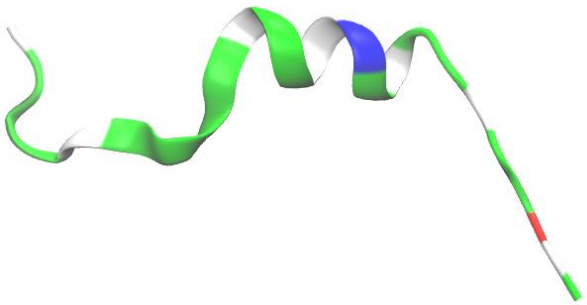


Figure 4.18). The peptide was only 37 % helical with a RMSD of 0.54 nm.

Osh5 ALPS Binding Free Energy to Membrane

The area per lipid increases as follows: DMPC< POPC< SDPC, and the trend still holds in our simulations that were mixed with POPS, as shown in Table 4.2. However, the area per lipid is similar between POPC-POPS and DMPC-POPS because the mixture's effects cause deviations from the trend for pure membranes. Other properties, such as the bending modulus and rigidity of the membrane, might track better with the free energy.

Table 4.2) Free energy of the Osh5 peptide, showing that the peptide had some variation and mostly preferred the DMPC and the POPC mixed with POPS. The SDPC appears to be consistently less favored compared to the others.

Membrane	Run	Osh5	ΔG (kcal/mol)	ΔG (kcal/mol) Boltzman Average	Area Per Lipid ($Å^2$)	% Helix
80% DMPC 1 20% POPS	1	WT	-36.4	-36.4	62.9+-0.3	48
80% DMPC 2 20% POPS	2	WT	-16.7		65.1+-0.3	76
80% DMPC 3 20% POPS	3	WT	-24.8		59.7+-0.2	68
80% POPC 1 20% POPS	1	WT	-23.5	-28.2	66.1+-0.3	80
80% POPC 2 20% POPS	2	WT	-28.2		65.1+-0.3	84
80% POPC 3 20% POPS	3	WT	-21.0		67.5+-0.3	56
80% SDPC 1 20% POPS	1	WT	-15.3	-20.2	69.4+-0.3	88
80% SDPC 2 20% POPS	2	WT	-20.4		69.0+-0.3	72
80% SDPC3, 20% POPS	3	WT	-19.6		70.9+-0.3	80

All of the lowest free energy simulations had the lowest area/lipid except for DMPC, which was very close between the first and third trials. Factors to further explore include variations of penetration depth, percent of folding, and area/lipid. The complete PMF profiles are shown in Figure A5 to demonstrate the full convergence and separation of the peptide from the membrane.

Osh5 ALPS Interaction Energy with Membrane per Residue

Perhaps the most critical question is the impact of the individual histidine amino acid on the overall interaction of the ALPS peptide with membranes. The interaction energy of Figure 4.19 can address this question. The results indicate that the histidine amino acid played a minor role in the interaction, resulting in less membrane interaction than tyrosine. Although we assume a protonation state for histidine consistent with a pH of 7 (neutral), the pH could be lower in certain cell parts. The pH ultimately affects binding as the pKa for histidine is around 6-6.5.

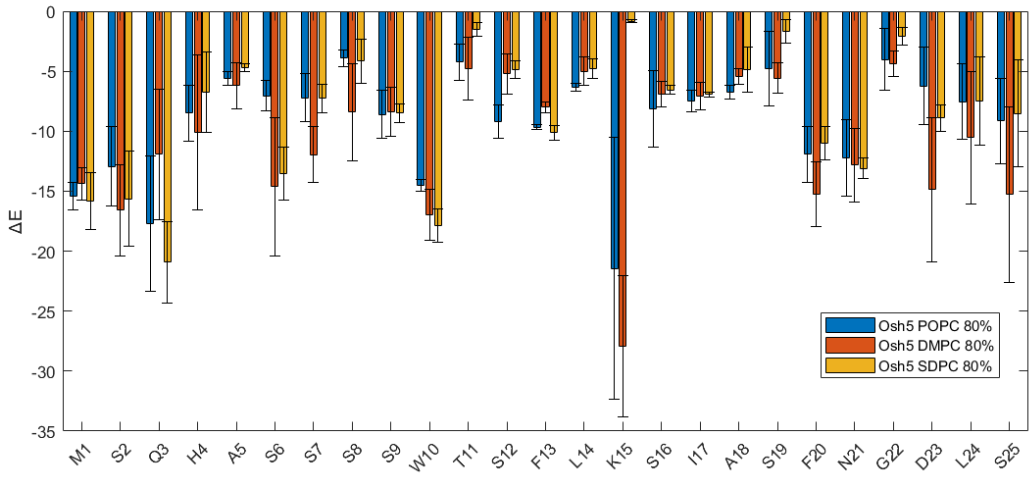


Figure 4.19) Interaction energy between the peptide and the membrane. The DMPC 80% POPS 20% had the lower free energy measured from the PMF. The standard error represents the three trials.

4.4 Discussion

In summary, the Osh4 ALPS structure bound relatively strongly to all the membranes tested. The results indicate inaccuracies in the force field because the peptides consistently attain

ΔG<-20kcal/mol. However, ITC results on the wild-type Osh4 from Dr. Amy Karlsson's lab indicate around -5 to -7 kcal/mol overall compared to some simulations that show -38 kcal/mol. The differences in the free energies are likely inaccuracies in the force field generated from a generic algorithm of predicted trends rather than actual experimental data. An essential contribution of this approach is to sample several structures using HMMM to accelerate the simulations. These results will play a role in future lipid-protein interaction parameterization because preliminary data on this peptide with ITC and other membrane surface interactions (others in the Klauda lab) suggest binding disagreements with the experiment.

Peptide folding is a crucial factor in the experiments. There was a typical "well" for all the peptides with complete disorder (0% helical). However, REMD with the CHARMM36m force field suggests a second "well," which indicates that part of the peptide is partially α-helical in the solution. Compared to the CD these results could make sense. A peptide that is 50% helical and 50% unstructured would appear 25% helical overall which is similar to the value found in CD. However, the overall average from summing the probability mass function is 37-39% helix for 0.15M ions and neutralizing ions respectively. The frames corresponding to the "well" centers are all folded or partially folded. The mutants that disrupted the amphipathic interface disrupted the folding, and the ionic strength of the solution played a role in the extent of folding. Other reactions are taking place, like the folding of the peptide and lipid conformational changes. Still, the folding alone would be at most 2 kcal/mol and does not fully resolve the mismatch between the simulation and the experimental results.

The Osh4 ALPS interaction with the membrane (Figure 4.10) shows that the first 1-10 amino acids are high in serine, threonine, tend to be more unstructured, and have higher interaction

energies. The other half of the peptide is more helical and inserts deeper typically. This helical segment tends to be more dominated by the presence of the lysine.

The Osh5 ALPS interaction energies show weaker interaction energies with the first 1-10 amino acids than Osh4 ALPS. The reason why different amino acids had different interaction energies could be a property of the peptide folding because the Osh5 was overall more folded than the Osh4. Another point of comparison was that the Y4H (Osh4 versus Osh5) simulation did have an interaction of dE~-10 kcal/mol compared to ~-20 kcal/mol for the lowest Osh4 simulations. One could argue that a neutral Y4H does not make enough of a difference in the binding to a more charged PM membrane. Instead, a charged Y4H mutation is required. The Osh5 peptide is also longer by five amino acids, but overall, the Osh4 peptide is stronger at binding. The peptide is 1.93 kcal/(mol amino acid) for the lowest Osh4 and 1.46 kcal/(mol amino acid) for the lowest Osh5. Just looking at the interaction energy per residue, the extra amino acids contributed more interaction energy than the middle 10-20 amino acids. The change in interaction energy might be caused by the lysine being so energy-dominant that it hinders the exploration of the other amino acids to find their energy minima.

The simulations with Osh4 in various PC lipids of SDPC, pHPC, and DMPC had similar binding free energies of -26 to -29 kcal/mol ignoring the minimal change due to unfolding in the water phase. It is important to note significant variations between the different simulation replicas. The Boltzmann average of the replications essentially selects the lowest free energy simulation. The lowest free energy simulation is vital for being highly favored in reality. For example, a -1.742 kcal/mol lower energy state populates 95% of the total states in a given system, as solved for the reference below.

$$\frac{p_i}{p_j} = \exp\left(\frac{\left(e_i - e_j\right)}{kT}\right)$$

$$\frac{p_i}{p_j} = \exp\left(\frac{1.730}{0.001985875 * 298}\right) = 18.6\%$$

$$p_i = 0.186 * p_j$$

$$\frac{p_i}{p_i + p_j} = \frac{0.186}{1.186} = 95\%$$

The force field appears far from the actual values found in experiments, so additional improvements will be needed to better model protein interactions with the membrane in Osh4 and Osh5 peptides. Examining the interaction energy per residue shows that the Y4H point mutation is not responsible for all energy differences. Instead, it combines all amino acids resulting from different folds, placements, and orientations. The POPC lipid is a dominant component of mammalian membranes commonly found in yeast membranes. Interestingly, this membrane mixture has the lowest free energy of the Osh4 ALPS. This membrane type with DMPC-POPS could mimic the packing of POPE and sphingomyelin that are prevalent in the membrane in the PM. DMPC and the lower free energy could represent the natural target for Osh5 and why it has the lowest ΔG. The bottom line is that the Osh4 and Osh5 ALPS can tightly bind to both membranes and that the histidine amino acid likely plays a role in placing the Osh5 peptide on a more charged plasma membrane. There could also be point mutations on the tryptophan and phenylalanine amino acids to determine if they prefer the tighter packing of the DMPC.

Chapter 5: WLBU2 Antimicrobial Peptide

The contents of this chapter are modified from the following publication:

- Reprinted (adapted) with permission from {Allsopp, R.; Pavlova, A.; Cline, T.; Salyapongse, A. M.; Gillilan, R. E.; Di, Y. P.; Deslouches, B.; Klauda, J. B.; Gumbart, J. C.; Tristram-Nagle, S. Antimicrobial Peptide Mechanism Studied by Scattering-Guided Molecular Dynamics Simulation. *The Journal of Physical Chemistry B* **2022**, *126* (36), 6922-6935}. Copyright {2022} American Chemical Society.

5.1 Introduction

While traditional antibiotics have long provided protection against bacterial infection and allowed surgical interventions to save lives, the worldwide problem of bacterial resistance178 motivates many researchers to explore alternatives. A comprehensive report assessing rising antimicrobial resistance has predicted that by 2050, over 10 million deaths will occur annually due to antimicrobial-resistant pathogens[173]. One approach to solving this problem is to use rational design to synthesize antimicrobial peptides (AMPs) as an alternative antibiotic because bacterial resistance is slower to develop. The Montelaro/Deslouches groups draw inspiration from the human cathelicidin, LL-37, a helical, broad-spectrum amphipathic peptide of 37 amino acids with 12 positively charged residues.[62, 174] A second inspiration was the naturally occurring AMP on the end of the C-terminal tail of the HIV-1 fusion protein, LLP1, which is also highly cationic, containing seven positively charged residues out of 28 residues.[64] The highly cationic nature of these peptides is thought to impart selectivity towards negatively charged prokaryotic bacterial cells and lower their toxicity to host eukaryotic cells.[175] The Deslouches lab has attempted to discover key aspects of AMP-caused bacterial killing by synthesizing simplified AMPs containing only three types of amino acids: valine (V), tryptophan (W), and arginine (R),[62, 176-179] instead of 15 types as in LL-37 or 11 types as in LLP1. Limiting the number of types of amino acids makes it easier to determine which physical properties, hydrophobicity, hydrophobic moment, length,

and charge of peptides are essential for permeabilizing cells and killing bacteria. The present work has focused on the rationally designed, cationic AMP WLBU2,[180-183], which is now in Phase II clinical trials for wound healing.[184]

The primary structure of WLBU2 is **RR**WV **RR**V**R** R**R**WV**R** R**R**VV**R** VV**RR** WV**RR,** with 13 R residues (shown in bold type) out of 24 total amino acid residues. If WLBU2 were perfectly α-helical, the helical wheel design would predict that the R's line the hydrophilic face while the V's line the hydrophobic face toward the lipid chains, with the W's close to the interface between these two faces (Fig. 5.1).

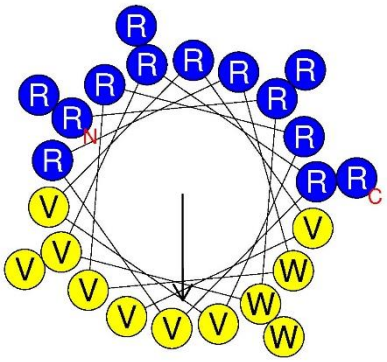


Figure 5.1. Helical wheel diagram of WLBU2 prepared using Heliquest website (heliquest.ipmc.cnrs.fr). Arrow shows the direction of the hydrophobic moment, μH.

The stability of the AMP improved with the addition of W in saline conditions like the human body.[182] The detailed secondary structure of WLBU2 in four different lipid model membranes (LMMs) and an aqueous solution was obtained using circular dichroism spectroscopy and was published previously.[65] While WLBU2 in water or 15 mM phosphate buffered saline (PBS) adopts primarily a random coil or β-sheet structure, the α helical content increases to ~80% in Gram-negative (G(-)) inner membrane (IM) or Gram-positive (G(+)) LMMs, and to ~40% in lipopolysaccharide (LPS)-containing LMMs. Although WLBU2 is not 100% helical by our determination,[15] it is still primarily helical when in contact with the inner membrane of G(-) and G(+) LMMs, partially confirming the locations of R and V on opposite faces in WLBU2's helical

wheel rational design. There was only a low level of α-helicity (~20%) when added to a eukaryotic membrane mimic.[65] Thus, the secondary structure of WLBU2 plays an essential role in avoiding toxicity to host eukaryotic cells while remaining toxic to prokaryotic bacterial cells.

Molecular dynamics (MD) simulations implement the visualization of the WLBU2 interacting with four different bacterial LMMs. The starting point for the comparison involved constraining the thickness of the different LMMs to those obtained using X-ray diffuse scattering (XDS) and the locations of the peptide to those obtained using neutron reflectometry (NR). Then, by simulating for an additional ~0.5 microseconds, the conformation of WLBU2 and the surrounding lipids were allowed to equilibrate to the final membrane-peptide structure. This structure produced an electron density profile (EDP) and, after a Fourier transformation, made a continuous X-ray form factor ($F(q_z)$). The accuracy of the simulation was verified by comparing the simulated $F(q_z)$ with the experimental $F(q_z)$ obtained using X-ray diffuse scattering (XDS). This comparison anchors the all-atom (AA) MD simulation to the experimentally determined lipid thickness and structure. Molecular details of WLBU2's conformation in the membrane give insights into the mechanism of bacterial killing by WLBU2's membrane perturbation.

5.2 Methods

MD simulations

G(-) inner membrane (IM). Each simulation involved membranes with lipids composed of POPE/POPG/TOCL in a 7:2:1 molar ratio, as in the scattering experiments.[15] Each simulation had one or two peptides placed in different locations, with either 100 or 160 total lipids. The simulations were created using 45-60 TIP3P waters per lipid[185] and 27-38 K^+ ions, depending on the simulation size. Simulations were run on the Extreme Science and Engineering Discovery Environment (XSEDE) using the Texas Advanced Computing Cluster (TACC) on Stampede2

and Comet and Expanse at the San Diego Supercomputer Center at U.C. San Diego. The simulations ran for 300 ns for the surface simulations and at least 420 ns for the inserted simulations to probe stability. It was known from the beginning that the peptide easily comes out of the membrane and that it was a challenge to sustain the peptide inside the membrane. Peptides remained in the membrane when surface tension was applied, and the structure agreed better with the experiment. The need for adding additional surface tension may have come from minor inaccuracies in the force field (lipid-lipid and protein-lipid interactions).[186] Additional details concerning the stability of the inserted peptide position are given in S.I. (see Figs. C1 and C2).

The G(-) trials required no local restraints. The only bias was additional surface tension. Reducing the surface tension helped optimize the fit by using the last 100 ns for analysis. The initial location and number of peptides were varied, as was the surface tension (from 0 to 15 dyne/cm). All surface simulations ran with NPT (isothermal-isobaric) or NPγT (isothermal-isobaric with lateral surface tension). The calculated form factors (see below) were used as a measure of success for the different simulations, focusing on the cross-over points (zero positions) on the x-axis and relative lobe intensities and their comparison to the XDS experiment. Additional details concerning the stability of the inserted peptide position are given in S.I. (see Figs. C1,2).

The peptides were simulated separately in two states (surface-bound or inserted) to reduce the system size, unwanted interactions, and complexity. The surface tension applied for the surface-bound or inserted states of WLBU2 differed because the initial optimization of the inserted peptide contained significant errors in some areas of low q-space when the surface tension was the same as the surface-bound state. Several surface tension combinations match the form factor crossing points (see below), but the best combination to lower the sum squared error was nine dyne/cm for the surface-bound and 15 dyne/cm for the inserted WLBU2.

MD simulations of G(-) membranes utilized NAMD 2.12-2.14 depending on the resource,[23] while also using the CHARMM36 force field for lipids[132] and CHARMM36m force field for proteins.[131] The simulations ran using a 2 fs time step along with long-range electrostatics interactions evaluated every other time step using the particle mesh Ewald method.[187] Short-range non-bonded interactions were cut off at 12 Å using a force-based switching function beginning at 10 Å. The temperature was maintained at 37 $^{\circ}$C using Langevin dynamics. The pressure was maintained separately in the membrane-planar (when no surface tension was applied) and membrane-orthogonal direction using a Langevin piston at 1 atm.

Electron density profiles (EDPs) from simulation. A computer-generated "SIM" file creates the form factors. The SIM file identifies all atoms and their bilayer positions using the SimtoExp software.[188] The Fourier transform of the EDPs is the X-ray form factor, also produced using the SimtoExp software. The SimtoExp software uses a chi-square goodness of fit test between the experimental and simulated form factors.

Robetta modeling. The CHARMM-GUI program requires an initial protein starting structure for simulations. We utilized the Robetta[189, 190] server to produce initial starting structures. Robetta is an online protein prediction server developed by the Baker Laboratory at the University of Washington. Robetta uses the Ginzu prediction protocol to match protein chains into putative domains with reasonable confidence. A homology modeling approach defines the 3-D structure using comparisons with other proteins that have solved structures. The *ab initio* structure prediction methods designed by the Robetta server involve running Monte Carlo simulations to make energetically minimized structures. Robetta returned with confidence two distinct models for WLBU2 depicted in Figure 5.2. Since both straight and bent helix models (see Fig. 5.2(a), (b)) were predicted with certainty using the Robetta server, both were considered potential starting

structures for the simulations in this work. However, the bent helix conformed better to the secondary structure obtained using circular dichroism (CD).[65] The simulations had no restraints to the helical secondary structure. Additional details concerning the Robetta modeling appear in S.I.

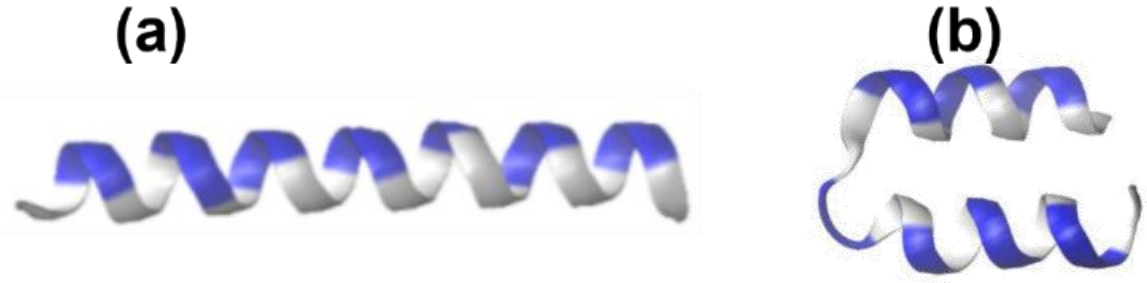


Figure 5.2. Structural predictions from the Robetta software. (a) Straight helix. (b) Bent helix. Colors: R, blue; W, V, white.

5.3 Results

Comparison of experimental and simulated form factors G(-) Inner Membrane (IM)

Initial efforts applied no surface tension to simulate the control G(-) IM model membrane. CHARMM-GUI's Membrane Builder[48, 132, 146, 191] built three replicas of an all-atom membrane used in the study. In Fig. 5.4(a), the agreement between the simulated form factor (black) and experimental form factor (red) is sub-optimal. The amplitudes of the diffuse lobes match reasonably closely; the value where F(qz) goes to zero (the cross-over points) does not agree. This comparison indicates that the simulated bilayer is thicker than the experimental data. Fig. 5.3(b) shows the simulated EDP, and Fig. 5.3(c) shows the VMD visualization.

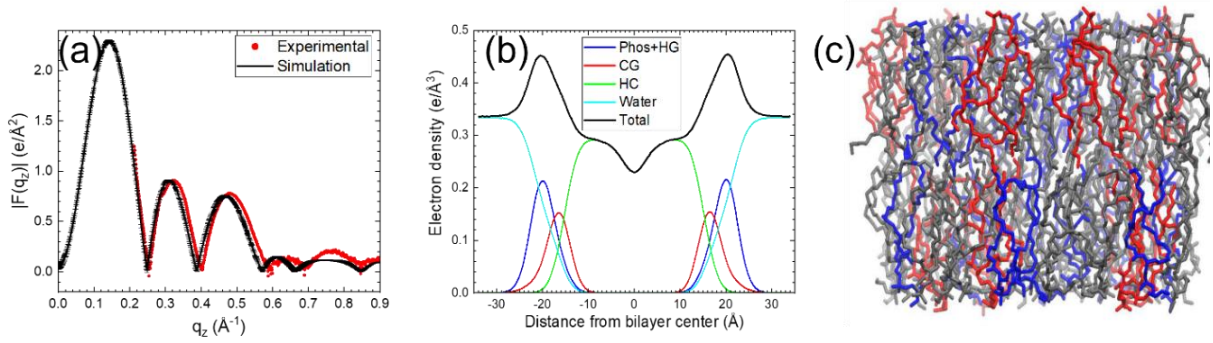


Figure 5.3. (a) An unsuccessful attempt to match the form factor for G(-) IM simulated without surface tension (black line) and experimental form factor (red circles). The simulated form factor data appear as an average of 3 simulations, with standard deviations. (b) Simulated EDP. Colors: Total, black; phosphate + outer headgroup, blue; carbonyl-glycerol, red; hydrocarbon, green; water, cyan. (c) VMD visualization of simulated G(-) IM bilayer. Colors: POPE, gray; POPG, blue; TOCL, red.

Various minor surface tensions were applied and simulated between 9 and 15 dyne/cm to match the experimental bilayer thickness more closely. The surface tension resulted in shifting the positions of the cross-over points. The best agreement with the experimental form factor for the control G(-) EDP occurred when a surface tension of 9 dyne/cm was applied (Fig. 5.4(a)). Fig. 5.4(b) shows the simulated EDP, and Fig. 5.4(c) shows the VMD visualization of control G(-) IM. The inclusion of potassium ions caused no change in the goodness of fit to the experimental data.

Table 5.1. The sum of squares error (SSE) for different G(-) IM/WLBU2 simulations at 37 °C. The amplitudes of the three diffuse lobes in the experimental form factor did not compare ideally to those of the simulated form factors. The SSE of each lobe was computed along with the total SSE of the entire curve as described in Materials and Methods. *Surface tension of the combined model: 15 dyne/cm for the bent inserted WLBU2 and nine dyne/cm for the bent surface WLBU2. The composition of the above G(-)/WLBU2 simulation had a molar ratio of 87:1; a second simulation at a molar ratio of 80:1 produced similar errors.

	1st Lobe	2nd Lobe	3rd Lobe	Total SSE (lobes + zeroes)
Combined* (shown in Fig 5.5(d,e))	0.53	0.16	1.00	3.07
Bent inserted 0 dyne/cm	1.36	6.71	10.04	19.72
Bent inserted 15 dyne/cm	1.04	1.74	1.60	5.66
Bent surface 0 dyne/cm	0.74	1.75	2.52	6.70
Bent surface 9 dyne/cm	0.37	0.18	1.13	3.13
Straight surface 0 dyne/cm	0.66	1.32	1.38	4.86
Straight surface 9 dyne/cm	0.39	0.20	1.34	3.52

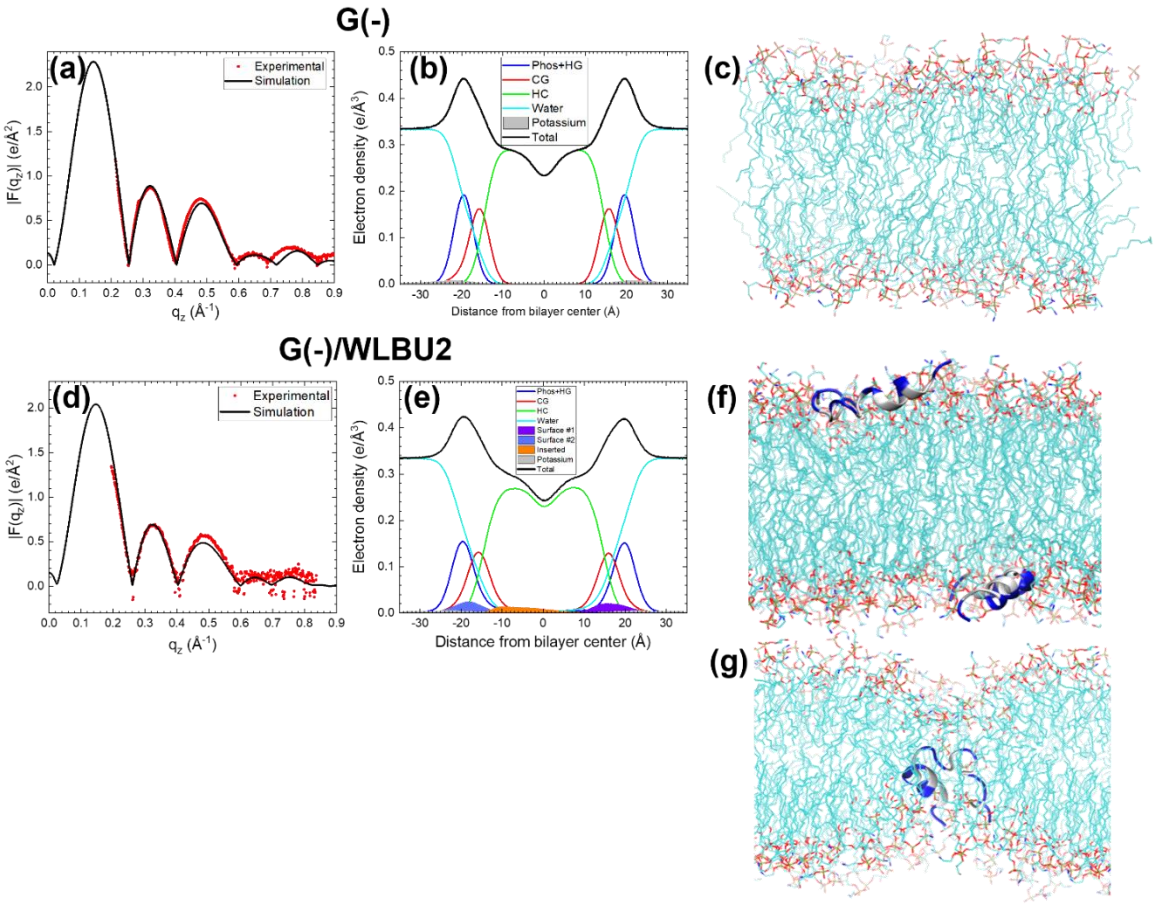


Figure 5.4. (a) G(-) IM control simulated (black line) and experimental form factors (red circles). (b) G(-) IM simulated EDP. Colors: Total, black; phosphate + outer headgroup, blue; carbonyl-glycerol, red; hydrocarbon, green; water, cyan; K^+, gray. (c) VMD visualization of G(-) IM control. Colors: Oxygen atoms, red; hydrocarbon chains, cyan. (d) G(-)/WLBU2, simulated (black line), and experimental form factors (red circles). (e) G(-)/WLBU2 simulated EDP. Colors: as in 6(b) with surface WLBU2#1, filled purple; surface WLBU2#2, filled blue; inserted WLBU2, filled orange. (f) VMD visualization of two surface states. Colors in 6(c) with WLBU2: R, blue; W, V, white. (g) VMD visualization of inserted WLBU2. Colors as in 6(f).

The best agreement with the G(-) membrane mimic experiment occurred after splitting the density of the bent surface model at nine dyne/cm and a bent inserted model at 15 dyne/cm (see Table 1). The weighting of the surface and inserted models was constrained to 37.2% inserted and 62.8% surface (see Fig. 5.4(a)) based on fitting the NR result to two Gaussians (Fig. S5(b)). Significant oxygen atoms from the glycerol-carbonyl, phosphate headgroups, and water enter the bilayer interior in Fig. 5.4(g). Table 1 shows the Sum of Squares Errors for different WLBU2

locations and surface tensions. Note that this procedure of dividing the peptide density into surface and inserted assumes there is no interaction between the surface and inserted states of WLBU2, as we have previously performed in the case of the lung surfactant proteins SP-B and SP-C.[192]

A visualization of a close-up of WLBU2 inserted into G(-) IM LMM appears in Fig. 5.5(a). The presence of the Rs in the center of the membrane draws phosphates, K+ ions, and water into the membrane. Fig. 5.5(b) shows a significant electron density of water in the center of the bilayer in the case of inserted WLBU2 but not for the two surface states of WLBU2.

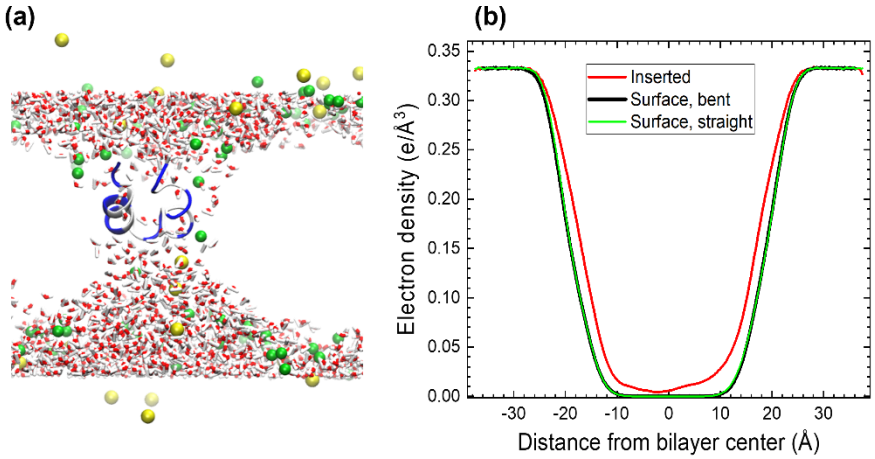


Figure 5.5. (a) VMD visualization of WLBU2 inserted into G(-) IM LMM. Colors: WLBU2, ribbon + licorice (R, blue; V, W, grey;) K^+ ions, yellow spheres; phosphate groups, green spheres; water, red and white sticks. The figure omitted the acyl chains for clarity. (b) Water electron density of G(-) IM bilayer with WLBU2 either inserted (red line) or on the surface in two different conformations (green and black lines (superimposed)).

The MD simulation gives information about the location of the three types of amino acids across the bilayer. To see these visually, we have plotted the groupings of R, W, and V in Fig. 5.6. These amino acid locations were obtained directly from the MD simulation using the SimtoExp program, which calculates density in 0.5 Å slices through the bilayer thickness.[188, 193] For the two surface models, R is closest to the aqueous phase, followed by V and then W, which is closest to the interface between headgroups and hydrocarbon chains (near the carbonyl-

glycerol group). These locations of amino acids follow their hydrophobicity, with W the most hydrophobic.[194] For the inserted model, R stretches the length of the hydrocarbon core and into the glycerol-carbonyl region. The W is near the interfacial region, and V is deep into the hydrocarbon region. Therefore, the amino acid residues in the inserted model do not follow their hydrophobicities. In Fig. 5.6(b), (c), the phosphate + outer headgroup and glycerol-carbonyl components from Fig. 5.6(e) are superimposed on the amino acid density to visualize their relative locations.

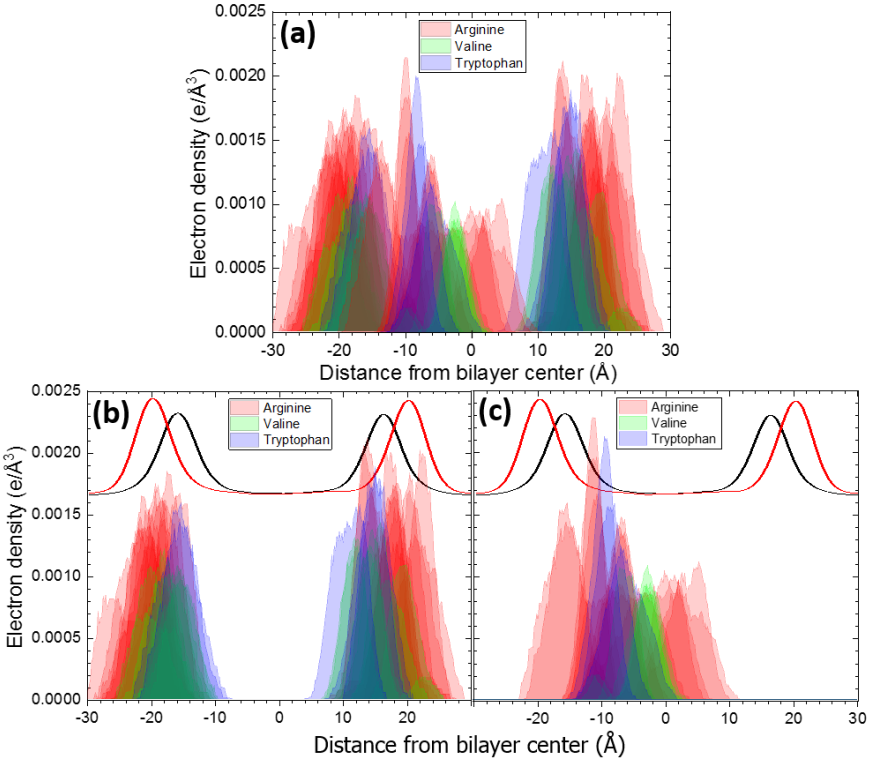


Figure 5.6. (a) The molecular locations relative to the bilayer center of the three types of amino acids in WLBU2 in the G(-) membrane mimic. Fig. 5.6(a) shows the surface and inserted models superimposed, Fig. 5.6(b) shows the two surface models, and Fig. 5.6(c) shows the inserted model. The positions of the phosphate + outer headgroup (red line) and glycerol-carbonyl (black line) are superimposed (lines not drawn to scale) on the amino acid electron density (not drawn to scale). The phosphate + outer headgroup (red line) and glycerol-carbonyl (black line) are superimposed (lines not drawn to scale) on the amino acid electron density (not drawn to scale).

5.4 Discussion

While MD simulation is a powerful tool to visualize molecules and measure molecular distances, equilibration of unbiased simulations to reach an equilibrated state of peptides in the membrane can require >0.8 μs.[36] One way to circumvent this problem is to conduct coarse-grained simulations on millisecond time scales.[36] Then, back-transforming (back-mapping) from coarse-grained to atomistic simulation can be performed to visualize the atomistic configurations of the lipid and peptide molecules at higher resolution.[37] This multiscale approach is valuable and has

yielded many innovative publications.[38-44] Another method to shorten the atomistic simulation time of adding peptides to membranes is to allow the peptide and lipids to self-assemble together into a lipid bilayer with bound peptides.[45] Another method is a steered MD simulation whereby a peptide is pulled into the membrane and allowed to equilibrate.[46] An alternative approach is to compare biased all-atom peptide simulations with peptide-membrane placement to some form of fully equilibrated experimental data. Studies have compared atomistic simulation to small-angle x-ray scattering (SAXS),[47] atomic force microscopy (AFM),[48] nuclear magnetic resonance (NMR),[49] circular dichroism (CD) and NMR,[50] analytical ultracentrifugation and ^{13}C NMR,[51] sum frequency generation (SFG) vibrational spectroscopy and ATR-FTIR,[52] fluorescence microscopy[53] and CD spectroscopy.[54] These experimental techniques can validate the percentage helical content in the peptide (CD and NMR), the amount of aggregation (analytical ultracentrifugation and ^{13}C NMR), the orientation of a peptide in a membrane (SFG and ATR-FTIR), or multimer formation (SAXS and AFM).

For WLBU2 to kill bacteria, it must translocate across the outer membrane, through the periplasmic space, and then perturb the inner membrane in G(-) bacteria. For our outer membrane mimic, NR in Fig. C3(d) indicates that WLBU2 is located only in the hydrocarbon interior. This interior location would facilitate self-promoted uptake[60] of WLBU2 through the outer membrane that must occur for it to reach the inner membrane. When the highly positively charged arginine residues embed deep into the hydrocarbon phase, they cause water to enter.

The outer membrane (OM) has traditionally been considered the major permeability barrier to antibiotics. Conventional antibiotics, such as beta-lactams, are believed to enter the OM through pores formed by porin proteins[61], but generally, only hydrophilic substances less than 600 Daltons can diffuse through the porins. In addition, the OM has an unusually low permeability to

hydrophobic molecules[62]. At the same time, cationic AMPs, like WLBU2, can permeate the OM similarly to cell-penetrating peptides due to their interaction with negatively charged lipid headgroups[63].

Although our study did not address the next step, presumably, WLBU2 then exits the OM and enters the periplasmic space. The periplasm has many functions, including protein secretion and folding, environmental sensing, peptidoglycan synthesis, osmoregulation, resistance to turgor pressure, and sensing and resistance to cationic antimicrobial peptides[64]. Within the periplasmic space is a layer of crosslinked sugars and amino acids termed peptidoglycan, linked to the outer membrane through covalent linkages to the outer membrane lipoprotein. Multicomponent protein complexes span the two membranes, such as the flagellar machine. Due to its constituents, the periplasmic space is osmotically active. A Donnan equilibrium controls the flow of water and ions from the cytoplasm or the extracellular fluid to the periplasm[63,] and the gel-like periplasm might be relatively permeable to small molecules like antibiotics and AMPs[63].

For G(-) bacteria, the AMPs next encounter the inner membrane (IM). Our NR results revealed in Fig. C3 that the bent conformation of WLBU2 finds two locations in the IM: 63% in the headgroup and 37% in the hydrocarbon interior. This dual location may be essential for WLBU2's function of perturbing the membrane and killing the bacteria. The surface states show that R spans the range from 11 to 30 Å, V spans 12 to 26 Å, and W spans 6 to 23 Å from the bilayer center. The outermost position for R is consistent with its smallest hydrophobicity[35]. R spans a region from the proximal monolayer's headgroup to the distal monolayer's hydrocarbon edge in the inserted state. Thus, even when R is in the hydrocarbon region, it has a portion that can bind to the phosphate headgroup. This electrostatic binding must be crucial in anchoring WLBU2 to the lipid headgroup region in the G(-) inner membrane while penetrating deep into the hydrocarbon

region. W has a component at the bilayer carbonyl-glycerol interfacial region at 14 Å when in both the surface and inserted states, which is a second anchor that keeps WLBU2 from fully penetrating the hydrocarbon interior.

G(+) membrane/WLBU2 interaction is similar to that of G(-) membrane, which is reasonable since WLBU2 kills both bacteria efficiently.[16] Our NR results revealed in Fig. C3 that the bent conformation of WLBU2 finds two locations in G(+): 54% in the headgroup and 46% in the hydrocarbon interior. W is again aligned with the interfacial region at ~ 15 Å from the bilayer center. When WLBU2 is in the bilayer interior, R overlaps with the surface R state, thus forming a continuous positive charge across the bilayer. The smaller Vs follow the positions of Rs and Ws. The continuous line of positive charges allows water to enter the hydrocarbon interior, creating a pathway for water and ions to leave the bacterial cell, thereby killing it.

In this work, we have not explored the question of AMP aggregation and how aggregated WLBU2s might interact with the bacterial membrane. Studies of the fungicide fengycin[65] searched for a role for aggregation in selectivity between eukaryotic and bacterial membrane models. Such studies used a much higher lipid: peptide molar ratio than this work since we constrained our peptide concentration to the highest permissible in the XDS experiment (76:1). In addition, fengycin contains a lipophilic tail which could cause aggregation via van der Waals attractive interactions. Similarly, studies of AMP polymyxin E (colistin) found that on planar lipid bilayers composed of LPS/PC, colistin, with its lipophilic tail, induces large-scale clustering as it segregates out LPS.[66] Since WLBU2 is highly positively charged (+13 *e*), it is unlikely that self-aggregation would occur, even when binding to phosphate headgroups. In addition, no MD simulations explored the interaction of WLBU2 with a LMM of the host eukaryotic membrane that contains phosphatidylcholine, phosphatidylethanolamine, sphingomyelin, and cholesterol.

5.5 Conclusions

To summarize, this work reports on using neutron and x-ray diffuse scattering to shorten the time required for all-atom MD simulations of the AMP WLBU2 interacting with bacterial LMMs. By constraining the thickness of the simulated membrane using surface tension to match that obtained by XDS experiments and by constraining the location of the peptide in the membrane to match that obtained by NR, micro- or millisecond simulations are not required. Importantly, equilibration is allowed for by removing constraints on peptide location. One significant molecular result in this investigation is the observation of water at the center of the bilayer when WLBU2 is in the inserted state. Other studies of the KvAP voltage-gated potassium channel[46, 67] and the HIV Tat protein[68] found water with charged amino acid residues and phosphate groups buried in the hydrocarbon interior. In the case of WLBU2, the internal moisture is continuous with the headgroup water due to a dual anchoring of WLBU2 in the headgroup and in the interfacial region, as well as penetration into the hydrocarbon interior. Arginine's binding to phosphate residues, plus tryptophan's location near the bilayer interface, may be critical anchoring mechanisms. WLBU2 causes a small (~1 Å) thinning at a lipid: peptide molar ratio of ~76:1 in both G(-) and G(+) LMMs. An observed increase in APL with WLBU2 resulted in G(-) and G(+) LMMs. Thus, the dual location of WLBU2 in the headgroup and hydrocarbon regions, the presence of water, phosphates, and ions in the interior, the location of arginines at the phosphate, tryptophans at the interfacial region, and the slight bilayer thinning all contribute to membrane destabilization thus leading to bacterial killing.

Chapter 6: 5-HT_{3A} Ion Channel Activation and Deactivation

6.1 Introduction

The 5-HT_{3A} ion channel is part of the cystine loop family of pentameric ligand-gated ion channels (pLGICs) that control essential functions in humans. These ion channels include nicotinic acetylcholine receptors (nAChR) that cause nicotine addiction, gamma-aminobutyric acid receptors (GABAR) critical for sleep, and glutamate receptors associated with attention. The interactions between neurotransmitters and how they induce varied allosteric signals are challenging problems that could lead to advances in treating neurological diseases. Understanding the allosteric pathways is critical to designing proteins to induce specific allosteric pathways. Toxins represent a wealth of evolutionary knowledge to trigger desirable allosteric pathways.

Past work with σ-conotoxin GVIIIA showed competition between the radiolabeled Zacopride and the σ-conotoxin GVIIIA. An alternative explanation is the receptor is allosteric, and it is not clear if σ-conotoxin GVIIIA could be closing the pocket or generating some allosteric signal to eject the Zacopride from the receptor. The response of Zacopride is biphasic, and the more potent phase occurs as low as an EC50 of 32nM[195] compared with an IC50 of 53nM for σ-conotoxin GVIIIA.[50] The IC50 of the conotoxin GVIIIA is a binding metric and does not indicate allostery, but if instead it were a measure of the conductance it would be easier to mimic the method the toxin uses to alter the allostery.

Pathways like these are essential because inhibiting the dopamine D2 receptors in different brain parts are responsible for various functions. The dopamine D2 receptors in one part of the brain reduced agitation; in another, it caused learning and memory issues.[196] A better atomistic understanding of pLGIC could lead to creative new approaches to treating complicated neurological diseases.

The survival of many animals, such as bees, spiders, snakes, snails, and jellyfish, depends on the toxicity of their venoms. Although, toxins are prevalent around the world and account for many accidental deaths annually, they are poorly understood. Studying small peptide venoms will help us to understand and improve the effectiveness of drugs if they reveal new mechanisms of disruption.

One aspect of this study is determining if these toxins have complex alternative allosteric signals. The motivation for this is that some toxins are excitotoxins that generate their allostery to over-activate a receptor. Is there any reason why an inhibitor could not do the same, and could that lead to new ways to treat addiction and diseases? This study has the following three hypothesis. **Hypothesis 1)** Determining the toxin structure and function in the 5-HT_{3A} ion channel is possible. It is crucial to have access to the ion channel complexes of the control, activated, and deactivated ion channels with toxins. **Hypothesis 2)** Achieving an unbiased pore opening with serotonin is possible. **Hypothesis 3)** If the toxin can bind to the serotonin pocket, that will more definitively prove that the toxin competitively displaces serotonin and interferes enough to prevent allosteric activation.

This chapter focuses on conotoxins from cone snails, an enigmatic animal venom group. There are over 8,000 peptides that target many of the most vital biological processes of fish and mammals, including ion channels.[82] They are so enigmatic because they are composed of unusual amino acids like bromotryptophan and hydroxyproline, and because of the high number of disulfide bonds. It is difficult to know how these unique amino acids evolved in the first place. The goal is to use these highly sophisticated peptides to learn how ion channels function and can be modified. The σ-conotoxin GVIIIA deactivated the 5-HT_{3A} ion channel and represents one of the first toxins found to target serotonin neuroreceptors. This specific toxin is somewhat unusual in

that it is 41 amino acids in length, so long that it borders between a short peptide and full protein. During discovery, the toxin proved to have a highly competitive binding coefficient comparable to Zofran (Ondansetron)[50], an anti-nausea medication. The sequence follows: G**C**TRT**C**GGOK**C**TGT**C**T**C**TNSSK**C**G**C**RYNVHPSGBG**C**G**C**A**C**S-NH_2, where B is for bromotryptophan, and O is for hydroxyproline. While antidepressants have numerous side effects, drugs like Zofran that act on the 5-HT_{3A} have almost no side effects.[197, 198] Although Zofran does not treat depression, it improves the symptoms in mild cases. Additionally, it could be combined with other drugs to reduce side effects overall. Using conotoxins for medicinal purposes is a potential application of conotoxins.

The 5-HT_{3A} ion channel is prevalent in the intestines and is essential in triggering nausea. Drugs like Zofran may mimic the effects of σ-conotoxin GVIIIA, assuming no detrimental toxin targets exist. There is evidence that drugs like Zofran improve mood and reduce depression and anxiety in patients.[199] This raises the question of why cone snails benefit from deactivating this specific ion channel. Since fish cannot physically vomit, that leads us to believe that the purpose of this peptide is to calm and relax the fish, preventing it from panicking. The fact that fish do not vomit indicates that there is a good chance that the peptide's target is inside the brain and the peptide passes the blood-brain barrier. The conotoxin was determined to competitively inhibit the 5-$HT_{3A,}$ meaning that the toxin binds to the same region as the serotonin, and there was evidence that the toxin displaced other inhibitors, supporting the competitive inhibition conclusion.

Past researchers have examined the 5-HT_{3A} receptor and proved that serotonin binds to a pocket that opens and closes.[4, 6] MD simulations indicate the pocket is open more frequently than expected. A salt bridge between two amino acids helps close the pocket, and that might help keep the serotonin trapped inside the pocket for several ms before the salt bridge breaks and opens the

pocket, as shown in Figure 6.1. The most recent crystal structure depicts amino acid E209 in a position to interact with the charged end of the serotonin tail, which previously did not orient close enough to interact.

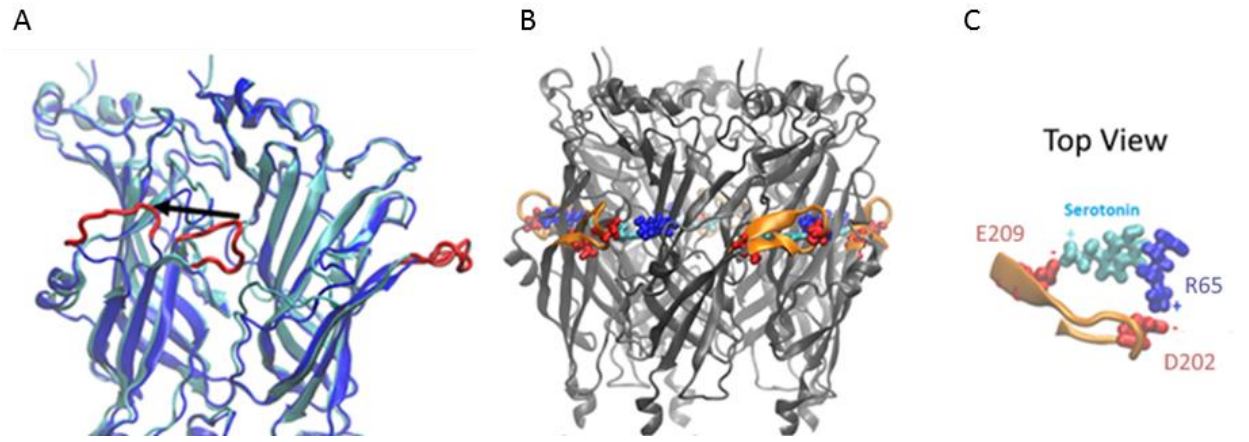


Figure 6.1) The extracellular domain of the 5-HT_{3A} ion channel with the orientation of the serotonin molecule in the binding pocket. A) depicts the well-known flexibility of the extracellular domain (ECD) β-sandwich, which has known two-state conformational changes.[200-202] B) depicts the serotonin pocket, highlighted in orange to show the significance of the ECD β-sandwich conformational changes. C) depicts a more detailed view of the serotonin in the binding pocket.

A significant milestone in understanding allostery has been defining the activated and deactivated ion channel structures and identifying the intermediate states and motions of the channels. This work aims to go one step further by researching the effect of unbiased serotonin binding to probe the protein's natural function and study the toxin's impact on the ion channel.

Another aspect of the research was to model the ion channel pore accurately. Recently, publications depicted open crystal structures of the 5-HT_{3A} channel pore.[203] One problem with the approach is the use of artificial restraints and extended equilibration times up to 300 ns to relax the pore with special restraints. Our approach relies more on the serotonin to maintain the pore in an open state. In one simulation the pore collapsed after 100 ns of restraint free simulations.[203] One of our goals was to test those structures' actual ion channel conductance. The ion channel

conductance data of extracellular ion mixture was available,[204] and point mutations increased the ion conductance. The first simulations tested NaCl solutions rather than the exact ion mixture. One reason is that the pore did not stabilize in the open configuration without careful testing. The highest conductance was around 37-38 pS, and typical voltages were -60 mV to -80 mV, but the aqueous phases usually have a mixture of several ions. Only NaCl systems fit in the scope of the study.

6.2 Methods

Several projects were attempted, including preliminary docking, extensions of the docking with Anton2, and unbiased binding with Anton2. The final part of the project was extending the docking with a more comprehensive ensemble. A comprehensive study first focused on docking the peptide to the ion channel and running unbiased docking simulations. The next round of studies involved generating several new peptide structures using advanced AlphaFold2 settings.

Table 6.1). The docked simulation used an older 6HIN crystal structure, and the control matched the docked simulation. The occlusion simulation used a more recently crystallized APO 6Y59 structure. The unbiased simulations all used the 6Y59 crystal structure.

	Protein Crystallization	Toxins	Lipids	Water	Serotonin	Ions	Time (μs)
Docked	6HIN	1	360	68400	0	POT 201 CLA 176	~15
Occlusion	6Y59	1	360	72000	0	POT 217 CLA 187	~15
Control	6HIN	0	360	68400	0	POT 206 CLA 176	~9.6
Unbiased 1	6Y59 Apo	6	480	110400	24	POT 288 CLA 307	~4.6

Unbiased 2	6y59 Apo	6	480	110400	6	POT 288 CLA 289	~2
Unbiased 3	6Y59 Apo	6	480	110400	36	POT 288 CLA 319	0
Unbiased 4	6Y59 Apo	6	480	110400	18	POT 288 CLA 301	~0.97
Unbiased 5	6Y59 Apo	6	480	110400	36	POT 288 CLA 319	~1.4
Unbiased 6	6Y59 Apo	6	480	110400	24	POT 288 CLA 307	~1.6
Successful Unbiased	6Y59 Apo	12	480	110400	36	POT 287 CLA 348	~1.8

6.2.1 Structure Prediction for Anton2 Work

Robetta[205] and AlphaFold[81] were used to predict the toxin structures shown in Figure 6.2. To ensure sufficient diversity in the simulated structures, we chose all possible structures from Robetta and the top AlphaFold structure. The selection helped to compare the two methods, as shown in Figure 6.2 below. One choice in the modeling process was to replace the bromotryptophan and the hydroxyproline with standard tryptophan and hydroxyproline. In nature, the peptide has these polar amino acids, but no force field is available to determine how they interact in the system.

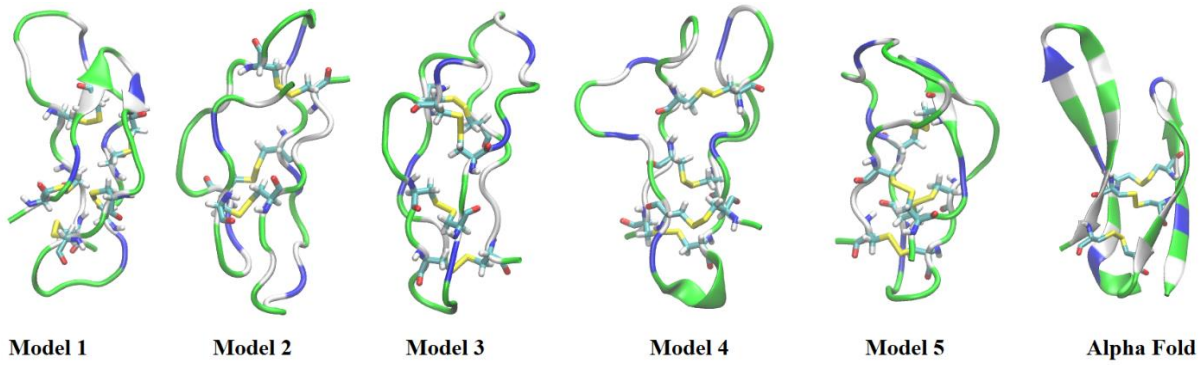


Figure 6.2). The different predicted structures produced by Robetta software with the AlphaFold structure at the end prioritized the β-sheet formation over a disulfide bond.

6.2.2 Structure Prediction for Extended Docking

Advanced settings included using AlphaFold2 with max MSA set to 16:32 to reduce the dependence on known examples and rely more on what AlphaFold has learned. Training is set to true to enable the stochastic part of the model to sample a more diverse set of structures. To prevent Alphafold from crashing, subsample_msa is set to true. The Num_relax parameter was set to "all" and the num_recycles parameter to "three." By setting the number of seeds to 32 and running five models for each seed, it was possible to generate 160 total initial structures. The selection was limited to 160 models because generating more reduces the accuracy by relying too little on the MSA. The ensemblator program[206] accepted 160 models created with reduced MSA as inputs and the initial AlphaFold peptide based on default settings that succeeded in binding unbiasedly. Ensemblator separated the peptides into two groups mainly based on the placement of the first 10-15 amino acid residues. Simulations of both clusters helped understand the impact of the order of these two loops.

6.2.3 Docking Approach

The Alphafold settings created an ensemble of structures for further study. Cluspro docking quickly generated an understanding of how different placements might interface with the allosteric binding pocket. The study used Autodock Vina[207] and ClusPro[208] but concluded using only ClusPro

to dock several peptides 5-HT_{3A} using the VDW+ELEC scoring function. CHARMM-GUI generated systems and NAMD ran the simulations to reach a new energy minimum according to the specific CHARMM and CUFIX force field used.

6.2.4 Simulations of Docked Toxins

MD simulations of conotoxin utilized NAMD 2.14,[23] while also using the CHARMM36 force field for lipids[132] and the CHARMM36m force field for proteins.[131] NAMD employed a 2 fs time step with long-range electrostatics interactions evaluated every other time step using the particle mesh Ewald method.[187] Short-range non-bonded interactions were cut off at 12 Å using a force-based switching function beginning at 10 Å. NAMD maintained the temperature at 37 $^{\circ}$C, and all simulations ran using 0.15 M NaCl. The simulations ran using Langevin dynamics, and pressure was maintained separately in the membrane-planar and membrane-orthogonal direction using a Langevin piston at 1 atm. CHARMM-GUI generated PMFs from snapshots of the 3D coordinates of the 5-HT_{3A} receptor extracellular domain (ECD) and fully solvated the ECD around a water box. The solution box for PMFs was run using the NPT ensemble.

Further studies focused on two docked structures based on ClusPro docking algorithm. The selection was based primarily on the ClusPro score and the desire to sample new configurations. Many new configurations were interesting, such as peptides blocking only parts of the pocket. The purpose of studying the ensemble was to focus on alternative docked structures and discover what other mechanisms were possible. The docked trials were then used to study the free energy of binding using a PMF. The literature value compared favorably to the PMF results. It was also not possible to know how successful different chosen peptides would be until after some preliminary testing. There could certainly be alternative binding peptides from both clusters that compared well to the literature value.

6.2.5 Free Energy Calculations

Initially, we tested a PMF approach that allowed for a flexible protein. The flexible protein approach worked fine as long as there was no unaccounted drift between the receptor and the protein. The Anton2 docked peptide benefited from having a more flexible receptor. For the other trials, fewer segments were unrestrained. NAMD restrained three amino acids in the pocket to prevent drift. An accurate PMF requires a uniform and controlled reaction coordinate. The flexible PMF approach initially inflated the PMF values. Additional restraints to the pocket seem to minimize drift. However, some structures had such low ΔG values that it was not worth re-running them with the improved fixed receptor. Anything higher than -2.4 kcal/mol was from the first round and re-processed. The input distance into WHAM was greater than the proper distance because the peptide would reach the target distances without being removed from the receptor. The high levels of variability identified this. The energy ranking eliminated several docked peptides during this phase of the PMF process for having PMF values higher than -2.4 kcal/mol.

The final binding free energy of the ligand-protein interactions was studied only on the two systems with free energies lower than -2.4 kcal/mol using a PMF with WHAM and NAMD.[22, 48, 141, 142] These final PMFs used a rigid receptor to prevent drift between the references as the toxin is being pulled off with the peptide and creating a non-uniform reaction coordinate. The PMF was implemented with the standard Grossfield[150, 209] approach using the centers of mass to track the positions for the spring to calculate the force.

6.2.6 Anton2 Simulations

The Anton2 portion of the work involved running full ion channel simulations in an NPT setup and a membrane mixture of POPC:SDPC:Cholesterol in a molar ratio of (7:7:6) with approximately 145 waters per lipid. The exact systems and compositions are shown below in Table

6.1. These simulations differed from the PMFs run in NAMD because the barostat was not Langevin but controlled by the Martyna, Tobias, and Klein (MTK) barostat.[116] The difference is that the Langevin barostat is more stochastic. The MTK algorithm more directly couples to the external bath. The simulations used the ACP and CNEU group tags and GLYP and CT2 for the peptide to match the exact terminals in the literature.[50] Although the exact size of the simulations varied, approximately 240 lipids were on one leaflet.

6.2.7 Asymmetry Measurement

The asymmetry measurement was calculated by first identifying the two segments of the core ion channel that were closest together and furthest apart. VMD calculated the distance between these two segments in each frame. VMD identified protein atoms with z coordinates >45 to measure only the upper one-third of the ion channel to get the residues of all amino acids above that selection. The selection was hard coded into the VMD distance routine to prevent the residues from changing between different runs or throughout the simulation.

6.2.8 Conductance Measurements

The ion channel conductance studies involved setting up an equilibrated membrane system and converting it to a rigid box to maintain a constant electric field. The NAMD units for the electric field are kcal/mol/Å/e, equivalent to 0.0434 V/A. NAMD applied restraints to the pore to prevent it from closing because it was unstable. Several simulations ran before we identified the proper protocol to minimum voltage while still observing ionic passage through the pore. The value of -0.015 kcal/mol/Å/e or ~ -40 to -150 mV in a 240 Å box. The voltage is reasonable because 40 mV is actually used in wet lab experiments, 150 mV might be on the higher end but it is close enough to extrapolate if needed. Even if 150 mV is too high the bigger point is that conductance has never been measured before and others can now follow this approach in the future. Point

mutations to the intracellular side were required to increase the conductance to observable levels. The simulation used physiological ionic strength of 0.15 M NaCl. The conductance is low overall, so point mutations, higher ionic strength, and slightly higher voltage were used to accelerate the ion counts and lower computational time. The goal was to determine if the ion channel is truly in the activated form and, if so, what conductance that corresponds to. The ion conductance without serotonin assistance used harmonic restraints on the entire ion channel, including the pore with 3kcal/mol/Å^2. The QDA point mutations resided on the intracellular side of the membrane and adjusted the ion conductance. The QDA mutation precisely corresponded to R416Q, R420D, and R424A in the 6y5A protein. The open serotonin pore ran using more minimal restraints of 0.04 kcal/mol/Å^2 on the NH3 of the serotonin and the disulfide bond of the 5-HT_{3A}. NAMD used a standard CHARMM36m force field for this work and did not implement CUFIX.

6.3 Results

By better defining where the conotoxin binds to the 5-HT_{3A} receptor a better understanding of the conotoxin function and MD force field performance can be evaluated. Additionally, there is a chance that the conotoxin could have allosteric importance to treat diseases in new ways. This hypothesis was not conclusively proven or disproven. Even though the toxin binds with similar strength to known inhibitors it is difficult to compare the two without measuring the conductance of the channel directly. To study the allostery, we predicted the bound state of the toxin and its stability with a docking/MD simulation relaxation approach. The docking approach allowed us to probe if the toxin binds to the allosteric binding pocket and induces its allostery. The PMFs compared the likely bound state against the previously measured values. Finally, VMD visualized the structural changes induced by long-timescale binding.

6.3.1 AlphaFold Structure Prediction

The AlphaFold2 software generated an ensemble of structures that bind the peptide to the 5-HT_{3A} ion channel. There was one main structural difference that resulted in two main clusters that appear in Figure 6.3. Ensemblator identified a difference in the toxin structures' first 1-10 or 1-20 amino acids. AlphaFold appeared unsure of the structure between 1 and 15 amino acids that can play an essential role in toxin binding.

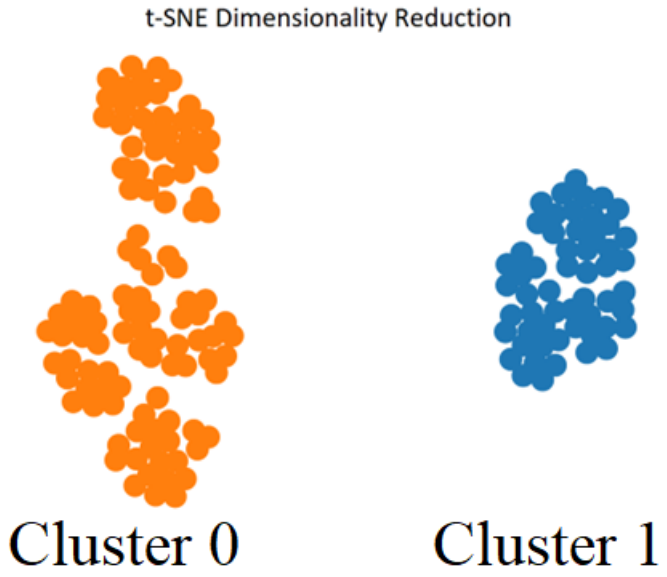


Figure 6.3) Shows the t-SNE dimensionality reduction on the Alphafold2 protein structures. This figure represents each protein as a dot; the color corresponds to the cluster. The docking studies attempted to sample both clusters, but only cluster 1 had strong binding. Two distinct clusters differed primarily in the first 13 amino acids, which can play an essential role in how the toxin contacts the 5-HT_{3A} receptor. The clustering program first aligns the peptides over a conserved region generated using an auto-defined 7.14 Å cutoff before further processing to sort the structures.

The clustering indicates that the differences in the first 10-15 amino acids could be significant because they appear to alter the binding surface. The structure may have multiple configurations. Additionally, it was interesting to identify a consensus on the shape of the peptide toxin. In general, AlphaFold2 predicted something similar to what appears in Figure 6.4. Figure 6.4 placed the two unusual amino acids (that would exist in nature, but we replaced them in the

simulation) at the top of the toxin, including the bromotryptophan and the hydroxyproline. The predicted binding orientations indicate that these unusual amino acids do not interact directly with the pocket. Instead, they may cause a unique folding because these amino acids start hydrophobic and become polar after posttranslational modifications.

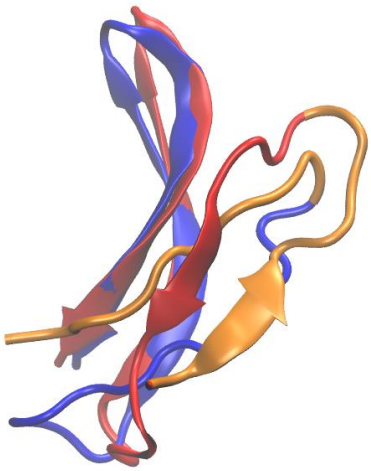


Figure 6.4) The two clusters (one red and one blue) overlaid on top of each other. Orange highlights the region with the most difference in the overlay.

6.3.2 Overview of Docking and Unbiased Binding Results

This chapter considered both unbiased binding and biased docked structures. The unbiased structure is shown below in Figure 6.5. In general, cluster 1 from Figure 6.3 might look similar to Figure 6.5, with two charged amino acids attached to the red loops of the protein that interfaced with the receptor. From the AlphaFold work, these loops have greater uncertainty in the fold. The unbiased binding peptide had the N-terminus of the toxin (shown in red in Figure 6.3) inserting and directly interfacing with the allosteric pocket, compared with loops facing away from the pocket in some docked peptides.

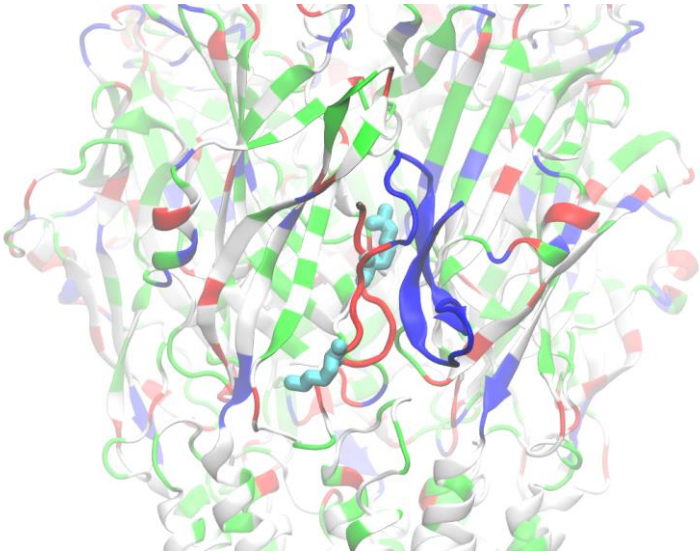


Figure 6.5) Shows the unbiased bound toxin in red and blue. The red section is the first 15 amino acids. The placement of the red section of the amino acids was uncertain and had two main clusters. The cyan color depicts the charged arginine and lysine amino acids. In the unbiased binding, that interaction with the receptor appeared important.

However, the top-docked peptide had uncertainty in the structure that faced towards the receptor, as shown in Figure 6.4. Further studies of the different configurations could help to understand if the loop plays a more significant role. It is clear from Figure 6.6 that both docked and unbiased binding poses can competitively block the serotonin from being inserted. To better understand the predicted structures, a t-SNE was applied to visualize the two clusters of toxin structures below.

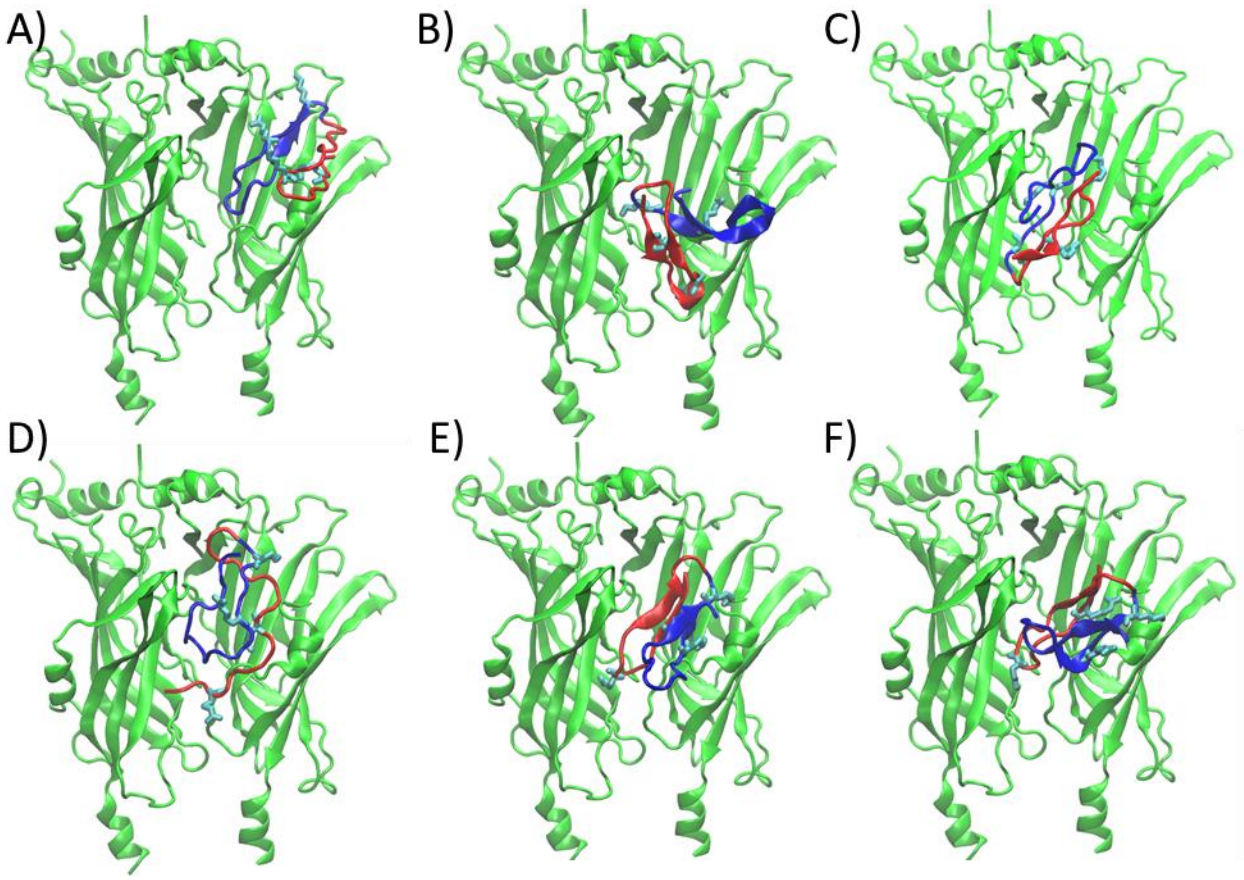


Figure 6.6) Alphafold docking study A is M1_C0_1, B is M3_C1_1, C is M3_C0_2, D is M3_C0_3, E is M0_C1_2, F is M0_C1_4

The Alphafold structures ran for 120-320 ns. An Anton2 docked structure ran for 15 μs and was studied to reveal that the toxin remained stable for the entire simulation length.

6.3.3 Conotoxin binding free energy

The thermodynamic binding free energy indicated that competitive blocking occurs. Past literature researchers calculated a Ki value by making assumptions about the mechanism of peptide binding. The Ki value is essentially an equilibrium coefficient that corresponds to the free energy of binding. ClusPro generated over 20 docked peptides, but fewer than 10 progressed to further analysis. One docked structure ran on Anton2 to determine if special allostery resulted from the toxin's binding.

A subset of four docked structures equilibrated for 240 ns to observe the final configuration. The docking exercise helped come up with several alternative binding arrangements. The most unexpected or unpredictable structures in Figure 6.6 include panels A, D, and F. Panel A is bound much higher and prevents the serotonin pocket from closing. Keeping the pocket open could still have an inhibitory effect. Panel D shows a much less folded structure that could still effectively block the pocket. Finally, Panel F shows that the toxin could bind on the side, allowing the unusual amino acids to interact with the allosteric binding pocket.

6.3.4 Thermodynamic Binding Free Energies

Several toxins rearranged over the first 120 ns of simulation time. Cluster 1 and Pose 29 rearranged significantly from the sideways pose to a more vertical pose. Cluster 0 Pose 1 also rearranged and moved closer to the β sandwich. We gleaned useful insights from these examples of rearrangements from the docking. The 15 ms docked peptide is essential to the results because this binding was less salt-bridge driven and more polar and hydrophobic. The results in table 6.2 indicate an unrealistic increase in the binding free energy of the Anton2 toxin. All of the trials were equilibrated for 240 ns with CUFIX, but the Anton2 docked trial had 15 μs of prior equilibration. The unrealistic free energy of the Anton2 docked pose suggests that the hydrophobic and polar interactions were more mismatched for that given pose. There were more hydrophobic contacts compared to electrostatic salt bridges in the Anton2 docked trial compared to other trials. The recent CUFIX improvements in electrostatics make the other (not Anton2 docked) poses agree well potentially because they are more electrostatically driven and best utilize the CUFIX improvements. Still, the force field may not be perfectly agreeable if the dominant binding is hydrophobic or polar. Free energy is the ultimate measure of how close a given docked pose is to

the experimentally observed one. The docked poses were relaxed for around 240 ns with CUFIX to reach equilibrium as shown in Figure 6.7.

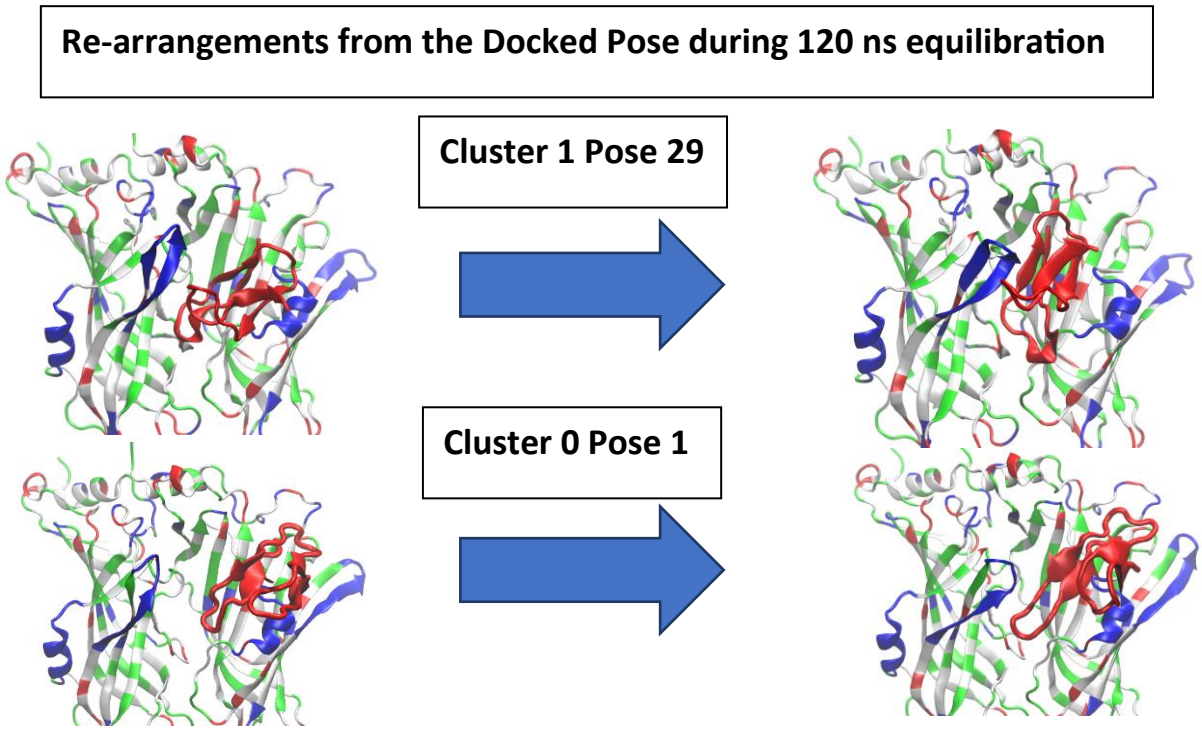


Figure 6.7) Shows the rearrangements of two simulations after running in the MD simulation to reach a new energy minimum. Some simulations interacted weakly even before the final round of PMFs. The final round of simulations involved another 120 ns of relaxation to end with 120-320 ns of equilibration. This additional equilibration time re-arranged some simulations so that only the M0_C1_4 was left.

The binding free energies of docked binding poses appear below in Table 6.2. The unbiased binding pose is within the experimental error provided by the Ki. The electrostatically driven (not the Anton2 docked) docked binding pose is close and might represent one of many alternative binding poses. An equal number of docked poses were sampled from both clusters, albeit the sampling was still low overall and is not enough to rule out potential roles for Cluster 0 completely.

Table 6.2) The Gibbs free energy of binding the peptide to the ion channel receptor. All of the simulations used the C36m with CUFIX force field. M# stands for the model, C# stands for cluster. The number helps differentiate the score of the model in the docking protocol. More than one binding pose came close to the literature value of -11.6 kcal/mol, and one of them appeared to have strong hydrophobic contacts that resulted in unreasonable binding.

Trial	ΔG (kcal/mol)
Literature Goal	-11.6 ± 0.74
Unbiased Cluster 1	-12.2 ± 0.02
Anton2 Docked	-31.9 ± 0.02
M0_C1_4	-9.9 ± 0.02
M3_C0_2	0.0 ± 0.02
M0_C1_2	-1.5 ± 0.02
M1_C0_1	-2.4 ± 0.02

6.3.3 Possible ECD Conformational Change

Even though the docked Anton2 simulation had an unrealistic binding affinity, this simulation showed an interesting large scale conformational change in the extracellular domain. The conformational change appears in Figure 6.8. Unfortunately, verifying the validity of this conformational change with the unbiased binding poses is problematic because it took several microseconds to achieve this. At the same time, the unrealistic binding affinity indicates that more work is needed to improve the parameterization of protein-protein interactions. This binding pose had some hydrophobic overlap, which might play a role in the unusually high binding affinity.

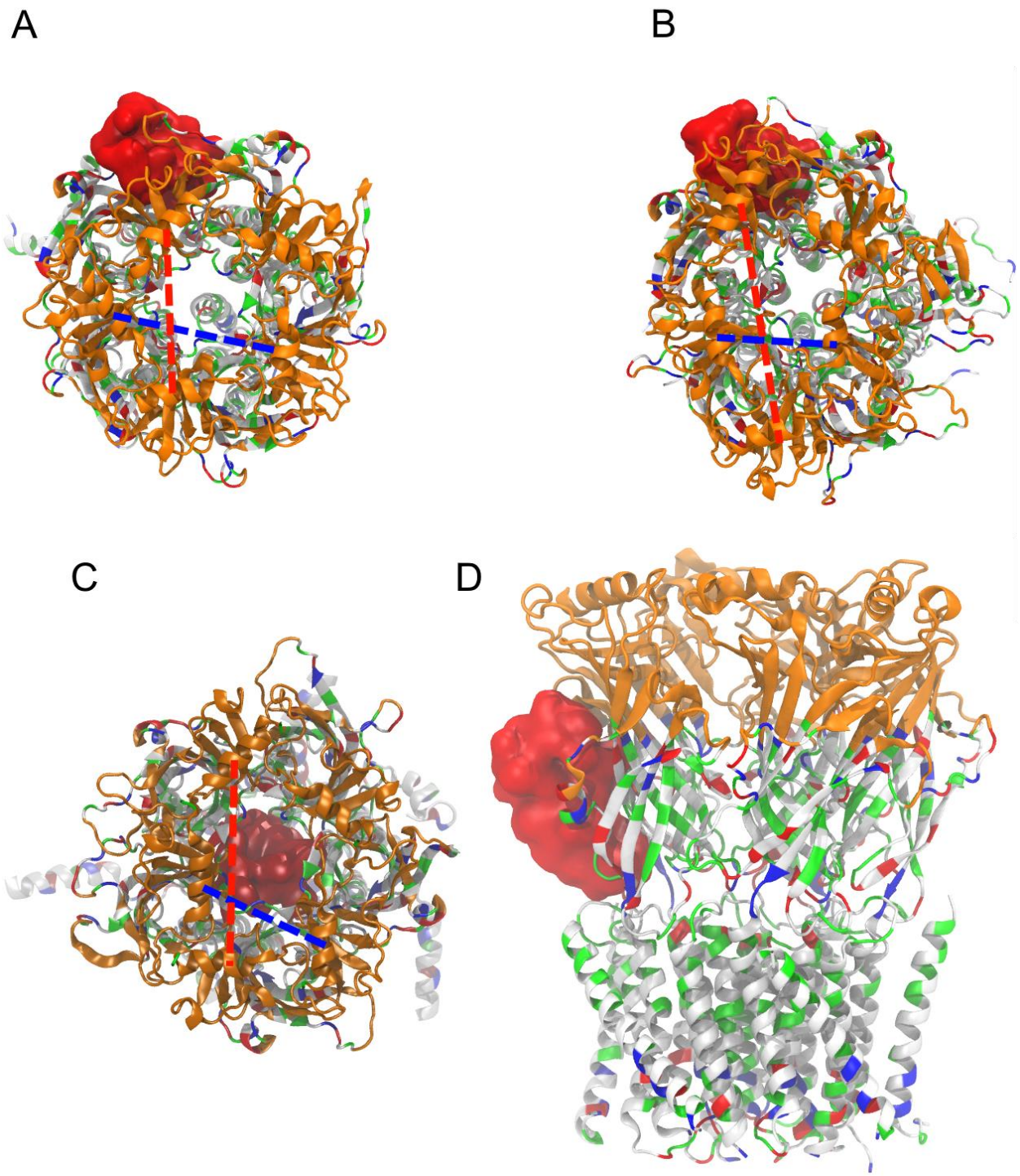


Figure 6.8) VMD generated asymmetry measurements by calculating the distance = B-A of the upper 1/3 of the protein ECD. In the end, VMD selected the upper 1/2 of the ECD for the calculation, and this is to give an idea of how the asymmetry was defined. Panel A shows a normal ECD that is unperturbed. Panel B shows a perturbed ECD with the toxin at the allosteric site. Panel C shows the perturbed ECD with the toxin in the pore. Panel D shows a side view of the ion channel.

One of the main motivations of the work was that the toxin would cause allostery beyond what was previously known. This could be a great hypothesis, but depends on confirming exactly where this toxin binds which is difficult because there are no crystal structures. We must rely on the thermodynamic data and the force field. One of the Anton2 simulations showed interesting asymmetry compared to the control trial. We wanted to investigate the asymmetry induced by the ligand binding. We calculated this by comparing domain distances as defined by section 6.2.7 and comparing the difference in the distances between the closest and furthest domains. The asymmetry data indicates that just one toxin could be enough to disrupt the ion channel as shown in Figure 6.9. However, this could also be the result of a force field problem where the toxin bound too tightly and altered the ECD via another mechanism. The result suggests an ECD allosteric effect due to ligand binding. If these results could be more robustly confirmed, they could indicate the involvement of additional factors beyond pure thermodynamics, such as an extra disruption efficiency. Different poses might disrupt the toxin more or less than other toxins.

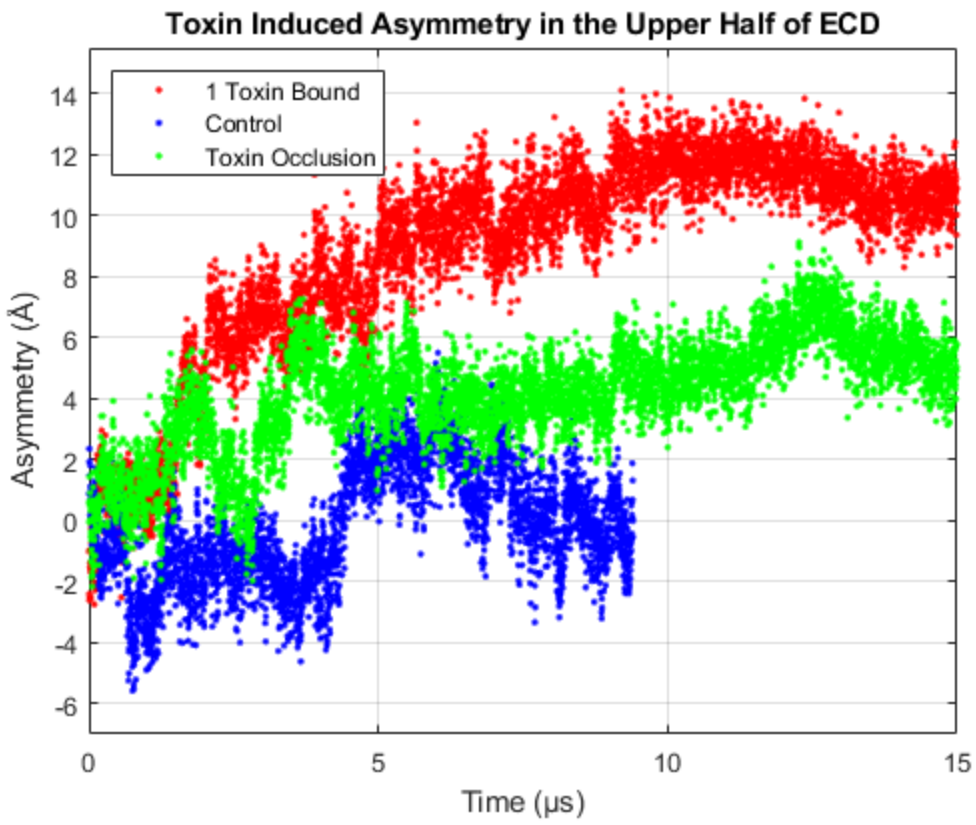


Figure 6.9) Asymmetry of the 5-HT_{3A} ion channel increases when a toxin binds to the allosteric domain.

6.3.4 Effects of Occlusion

An alternative mechanism of inhibition is peptide occlusion of the 5-HT_{3A} pore. We wanted to test all possible mechanisms. This binding pose was run for 15 μs and such structures do reduce the ion conductance of the pore. We did not run an exhaustive analysis of the conductance because the pose no longer seems as likely because the PMF of the conotoxin bound to the allosteric pocket agreed so favorably. Toxin blocking the allosteric pocket makes so much more sense because the free energy agrees and could occur instantaneously. The PMF of the occluded toxin would include breaking the ion channel complex apart which seems complicated. One of the unbiased simulations shows a toxin interacting with the opening of the 5-HT_{3A} ECD which could count as occlusion.

The simulation depicted in Figure 6.10 ran on Anton2. In this simulation the ECD of the 5-HT_{3A} was also perturbed by 3 Å on average. The control appears average around 0 Å of asymmetry. It would be interesting to know if the asymmetry is linked to some kind of overlapping perturbation of the structure.

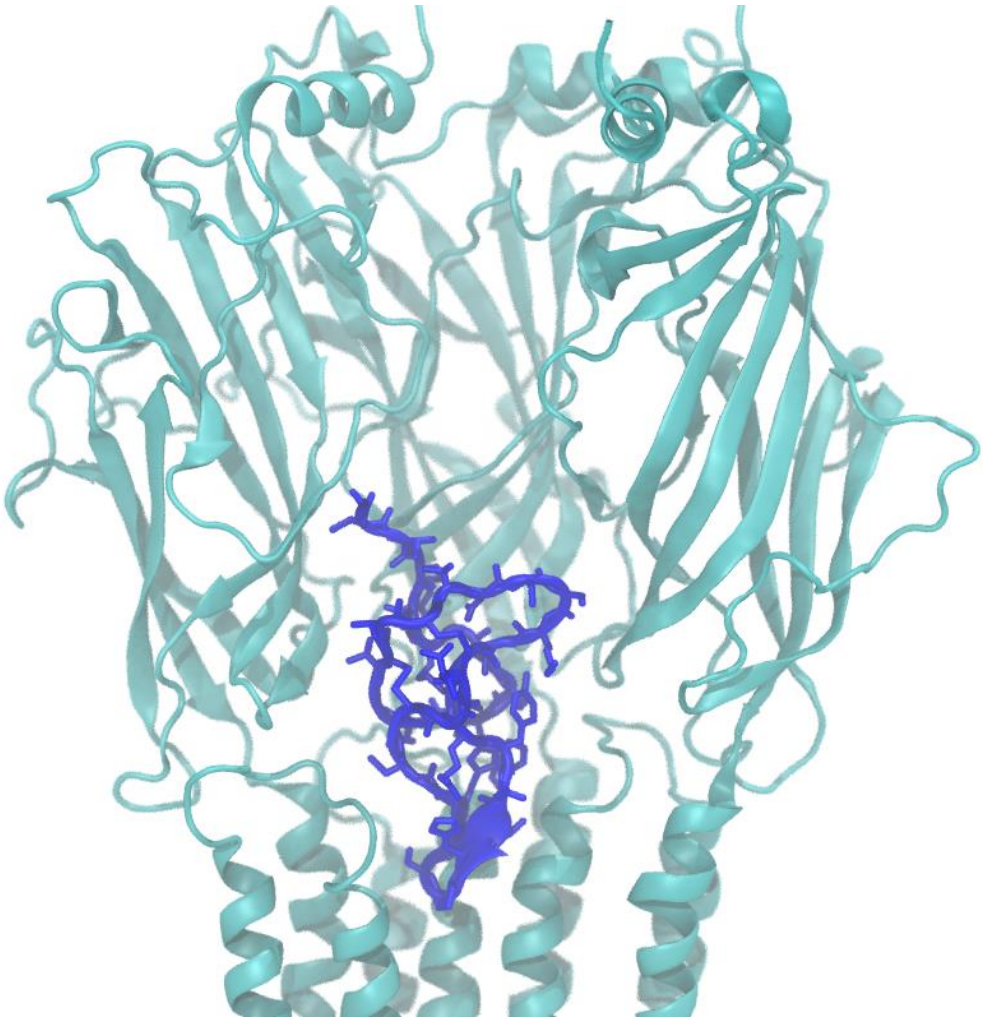


Figure 6.10) An image of the occluded 5-HT_{3A} ion channel complex. This complex perturbed the ECD on average 3 Å but the maximum deformation was almost 8 Å. Such perturbations speak to the allosteric nature of the channel even if they may not be the primary mechanism of deactivation.

6.3.5 Competitive allosteric toxin blocking.

Unbiased Binding of Conotoxin Peptide Compared with Docked Models to 5-HT_{3A}

Anton2 generated several microseconds of simulation time to obtain the unbiased binding events where the peptides diffused around the simulation box. We only identified one peptide impartially bound to the serotonin binding pocket. In the successful binding approach shown in Figure 6.11, the peptide bound to the membrane first and then transitioned to the binding site from the membrane rather than from the aqueous phase. The successful binding approach of the peptide to the desired allosteric binding site supports the surface contacting mechanism, which is a prerequisite to understanding the allostery of the protein. A time series progression is shown below, with the peptide starting from the membrane surface and diffusing into the 5-HT_{3A} binding pocket. The membrane appears to have played a critical role because the peptide binds first to the membrane before attaching to the ECD. The binding sequence appears in Figure 6.11.

Additionally, the ECD appears covered in negatively charged amino acids, and the peptide eventually binds to a specific pocket. The overall binding appears initiated by first binding to the membrane, which is much more plentiful in the cell than the 5-HT_{3A} receptors. The membrane plays a key role by limiting the direction of approach to the 5-HT_{3A} binding pocket, thereby increasing the selectivity of the toxin. An aqueous system might favor off-target binding interactions over the functional allosteric site.

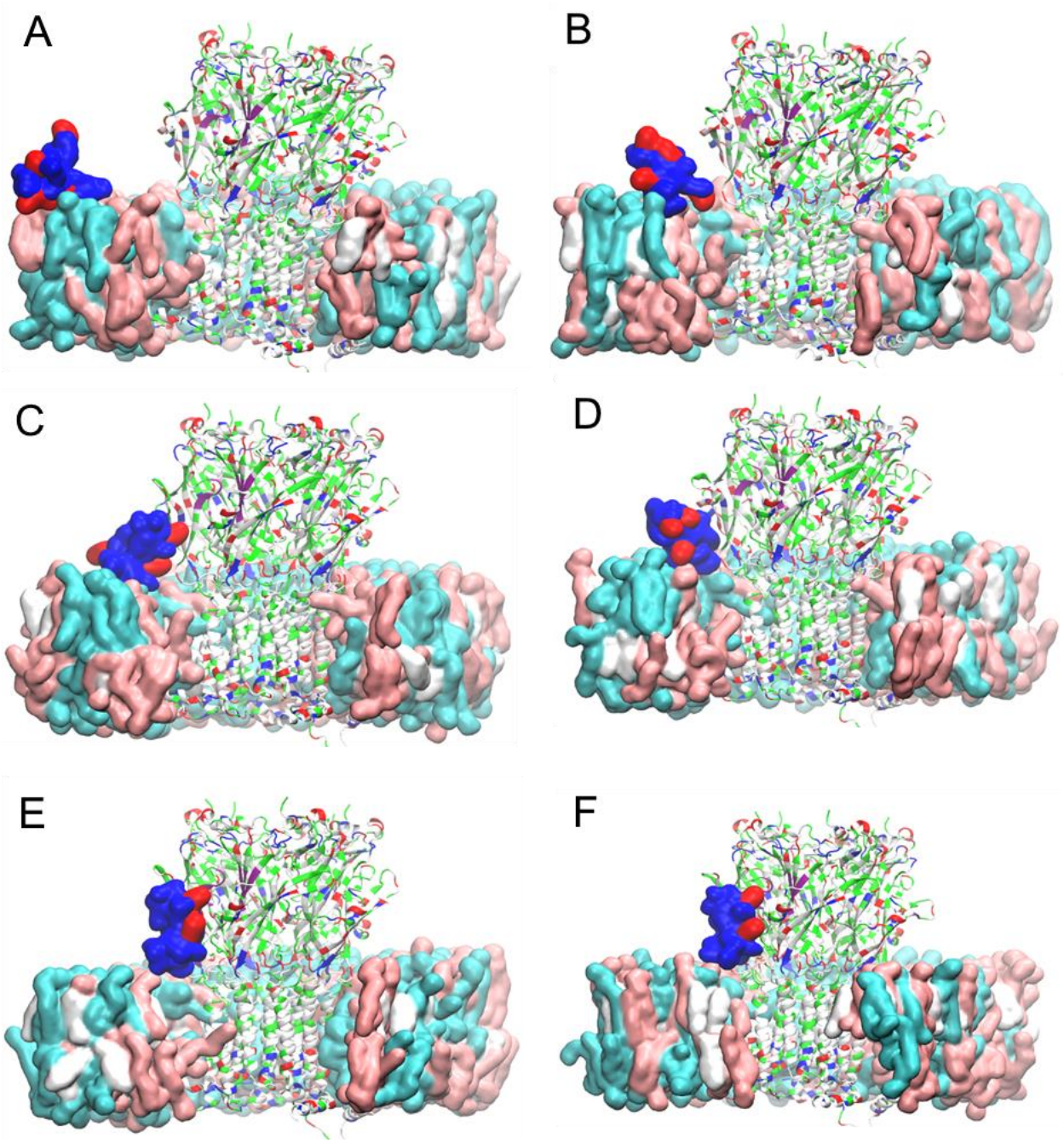


Figure 6.11) Shows the time series of the blue (not electrostatic amino acids) and red (electrostatic amino acids) peptide attaching to the ion channel receptor unbiasedly. The cholesterol is white, and POPC and SDPC are cyan or pink. The Δt means we started part way through the simulation right before it bound. It took much longer than 170 ns to bind. A) Δt=0 ns, B) Δt=2.4 ns, C) Δt=4.8 ns, D) Δt=7.2 ns, E) Δt=9.6 ns, F) 168 ns makes first contact in panel E.

The unbiased binding resulted in approximately the same placement as the docked poses shown in Figure 6.12 below and indicated that the unbiased binding position would mimic any findings generated from the docked results.

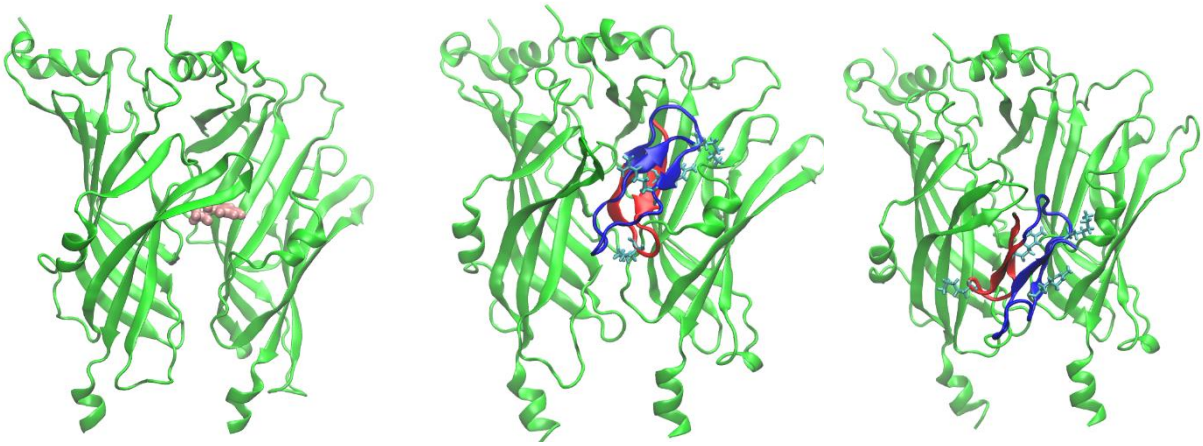


Figure 6.12) Shows the similar placement of the docked and unbiased peptides compared to the serotonin position. Red is the docked position, and blue is the unbiased binding position. Unbiased is on the farther right lower placement. The cyan amino acids show the charged lysine and arginine placements that likely play an essential role.

Figure 6.12 shows that there are some differences in the binding position of the toxin. Unbiased binding pathways will likely start from the membrane and transition to the receptor resulting in the toxin being placed lower and closer to the membrane. Over time the toxin could "crawl" up the receptor to end up populating a higher position similar to the docked placement.

6.3.6 Unbiased serotonin movements alone open the pore

One of the objectives of the unbiased simulations was to observe the competitive interactions between the toxins and the serotonins. Even though it was not the main focus, a

significant amount of serotonin data was available. During the one successful unbiased binding simulation, the toxin bound to a pocket with a serotonin bound shown in Figure 6.13. The serotonin was displaced, and the toxin remained bound on top.

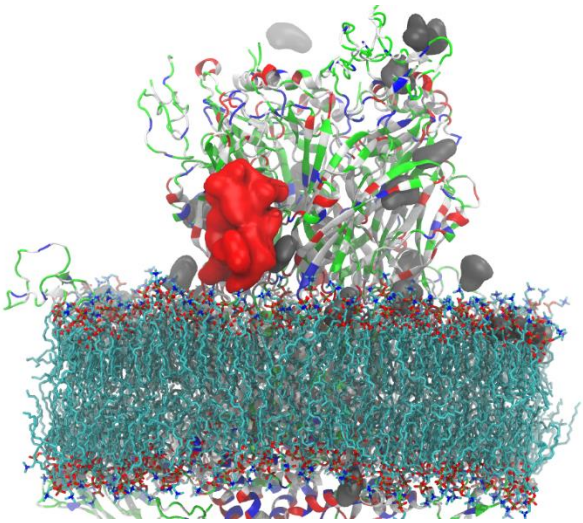


Figure 6.13) Shows the serotonin bound in the pocket and the toxin binding on top of the serotonin

The simulations generated several poses of serotonin and demonstrated the competition for the binding site. The serotonin binding pocket remained partially open for the majority of the pockets. The configurations also changed spontaneously, indicating that the serotonin naturally re-configures in the pocket and binds and unbinds. The principal components of serotonin, shown in Figure 6.14, were used to study and cluster the conformations. VMD also tracked the time series of the serotonin to determine which 5-HT_{3A} ion channels had the most interactions.

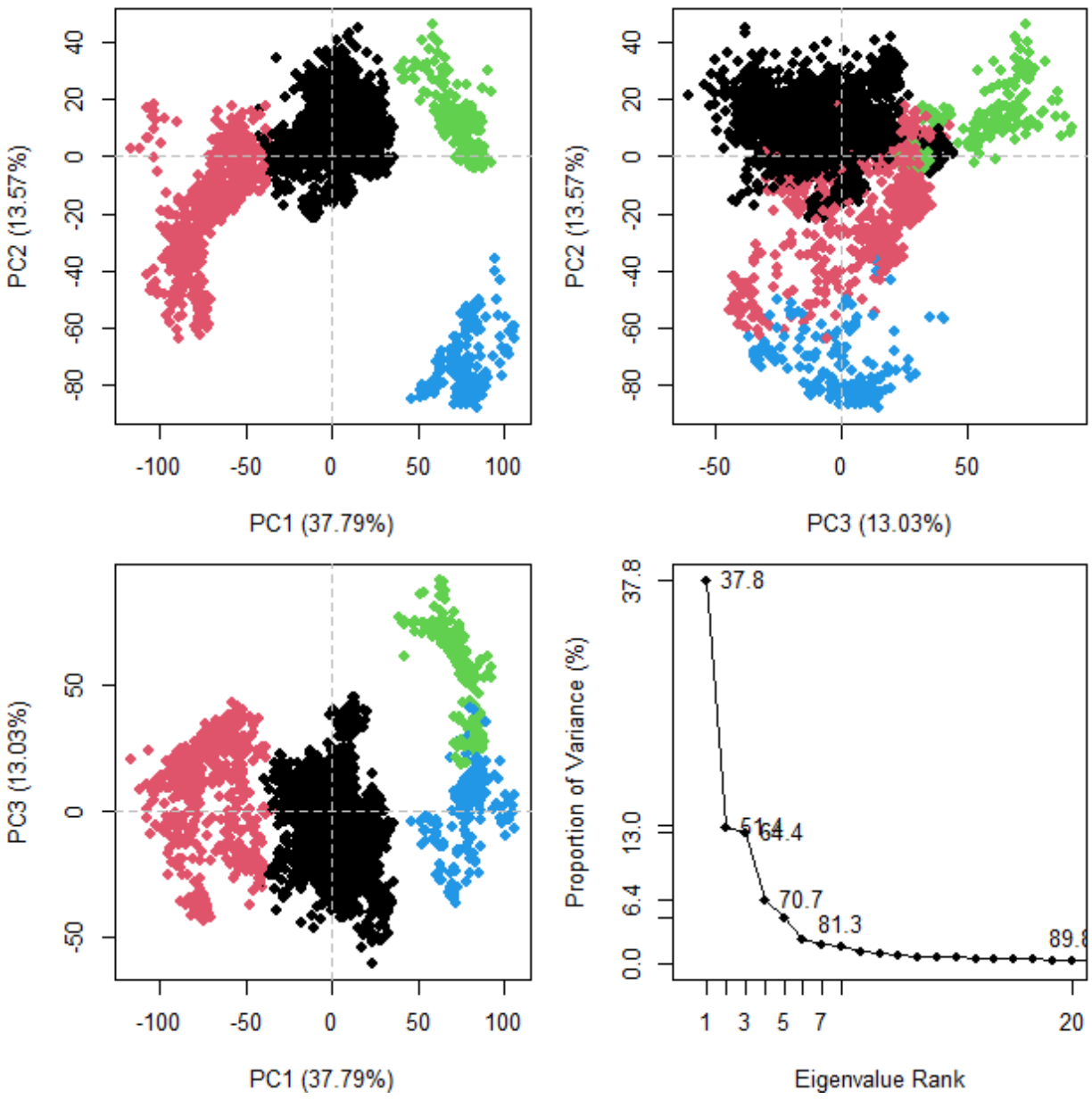


Figure 6.14). The principal components of serotonin indicate the serotonin molecules have differences in orientation and coordinate position. They were aligned and superimposed so that differences in the X, Y, and Z coordinates made a difference in the clustering, and technically, parts of the pocket also had an impact. Figure 6.15 does not show the differences in the pocket.

The serotonin clustering identified four main serotonin clusters and these are depicted in Figure 6.15. The most populated clusters had the least semblance of the crystal structure serotonin. As shown in panel D, the serotonin conformation resembled the crystal structure serotonin, indicating that interaction was accessible. However, it is unclear if the serotonin interactions shown

in Figure 6.15, especially panels A and B, where the serotonin resides in the lower sections of the binding pocket, affect the channel opening.

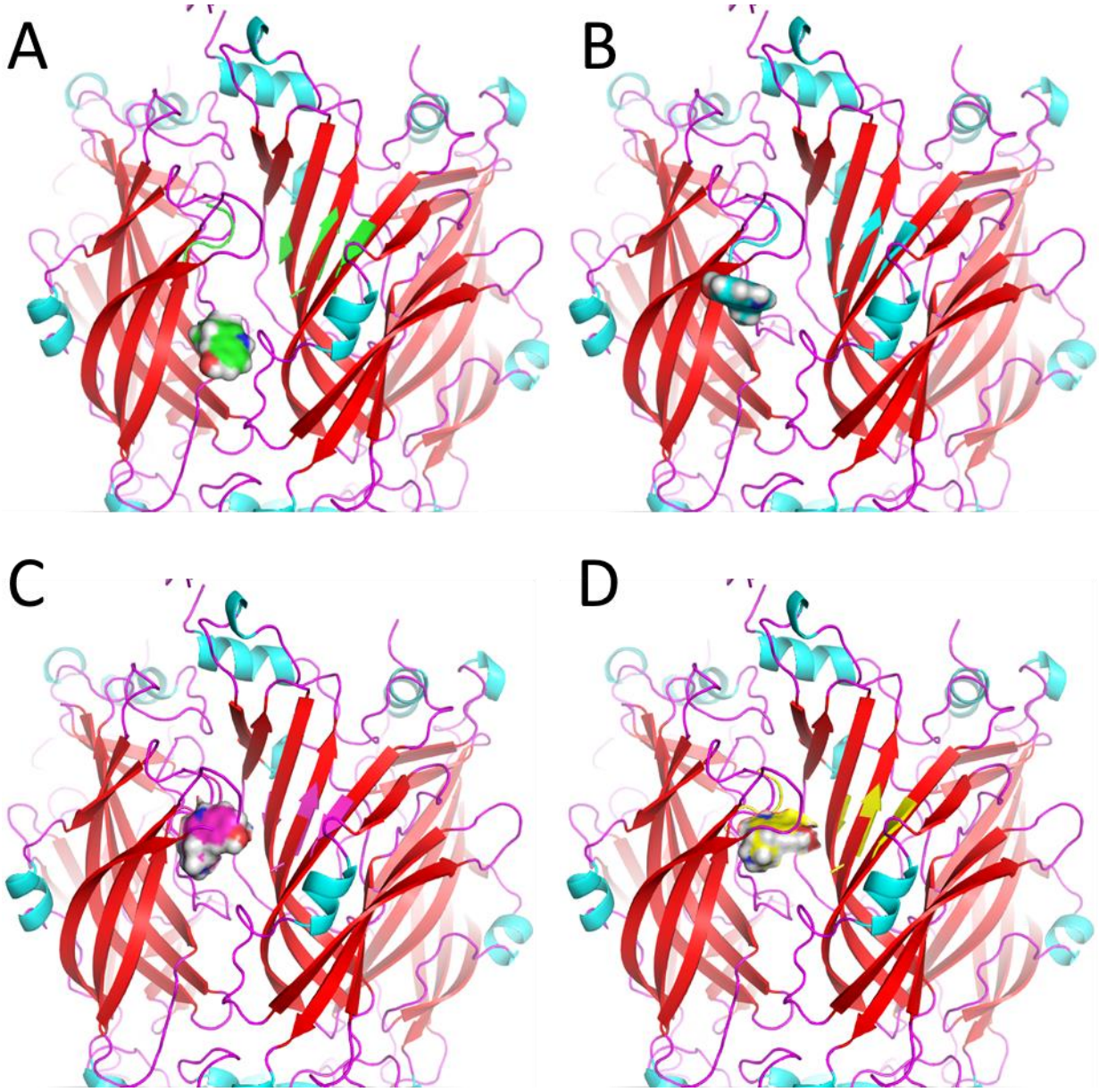


Figure 6.15) Visual representation of the serotonin in the binding pocket. The more populated serotonin binding sites were in panels A (80%) and B (13%) because tighter binding into the pocket requires the pocket to be open and accessible and the right enthalpic bonds to form to hold the serotonin in place, while a looser binding just out of the pocket can likely occur with or without the strong bonds. Panels C and D were much less populated with 3.8% in C and 3.5% D.

6.3.7 Ion Channel Pore Opening

Simulations of the 5-HT_{3A} helped better understand how the serotonin might interact in the different pocket sections. Based on the open configuration, snapshots of the Apo structure briefly filling with water were also identified, shown in Figure 6.16. Simulations from an Apo 5-HT_{3A} ion channel crystal structure bound many serotonin molecules. Even though the crystal structure started in the Apo closed state it would briefly fill with water for a few ns before de-wetting. Additionally, water filling is not the same as ion conductance which is the function of the ion channel. At one point, the serotonin molecule entered the pore from the intracellular side. It would be interesting if the serotonin entered the pore because the serotonin molecule is charged and technically counts as ion transport. The serotonin entered the pore after 561 ns of simulation time. The serotonin left the pore after 991ns.

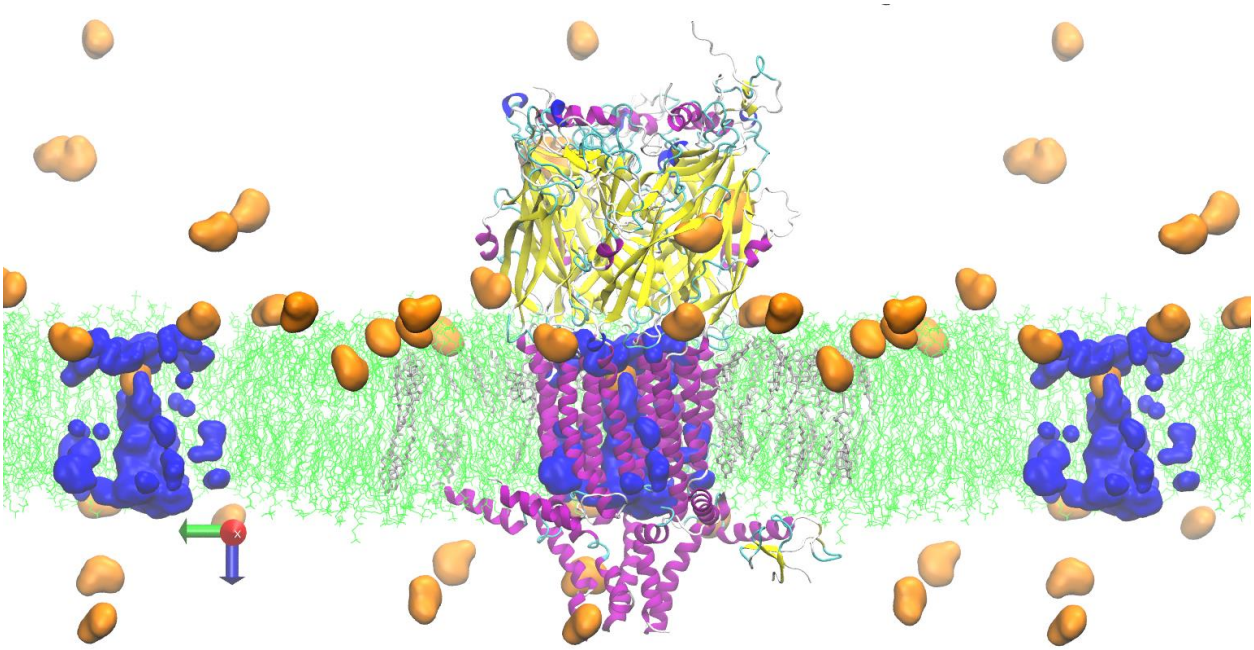


Figure 6.16). This figure shows how even the apo crystal structure of the 5-HT_{3A} can fill with water. There are several times where the ion channel partially fills with water, but the majority of the fillings occurred around 0.7 µs which is where this snapshot is taken. The 5-HT_{3A} is color coded so that yellow corresponds to beta sheet, purple corresponds to alpha helix, and turquoise corresponds to random coil. The membrane is primarily green with cholesterol colored grey. The blue in the membrane is the water that has filled the ion channel pore. The orange molecules represent the serotonin diffusing throughout the simulation box. The ion channel is only displayed in the center of the image while mirror images of the membrane and serotonin and water filling the ion channel are shown on the sides. The mirror image of the protein was not shown to highlight the water filling the ion channel. This channel has not undergone the conformational changes to stabilize the pore in the open conformation, but it still opened the pore.

6.3.8. Ion Channel Conductance to Model Open Pore Properties

The serotonin bound crystal structure of the ion channel was used to study conductance. These simulations continued the crystal structure trajectories using serotonin force field. Previously, researchers proposed extended equilibration time including 300 ns to relax the pore with restraints in order to simulate the 5-HT_{3A} with an open pore.[203] However, our approach limited equilibration time of the pore to around 10 ns and relied more on the interactions of the serotonin molecule in the pocket to maintain the pocket open. These simulations ran longer than 0.5 µs to ensure there was enough data and were equilibrated long enough to ensure that the pore would not

collapse. There were restraints on the serotonin and the disulfide bond of the ECD used to keep the serotonin and the ECD centered in the box. The simulations proved that it was possible to study conductance at similar voltages and ion concentrations as are found in the literature. Although no experiments were found using only NaCl ions, this is a big stepping stone for proving that it is possible to simulate the open ion channel accurately. As mentioned in the methods, ~40-150 mV produced reasonable flux.

A custom serotonin force field might better represent the proper pore configuration. The force field may help stabilize the pore in a lower free energy configuration, although no exhaustive analysis exists. Instead, the focus has been on equilibrating the structures and developing protocols to stabilize the pore without using serotonin but just equilibrating the channel with restraints.[203] Our approach did not need special restraints on the pore itself beyond the standard CHARMM-GUI equilibration. However, from past work in the lab, the serotonin would bind and unbind from the pocket throughout the simulation. So, we defined minimal restraints in NAMD on the NH3 tail of serotonin, and the ECD cystine disulfide bond keeps the ion channel in place and the serotonin in the pocket. The force constant was relatively low at 0.04 kcal/(mol * Å^2) and was only used to prevent binding and unbinding.

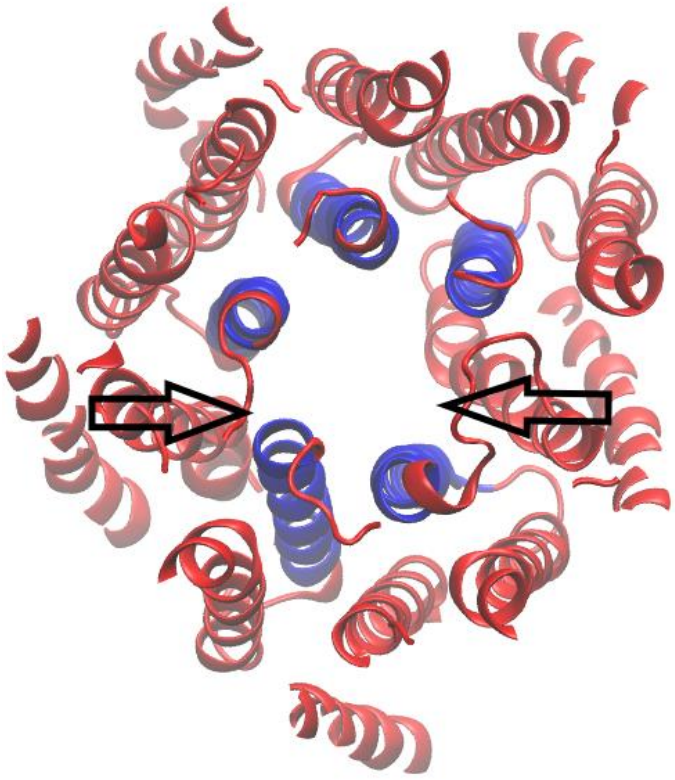


Figure 6.17). This figure depicts the conformational change needed to stabilize the open conformation of the ion channel. The figure may not represent the final fully open pose of the ion channel, which would require extensive testing to match the conductance measurements. The blue helices are normally the true center of the pore, while the red helices are the protein selection within some cutoff of the central helices (resid 250 to 270)

Obtaining the pore in an open configuration without any special equilibration protocol and using just the serotonin and the minimal restraint on the ECD disulfide bonds appears to be a significant accomplishment. The most exciting part was observing a pinching motion of the TM helices that line the inner pore itself. Compared to the Apo structure, the relaxed open structure has one of the outer TM helices pinching inward to wedge between the main pore-lining helices, as shown in Figure 6.17. Technically, two are pinching inward, or I suspect two should be, but they are not quite symmetrical. The helices with the arrow pointing inward are ordinarily not in contact with the inner pore. The helices are also not symmetrical and could rearrange inward to wedge the pore open better. I believe that the role of the outer helices is to pinch inward and prop the pore open. However, this is not clear in the original crystal structure. The most recent studies of the 5-HT_{3A} could have missed the conformational change.

Table 6.3) The standard wild-type 5-HT_{3A} ion channel stabilized in the open conformation with serotonin in the pocket was simulated with different voltages. The results prove that it is possible to simulate the pore at physiological ion concentrations and voltages.

	pS	Error	Bin Size	Simulation Length (ns)
0.15M 40 mV	9.6	19.5	100 ns	628
0.15 M 150 mV	47.4	8.0	100 ns	740

These first MD-based conductance measurements for the 5-HT_{3A} are a good starting point and show that the ion channel remains stable in an open position with a less complicated equilibration process than previously reported. Based on an ion mixture representing extracellular fluids with Mg^{2+} and other ions, the conductance was 36 pS for QDA and ~1 pS for the wild-type 5-HT_{3A}.[204] These conductance results seem reasonable and can be explained by the varied ionic mixtures and voltages used in ion flux experiments.

6.4 Conclusion

The unbiased binding process was reasonably quick, and observing this process in an MD simulation for the first time was possible. This simulation motivated a more comprehensive analysis of possible toxin venom structures in AlphaFold, leading to the discovery of two main systems that differ in the placement of two loops. The general goal of the work was to begin obtaining 3D structures and organizing information about different activated and deactivated structures. **Hypothesis 1)** Determining the toxin structure and function in the 5-HT_{3A} ion channel is possible. It is crucial to have access to the ion channel complexes of the control, activated, and deactivated ion channels with toxins. **Hypothesis 2)** Achieving an unbiased pore opening with serotonin is possible. **Hypothesis 3)** If the toxin can bind to the serotonin pocket, that will more definitively prove that the toxin competitively displaces serotonin and interferes enough to prevent allosteric activation.

The data proves **Hypotheses 1** and **2**, and evidence supporting **Hypothesis 3** is validated. The data presented to support **Hypothesis 1,** identified several structures that could play roles in competitive inhibition that match the thermodynamic constant. The data supporting **Hypothesis 2** identified a successfully opened pose of the 5-HT_{3A} ion channel and was used to quantify the first conductance measurements. In **Hypothesis 3,** the toxin is shown to act as a competitive inhibitor, but it is difficult to rule out allosteric effects as well. It would be easier to rule out allostery if the conductance of the receptor was measured with a comparable drug like Zofran with a similar IC50. If inhibitors with the same IC50 have the same ion conductance that data would conclude if there was additional allostery. Errors in the force field could cause conformational changes to the ECD. The force field argument is worth considering because the specific trial's PMF was unrealistically low.

Several new structures were then docked from the ensemble of structures and relaxed to determine the correct bound poses for free energy analysis. M0_C1_4 had the best free energy value of 9.9 kcal/mol compared to a literature value of -11.6 kcal/mol. The Anton2 docked structure had a value of -31.9 kcal/mol which is too high. The other docked structures were too weak in binding because their free energy was higher than -3 kcal/mol which did not agree well with the literature. The free energy analysis indicated that two of the bound structures had matching free energies to the value in literature. One docked peptide represented 50% of the reasonable structures; the other came from the unbiased binding. The comprehensive structural analysis revealed that the toxin prefers to bind vertically. Different deactivation strategies could involve blocking the salt bridge that closes the serotonin binding pocket. Finally, the serotonin was studied in detail, revealing that the toxin can bind on top of any given serotonin molecule already in the pocket. Eventually, the serotonin should dislodge from the pocket. A PCA was used to visualize

the main binding configurations of the serotonin. The PCA identified four main binding configurations. The most common were the two that were below the β sandwich. However, the interactions below the β sandwich still could have interacted partially with the pocket and might be necessary.

Chapter 7: Cellular Protective Mechanisms of Neural Cell Death

The contents of this chapter are modified from the following publication:

- Allsopp, R.; Klauda, J. B.; Non-Canonical bFGF Interactions with Polysaccharides to Treat Hypoxic TBI **2023** (Under Review).

7.1 Introduction

Concussions cause a serious health condition that needs to be better understood to minimize brain damage. Concussions impact young, growing adolescents as well as adults and cause a range of symptoms, including inattentiveness, blank stare, slow/incoherent speech, impaired verbal learning, memory impairment, and emotional lability.[210] Concussions are serious enough that all 50 states and the District of Columbia already have legislation related to concussion management.[211] Hypoxia is a different condition that can occur after mechanically induced TBI caused by the accumulation of fluid in the brain[212, 213]. Hypoxia refers to any lack of oxygenation in tissue. A better alternative to the now popular retroactive treatments is to develop a small molecule drug that prevents or minimizes brain injuries that are scientifically grounded and all-purpose. For example, Glycosaminoglycans (GAG) are released after TBI to protect and repair the brain, but when similar pathways are activated before the injury, there are better outcomes for hypoxia,[2, 51, 52, 95, 96] specifically for mechanically induced TBI.[97] The purpose of this manuscript is to provide a scientific understanding in the fundamental working of FGF and activity connected to various GAGs.

Acidic FGF (aFGF) and Basic FGF (bFGF) have emerged as critical proteins to prevent brain damage. These proteins can be activated by heparin and heparan sulfate. Inositol Hexa Phosphate (IP6) and Inositol tetra phosphate (IP4) are also highly charged and also interact strongly with bFGF and aFGF. The location of the strong binding site is relatively well known

from aFGF (PDB ID 3UD8, 3UDA, 3UD9, 3ud7, 1AXM, 2AXM)[214, 215] and bFGF (PDB ID 1BFC and 1BFB).[216] Although the primary binding site for these GAGs is the most well studied, there is evidence that point mutations to other sites play a role in improving thermal stability. There is a missing gap in our understanding of how such varied mechanisms result in activation of FGF.

Thermal stability tests or thermal engineering is the process of making point mutations to a protein and determining if they increase or decrease the equilibrium coefficient for folding. In other words, how did a mutation adjust the tendency of a protein to fold or unfold. Thermal engineering work on FGF was recently conducted in 2018[217] and verified in 2019[218] with cellular expression to show the decreased dependance on heparin. Point mutations were done using the FoldX program to predict which amino acid point mutations were most effective for promoting folding. The results showed that the point mutations improved folding to the point where heparin and heparan sulfate were no longer required.[217]

Deactivating point mutations have also been studied on amino acids in the loop connecting the 9th and 10th β strands.[219, 220] Most importantly, two-point mutations could dramatically reduce mitogenic activity without altering the binding affinity. ITC[220] of the FGF2 confirms that the free energy is slightly effected, with a 2 kcal/mol increase in binding energy. However, the amino acid sequences of FGF2 (bFGF) and FGF7 could be exchanged to form a chimera that was still functional.[219] Additionally, detailed studies of radiolabeled FGF7 showed that equivalent mutations did not affect the percent of radiolabeled binding of FGF7, but significantly reduced mitogenic activity.[219] These results were very interesting, and the effect of the mutations on the dynamics of the protein could be studied by MD and REMD.

New research raises questions regarding the thermal stability of the bFGF protein. Past work with differential scanning calorimetry should be viewed with skepticism because it was

conducted in an artificial laboratory environment. [221, 222] For example, the 10 mM phosphate buffer or 1 mg/ml artificially elevates the melting temperature (T_m) which leads one to conclude the protein is more stable. The phosphate buffer used is higher than physiological conditions that are maintained in the interstitial brain fluids at ~1 mM [223, 224] and improves the thermal stability of many proteins likely including the FGF protein.[225] It is not clear if the FGF protein is truly folded at physiological conditions. In the lab, the measured T_m at pH 6.5 phosphate buffer with a ligand/bFGF weight ratio of 0.5 is 80ºC with IP6 , 90°C with low molecular weight heparin, and in the absence of a ligand is 59°C.[222] The thermally engineered bFGF tested at pH 7.5 phosphate buffer had a ΔT_m of 18.7°C shifting from 53.5°C to 72.2°C.[217]

Past research showed that non-sulfated di/trisaccharides performed similarly to standard heparin in biological activation,[226] e.g., popular compounds that have been studied with bFGF include heparin, IP6,[108, 227] and IP4[228]. It is surprising that short GAGs that are minimally charged can influence signaling so drastically. The thermal engineering mutants (Figure 7.1A) have remarkably similar overlaps with the placement of the uncharged trisaccharide interactions (Figure 7.1B). Evidence has shown that pre-treatment with fructose-1,6-bisphosphate (F-1,6-BP) significantly protects tissue against hypoxia by activating the Erk1/2 pathway and dramatically improves outcomes in hypoxic conditions.[2, 51, 52, 95, 96] The combination of DMSO with F-1,6-BP extends protection against mechanically induced TBI.[97] A better atomic-level understanding of how the chemicals released during the TBI impact the brain are needed to develop better treatments.[86]

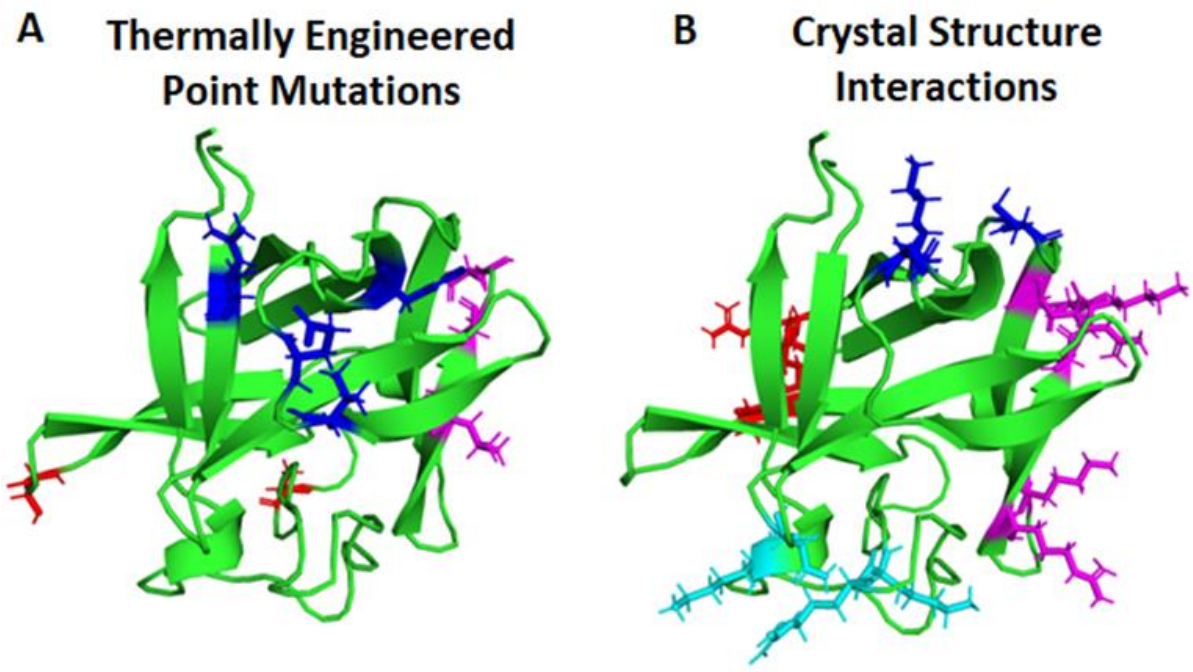


Figure 7.1) These two identical views of the bFGF protein are overlayed where cyan is site 1, red is site 2, blue is site 1', and magenta is site 2'. (**A**) Displays the thermally unstable amino acids which are color coded to match B. (**B**) Shows the amino acids that interacted with the sugars in the crystal structure. The exact amino acids that were thermally engineered via point mutations are shown in Table D1, along with the corresponding number in the wild type used in this work.

Our hypothesis is that sugars or polysaccharides (PSs) including F-1,6-BP, and tri-3 improve thermal folding leading to increase productive receptor activation. A critical component of the thermal folding is that some polysaccharides, like heparin, will bind to the primary site permanently while other sugars like F-1,6-BP, tetra-1, and tri-3 will bind and unbind over the entire protein surface.

A deeper understanding of the mechanisms and thermodynamics of how heparin interacts with each protein could lead to improved methods to treat TBI. Although full understanding would be an extensive effort, we focused our efforts on studying model compounds that interact strongly with bFGF in terms of Kd or in terms of mitogenic activity compared with heparin. The GAG

fragments were used to compare and contrast the thermophysical properties that explain how differences in the structure impact the performance of the compound toward activation and protein dimer stability. Our hypothesis that di-, tri-, and tetra-saccharides will bind to various FGF sites and still elicit a response even though they do not bind permanently or in some cases have no sulfate.[226]

7.2 Methods

7.2.1 General Methods

Simulations were run using NAMD 2.14 and Gromacs version 2020-2022 [22, 142, 229-234]. The simulations were setup using the CHARMM-GUI[142, 148, 235, 236] *Solution Builder* with 0.15 M K^+Cl^- neutralizing ion concentration and 310 K to mimic physiological conditions. The heparin sugar was interpreted using the *Glycan Reader and Modeler* [235, 237, 238] while the inositol phosphate and F-1,6-BP molecules were created using CGenFF [137, 138, 239, 240] from cgenff.umaryland.edu website as well as CHARMM-GUI's *Ligand Reader and Modeler*.[148, 241] Terminal group caps were used and included NNEU and CNEU to neutralize the protein ends. Systems were always setup using CHARMM-GUI and typically setup using HMR conditions in larger boxes to relax the initial crystal coordinates for 200-400 ns.[242] All of the aFGF dimers were initialized from the 2AXM PDB, which were 6 sugar monomers long and had -12 charge. The simulations of the IP6 involved replacing the heparin with IP6 and the simulation box was reduced in one dimension to generate a cuboid for the PMF. The simulations of the IP6 were restricted from unwanted interactions. Additionally, systems were run using HMR to relax the structure before it was converted back to standard hydrogen masses and 2 fs time steps for the potential of mean force (PMF). All of the simulations used the CHARMM36m additive all-atom protein force field[131], TIP3P water model.[143] Some simulations used the CUFIX[113] improvement to the original ion force field so that

the results were more capable of predicting ionically driven complexes. Heparin was simulated using the already optimized PS force field.[134, 243-245]

Van der Waals and electrostatic forces were computed using a force switching algorithm. The van der Waals have no options other than to set "vdwForceSwitching to yes", but the long-range electrostatics are computed through a particle mesh Ewald (PME)[139] with a force-based switching function set to switch at 10 Å and to cutoff at 12 Å.[140] This is done because it is too costly to compute all the effects especially for short range Lennard-Jones. The most primitive approach is to simply cutoff the forces beyond a certain range, but this causes discontinuity at the cutoff radius. In the switching method the potential is gradually reduced beyond a certain cutoff to more accurately represent interactions without creating discontinuities. The temperature was set to 310 K to ensure physiological conditions, all of the simulations used Langevin dynamics to control the temperature.

For simulation analysis, the bFGF sugar binding sites were overlayed on top of each other using the Pymol align feature. The FGF interaction profile diagrams that show the hydrogen bonds and other contacts were created using a server named Protein Ligand Interaction Profiler (PLIP) which was originally developed for DNA.[246]

The RMSDs of dimerized FGF complexes were calculated and presented for both proteins together. The focus of the analysis was not on the conformational changes but rather the global changes in dimers. The RMSD for aFGF shows only the backbone relaxation of the dimer, not the individual aFGF by themselves because the goal was to observe the full dimer stabilize in an unconstrained aqueous solution. For the single bFGF with it, many sugars interacting with it the goal was to study the individual protein backbone RMSD rather than the changes in the dimers as was done in the aFGF.

7.2.2 PMF Approach

For the PMF analysis of the protein dimers, one of the protein fragments was held with restraints on all of the C_a of the protein and the C1 atoms on the sugars. The other protein was pulled specifically in the +x or -x directions using the selection of C_a between amino acids 112-129 since this region consists of the interface and prevents unfolding of the protein using NAMD collective variables. Then, the standard umbrella sampling techniques were used to sample the regions as the protein fragment separated. This involved using 1 Å differences in the position of the peptide over time. The protein was pulled at the rate of 1 Å every 0.1 ns to generate the seeds, which were then relaxed for 5-15 ns before data was included in the final calculations for a total of 52 ns for each window. A replica of site 2` was used to calculate a standard error for the trials.

A similar approach was used for the unbiased sugars except the Orient package in VMD[141] was used to define the principal axis of the sugar to find the best vector to pull the sugars off the various binding sites after the system was created in CHARMM-GUI. A reaction coordinate was chosen for each specific binding site and the pathway was divided into 6-7 windows, where each window had a different location for the minimum of the restraint. A collective variable was applied to the sugar to pull in a specific direction using a force constant of 5kcal/(mol·Å^2). The protein as a whole was also restrained with force constants for the harmonic restraints set to 10 kcal/(mol·Å^2). An initial pulling phase was set up with a moving setpoint that adjusted continuously at a fixed rate to cover 1 Å every 0.1 ns. After an initial pulling phase, the coordinates, velocities, and system sizes were used to seed the new windows along the reaction coordinate pathway. These were then run in production mode that has no changing spring position but instead fixed positions and the output was collected including the “centers” variable and the actual position of the selection over time to process later with Grosfield implementation of

WHAM.[150, 209] The Grosfield implementation of WHAM was used with 200 bins and stop criteria of a tolerance of 10^{-28} or 100,000 iterations whatever is reached first with and a temperature of 310 K.

7.2.3 Alternative Binding Sites

The unbiased binding simulations were inspired by the crystal structure generated by Ornitz,[226] and the knowledge that the thermally unstable amino acids were scattered across the protein and not localized to the primary binding site. To study this in greater detail, unbiased binding simulations were required to recreate similar bound orientations to those reported in the crystal structure that are not available in RCSB. Initial attempts to find the ideal crystal structure (2FGF) failed because many structures such as 1BLA and 4XKI have point mutations. To best utilize our time, AlphaFold [80, 81] was used to obtain the equivalent structure. If we compare the AlphaFold2 structure to the 2FGF crystal structure using the entire selection its RMSD was 0.386 Å. AlphaFold2 was used for simplicity because the crystal structures that were identified had a few point mutations from the wild type. Rather than apply the mutations manually, Uniprot was used to get the exact structure. Alphafold2 converted the Uniport ID P09038 for Human bFGF to generate the 3D structure, of which residues 160-288 were extracted for the MD simulation. AlphaFold2 was used for convenience to correct for point mutations in the original crystal structures. The structure was double checked to see that the prediction matched the past crystal structures with an all atom RMSD of 0.386 Å. The sugars were then generated in CHARMM-GUI *Glycan Reader and Modeler*.[235, 237, 238] The exact sugar sequences and connectivity are shown in Table D2. These human sugars are shown in Figure 7.2 below. Pymol[247] software was used to duplicate the sugar and create a high concentration of sugars, i.e., 87 mM, so that the sugars would bind around the bFGF protein. A second lower concentration of 25 mM sugar used only three

sugars each docked with Autodock Vina[207] docked using Kollman charges. For the sugars, the Kollman sugars resulted in 0 charge, so Autodock tools overrode my input selection and provided Gasteiger charges. The other variable is the torsions. The torsions were automatically set to 16 of 17 total bonds, and we manually maxed it out because we wanted to explore all the conformational flexibility so there were 17 torsions. Autodock Vina was not originally used to generate heparin bound configurations. However, it successfully generated these configurations. This outcome ranked in the top 6-7 results and is reproducible over the two trials simulated. In other words, Autodcok Vina can produce bound structures of the sugars to the primary site. However, the tri-3 sugar is more difficult to obtain bound structures to the primary site but exist after 3-5 trials. Additionally, the states did not fall in the top position but within the top 9. Out of three docking trials of tri-3, only one of them obtained a bound pose on the primary site. There was not a clear binding site for any of the minimally charged sugars, so Autodock Vina had to have a reduced search space to test the different binding sites. Initially, Autodock Vina[207] was used to generate docked orientations, but these orientations quickly unbound in the simulation. These simulations were left to continue to generate a variety of interactions that closely resemble the Ornitz crystal structure[226] using unbiased binding. The combined high concentration Pymol system[247] was uploaded to CHARMM-GUI interface[148] and the *solution builder*[142, 236] was used to add a neutralizing 0.15 M K^+Cl^- and enough water to solvate the system.

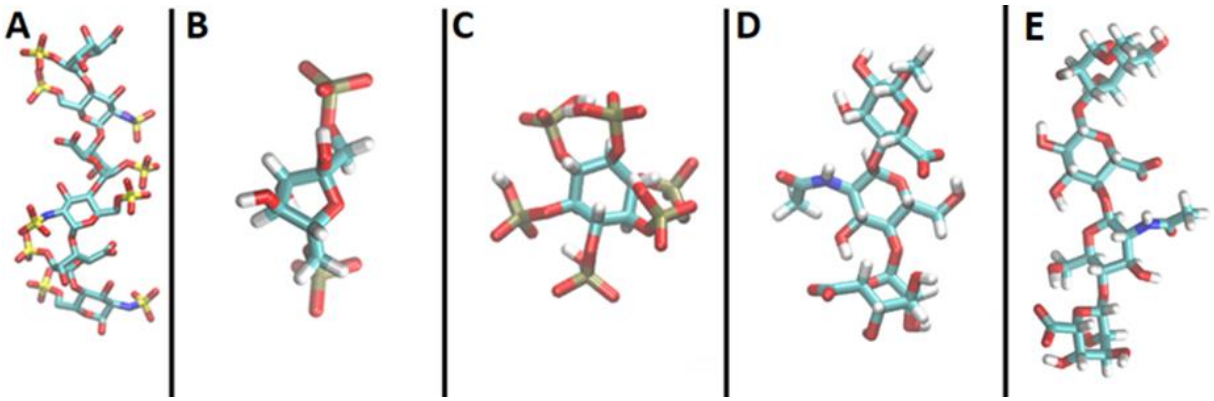


Figure 7.2) Panel A shows the commonly crystalized heparin, panel B shows the fructose-1,6-bisphosphate, panel C shows inositol hexaphosphate, and panel D shows desired sugar α***-L-IdoA-(1->4)-*** α***-D-GlcNAc-(1->4)-*** *β***-D-GlcA-1->OMe** generated with the appropriate chemical modifications using the CHARMM-GUI Glycan Reader and Modeler, and panel E shows Tetra-1 which has a chemical name of α***-L-IdoA-(1->4)-***α***-D-GlcNAc-(1->4)-*** *β***-D-GlcA-(1->3)-** *β***-D-Gal-1->OMe** and was bound with unbiased binding simulations to regenerate the alternative binding sites.

7.2.4 FGF simulations with Fructose-1,6-Bisphosphate, IP6, and Heparin

Another objective was to understand if there were any interactions between F-1,6-BP and aFGF. It is possible to predict which molecules activate the growth factors by calculating the binding free energy of dimerization induced by molecules such as the crystalized heparin, IP6, inositol 1,4,5,6 tetra-phosphate, and F-1,6-BP. The aFGF dimers were taken from PDB ID: 2AXM. The charge choice of the IP6 and the inositol-1,4,5,6-tetraphosphate was carefully chosen to be -1 charge per phosphate. This choice was made as there is some evidence that the overall charge of the molecule is closer to -8 instead of -12 at physiological pH 7.60 due to the already high amount of charge contained in such a small space.[248] Additionally, the overall objective was to study hypoxic induced TBI which is known to cause higher levels of CO_2 and lower the pH below the usual value. This model is most representative of the IP6 at pH 7.0 and there was not easily obtainable data for the inositol-1,4,5,6-tetraphosphate, but the -1 charge per phosphate was maintained for consistency.

The initial box that ran fructose 1,6-bisphosphate was a square that had 81 Å for X, Y, and Z dimensions. The boxes that were used for the PMF simulations were 110 Å by 55 Å by 55 Å and employed X, Y constraints on all of the proteins and sugar molecules to prevent unwanted rotations, while fixed atoms were chosen on the sugar to prevent it from peeling off. The F-1,6-BP ran for a total of 400 ns with the standard force field before it was switched to the CUFIX force field.[113] The CUFIX has weaker electrostatic interactions and was developed to reproduce the electrostatic interactions more realistically for aggregation and dimerization. The dimer that was stable for the original force field but was not stable for the CUFIX.

7.2.5 FGF Unfolding with Replica Exchange

Replica exchange MD (REMD) was used to study the partial unfolding of the bFGF protein. Various sugars were also studied in the REMD to elucidate if they impact protein unfolding. All of the REMD simulations ran 64 windows and there were two temperature ranges. The systems were around 75,000-80,000 atoms with 0.15 M NaCl ions to mimic physiological conditions. The sugar concentrations are shown in Table D3. The lower temperature range went from 293 K to 400 K and mostly focused on the 310 K window for study of physiological effects. The higher temperature trial focused on temperatures ranging from 347 K to 450 K. The reason for using 347 K was because that was the T_m for the protein in 10 mM phosphate buffer. This will ensure that the unfolded structure is sampled.

The population of the folded state was determined by binning the probability of different RMSD and radius of gyration. The first 60 ns of data was discarded in this analysis. Higher RMSD and radius of gyration indicate that the protein is more unfolded. To make the folding landscape the formula -R*T*ln(p_i), where the probability histogram bins (p_i) were generated in MATLAB.

The REMD simulations were also used to study the sugar contacts to determine what interactions were most dominant. The contacts were studied using MDAnalysis contacts function and tracking the number of contacts with different amino acids and printing the output. From the output data a normalized average was assigned for each amino acid. This was defined as the average of each amino acid divided by the sum of the averages. These were applied as b-Factors in VMD and saved as a ".pdb" for developing analysis plots.

A clash analysis was performed as well using VMD and Pymol.[141, 247] based on all frames in the REMD. The protein coordinates were extracted from the .xtc file and converted to a .gro file that was loaded into pymol. From there the C_a atoms were aligned over all the states and the carbon alpha of each trajectory was aligned over the receptor. Then the receptor and the superimposed bFGF trajectory were loaded into VMD and the coordinates of the receptor were extracted and used to test if any atoms in the selection between resid 267 and 276 fell beyond the X, Y, and Z coordinates of Asp 200. The number of bFGF atoms that crossed outside all three X, Y, Z planes were counted in each frame and stored to make the box plots.

7.3 Results

7.3.1 Summary of Results

The aFGF and bFGF interactions with heparin, IP6, IP4, F-1,6-BP, Tri-3, and Tetra-1 showed that heparin-like compounds had strong interactions that compare well with ITC data. While attempting to classify all these sugars, the F-1,6-BP was found to behave very differently causing all sugars with less charge to be grouped separately. F-1,6-BP, Tri-3, and Tetra-1 all did not fit previous models of interaction. This result proves that strong binding is not a requirement for stabilization. For these three sugars, binding occurred all over the protein surface as summarized in Figure 7.3. Tri-3 and F-1,6-BP were shown to have much weaker binding. Binding

poses for each of the major binding sites shown in the original crystal structure were re-created with MD simulations. The PMFs showed that the Tri-3 and Tetra-1 sugars have affinity for specific interactions with the thermally unstable point mutations.

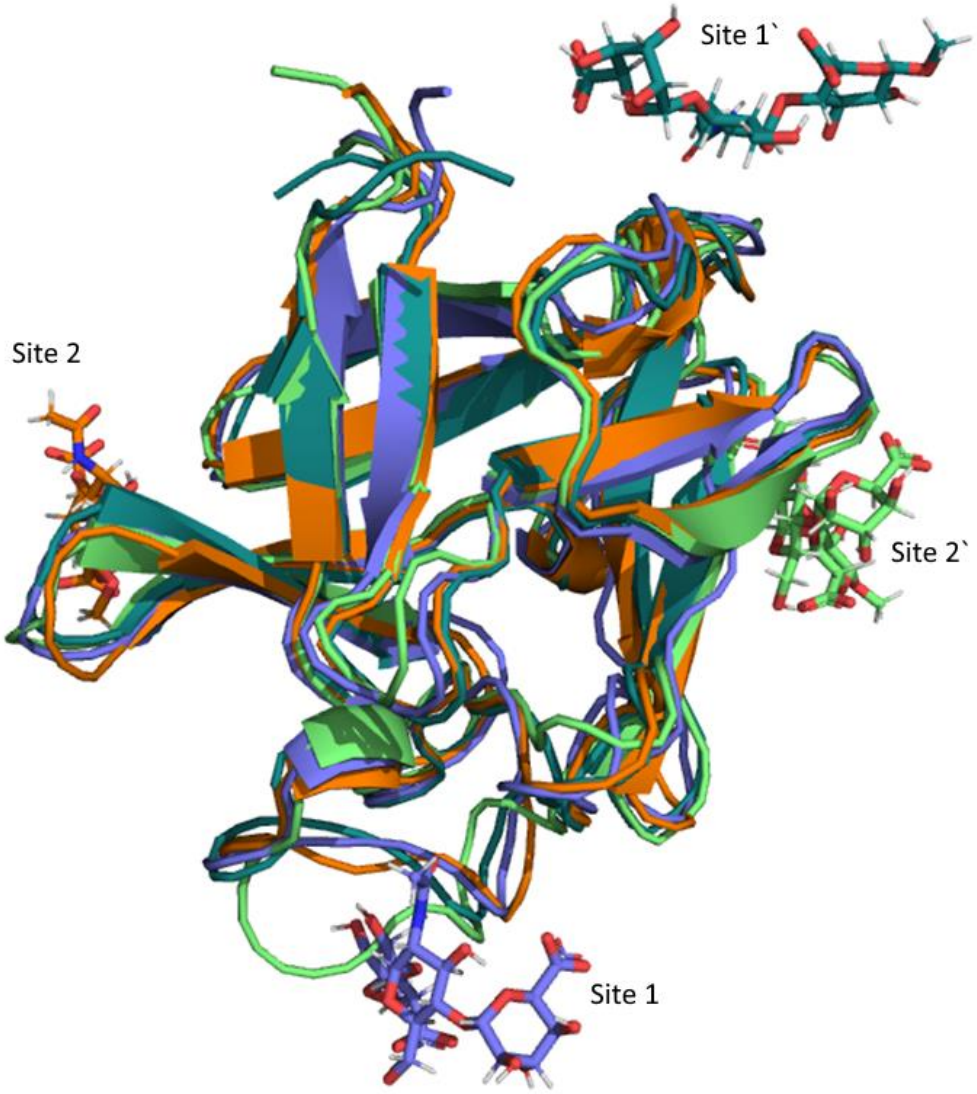


Figure 7.3) Refined sugar binding orientations on FGF using MD simulations confirmed that there was affinity for these sites. There are minor differences in the MD derived sugar orientations compared to the crystalized sugars pictured in [226] caused by the need to have crystal symmetry. Each of the four different colors correspond to a set of protein and sugar coordinates that were aligned over each other. This shows that each simulation was started from a slightly different set of coordinates. Three unstudied sites were discovered, and a more detailed description of each site follows.

These results raised questions as to the role of the weaker binding carbohydrates and why they bind to the thermally unstable amino acids. REMD was used for the first time to clarify their role in folding. The primary binding site was shown to be the most unstable point and most likely to unfold. Furthermore, both the trajectories in the presence of sugars and the thermally engineered FGF protein shifted the RMSD quartiles toward the folded state.

Surprisingly, the engineered point mutations mainly impacted the primary site which had no direct thermal engineering. Using PDB ID: 1CVS and 1FQ9 as models, a less obvious unfolding was observed to clash between bFGF and the receptor. The MD simulations showed that the clash was more prevalent in the control (wild type without sugar) and the destabilized point mutations. The clash was still more prevalent even when sugar was added to the simulation.

7.3.2 Models of F-1,6-BP, Tri-3, and Tetra-1 versus Canonical Heparin, IP4, and IP6

Binding of FGF to Heparin, IP4, and IP6

This section focused on classifying other sugars with varied functional groups based on binding strength. Understanding the interactions of small-charged sugars with FGF are fundamental to improve TBI treatments. Even though the bFGF has been extensively studied with point mutations, the aFGF protein was chosen to evaluate strong binding interactions with molecules like heparin because dimerized crystal structures were available. Dimerization was thought to be an indicator of stronger interactions making aFGF more suited for this specific evaluation. When IP6 replaced heparin, the dimers remained in strong contact and the RMSD of the different stable inositol dimers resulted in 5-6 Å of deviation (Figure D1). The simulations revealed that the energy to break the dimer with two IP6 sugars was -20.4 kcal/mol (Figure D2, Table 1). However, when only one sugar was present the energy was -16.3 kcal/mol indicating

some kind of saturation effect. Additionally, IP4, a similar molecule, resulted in dimerization energy of -17.8 kcal/mol and contributed less energy to stabilize the dimer (Figure D2). These simulations of the small-charged sugars were then compared with the heparin molecules. See Table D4-S5 for the system builds and the PMF analysis.

Table 7.1) Free energy aFGF dimers in the presence of IP4, IP6 or heparan sulfate (HS). The standard error estimate was based on errors from the plateaus of the free energy plots (Figure D2). HS strong/weak side are illustrated in Figure D3. Heparin binding interactions are not equal on both sides of the dimer as supported by Table D6 and Figure D3.

Trial	Restrained Orientation (kcal/mol)	Number of Sites	Overall (kcal/mol)
Literature bFGF-IP6 Dimerization	-7.3	2.2	-16.2
Literature bFGF-FGFR Heparin	-19.4	0.95	-18.43
IP6 aFGF	-9.9±0.5	2	-19.7±1.0
IP6 aFGF Replica	-10.6±0.5	2	-21.3±1.0
IP6 aFGF	-16.3±0.5	1	-16.3±0.5
IP4 aFGF	-8.9±0.5	2	-17.8±1.0
HS Weak Side aFGF	-15.6±0.5	1	-15.6±0.5
HS Strong Side aFGF	-34.2±0.5	1	-34.2±0.5

The heparin stabilized FGF dimers resulted in similar RMSD values of 5-6 Å (Figure D1). The free energies were -34.2 and -15.6 kcal/mol when applying the collective variable to the two different aFGF proteins in the different systems (Figures D8, Table 7.1). There are two different binding sites, the first one has a value of -15.6 kcal/mol and the second has a value of -34.2 kcal/mol. The past simulated literature value is close to the -15.6 kcal/mol value. The -34.2 kcal/mol free energy may only be accessible to aFGF dimers.

7.3.3 Important Sugar Binding Interactions with FGF

Figure D4 shows the backbone RMSD of the bFGF protein with 3 tri-3 carbohydrates, with 9 tri-3 carbohydrates, and with 9 tetra-1 carbohydrates. The figure indicates some variability in the RMSD related to the type of sugar and the concentration. With tri-3 potentially having the most conformational change. At this stage of the research, it was difficult to assess if that meant the bFGF protein was activated or deactivated.

Site 1

We studied how the sugars interacted with mutated amino acids because past mutations have been identified and the weaker binders could bind anywhere on the bFGF protein. The site 1 did not have any point mutations because it already had a previously known significance for binding heparin. Even though this region was not researched for thermally unstable point mutations, they likely still exist on this site as was shown in a previous article [221, 249, 250]. The free energy data indicates this site populates a significant fraction of the thermodynamic states. An individual sugar that was bound to site 1 (Figure D5) and site 2' was simulated to gauge the time for unbinding. The data revealed that site 1 remained bound for >35 ns and the site 2' only remained bound for ~5 ns.

Site 2`

The most promising alternative to site 1, the established binding site, was site 2` because of the low free energy and close proximity to the thermally unstable amino acids C229, S227, and L225. Figure 7.4 indicates that there are physical interactions in the vicinity of C229 and S227 that are thermodynamically unstable. This pose shown in Figure 7.4 could directly stabilize the unstable regions. While the interactions that are depicted do not directly form a hydrogen bond

with the side chains, they do directly interact with adjacent amino acid backbones in at least two (S227 and C229) of the three unstable amino acids. These amino acids could increase the stability of the region. An alternative conformation is that the N-acetyl group rotates to interact with the L225 backbone.

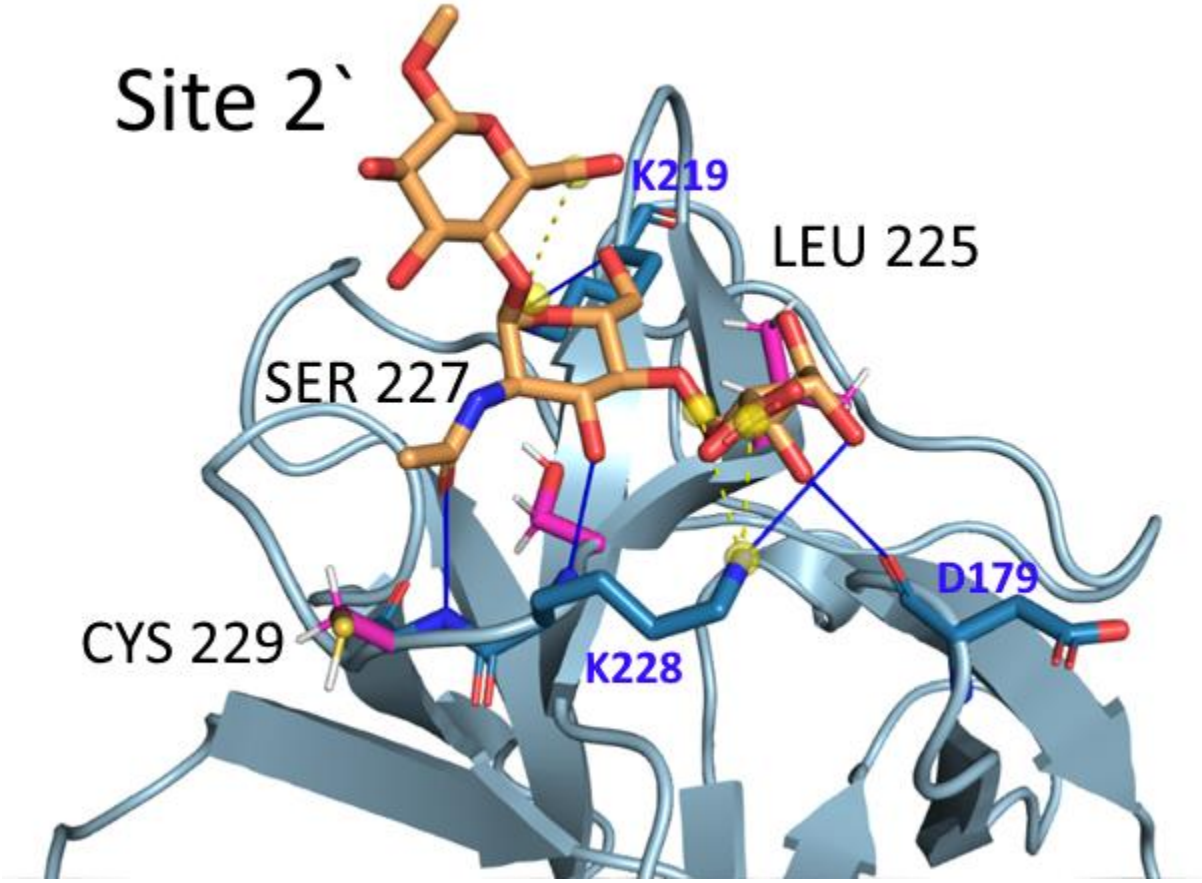


Figure 7.4) Site 2` interactions were depicted using the PLIP tool to show the thermally unstable amino acids in pink in relation to the rest of the FGF protein in cyan. The hydrogen bonds are shown with solid blue lines, the predicted salt bridges are shown with dotted yellow lines, and two of the charged centers are shown between two oxygens as yellow spheres that are partially charged.

Site 1`

This binding site had four of the eight total unstable amino acids including H192, R164, V185, and E187. Tetra-1 might reinforce the fold as shown in Figure D6-7 with specific

interactions by connecting with multiple strands. Interactions were identified on the β-strand with R181 and R175. The R175 is between amino acids V185 and H192. Another point of contact was on R223 which is adjacent to the backbone of the unstable amino acid L225 from site 2'. From these interactions, it appears that the tetra-1 sugar can indirectly stabilize several amino acids that were previously identified to be thermally unstable.

Site 2

The tri-3 sugar had some weak binding affinity to site 2, and the possible effect on locally unstable amino acids is shown in Figure D8. The SER 242 mutant was a rationally designed mutation likely offering a salt bridge to a neighboring (+) charged amino acid. SER 242 is likely the closest direct interaction and THR 254 appears to be the furthest direct interaction from the sugar in this region.

Binding Free Energy of Sugar to FGF

The binding free energies were used to gauge the CUFIX force field interpretation of the sugars and how strongly the sugars interacted in various parts of the FGF protein (Table 2). The binding free energies are an important metric because the lower free energies correspond to stronger interactions that are more likely to make a reasonable contribution to protein folding. The sugar interactions are displayed in Figures D2-D5, and the PMF values are displayed in Figure D9. The PMF indicates that this specific polysaccharide is roughly equally likely to bind to the secondary site as it is to bind to the primary site.

Table 2) Lists the free energies of binding of the α-*L-IdoA-(1->4)-* α-*D-GlcNAc-(1->4)-* *β-D-GlcA-1->OMe* to various sites (Figure 7.3) of FGF. The standard error (±value) was estimated from replicas of the site 2' PMF using unrestrained orientations which were thought to have the higher variability.

Site	ΔG (kcal/mol)
Site 1 (tri-3)	-5.35±0.07
Site 2 (tri-3)	-1.28±0.07
Site 2' (tri-3)	-5.82±0.07
Site 1' (tetra-1)	-2.37±0.07

7.3.4 Binding of tetra-1, tri-3, and fructose-1,6-bisphosphate FGF (Replica Exchange)

The previous results were a good start to understand the sugar interactions and proving that some sugars behave differently from others. However, this approach offers a limited understanding of how the sugars impact the binding. The REMD takes advantage of the faster kinetics at higher temperatures to better explore different states which can prove if certain sugars or mutations alter the folding. The sugar contacts shown in Figure 7.5 revealed the location of hot spots of interaction. The results are comparable to the PMF values that indicate strong interactions at multiple sites. The PMF used a different force field and ranked the binding sites slightly differently from the REMD. For example, the tetra-1 sugar did not show up strongly in the fraction of contact at site

1'. Based on the REMD, the primary binding site has the strongest interactions, site 2' is strong, and site 1' was weakest which is in line with the ΔG results (Table 2).

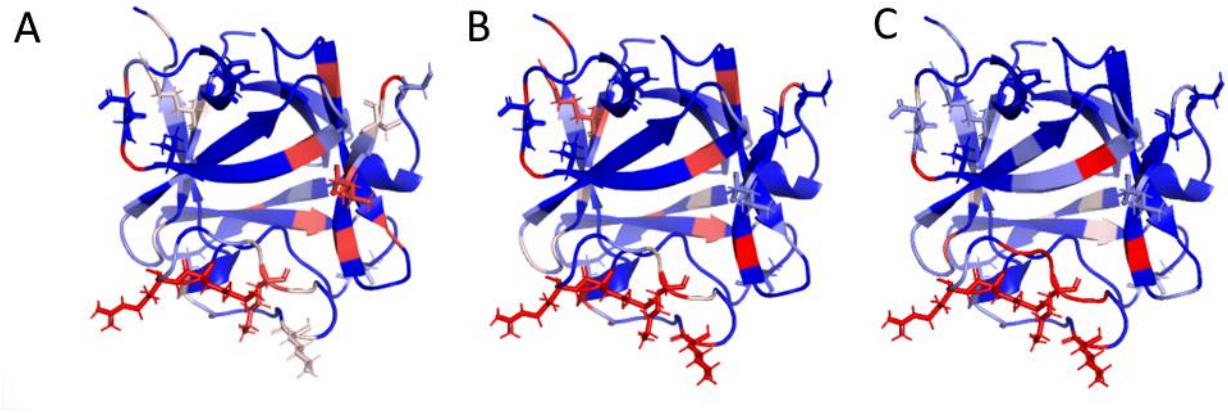


Figure 7.5) Shows the total sugar contacts. Panel A shows the interactions with the tri-3 and tetra-1 polysaccharide mixture, and panel B shows the contacts of the 5 F-1,6-BP. Panel C shows the contacts of only tetra-1 within the simulation from panel B. All of the contact data are normalized as a percentage of the total average contacts. The colored images were created in Pymol using a blue-white-red coloring based on the b-factor of min=0 contacts and max=0.035 contacts. The 0.035 b-factor value is about half the true max b-value.

7.4 FGF Protein Folding with Replica Exchange

The first critical step was to use elevated temperatures of at least (347 K) in REMD systems to define the unfolded FGF protein structure. The temperature 347 K was chosen because the differential scanning calorimetry (DSC) previously determined that this temperature was the T_m with 10 mM phosphate buffer. Without the phosphate buffer, the T_m will lower, making 347 K an ideal temperature to study the unfolded structure. The unfolding of the FGF was determined to primarily occur at the primary binding site but the exact reason why that affected the receptor will be addressed later in this paper. This was confirmed by aligning the centers of the deepest free energy wells from the control and the Tri-3/Tetra-1 simulations as shown in Figure 7.6. The unfolding was confirmed when the folding in this region was improved after these sugars were

added. It is difficult to know how much functionality was regained after adding the sugars without knowing more about how to deactivate bFGF.

The sugar concentration of 8.5 mM tri-3 and 2.1 mM Tetra-1 was used in the simulations as it improved the folding of the FGF protein even at extremely high temperatures as shown in Figure 7.6. The energy well was certainly shifted by the sugar interactions to a lower RMSD value and broadened to include lower RMSD configurations. This shift indicates that the sugar promotes better folded FGF structures. The visual structural representations comparing the control with the FGF in sugar are shown in Figure 7.6.

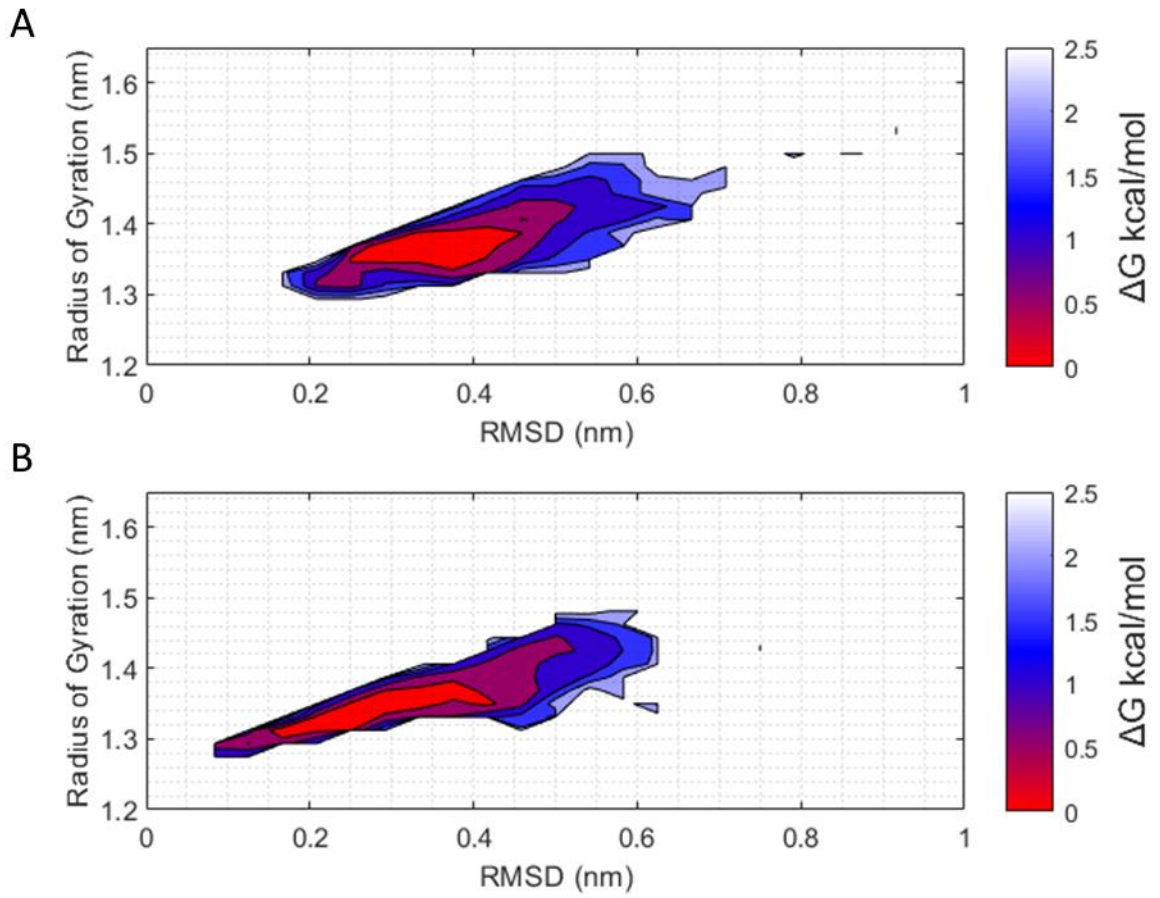


Figure 7.6) These free energy landscapes are derived from 347K REMD. The top shows the control and the bottom shows the experimental Tri-3 and Tetra-1 sugar combination. RMSD=0 indicates the structure before any simulation that was generated form AlphaFold2 and has no notable difference from past crystal structures.

Figure 7.7) Overlays the most populated bins at 347 K for the control (cyan), experimental sugar (pink), crystal structure (green). Cyan color is the control with an RMSD=1.37 nm and a radius of Gyration = 0.32 nm. The experimental sugar simulation had an RMSD=0.26 nm and radius of gyration=1.325 nm. The FGF is around 8.5 mM tri-3 and 2.1 mM tetra-1. It is interesting that the primary site unfolds.

Folding Effects at Physiological Temperature

At physiological temperature, the control FGF is similar to the control at 347 K. However, the folded "well" of the experimental sugar is shifted to the left toward a better folded structure (Figure 7.8). The presence of sugars drastically changes the folding landscape by making the highly unfolded structures inaccessible. Even the mildly populated purple bins with RMSD of 0.5 nm are now completely inaccessible in the presence of the sugars used in the experimental simulations (tri-3/tetra-1 and F-1,6-BP). The mutants also do a remarkable job of shifting the RMSD towards lower values indicating better folding, although there were no engineered point mutations on the primary site. All of these techniques to activate or stabilize the growth factor, supported the interpretation that the primary binding site is what is unfolding/misfolding in the

absence of the carbohydrate ligand. Even though the point mutations are not in the direct vicinity of the primary binding site they still have some allosteric effect on the folding or the motions of the bFGF protein.

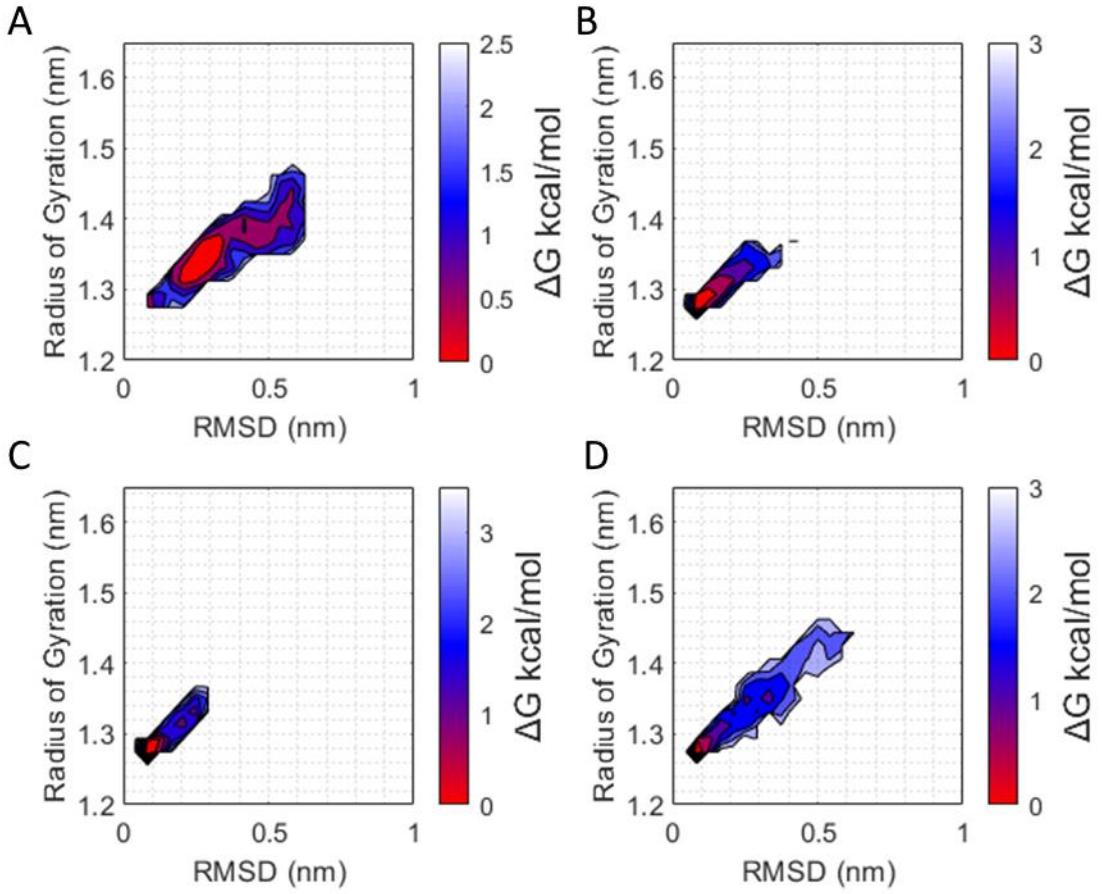


Figure 7.8) The effect of a variety of sugars on the folding of FGF at 310 K. A) Shows the control trial at 310 K indicating a broad energy landscape that accesses folded and unfolded states of the protein. B) Shows the mixture of 8.6 mM tri-3 and 2.1 mM tetra-1 which effectively eliminated the existence of the unfolded structures. C) Shows the effect of F-1,6-BP on the FGF protein with similar results. D) Shows the effect of applying thermal engineering mutations which surprisingly shifted the folding as well. RMSD=0 indicates the structure before any simulation that was generated form AlphaFold2 and has no notable difference from past crystal structures

Structural Differences Between Folded and Partially Unfolded bFGF

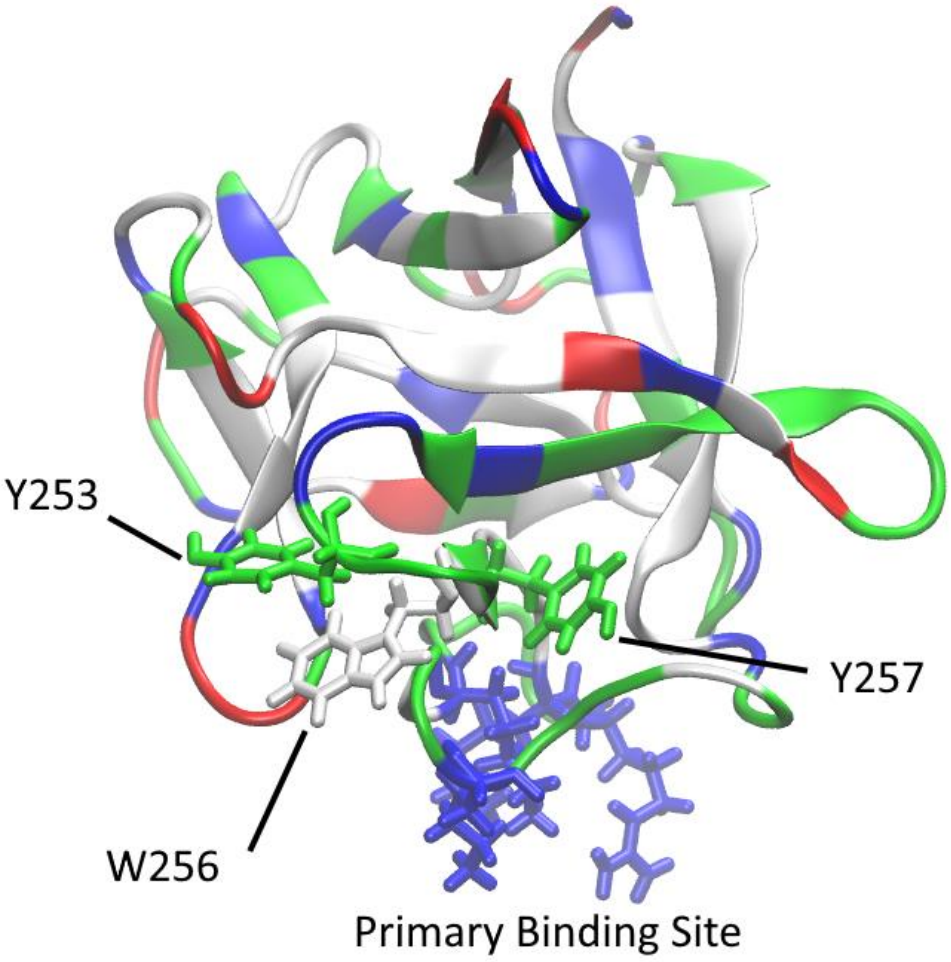


Figure 7.9) Snapshot of bFGF focused on the locations where Y253A, W256A, and Y257A were mutated in the alanine scan.

Additional insight was gained when looking at three point mutations (Y253A, W256A, and Y257A) that were tested (Figure 7.9) with ITC and other experiments in the past and these resulted in minimal changes in binding K_d.[220] The results of the point mutations were a slight increase in the lower bound of RMSD indicating a slightly better folded ground state compared to the wild type and the upper unfolded states were more unfavored (Figure 7.10). However, the effect of the F-1,6-BP was also diminished. The amino acid positions Y253 and W256 are on the same loop and W256 is labeled in Figure 7.12. FoldX was used to calculate a 2 kcal/mol destabilization for

each point mutation and helped identify a third neighboring mutation with Y257A. Surprisingly, the energy minima only shifted by 1-2 Å in RMSD, and the overall the energy landscape has much more overlap with folded structures than expected. A closer examination is needed to understand if the 2 Å is spread out or concentrated in one area where it might disrupt receptor signaling. Some π–π stacking likely existed between the tryptophan and tyrosine rings.

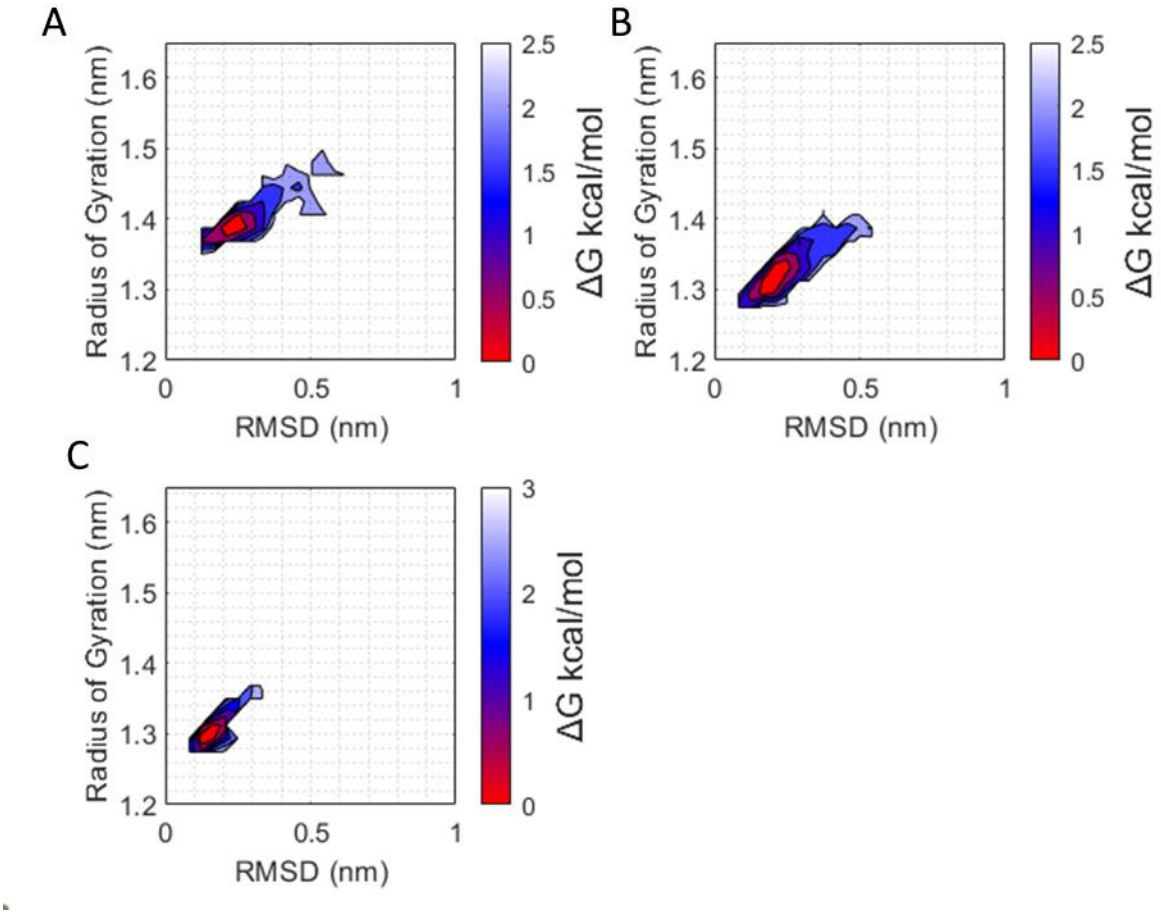


Figure 7.10) Panel A shows the simulation with Y253A, W256A, Y257A point mutations, while Panel B shows only Y253A and W256A. Finally, Panel C shows the Y253A, W256A, and Y257A in a solution of F-1,6-BP sugars. RMSD=0 indicates the structure before any simulation that was generated form AlphaFold2 and has no notable difference from past crystal structures

bFGF is known to activate the receptor without heparin as proven with the stabilizing point mutations. Significant re-arrangements in the bFGF interaction with heparin are required to adjust

how bFGF interacts with the receptor (Figure 7.12). The clash between the red helical turn of the receptor and the pink loop could alter the binding interaction (Figure 7.12).

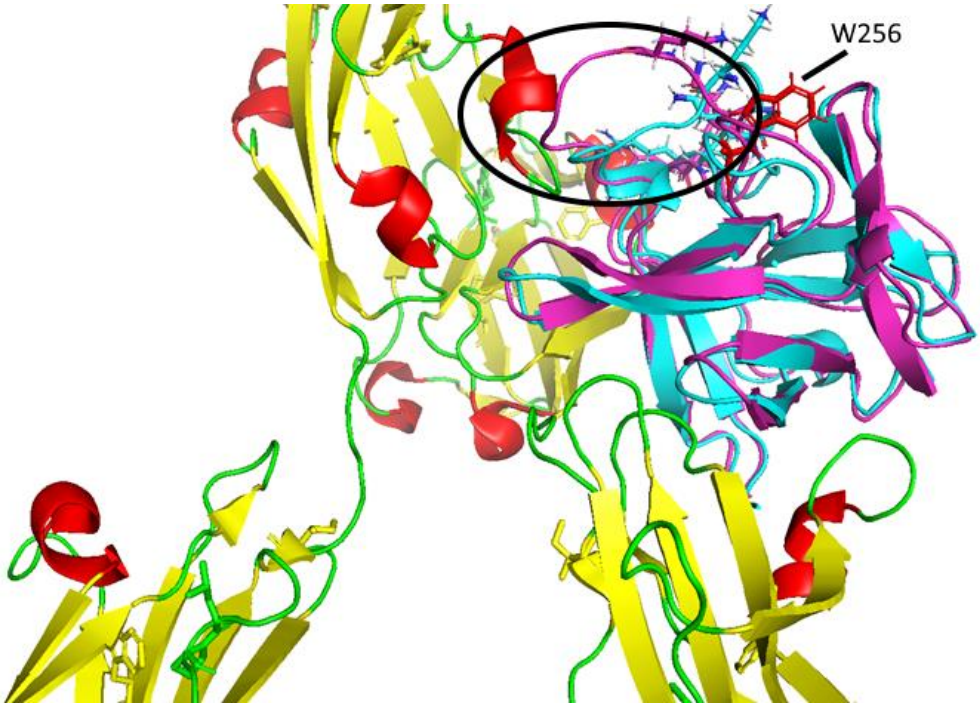


Figure 7.11) Shows the clash with the receptor caused by the point mutations Y253A W256A Y257A. The yellow, green, and red protein is the FGF receptor (FGFR). The cyan and pink bFGF proteins are aligned over each other. The pink bFGF protein clashes with the receptor while the cyan does not. This snapshot was representative of the difference between the stabilized and the destabilized bFGF proteins. The stabilizing point mutations did not perfectly fold the primary binding site, but it did reduce the clash compared to the F-1,6-BP and the Tri-3/Tetra-1 carbohydrates.

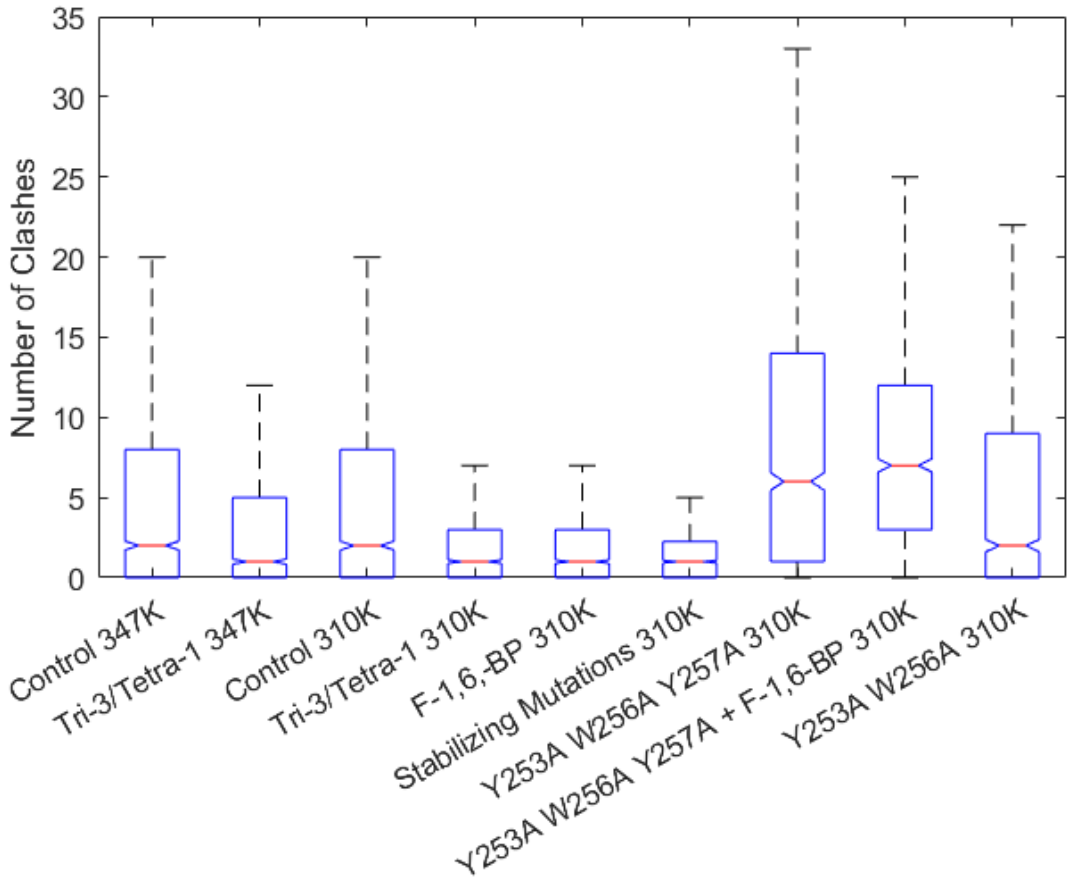


Figure 7.12) Box plot of the number of clashes measured over time. There is a clear difference in the upper quartile of the box plots that are stabilized and those that are not stabilized.

Figure 7.12 shows that increasing the number of clashes correlates with the expectation of activation or inhibition. The same box and whisker plot is shown in Figure D10 with all the outliers. The severity of the clash is correlated with the number of clashes. The upper quartile appears to be the most significant in differentiating the activated and the deactivated. This could be explained if there was some tolerance and flexibility in the receptor to accommodate the movement of the ligand, but beyond a certain deviation the ligand does not make effective contact. Figure D11 supports the validity of the results because it shows the consistent sampling and transitions between replicas at different temperatures. This value could be better defined with

experimental validation of the mitogenic activity at different temperatures. The clashes seem like a reasonable explanation and still make sense given the ITC data in Table 3 that indicates a slight decrease in binding free energy caused by the mutations which raised the overall binding free energy from -7.7 to -6.0 kcal/mol.

7.5 Discussion

Our simulation results show proper folding of bFGF is required to yield activation for receptor binding. One of the most interesting findings was that all of the different strategies of activation converged on similar structures with the deepest well in the free energy landscape, and these decreased the upper quartile of clashing. The simulated PMFs show that the primary binding site agrees with the ITC data for dimerization and for receptor binding. However, the destabilizing point mutations only changed the primary binding site ITC[220] value by around 0.5 kcal/mol which is close to the margin of error. The REMD quantified two main conformational changes of bFGF that could potentially clash with the receptor and might explain why bFGF binds without activating the receptor. To better rank the importance of the conformational changes, the crystal structures 1CVS and 1FQ9 were examined. After alignment, the first conformational change would clash with heparin but might also bind to the primary binding site of bFGF even with clashes. The ITC and PMF support a binding model where the dominant interaction value of ~10kcal/mol is not inhibited by minor clashes. The radiolabeling of the bFGF supports the theory that binding is not inhibited by minor clashes. The ITC interactions with the rest of the receptor indicated 1.7 kcal/mol difference between wild type and mutant. This difference was significant and supports the observed clash between the primary binding loop and the receptor. When the upper quartile of clashes of the wild type and destabilized mutant (Y2523A, W256A, and Y257A) were simulated in the presence of F-1,6-BP, the destabilized mutant results were 0.5 Å higher RMSF than the

stabilized. The 0.5 Å might cause a clash with the receptor near the primary binding site that prevents the receptor activation. These data suggest that the receptor is highly sensitive to the exact conformation of the bFGF protein, how it folds, and that even a slight shift of a few angstroms can disrupt the signaling.

This work also demonstrates other sugars (F-1,6-BP, tri-3, and tetra-1) have similar effects on the FGF folding. Tri-3/tetra-1 is not specifically confirmed to be neuroprotective but has a well-defined mechanism that one would expect to trigger neuroprotection. However, the binding of these sugars is weaker than heparin and it is difficult to localize the interaction to just one binding site. Possible mechanisms of the carbohydrates include: 1) directly stabilizing the backbone amino acids via hydrogen bonds, and 2) perturbing the water activity to greater penalize hydrophobic exposure to impact folding. Inositol phosphates on the other hand are permanent binders and even though they raise the T_m they somehow deactivate the receptor possibly by forming aggregates that are otherwise prevented in the presence of heparin, where signaling is restored.[108, 227]

The thermally stabilizing point mutations also improved the folding similar to F-1,6-BP and tri-3/tetra-1. Even though the thermally engineered point mutations avoided the primary binding site, they still altered the folding landscape to favor the most folded conformation. Our hypothesis is that the small sugars triggered Erk-1/2 by better folding bFGF proteins which can bind to the FGFR without heparin. The “destabilizing” point mutations result in slight re-arrangements of a loop that could clash with the receptor and slightly alter the binding (Figure 7.11).

The binding strength of the tri-3, tetra-1, and F-1,6-BP differ from previously studied sugars that were much tighter binders like IP4, IP6. The exact data for the IP6 versus heparin is not completely conclusive, because the wet-lab ITC predicts IP6 to be weaker than the heparin.

Some uncertainty could arise because the heparin was bound to a receptor but the IP6 was not. The simulations indicate the IP6 is a stronger aggregator which could explain why it inhibits the activation of the receptor. The dimerizing energy of IP6 from our free energy calculations was similar to that in literature at -7.31 kcal/mol[108] with bFGF, as compared to the simulation value of -10 kcal/mol. The aFGF research with IP6 could be interpreted as having stronger aggregation or receptor blocking as briefly discussed in the literature.[217, 218] IP5 and IP6 have the same charge, but IP5 could be the strongest aggregator because there is less resonance to distribute the charge. It should also be pointed out that work done with endothelial cell lines exposed to serum[228] should be analyzed as if they generated heparan sulfate which is functionally equivalent to heparin[85]. The IP4 was another experimental compound [228] that showed the highest cellular survival during hypoxia in certain tests. The improvement is not completely understood, and may have to do with Ca^{2+} receptors interactions with IP3/IP4[104] , or less aggregation causing more functional FGF. The ITC of bFGF binding to the receptor is assumed to have heparin which would agree with the simulated value of binding two bFGF to one heparin molecule (ΔG=~20kcal/mol). The differences in these energies could arise from CGen Force Field parameterization or differences between aFGF and bFGF (bFGF would raise the K_d/lower ΔG).

Preventative TBI research has been ongoing for several decades and identified at least three molecules that pass the blood brain barrier and trigger neuroprotective mechanisms.[51, 52, 93-96] It is assumed in this manuscript that the catastrophic rise in intracellular Ca^{2+} is the cause of neural cell death. See [2, 51, 52, 94, 100, 101] for a complete list of references of neuroprotective compounds. It is known that Ca^{2+} concentration decreases when cells are protected by F-1,6-BP. Several possible explanations exist as to where the calcium goes. One explanation is that calcium ion channels on the endoplasmic reticulum are activated to uptake [251] and store the excess Ca^{2+} ions in the

endoplasmic reticulum[252]. The ER can hold 5,000 times the concentration of Ca^{2+} as the cytoplasm.[253-255] This could be achieved in the Erk1/2 signaling pathway during oscillations in Ca^{2+} concentration creating a Ca^{2+} sink. However, it is important to remember that the Ca^{2+} oscillations are rather complex [251] with a desensitization phase, re-sensitization phase, and roles for IP3 to modulate these oscillations.[256, 257] For reference, FGF receptors are prevalent in adult neurons and astrocytes.[98, 99] The stabilization might be connected to IP3/IP4 molecules interacting with Ca^{2+} ion channels.[2, 96] Fahlman et al. reference [2] demonstrates how Erk 1/2 triggers a complex protective pathway involving Ca^{2+}, and by extension IP3 receptors that play important signaling roles.

7.6 Conclusions

Our research gathered data on a potential trigger of Erk 1/2 to better understand how a different class of molecules can trigger FGF protein folding. The data will help protect the brain from TBI by revealing the fundamental dynamics of bFGF. The fact that the short sugars like F-1,6-BP, Tri-3, and Tetra-1 improved thermal folding at the primary site supports the hypothesis that bFGF is the trigger for Erk1/2. Additionally, novel interactions at site 2' (Fig 6 and Table 2) showed an overlap with the thermal engineering point mutations, indicating potential allosteric effects on the folding. REMD simulations determined what the unfolded bFGF looks like and where it unfolds. The unfolding was confirmed by thermally engineered mutations to secondary sites known to increase thermal stability. The dynamics of the mutated protein proved that the primary site was stabilized. It was shown that therapeutic sugars trigger Erk 1/2 by activating bFGF through previously unidentified interactions, that some sugars interacted strongly in new binding sites, and a new force field was tested.

Deactivating point mutations were not found to dramatically impact the structure of the FGF protein. 1-2 Å level changes in the lower quartile were detected that were consistent with other activated FGF REMD simulations. Previous studies with FGF2 and FGF7 protein showed that radiolabeled FGF7 with equivalent point mutations remained successfully bound to the receptor but had dramatically reduced mitogenic activity. These results indicate that slight changes in the primary sugar binding site are important for mitogenic activation of the receptor.

Three (L225, S227, C229) of the eight total unstable amino acids on FGF, were accessible at site 2'. Two (S227 and C229) of the three are visited more frequently than the L225 due to conformational changes that repositioned the N-acetyl group. L225 is not stabilized by the tri-3 sugar because the free energy of that pose was higher than -0.5kcal/mol. However, there was a tetra-1 sugar that did interact with the adjacent side chain which would reinforce the structure at the site 2' FGF fold while simultaneously interacting with β-strands around site 1'. Figure 7.3 depicts the accurate bound states that are similar to crystal symmetry binding sites from a previous study of the most Tri-3 and Tetra-1 (optimized heparin fractions) Inositol Hexa Phosphate (IP6) and Inositol tetra phosphate (IP4) are also highly charged and also interact strongly with bFGF and aFGF. The conclusion is that the alternative sites were visited by the PS to stabilize them. The importance of these fundamental findings on GAGs binding to a new region is highlighted by the protective properties these proteins have on hypoxic TBI. This sulfate free activation may be unique among growth factors because dimerization is unnecessary and has led to the first effective treatments. At the same time, future studies with enoxaparin that combine fragments of heparin, tri-3, and tetra-1 in an optimized manner may better activate FGF, because each one reinforces different regions in ways that the others do not. The conclusion is that short carbohydrate molecules are neurologically important to pass through the blood brain barrier and activate

protective signaling[106, 107] especially for TBI. However, preventing high concentration, blood clotting, and alternative immune signaling are still relevant concerns. This work has proven that different fractions of heparin are optimized to interact with previously unidentified sites on bFGF that are known to play a role in folding and might be exploited to control signaling.

In addition to the strong binding energy, there are likely several candidates for further research that physically interact with the thermally unstable regions in ways that the primary binding site does not. The new binding interactions could be studied to determine if there are allosteric effects due to the direct contacts with the unstable amino acids. Complex physiological processes such as a cellular response to TBI need to be studied at multiple levels to assess the pros and cons of different approaches and identify new improvements.

Chapter 8: Conclusions and Future Work

The Osh4 project has yielded a method to calculate the free energy of the peptide binding to the membrane. There is reasonable confidence in the approach because it mimics a protein-protein PMF approach developed to study the conotoxin binding to the 5-HT_{3A} receptor using the recently improved CUFIX force field. Additionally, wet lab experimental data is already internally available to begin improving the parameterization of the force field. One hypothesis is that the Osh4 peptide has membrane-sensing properties because it senses lipid packing during binding. The alternative hypothesis is that the peptide folds differently in different lipid-packing environments. Either view yields valuable data to improve the modeling. The first improves the force field directly by providing data to fine-tune the lipid-protein interactions further. The second provides an alternative hypothesis, which better explains how the Osh4 protein function.

The WLBU2 research project seeks to identify the ideal features of AMPs. Our research furthered the overall goal by proposing the position and orientation of the WLBU2 peptide to understand how different amino acids interact with the bilayer. It could be further improved if more research data was provided about the role of different amino acids to increase the accuracy of the peptide. As mentioned above, improvements to the peptide-membrane forcefield could be made. Exploring other techniques like the implicit membrane models would also be helpful. The implicit membrane model uses a Generalized Born model to solve the equations and could be beneficial in predicting the orientations of different thicknesses and membrane properties like dielectric constants.

The 5-HT_{3A} project predicted different conotoxin structures and placements. The placements were verified with binding free energy and unbiased binding experiments. Caution should be applied when predicting structures to ensure the toxins do not form complexes linked

with disulfide bonds. This approach was successful because the toxin docked well into the allosteric binding pocket. Thanks to the Anton2 computer, it was possible to show how the toxin modified the 5-HT_{3A} extracellular domain. This seems to confirm that the toxin venoms alter the structure of the protein allosterically and could cause the pore to close. However, stable structures of the open pore conformation are still under review. With new crystal structures, it was possible to show that the pore remains open in the presence of the serotonin in the pocket.

Additionally, the restrained crystal structure had low conductance. The mutation was supposed to have a higher conductance. However, the asymmetric open conformation reveals how the pore might be more fully open. Future work could prove the pore's true open conformation and that the crystal structure is only an intermediate. Then, with the open pore, bind the toxins to the ECD and show that they close the pore. Or show if the serotonin can recover the signal from the toxin if only a few toxins are placed on the receptor.

Initially, there was a lot of interest in studying the TLR proteins and how they interacted. However, the force field appeared to produce too weak interactions with polysaccharides. There is the opportunity to re-parametrize the force field and confirm that the polysaccharides bind in the correct places on the TLR4-TLR4 protein and the TLR2-TLR2.

This research also led to the FGF research. Future work on this project could include progressing the research into an actual therapeutic. Much research with cells and a methodology would need to be developed to produce the specific sugars from scratch. It seems like such simple di- and trisaccharides could be made synthetically, but research still needs to be done to think all those steps through. Medical products have a lengthy approval process and require a significant investment of resources. I also believe there is something to be learned about stabilizing pharmaceutical proteins from thermal engineering work. If this can first be proved with hard

science from the wet lab, it will lead to a better understanding of how to engineer proteins to best interact with the sugars in the surrounding environment. This is currently more of an art than a science. Still, with inspiration from nature, it should be possible to learn motifs and patterns in the amino acid sequences that could be incorporated to improve the protein stability for formulation. If nothing else, it could become a problem for protein engineers who use the power of evolution to search for new motifs and applications. Proteins could be evolved to become more responsive to external stimuli like sugars to grow the list of motifs that could be engineered onto a given protein. The goal is for this to translate into better tolerance and shelf life for pharmaceuticals. Another area of research is to develop databases for how proteins behave at extremely high sugar concentrations. While the protein folding problem has received much attention recently, there is still a gap in understanding how ligands assist the folding process. I suspect that it would also be wise to try studying specific parts of an antibody that are conserved to find motifs that stabilize those segments first.

Appendix A

Table A1. Displays the distance in (Å) between the peptide and the phosphate component of the membrane with "avg" being the average distance and "min" as the minimum distance, and the different initial peptide orientations that were horizontal (H) or vertical (V) to the membrane.

Distance from Membrane Phosphate to Peptide Center of Mass	H1		H2		H3		V1		V2	
	Min	Avg	Min	Avg	Min	Avg	Min	Avg	Min	Avg
non-PIP2	8	16	8	16	17	19	6	22	3	28
PIP2	15	21	14	21	16	21	13	30	>25	36

Table A2: Details of the peptide binding: the time at which the peptide crossed the main phosphate plane, the distance that the deepest residues travel beneath the phosphate with ±standard error (z_{min}), and the average center of mass of peptide are relative to the phosphate plane with ±standard error (z_{COM}).

		H1a	H2a	H3a	V1a	V2a
Non-PIP2	Time (ns)	25	20	110	11	--
	zmin (Å)	-4.2 ± 2.3	-4.8 ± 1.8	-4.4 ± 2.0	-3.3 ± 1.9	--
	z_{COM} (Å)	5.6 ± 1.4	4.2 ± 1.3	4.9 ± 1.4	7.1 ± 1.5	--
PIP2	Time (ns)	73	108	67	45	42
	zmin (Å)	-6.3 ± 2.3	-5.3 ± 1.9	-4.9 ± 2.1	-7.0 ± 2.2	-5.1 ± 3.1
	z_{COM} (Å)	6.4 ± 4.0	6.1 ± 1.2	5.7 ± 1.5	2.6 ± 2.3	7.5 ± 1.7
		H1b	H2b	H3b	V1b	V2b
Non-PIP2 set 2	Time (ns)	10	40	30	50	25

	zmin (Å)	-4.2 ± 2.3	-4.8 ± 1.8	-4.4 ± 2.0	-3.3 ± 1.9	--
	z_{COM} (Å)	5.6 ± 1.4	4.2 ± 1.3	4.9 ± 1.4	7.1 ± 1.5	--
PIP2 set 2	Time (ns)	30	135	95	55	50
	zmin (Å)	-6.3 ± 2.3	-5.3 ± 1.9	-4.9 ± 2.1	-7.0 ± 2.2	-5.1 ± 3.1
	z_{COM} (Å)	6.4 ± 4.0	6.1 ± 1.2	5.7 ± 1.5	2.6 ± 2.3	7.5 ± 1.7

Table A3) This is the peptide-peptide interaction summary for the non-PIP2 HMMM data

HMMM Non-PIP2	D319-K315	K315-E338	E338-K334
H1a	0.9797	0.1649	0.3553
H2a	0	0	0.3511
H3a	0.0001	0.0005	0.4291
V1a	0.0001	0.0003	0.2773
V2a	0.5249	0.0012	0.3883
H1b	0.7215	0.0068	0.2839
H2b	0.8774	0.0001	0.3235
H3b	0.6053	0.0008	0.3159
V1b	0.9463	0.0069	0.4049
V2b	0.7451	0.0091	0.1936
Total Average	0.54004	0.01906	0.33229

Table A4) This is the full-length summary of the peptide-peptide salt bridge breakdown, indicating that H1a was the only simulation with high interaction across 315-338. However, it is still possible in other simulations, but only briefly.

Full Length	D319-K315	K315-E338	E338-K334
H1a	0.5229	0.3462	0.1565
H2a	0.7858	0.0001	0.2746

H1b	0	0.0911	0.2474
H2b	0	0.0003	0.2670
H1a PIP2	0.4279	0	0.3329
V2a PIP2	0	0.0012	0.3312
V1b PIP2	0.3042	0	0.2903
V2b PIP2	0	0.0012	0.2302
Total Average	0.2551	0.055013	0.2662625

Table A5) This table shows the total fractional breakdown of the peptide-peptide salt bridges between residues that may help to stabilize the hydrophobic region and also stabilize the charged amino acids that could enhance the dwell time on the liposome surface when they are not interacting with charged headgroups

HMMM PIP2	D319-K315	K315-E338	E338-K334
H1a PIP2	0.3204	0.0005	0.1678
H2a PIP2	0.6671	0.0315	0.3156
H3a PIP2	0	0.0573	0.4054
V1a PIP2	0	0.0113	0.5021
V2a PIP2	0.2250	0.0183	0.3141
H1b PIP2	0.1099	0.0043	0.4530
H2b PIP2	0.3748	0.0014	0.4916
H3b PIP2	0	0.0042	0.2299
V1b PIP2	0	0.0047	0.3954
V2b PIP2	0.4004	0.0135	0.4139
Total Average	0.20976	0.0147	0.36888

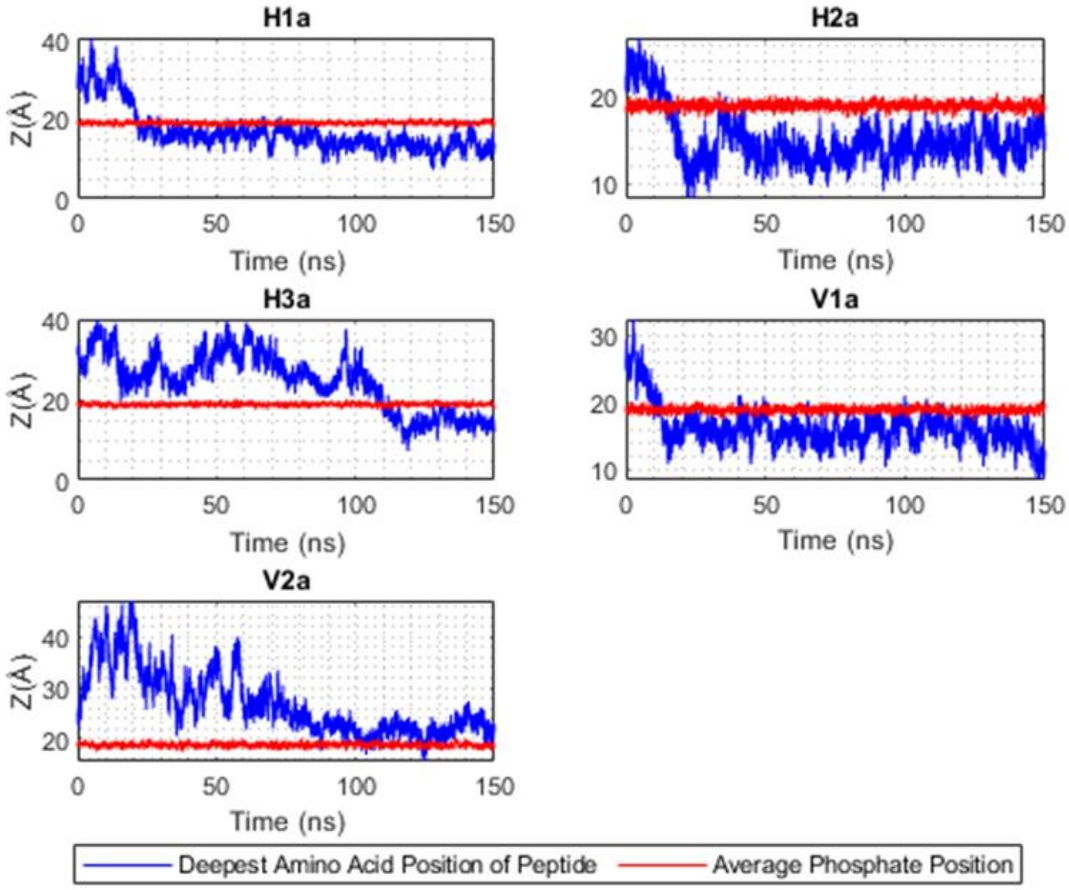


Figure A1. Time series of deepest (when bound) or closest (before binding) residues to the membrane. Position is relative to the phosphates, and the red line is the main phosphate position. The main point is that the peptide starts to detach from the membrane and connects deeply to the membrane beneath the main phosphate layer in all trials except V2. A natural fluctuation is expected, and there may also be some variation in the average position, but nothing significant. The position is defined as the center of mass of the deepest residue in the membrane.

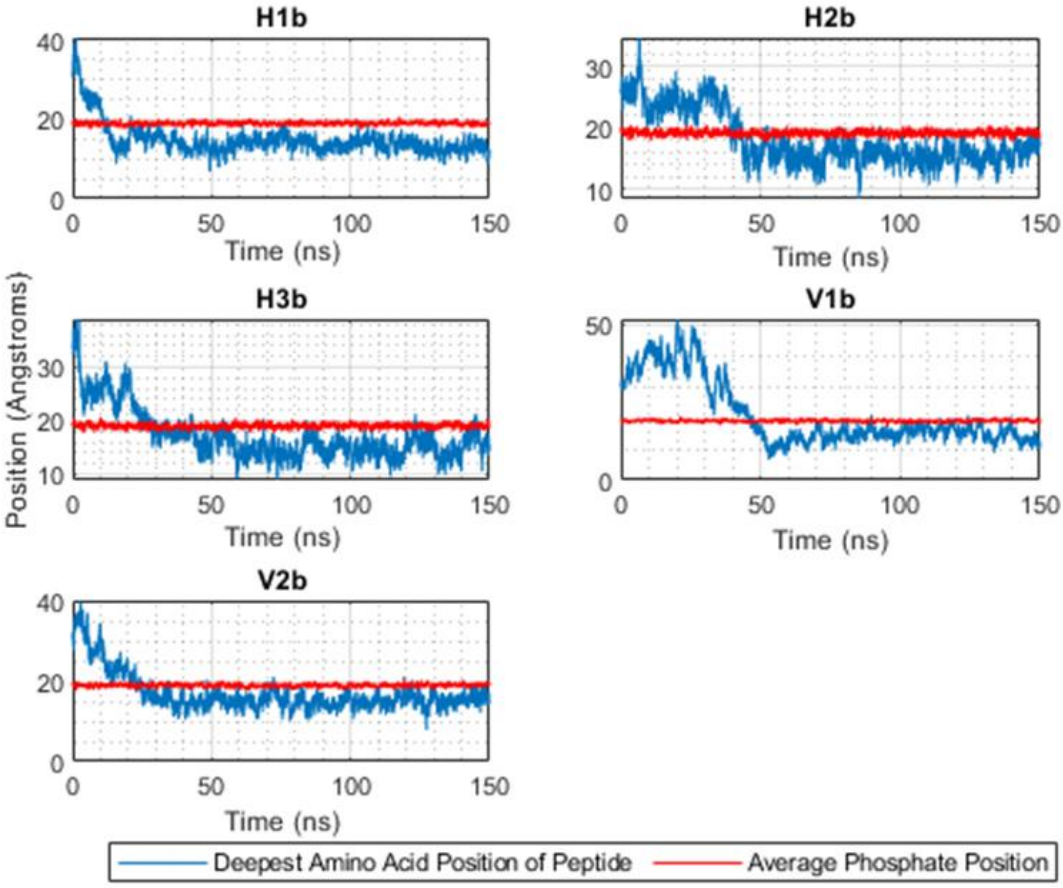


Figure A2. Showing the deepest amino acid residue COM for the second set of HMMM trials. These show that all simulations are bound within roughly the same amount of time.

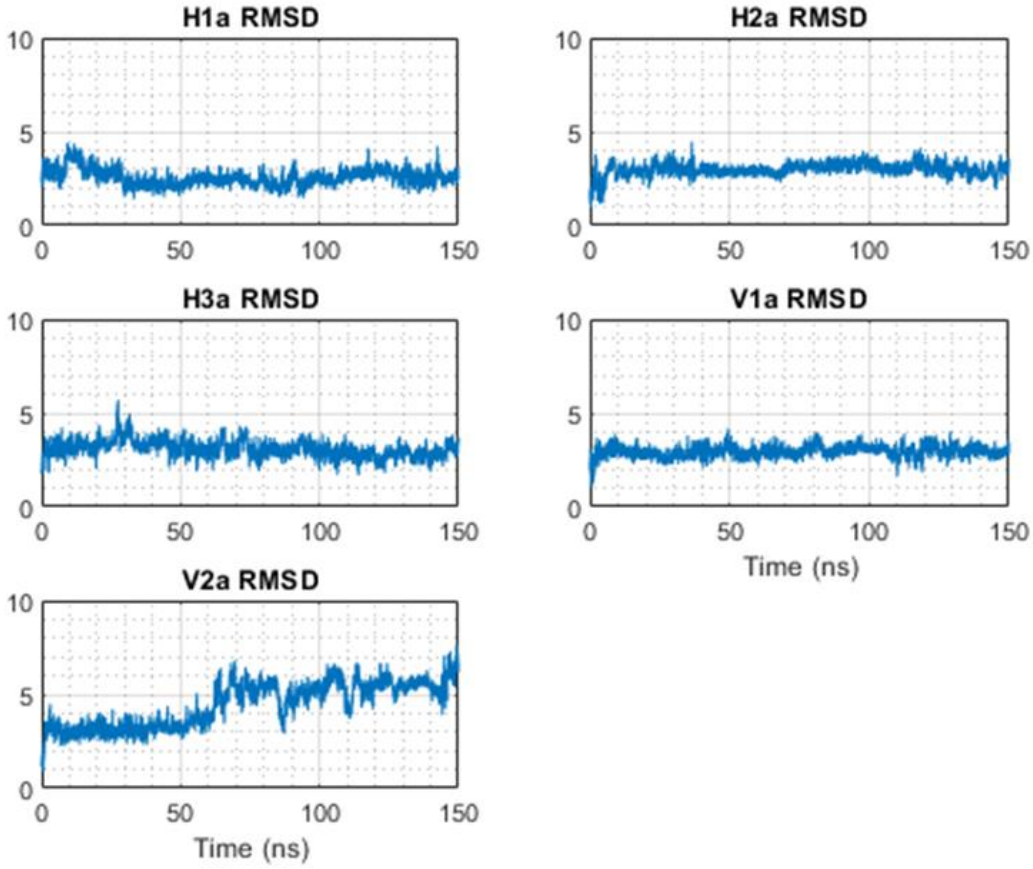


Figure A3. RMSD of the HMMM section of the non-pip simulations with an average value of RMSD around 3-4. The main point of the figure is that all of the simulations in this series converged to a stable state except V2, which has a higher RMSD than the others due to some deformations on the α_7 helix.

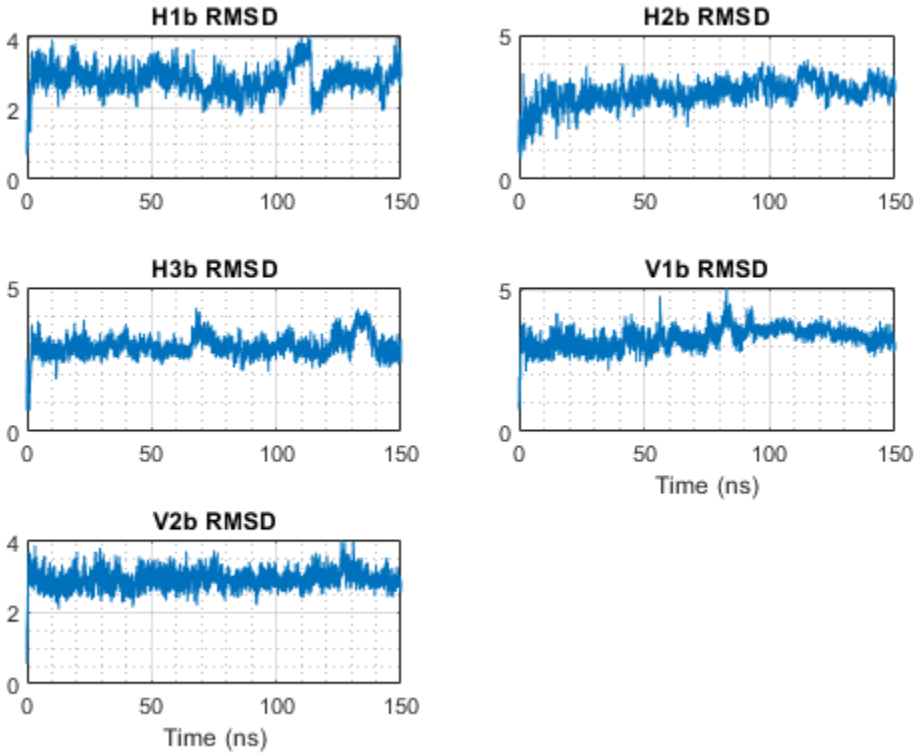


Figure A4. Showing the RMSD of the second set of HMMM trials. The main point of the figure is that all of the simulations in this set have approximately the same RMSD, showing convergence to a stable state.

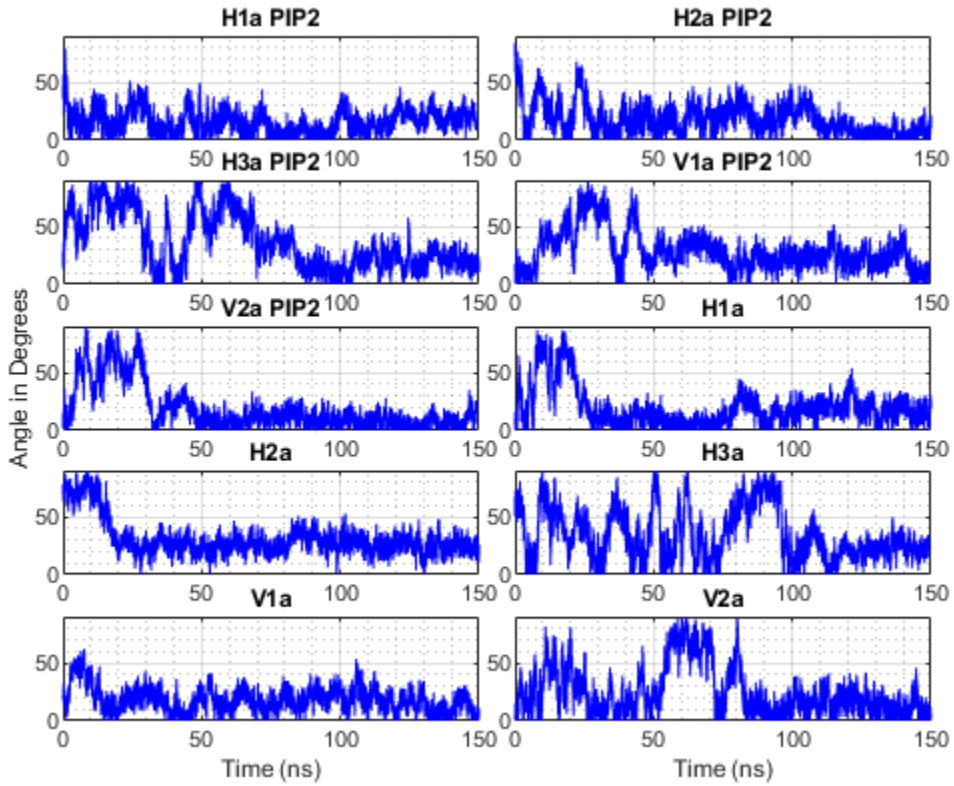


Figure A5. Showing the ϕ angle of the HMMM trials throughout the simulations. The main point is that these trials converge to a stable low angle and that there is lots of rearrangement before that moment.

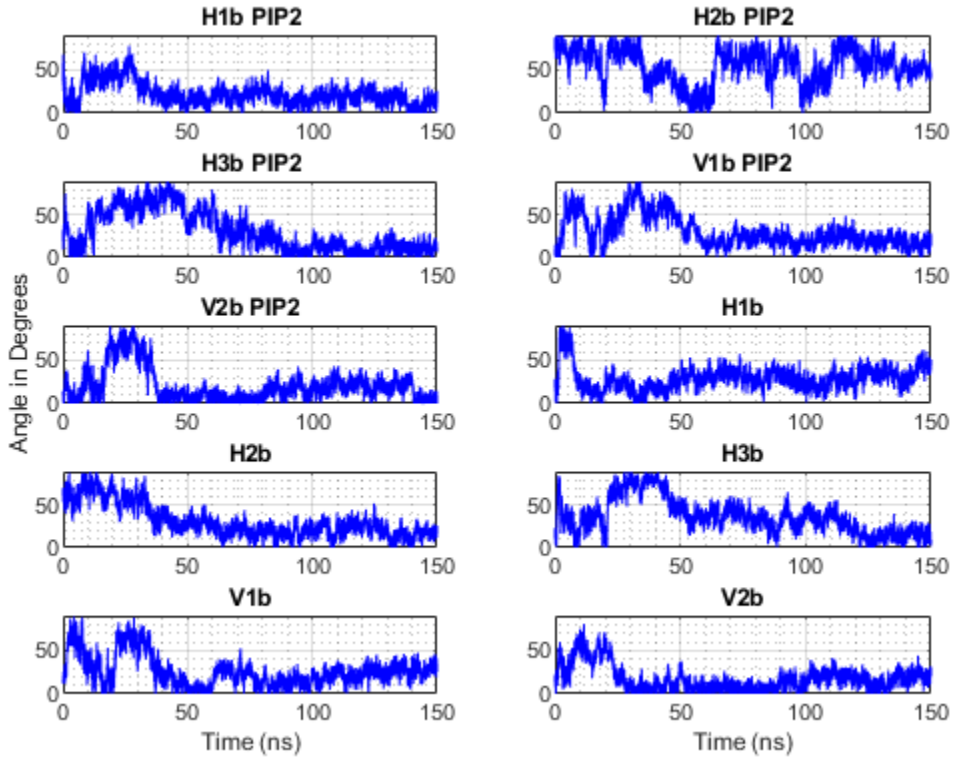


Figure A6. This is the second set of trials for the ϕ angle. The main point of this figure is to show that there are a variety of different angles that exist at binding.

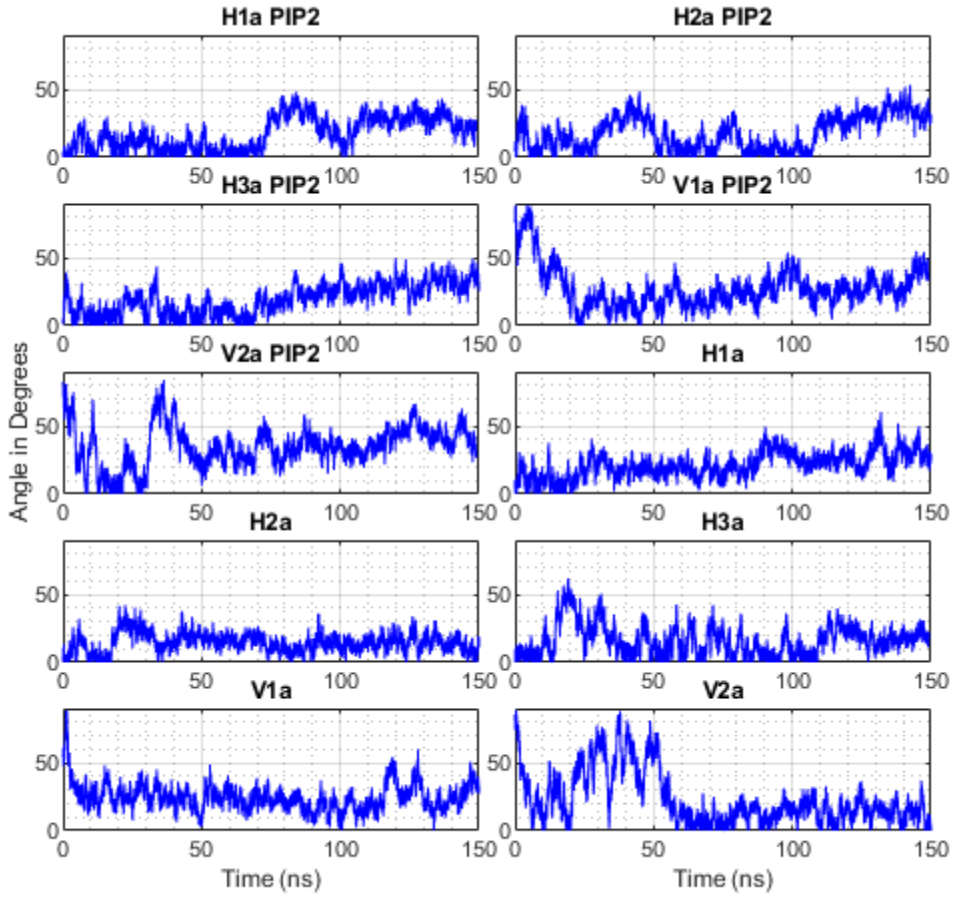


Figure A7. The individual θ_2, defined in Figure 1, shows the individual angles formed between the α_6 and helix and the membrane throughout the non-PIP membrane simulations. Two states were identified with this angle: a high and low angle. The low angle can be seen whenever the angle frequently reaches 0 degrees, and the high angle is observed when the angle never touches 0 degrees for some time.

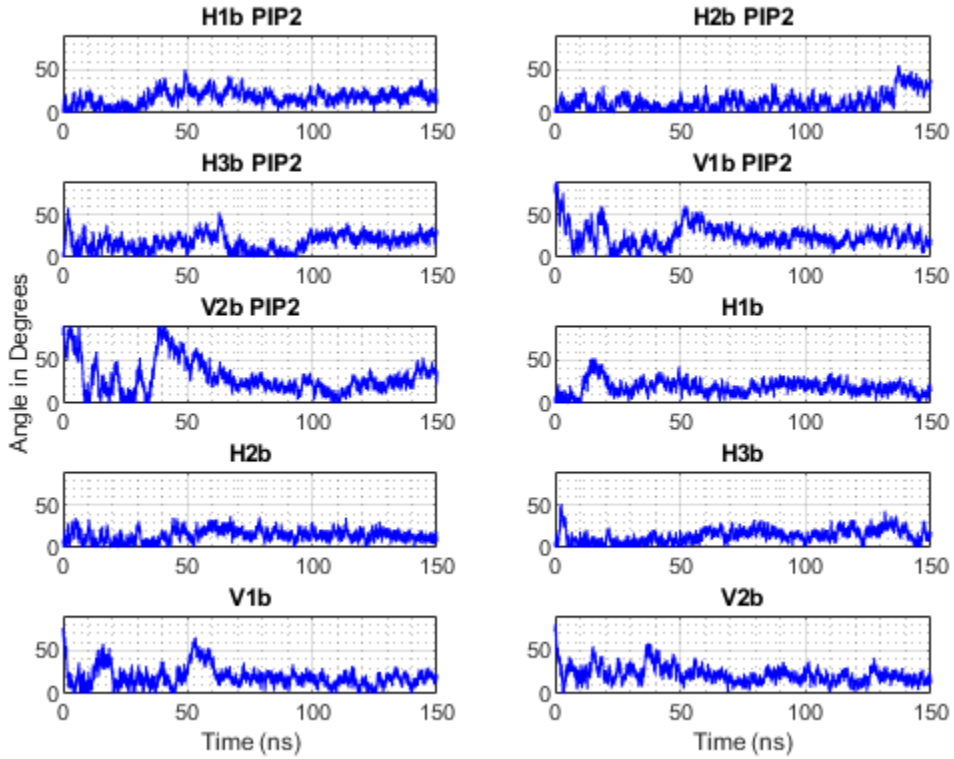


Figure A8. Showing the individual θ_2 angles for the second set of trials. It was noticed that some of the simulations touched down to zero and others didn't, such as H2b touching zero and H1b not reaching zero as frequently after binding.

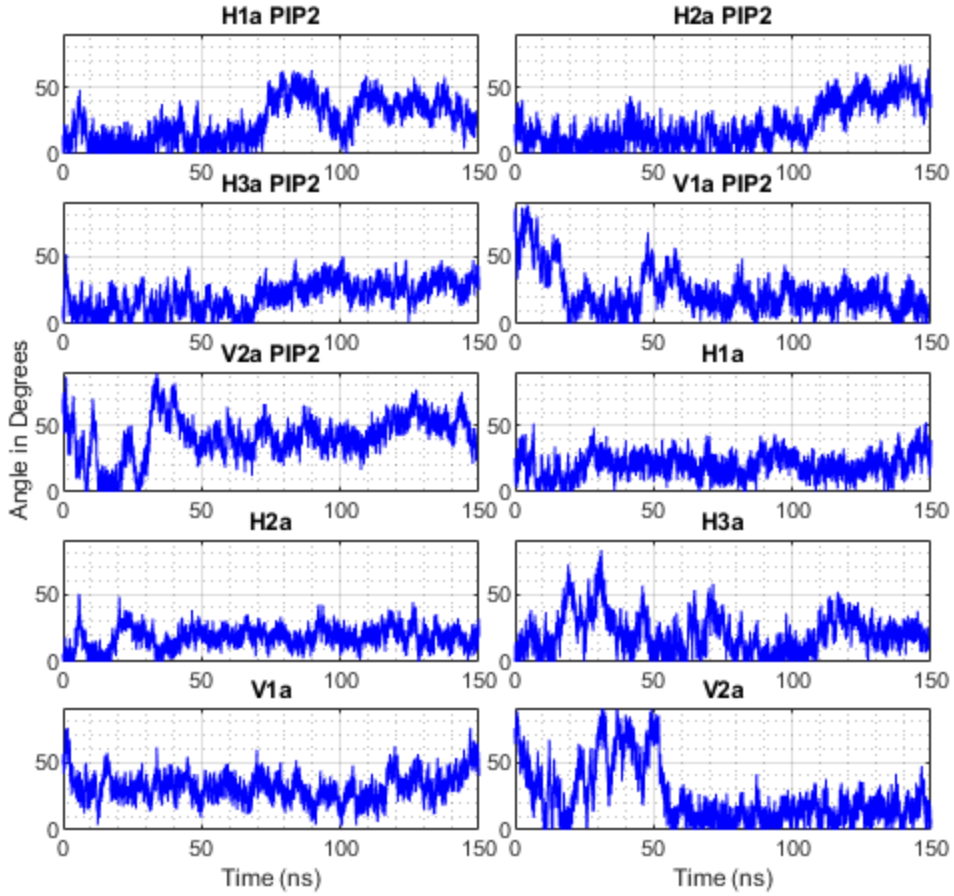


Figure A9. The individual $\theta_{i,}$ defined in Figure 1, shows the individual angles formed between the α6 helix and the membrane over the first round of membrane simulations. This variable also shows a difference in high and low angles. Still, it is not expected to impact the interaction energy because the α_6 is shorter and remains in better contact with the membrane compared to the α_6 helix.

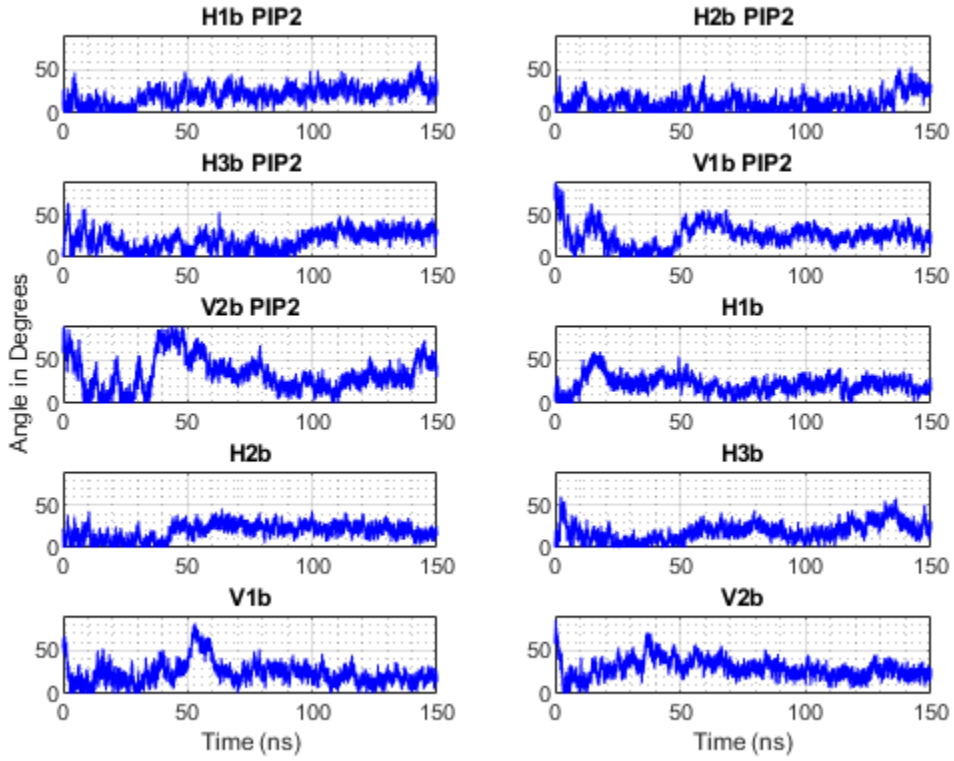


Figure A10. Showing the individual θ_i angles. There were differences in this angle, and they were calculated to help visualize the simulations.

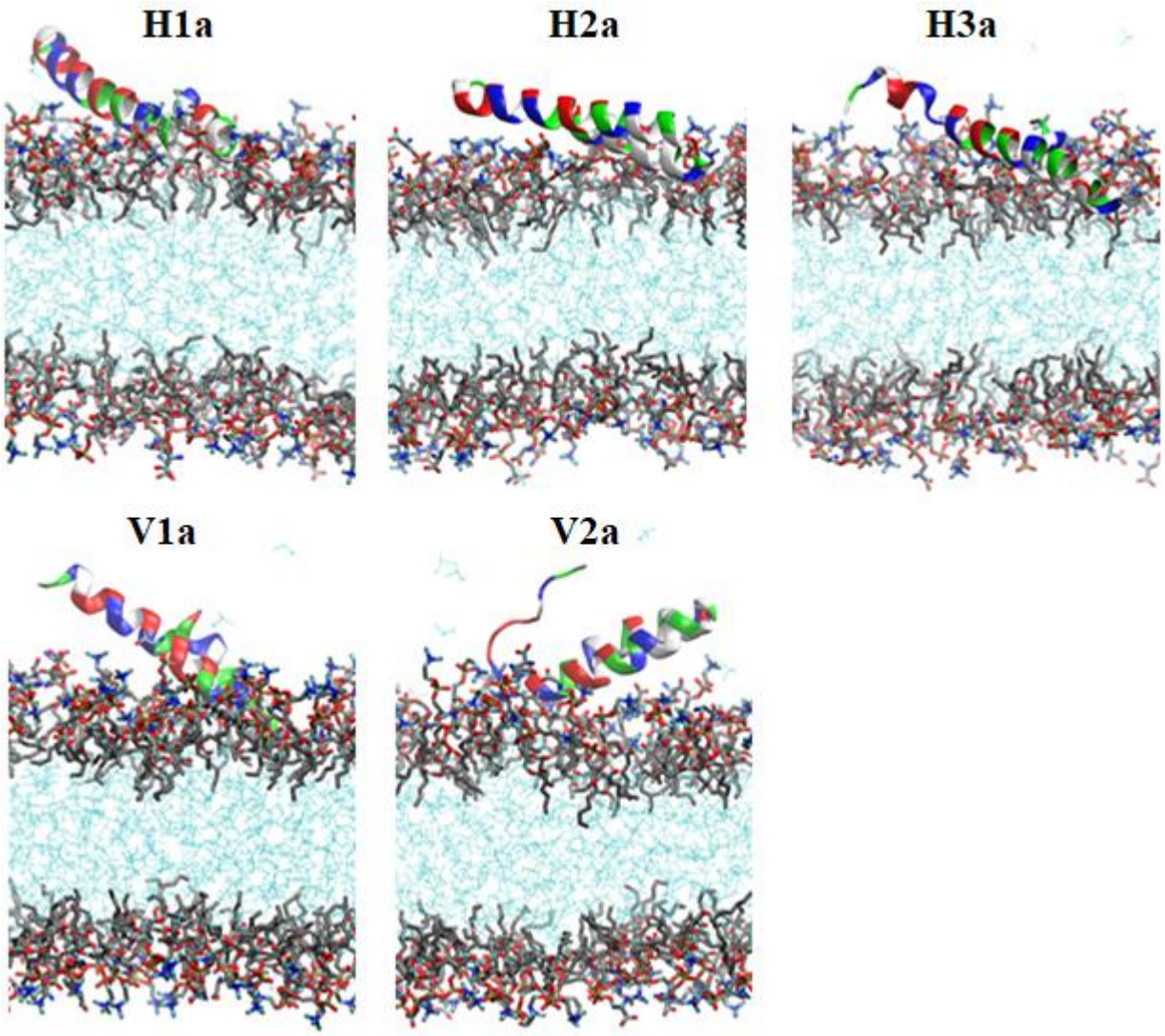


Figure A11. This shows the final bound states of the HMMM simulations. It is possible to see a slight difference in the angle of the α_7 helix (the longer helix) between H1 and H2 where the higher angle is on H1 and the lower angle is viewed on H2, and simulation V2 didn't bind properly.

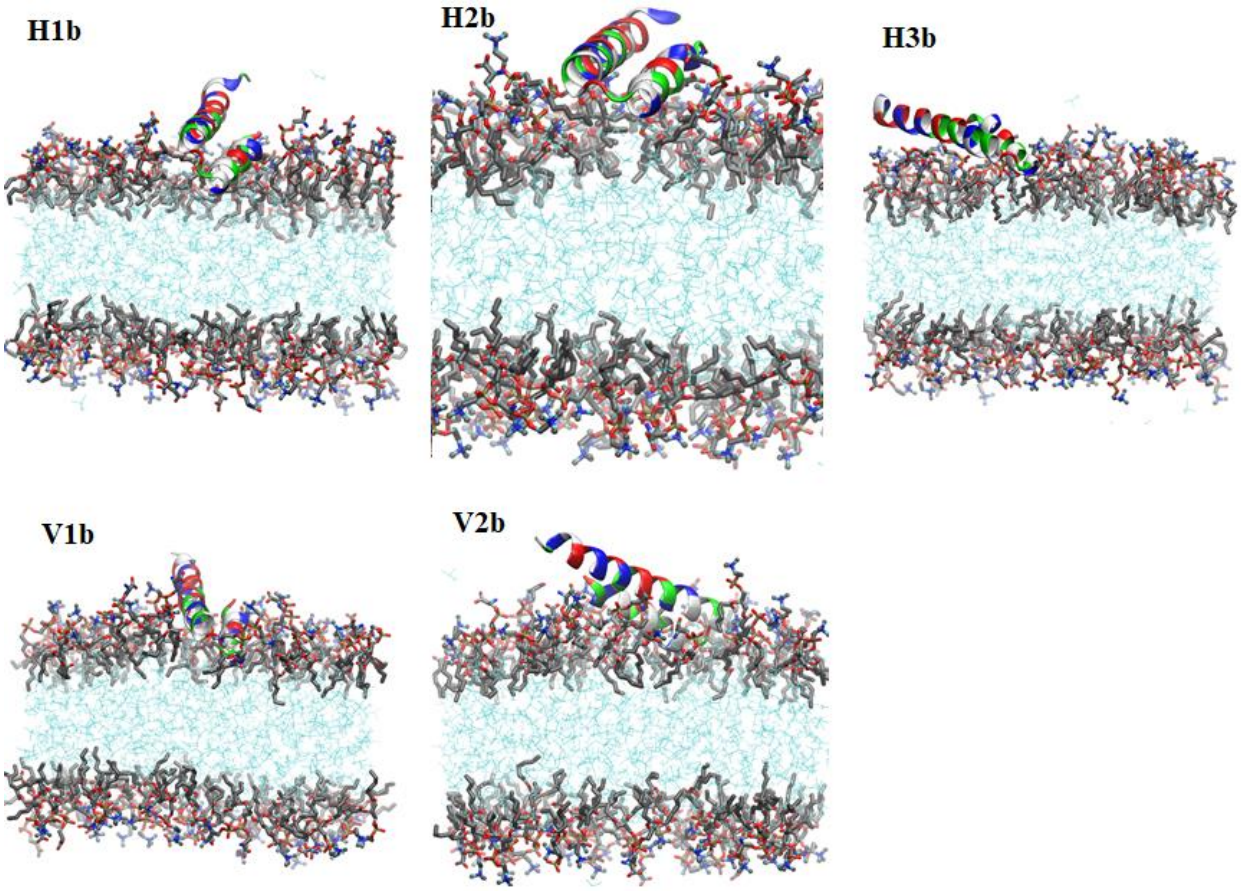


Figure A12. These images of the second set of simulations show different views of the final bound states. The main point is to show that all of these simulations are bound to the membrane and to show the orientation of the peptide in the membrane.

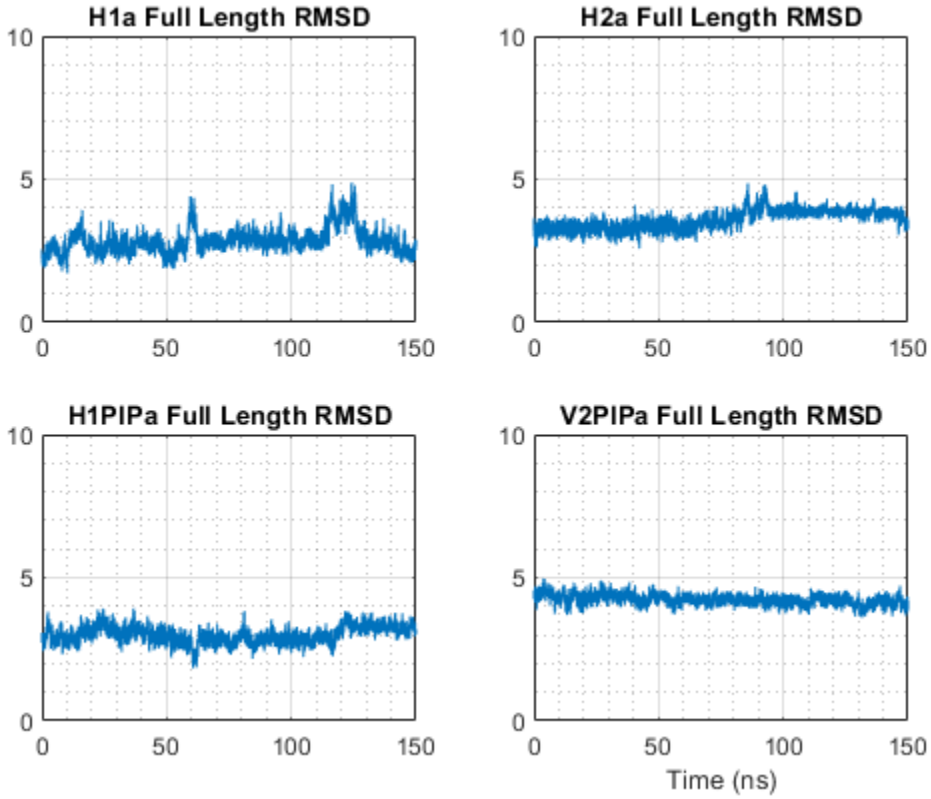


Figure A13. RMSD of the all-atom section of the simulations. This indicates that the full-length simulations had RMSDs within the expected range.

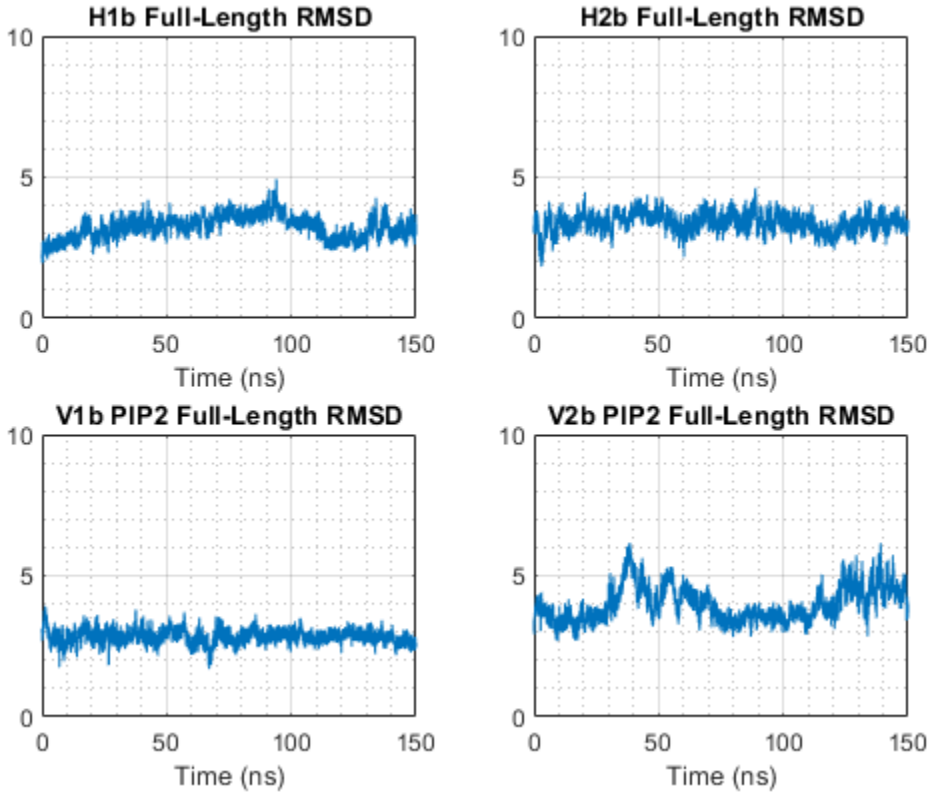


Figure A14. RMSD of the second set of full-length trials shows that the simulations also fell into an expected range and that none of these simulations deviated much above 5 angstroms.

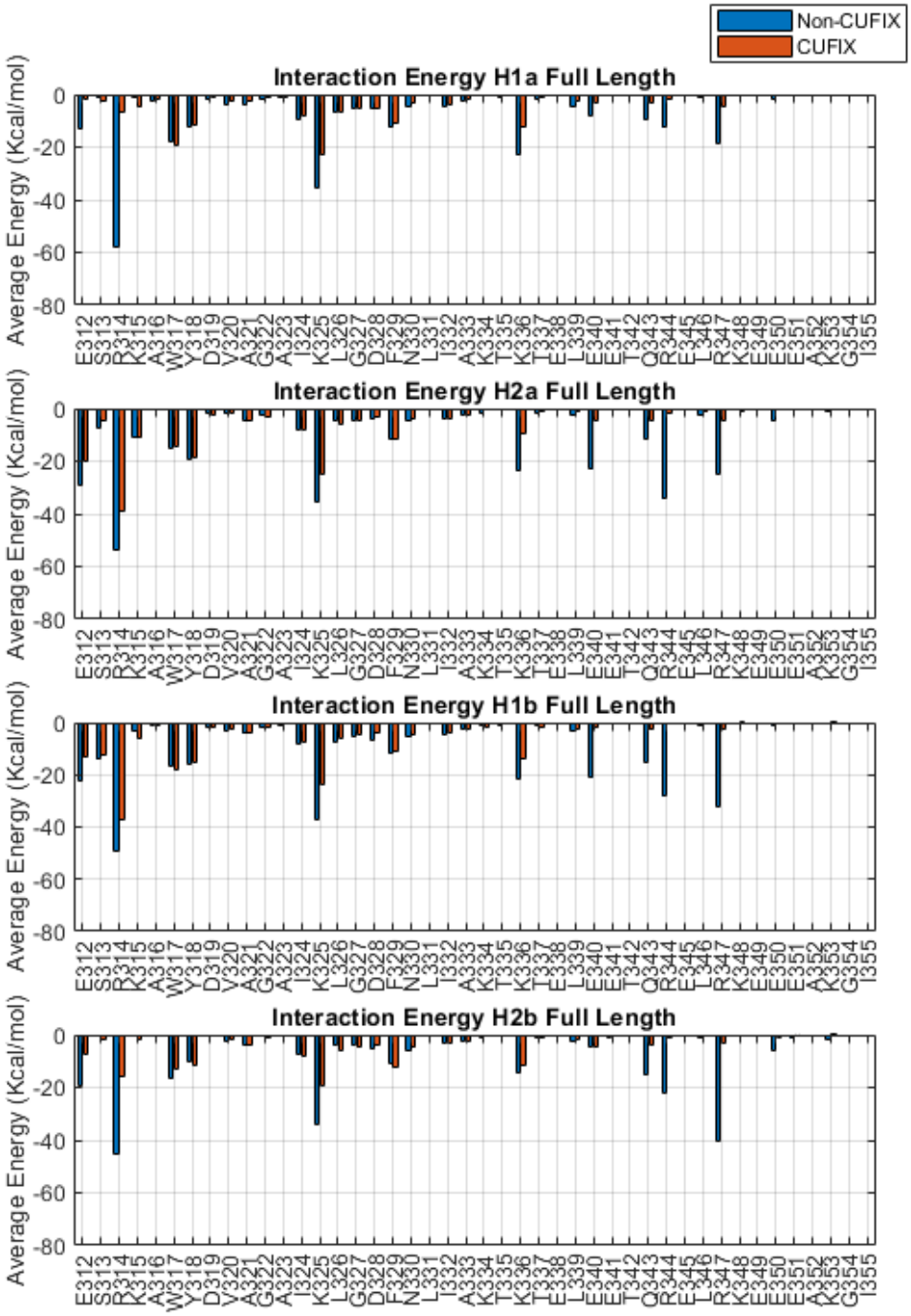


Figure A15. Displays the energy per residue for full-length membrane simulations; interestingly, R314, E312, and K325 have much higher energies in H2 than the ones seen in H1. Additionally, residues R344 and R347 may have higher energy because of closer proximity to the membrane. It can also be seen that the residues on the $\alpha_7$helix (beyond K325) of H1 have lower interaction energy than H2 due to the higher angle.

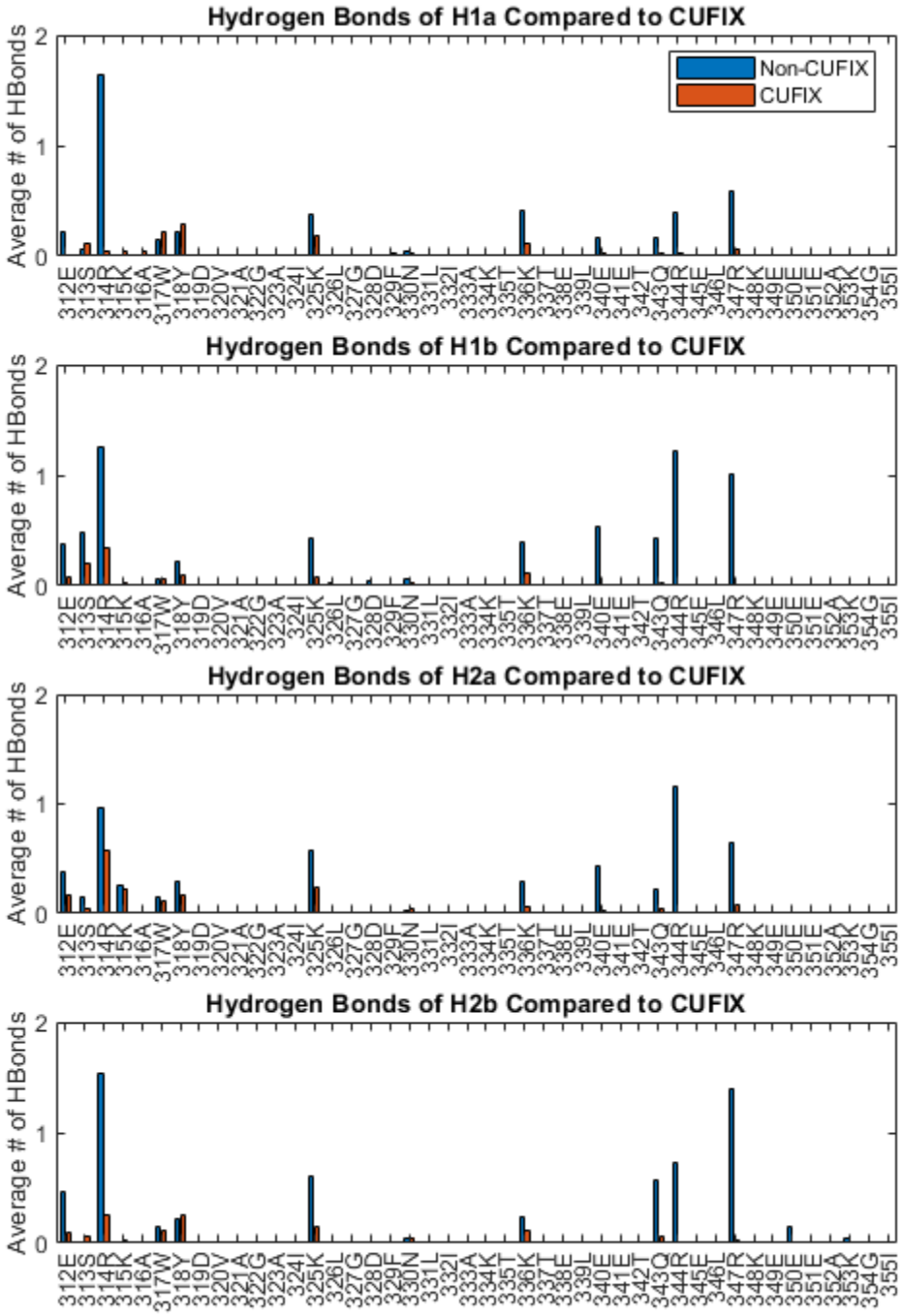


Figure A16. This shows the hydrogen bonding of the two non-PIP full-length simulations.

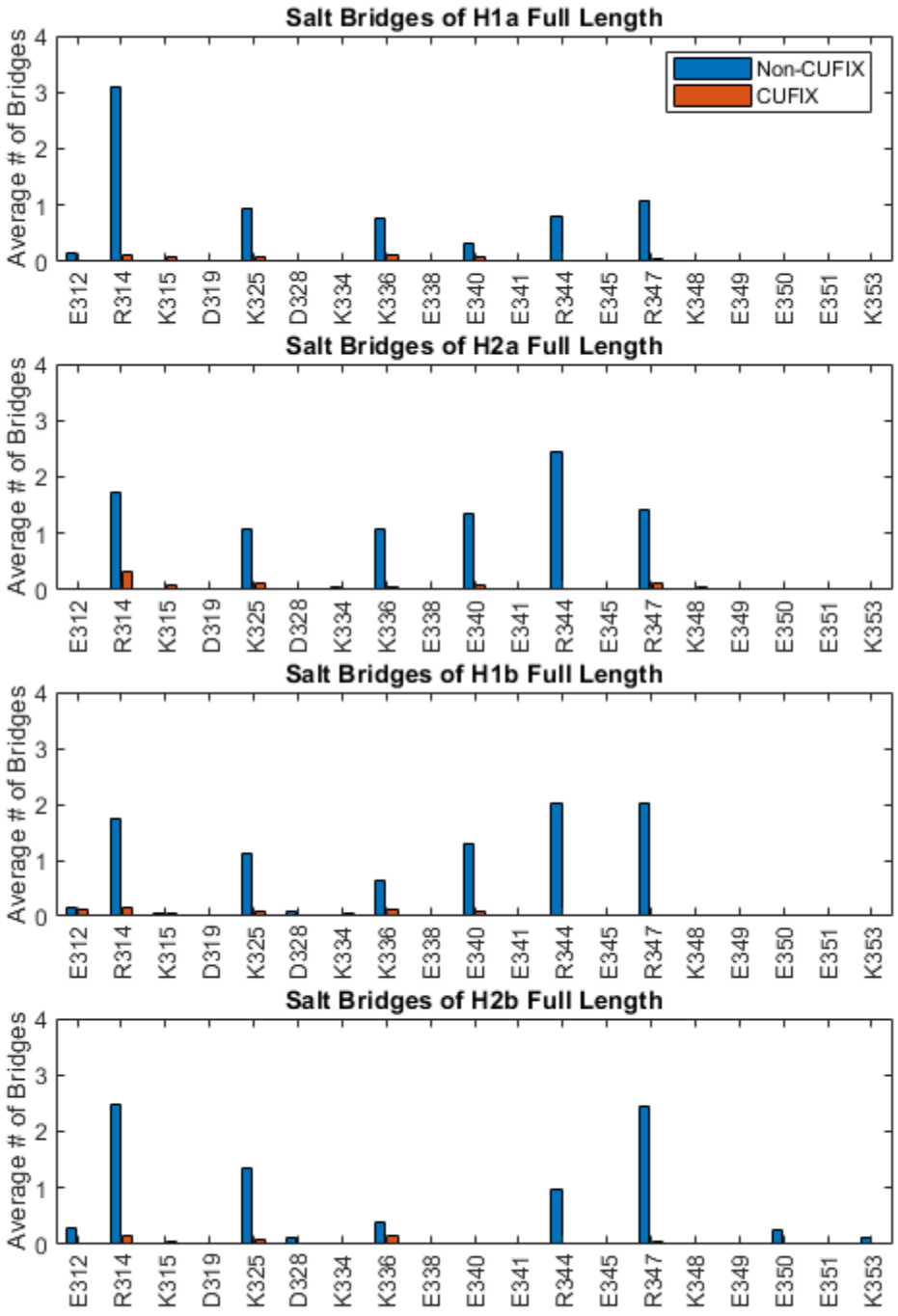


Figure A17. Displaying the number of salt bridges for the full-length non-PIP2 simulations, simulation H1 had fewer salt bridges than simulation H2; additionally, supplementary data in Figure A4 shows that the hydrogen bonds cause the lower angle.

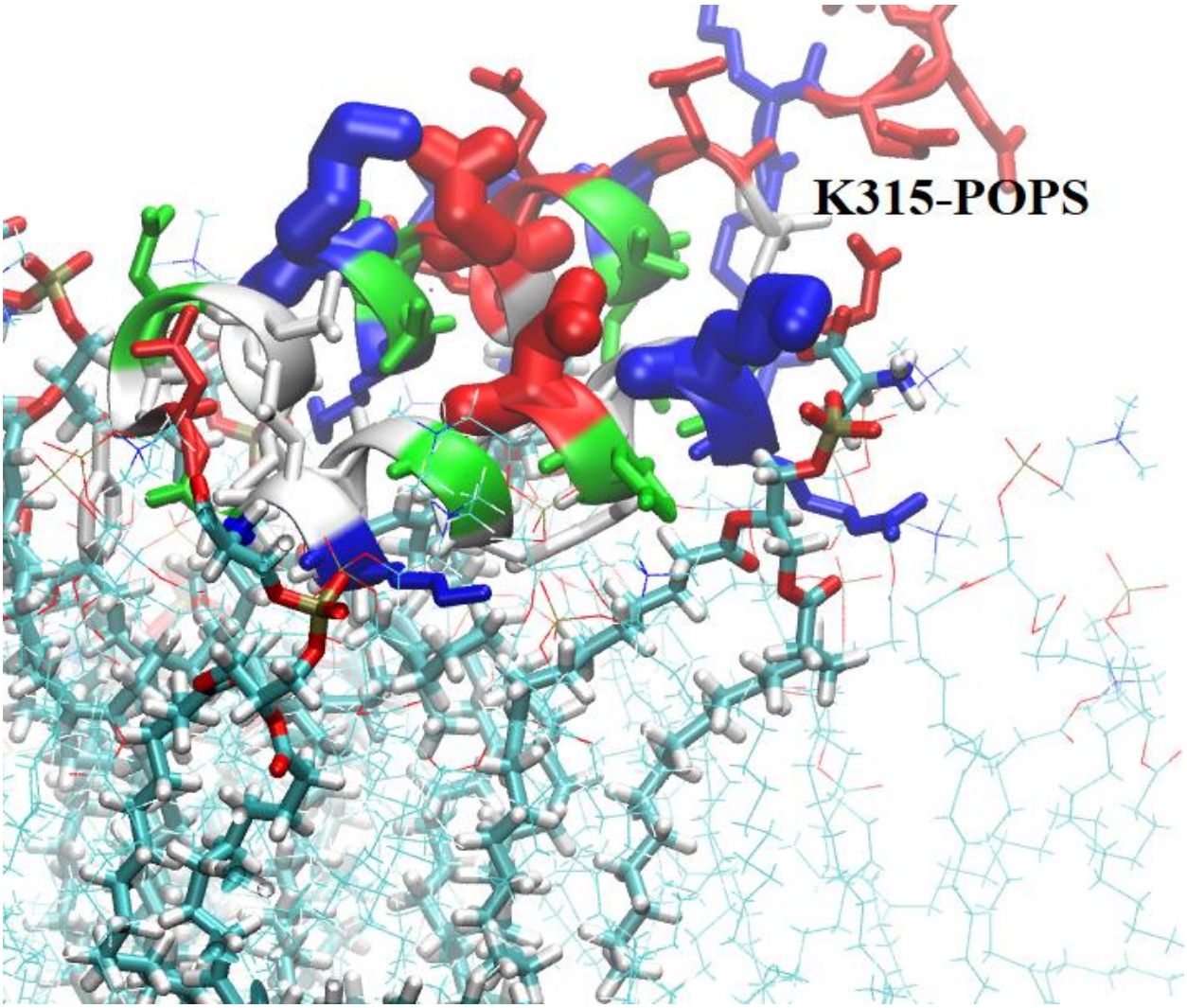


Figure A18. H2a shows the residue K315 interacting with a POPS lipid. Much of the time, this was visualized. It does not appear to be in contact with the POPS lipid, but the salt bridges indicate high numbers for both force fields. H1a was observed to interact with POPS in the second force field, which is consistent with the salt bridge results. While there was a POPS lipid close by in simulation H1b, it was not detected in the salt bridges. Something similar happened with H2b, where there were POPS lipids around, but the K315 didn't interact with them, at least in H1a, and H2b residue K315 was higher above the membrane from the placement graphs, maybe placing them out of reach.

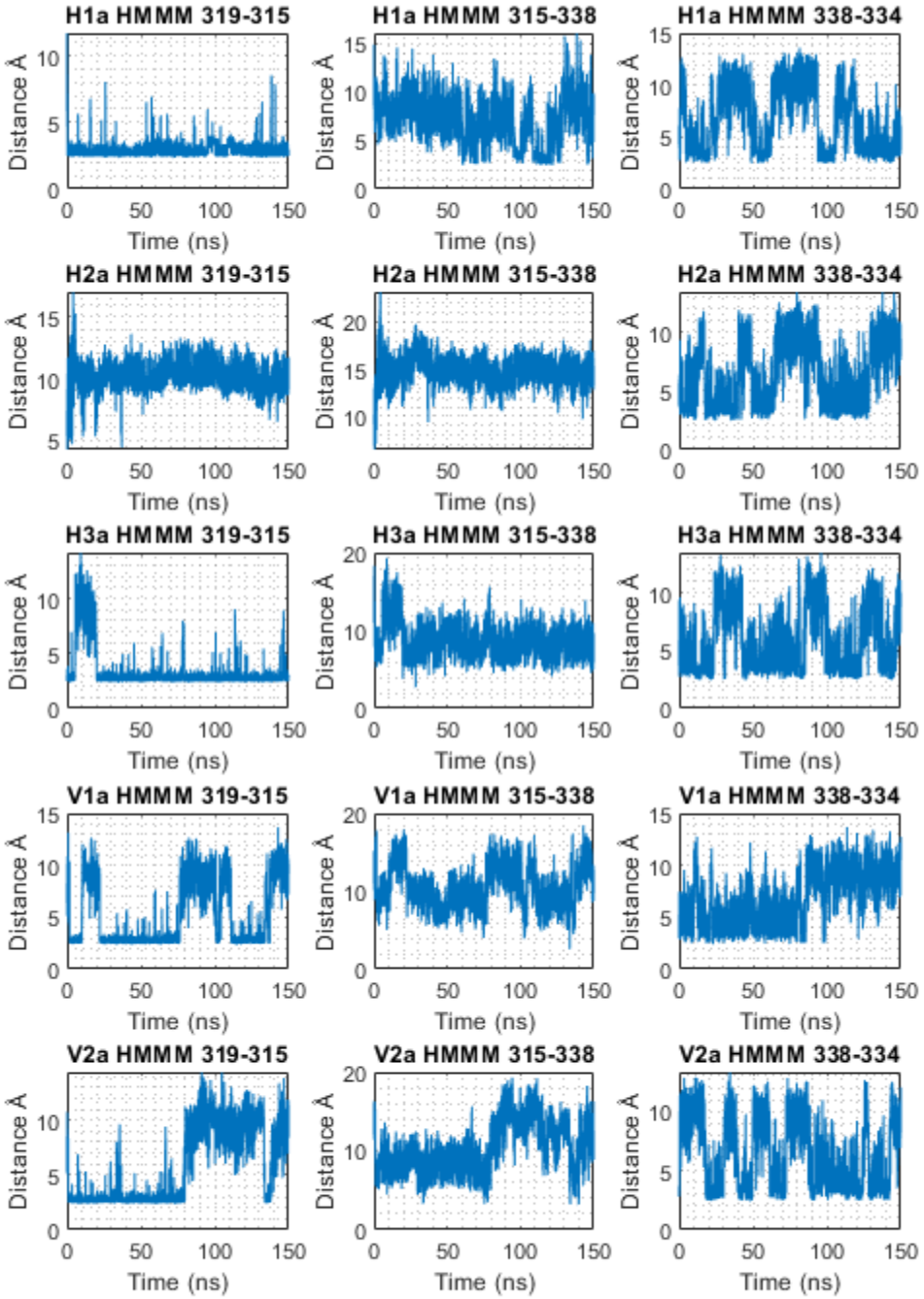


Figure A19. Image of the distance between residues in HMMM simulations that form salt bridges. These are important because these long-range water-mediated interactions help to stabilize the α_6-α_7 helices, and they are in the range of critical.

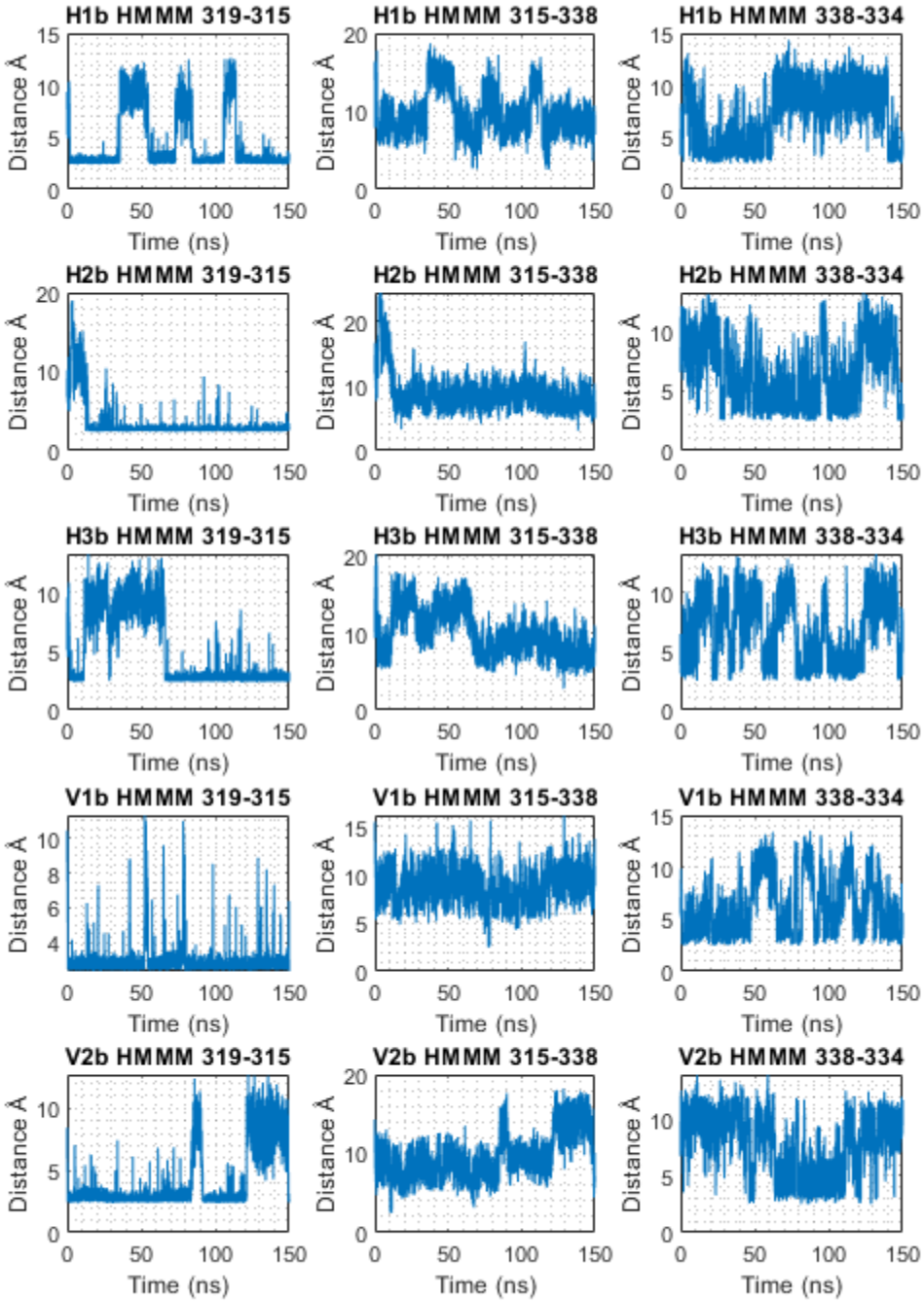


Figure A20. Image showing the second set of distances between many of these simulations were closer than 4 angstroms but not frequently observed for the bridge between 315-338 that connects the two helices.

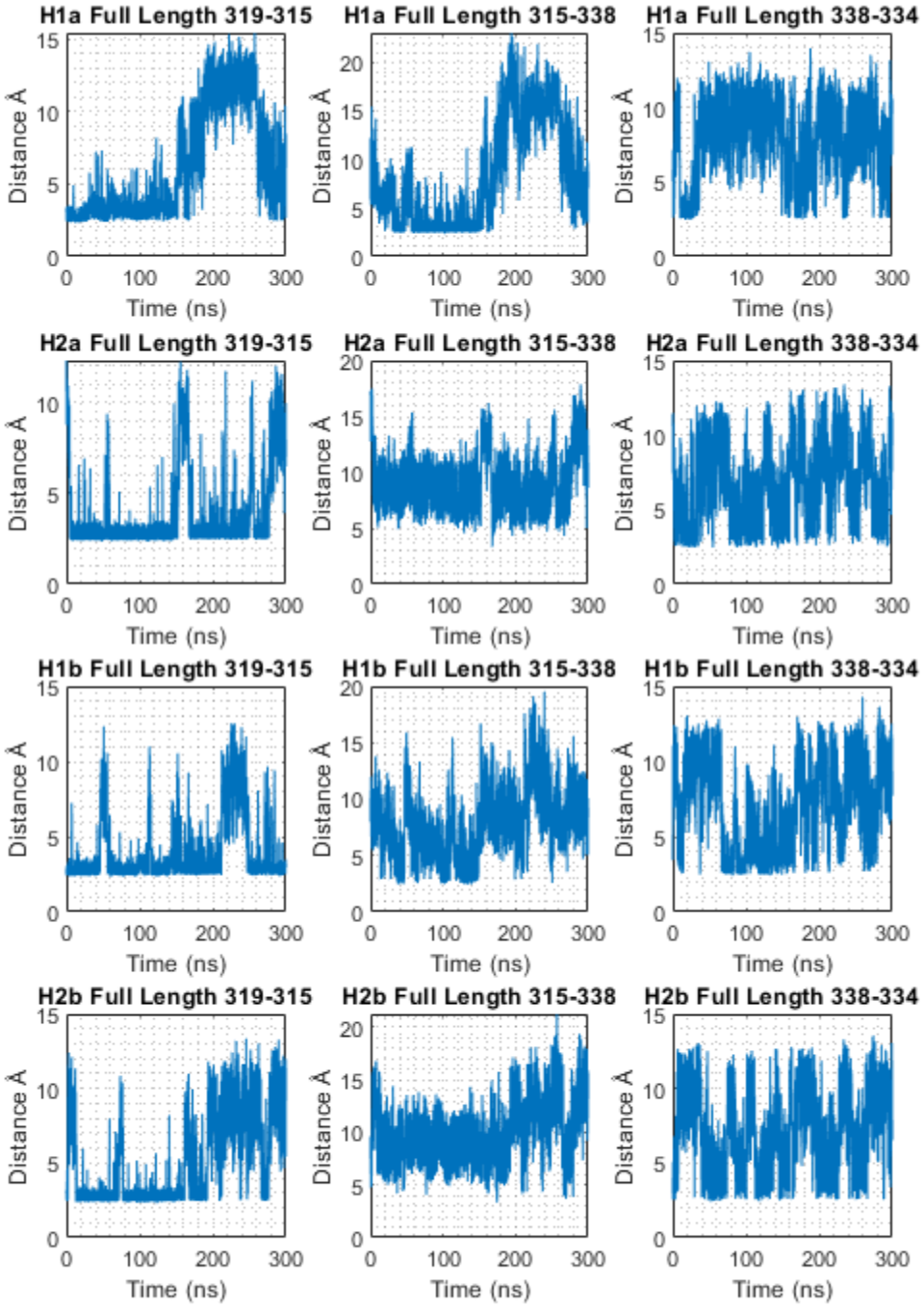


Figure A21. Image showing the full-length simulation distances vs time. Again, this shows high salt bridges for 319-315, 338-334, and 315-338, but only in simulation H1a.

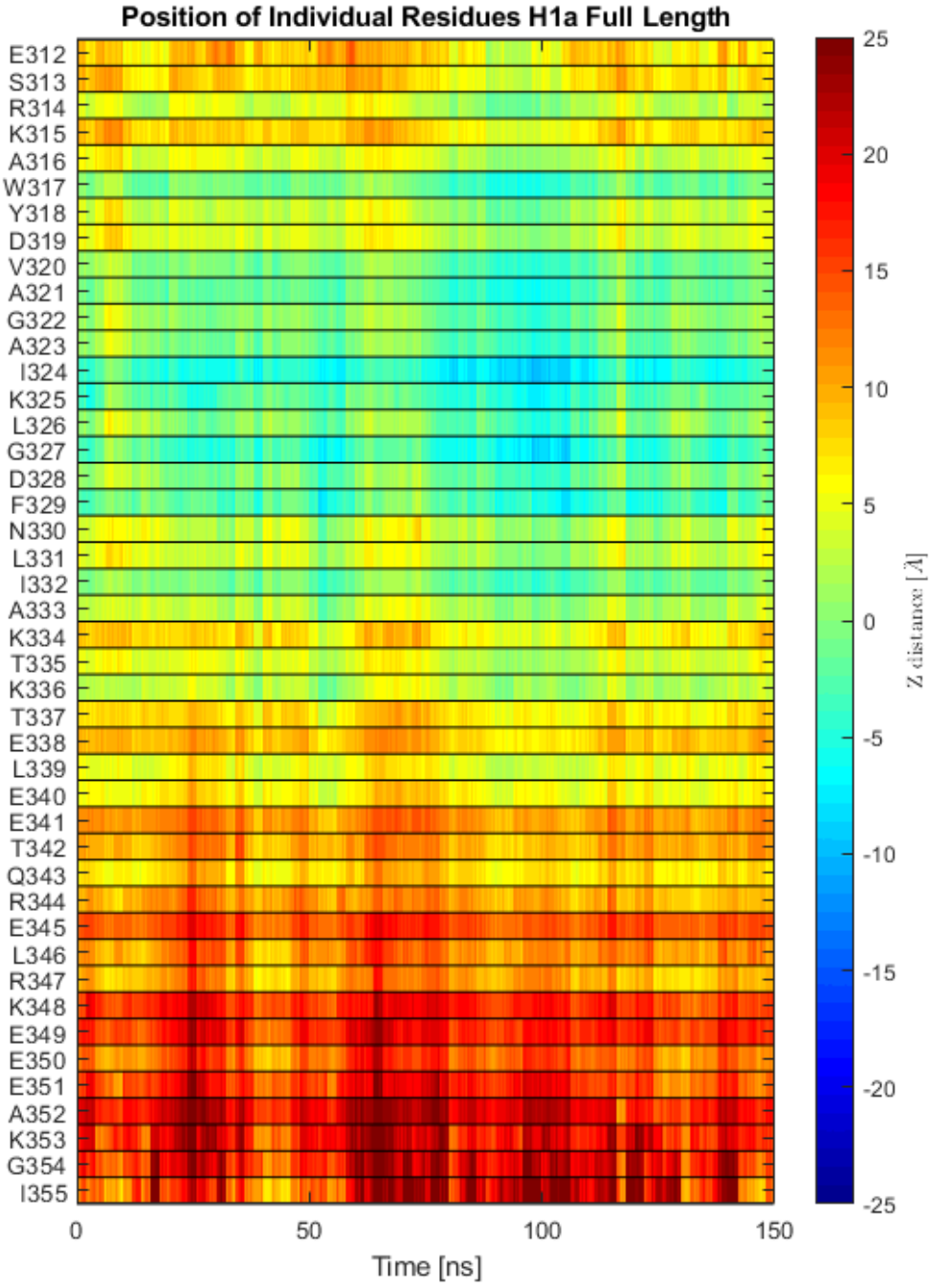


Figure A22. Image describing the average position of individual peptide residues location relative to the average membrane phosphate. It is seen that the peptide begins bound and remains bound throughout the simulation and that the residues in the middle are the deepest below the phosphate.

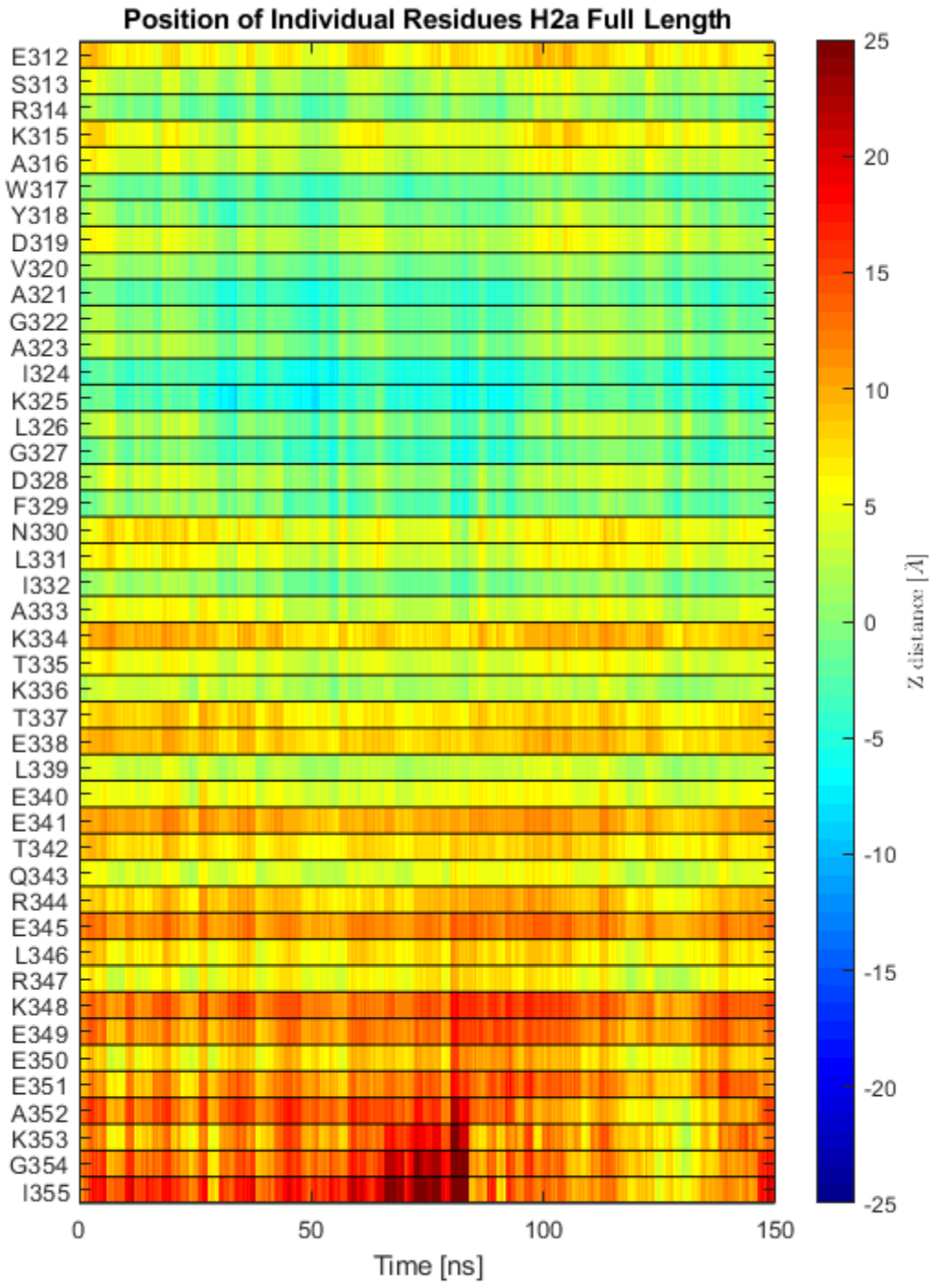


Figure A23. Image describing the average position of individual peptide residues location relative to the average membrane phosphate. Residues in the middle of the peptide are the deepest.

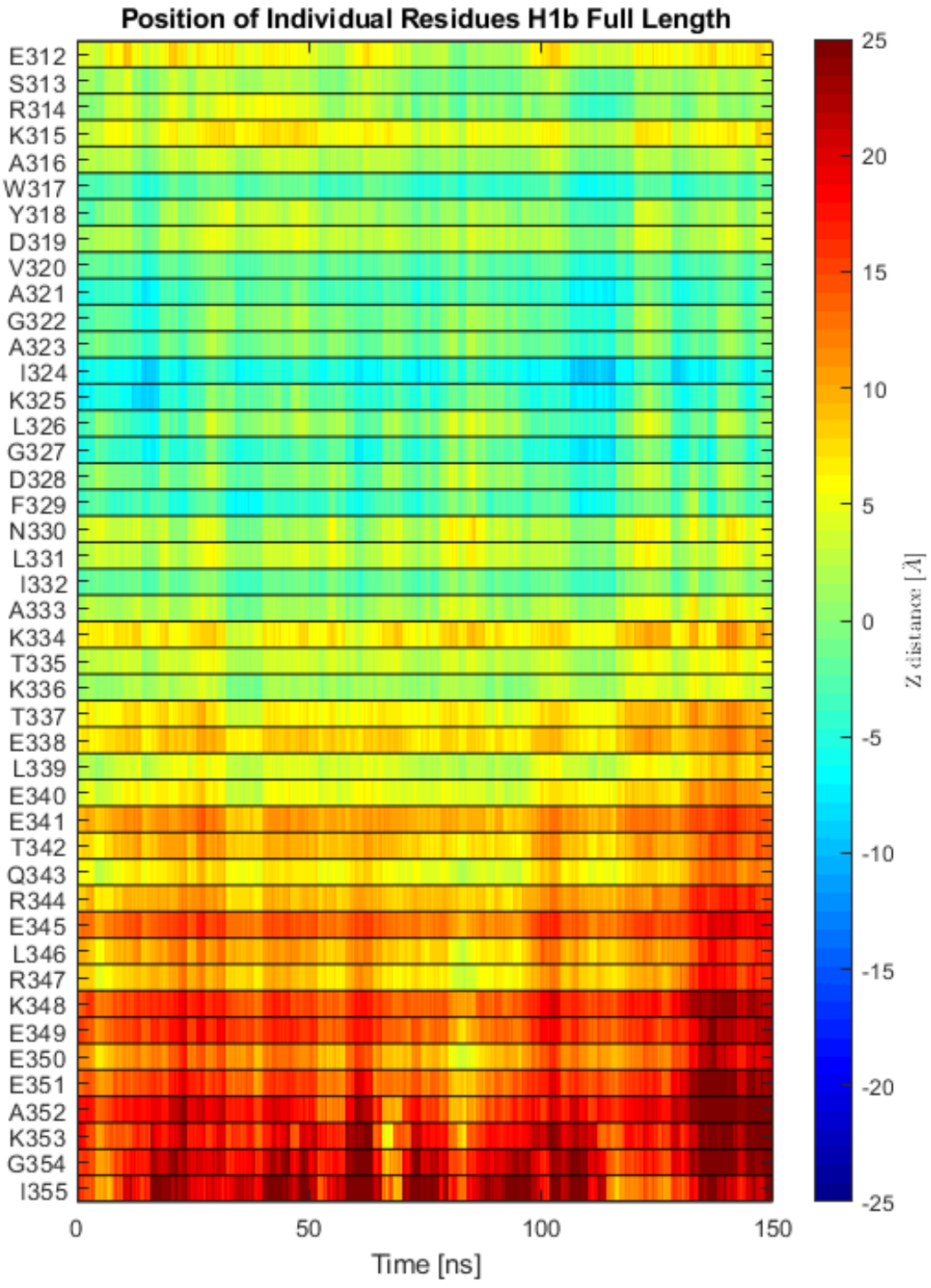


Figure A24. The image shows the placement of H1b follows the same trends as the other simulations, except that K315 is a little deeper than in other simulations.

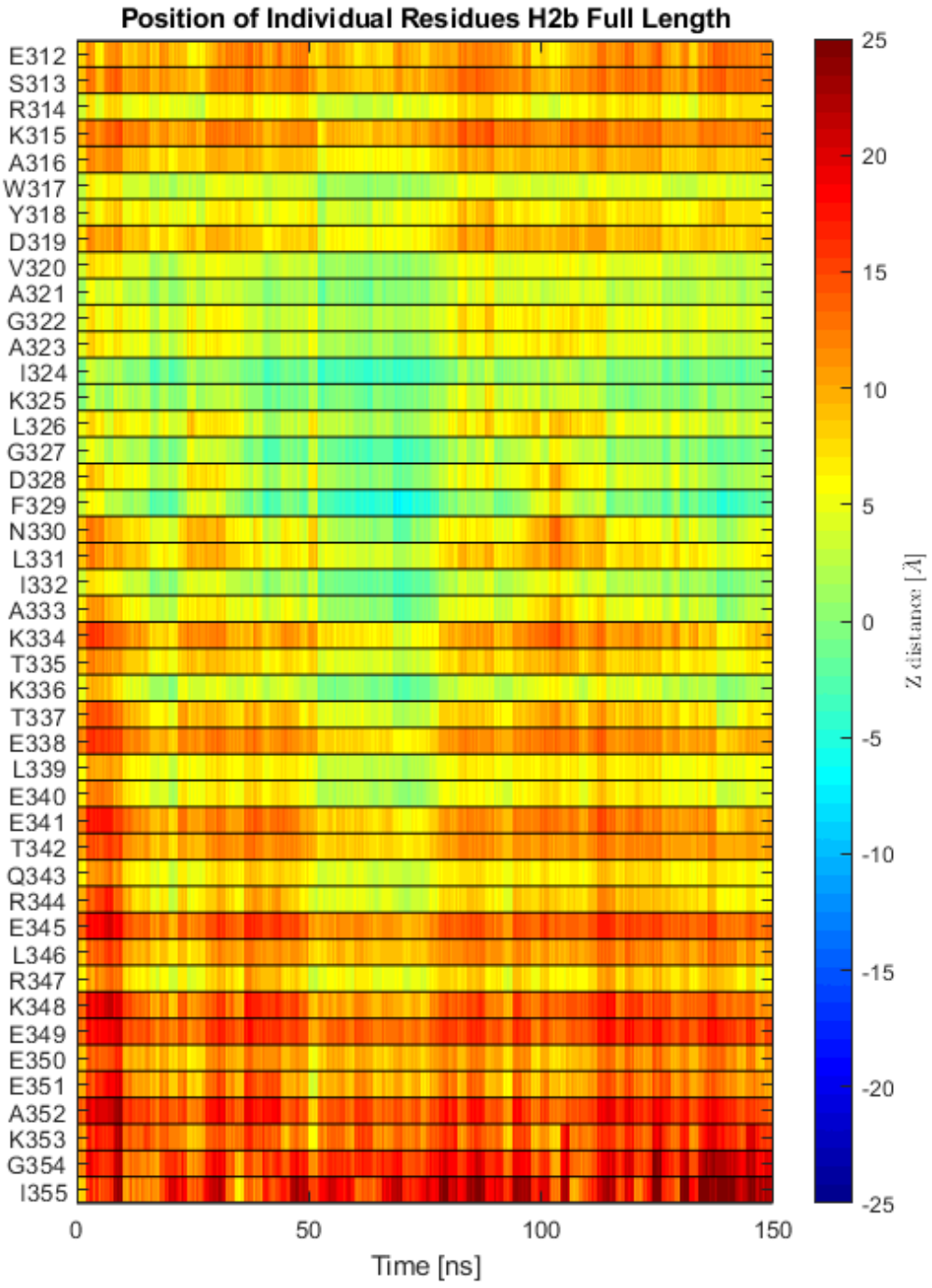


Figure A25. Image showing the placement of the H2b simulation following the same trends as the other simulations except that k315 is not as deep, possibly explaining why this simulation had less interaction energy on residue K315 even with POPS lipids close by to interact with.

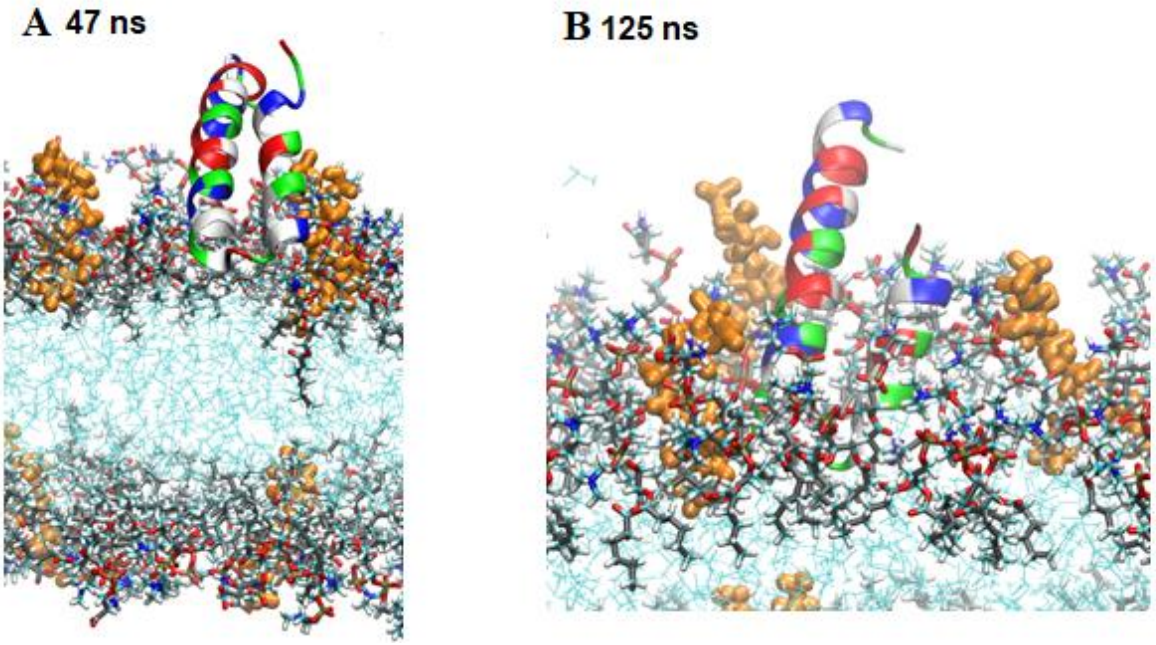


Figure A26. (a) simulation V2a, the only simulation bound with the α7 helix deepest, is depicted moments before binding, and (b) simulation H1a shows the bound state with PIP lipids interacting with the exposed residues inaccessible to the main membrane lipids.

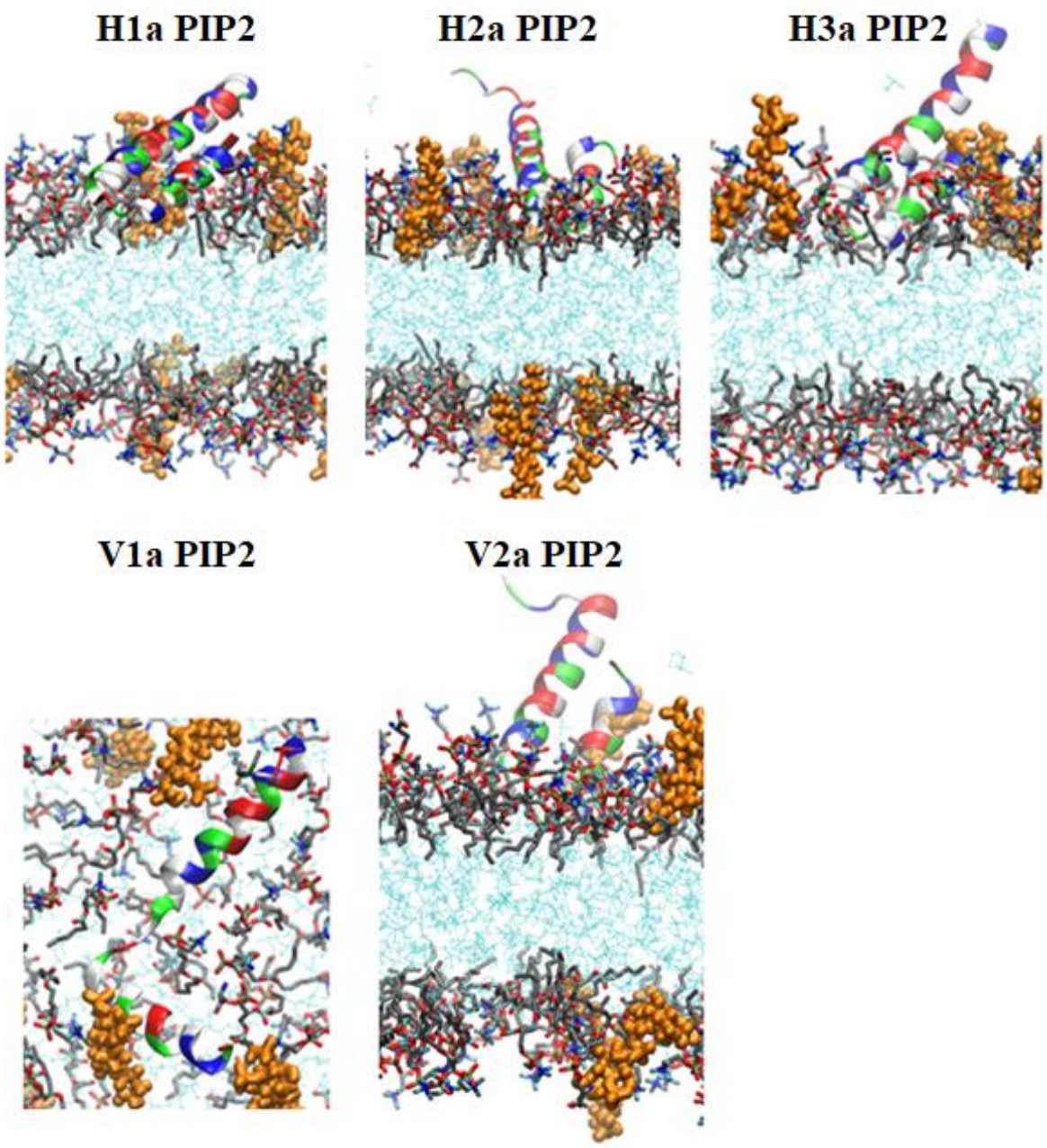


Figure A27. Shows the final bound states for the HMMM simulations of PIP2 lipids. These generally have higher angles compared to the lower angles seen in the non-pip trials, likely due to the charges of the PIP2 lipids. The trial V1 PIP2 had the α_6 and α_7 helices split apart.

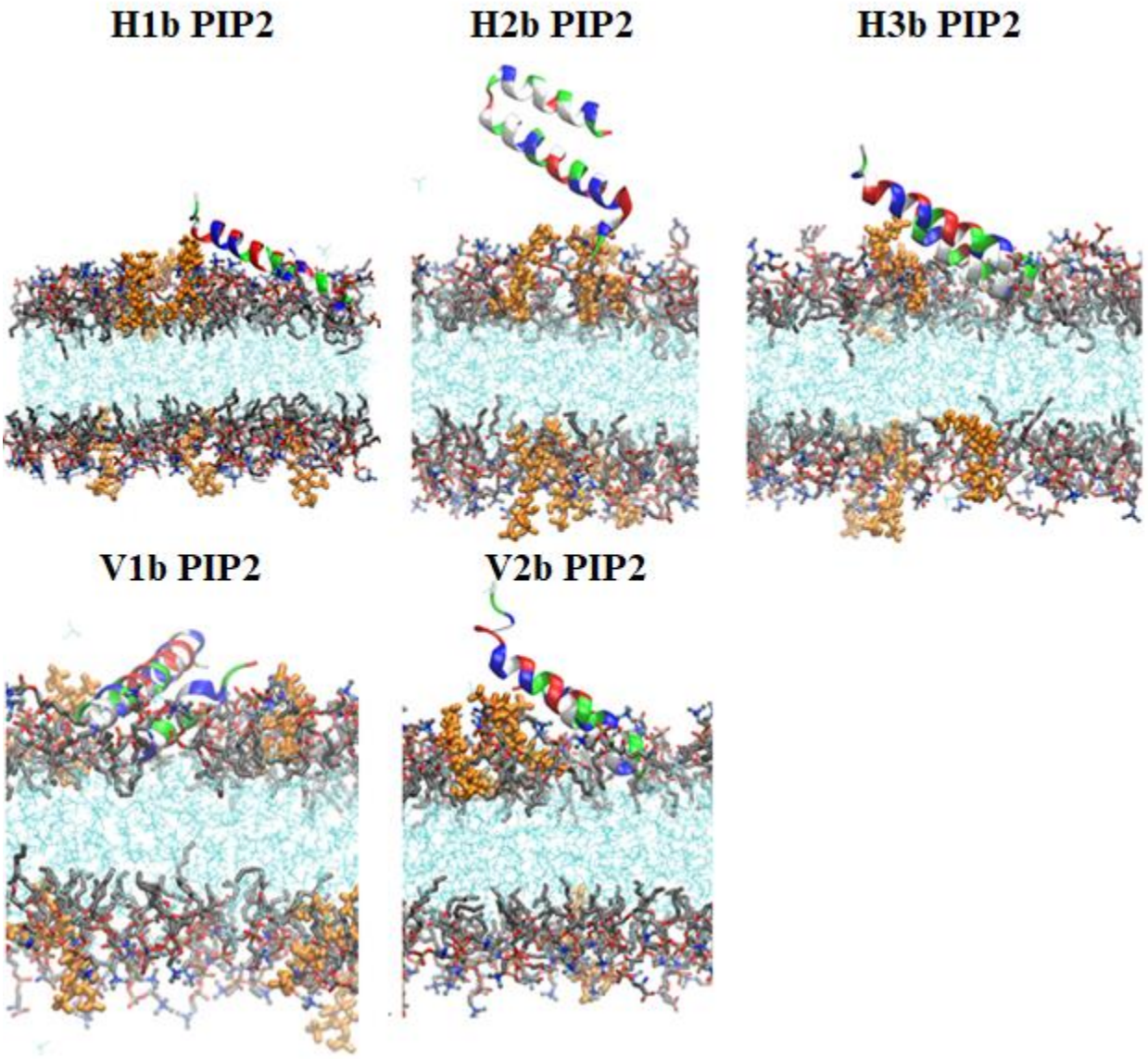


Figure A28. Image showing the final bound states of the PIP2 b trials. The results indicate that all of the simulations bound to the membrane except one of them, which was H2b.

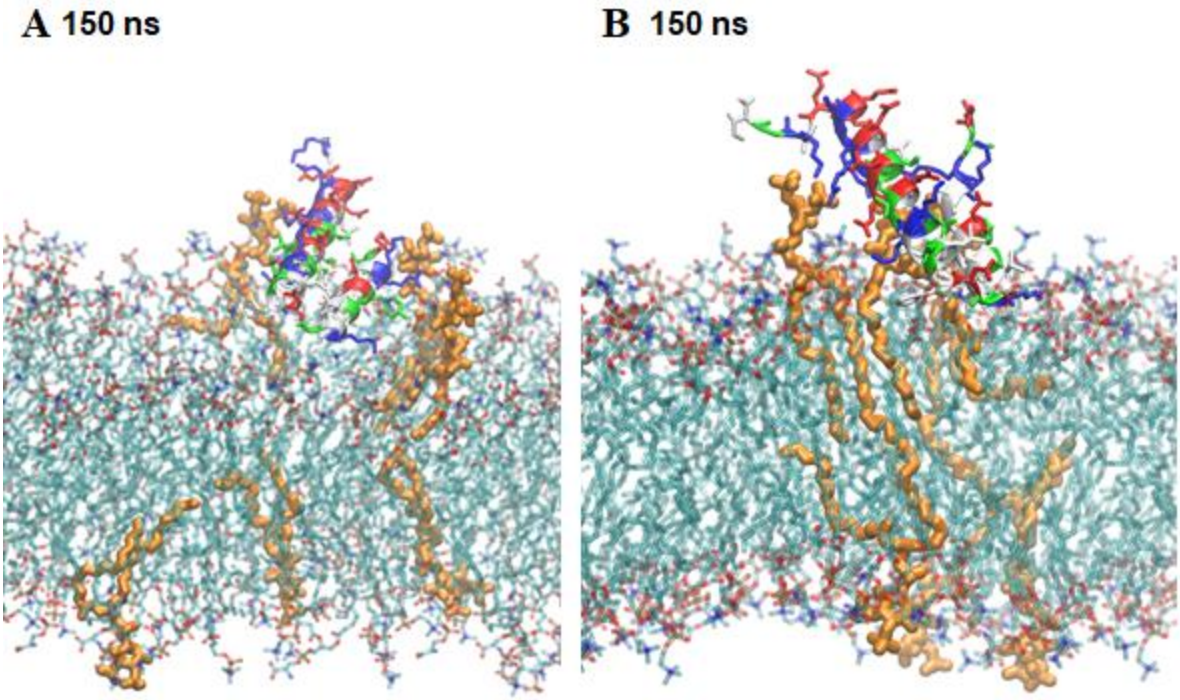


Figure A29. The **l**eft shows the placement of the peptide during the H1a simulation with PIP lipids on both sides of the peptide, and the right shows the peptide placement during simulation V2a with PIP lipids directly underneath the peptide. These are the images of the equilibrated full-length membranes with PIP2 peptide interactions.

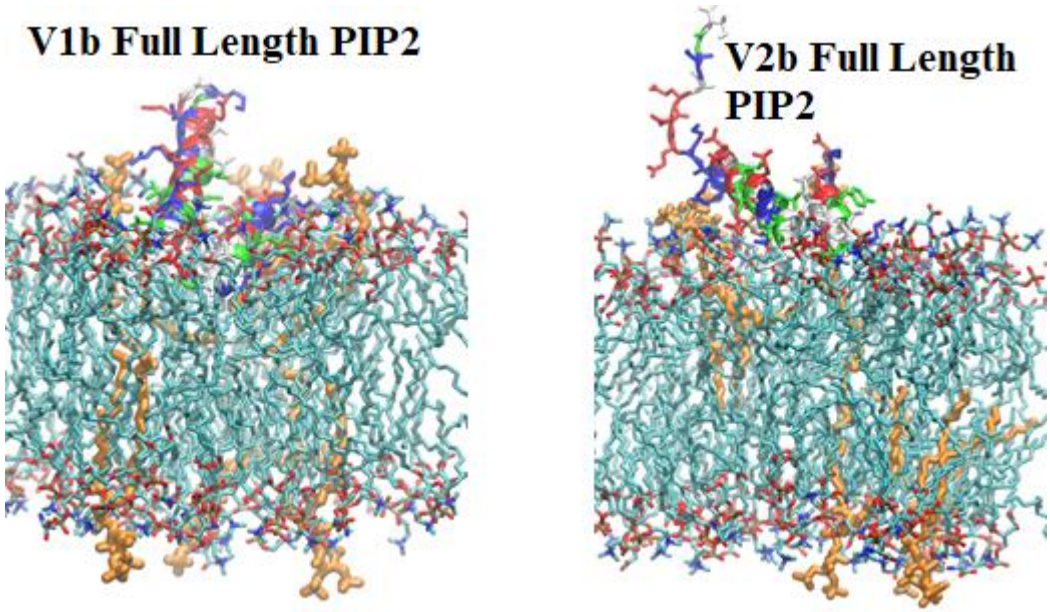


Figure A30. This image shows the two PIP2 simulations with the PIP2 lipids interacting on the sides of the membrane. The main point is that these simulations were chosen because they had PIP2 lipids interacting with a portion of the membrane.

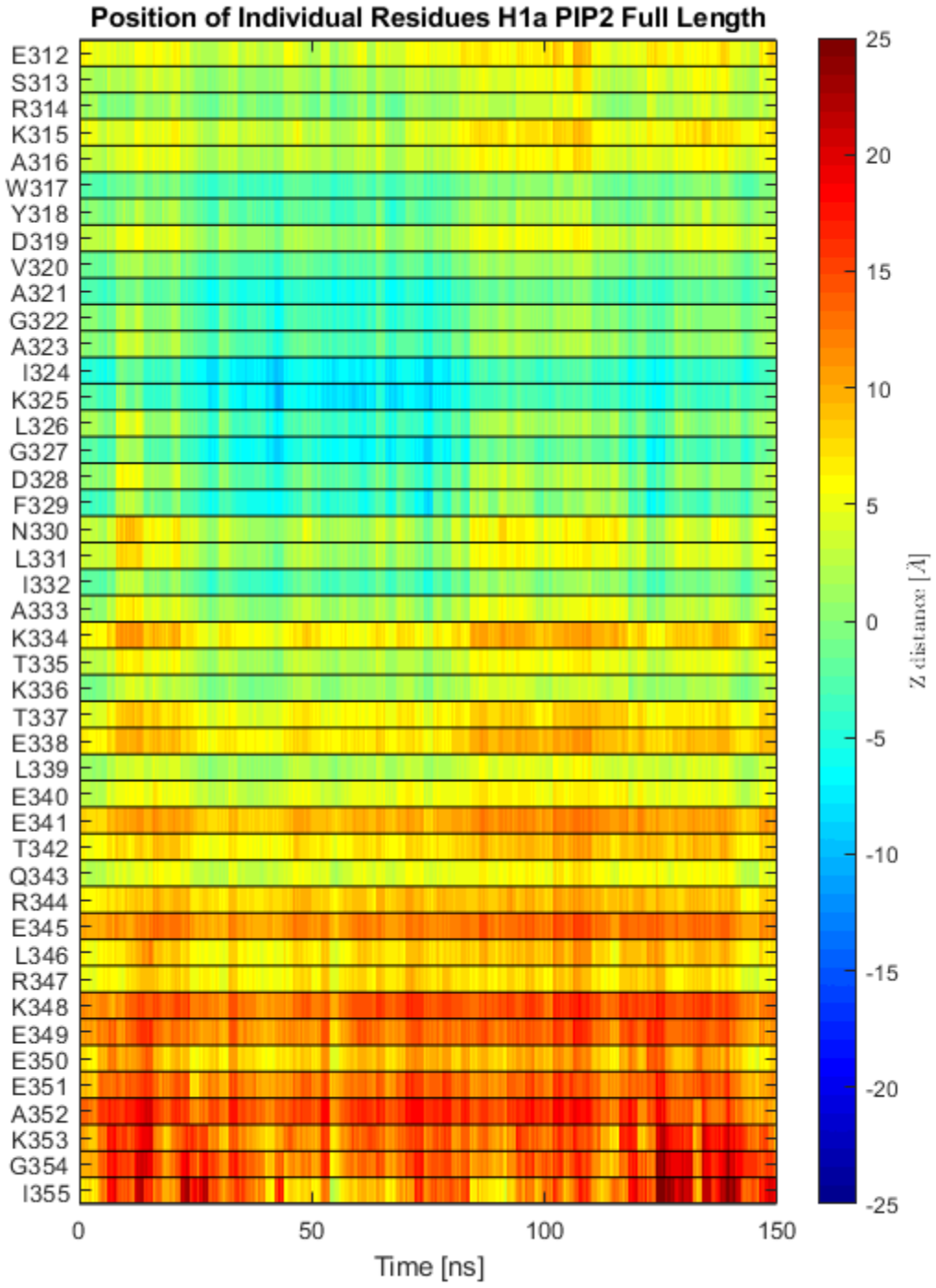


Figure A31. Image describing the average location of individual peptide residues relative to the average membrane phosphate. The peptide begins in the full-length membrane and is fully bound to the membrane. The deepest residues on the peptide are between V320 and F329.

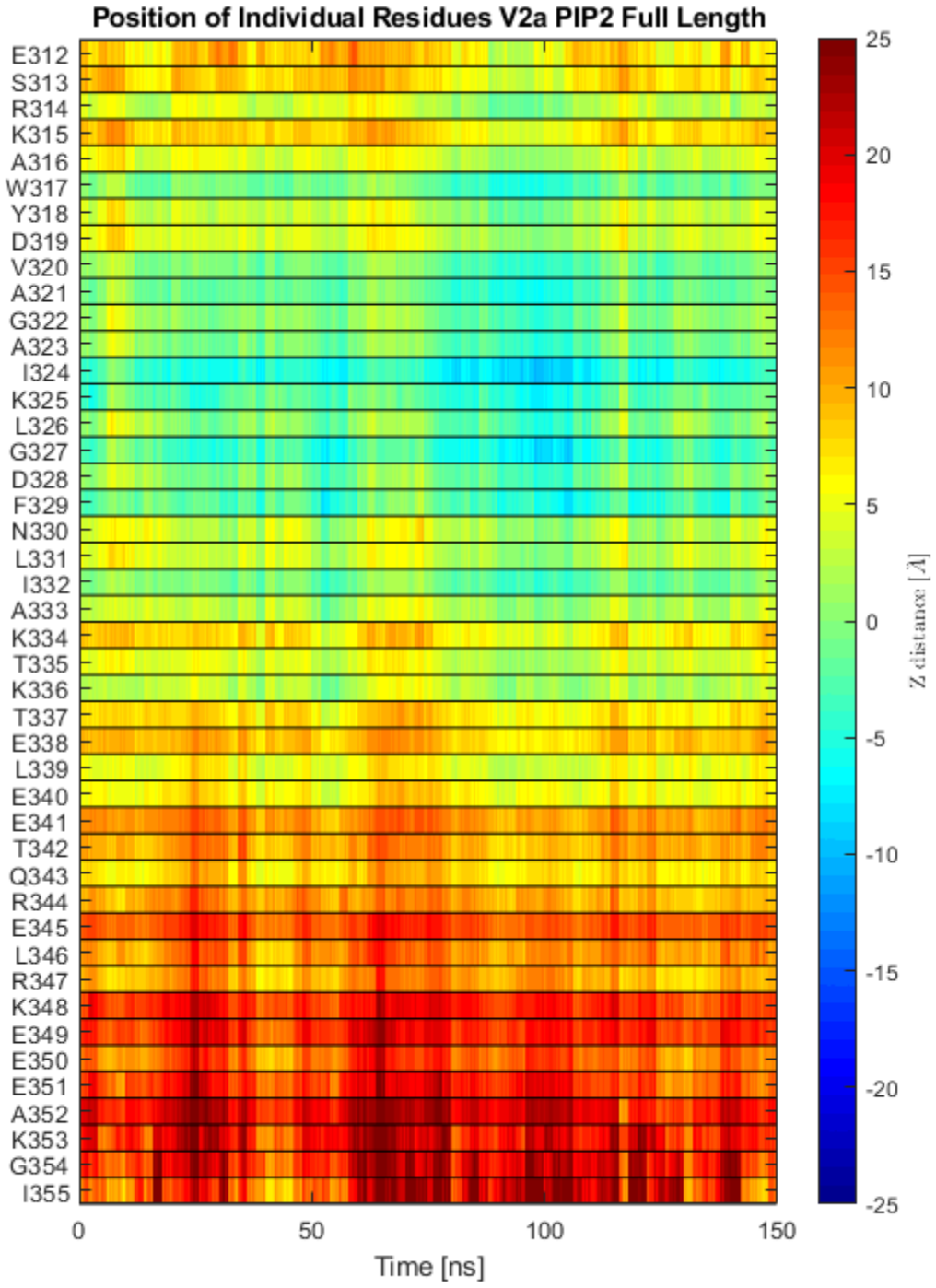


Figure A32. Image describing the average position of individual peptide residues location relative to the average membrane phosphate. This shows that the Full-length simulation begins with the peptide bound to the membrane and that the deepest residues on the peptide are between V320 and F329.

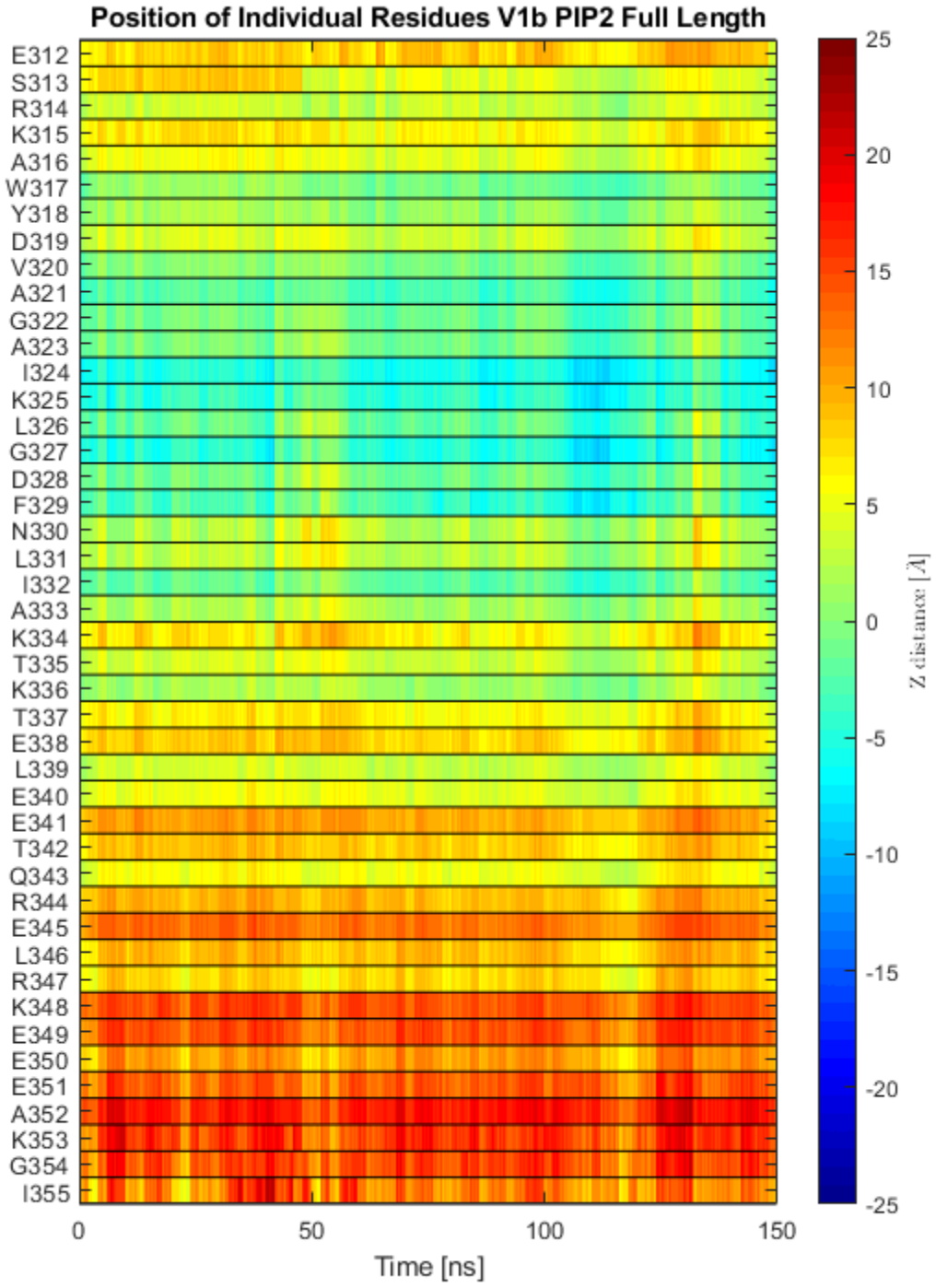


Figure A33. Image showing the placement of all the amino acids of the peptide over time, and how there is a range of residues that are deeply embedded into the membrane and some that fluctuate in a position such as K315 that is deeper in this simulation but higher in other simulations showing some variability in placement by different bound states.

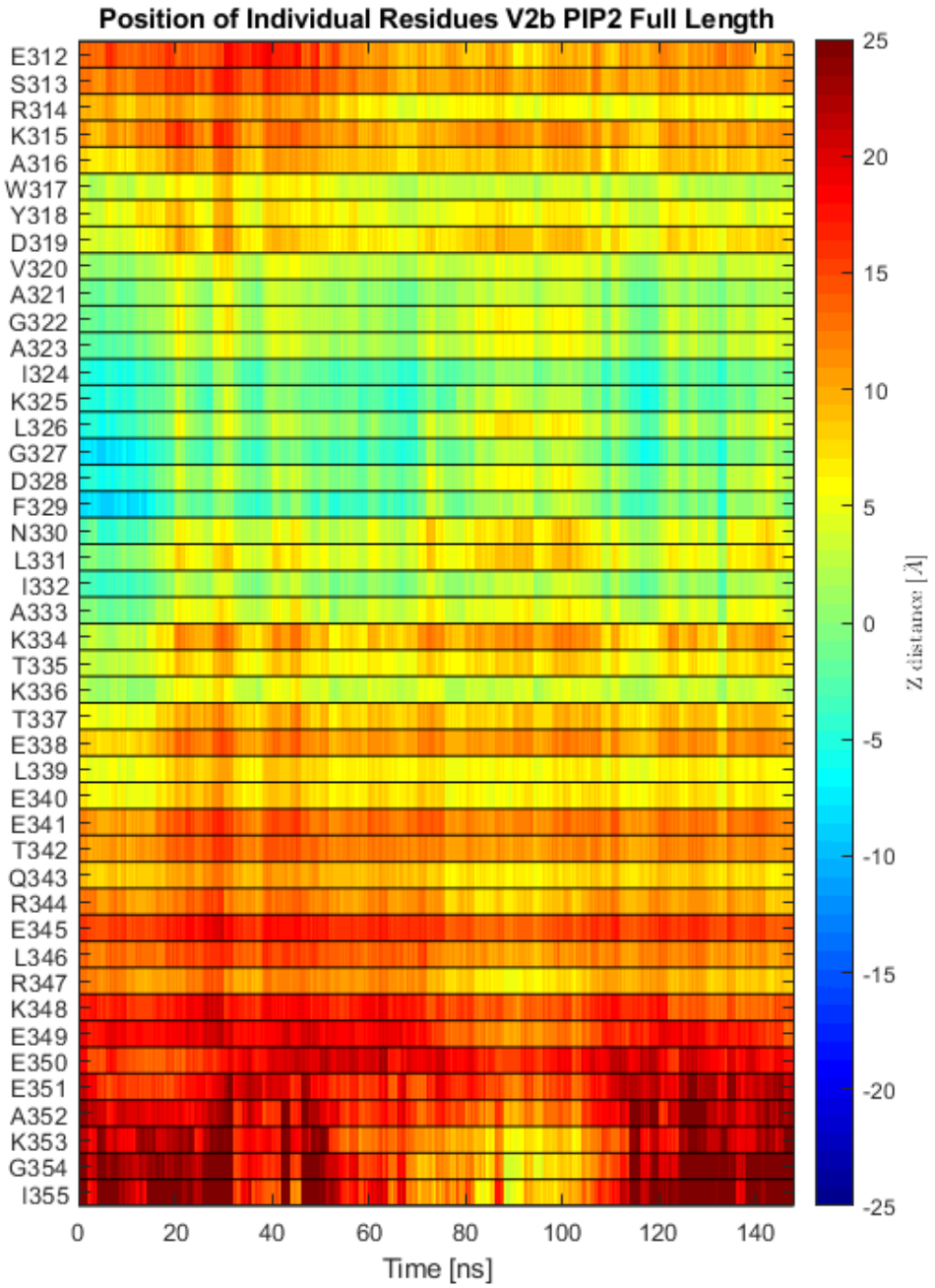


Figure A34. This image shows the placement of different amino acid residues over time. It shows that the amino acid residue K315 is higher than observed in the simulation V1b PIP2, which may account for the difference in interaction energy of this amino acid residue.

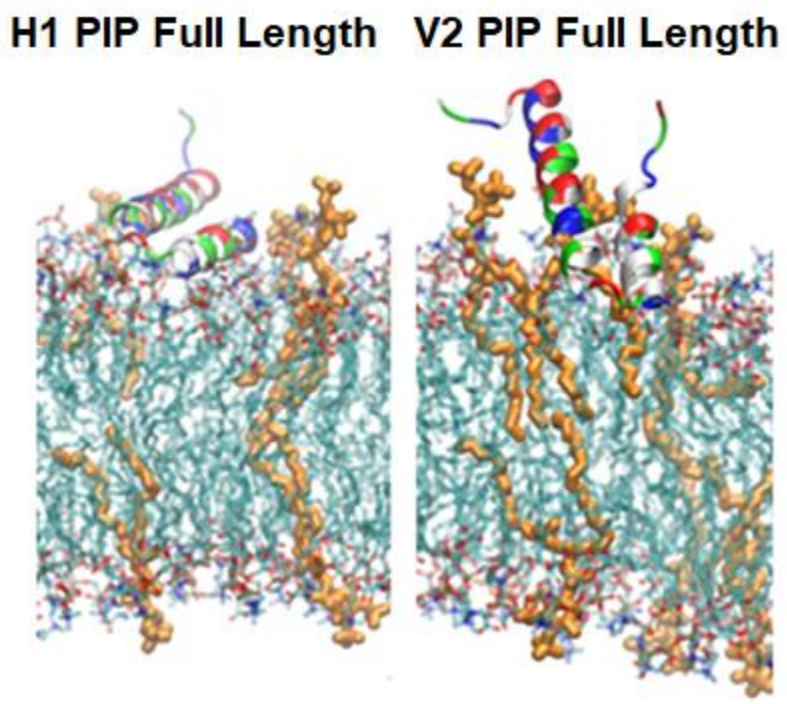


Figure A35. This image shows the final bound states of all-atom simulations, and the main point is that the PIP2 lipids interact with the peptide in the upper region of the peptide.

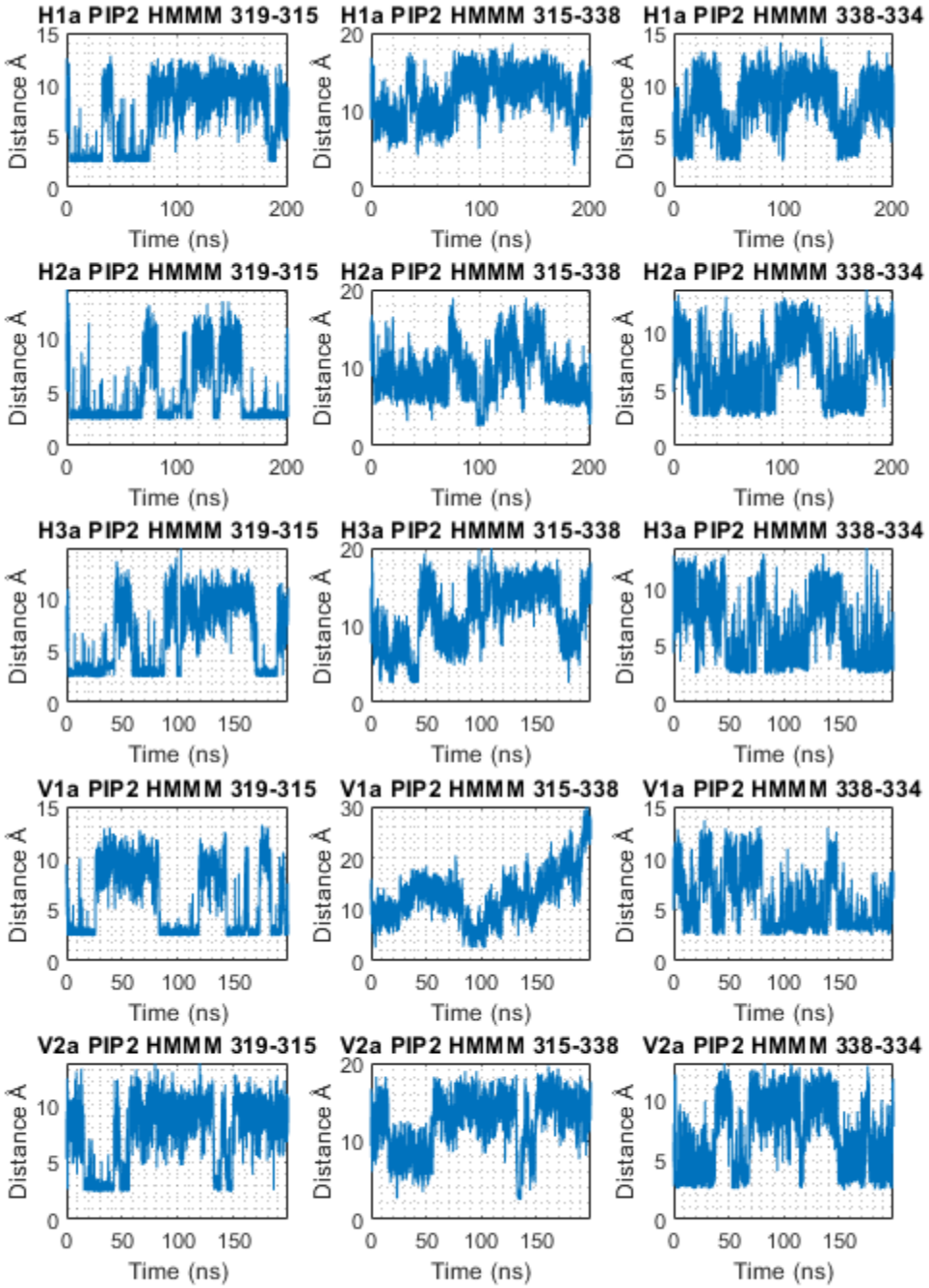


Figure A36. This figure shows the distances between the critical amino acid residues on the top of the peptide. It shows no interaction between 315-338. However, there is undoubtedly some interaction below 4 angstroms for the other bridges. Additionally, V1a split apart between the α_{6}-α_{7} helices.

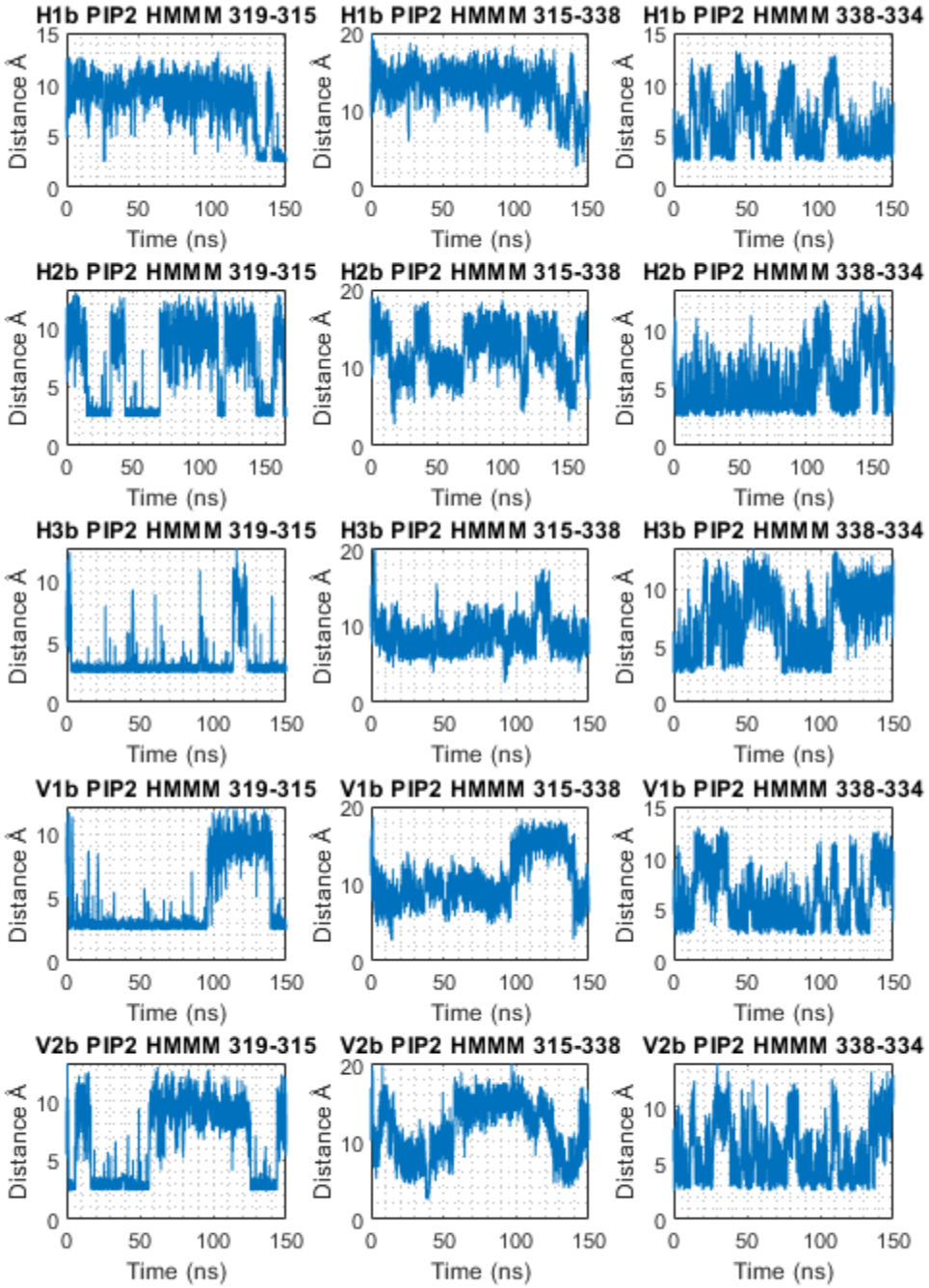


Figure A37. Image of the distance between the key residues that stabilize the membrane binding region from HMMM simulation. These are important because these long-range water-mediated interactions help to stabilize the $\alpha_6\alpha_7$ helices, and they are in the range to be important.

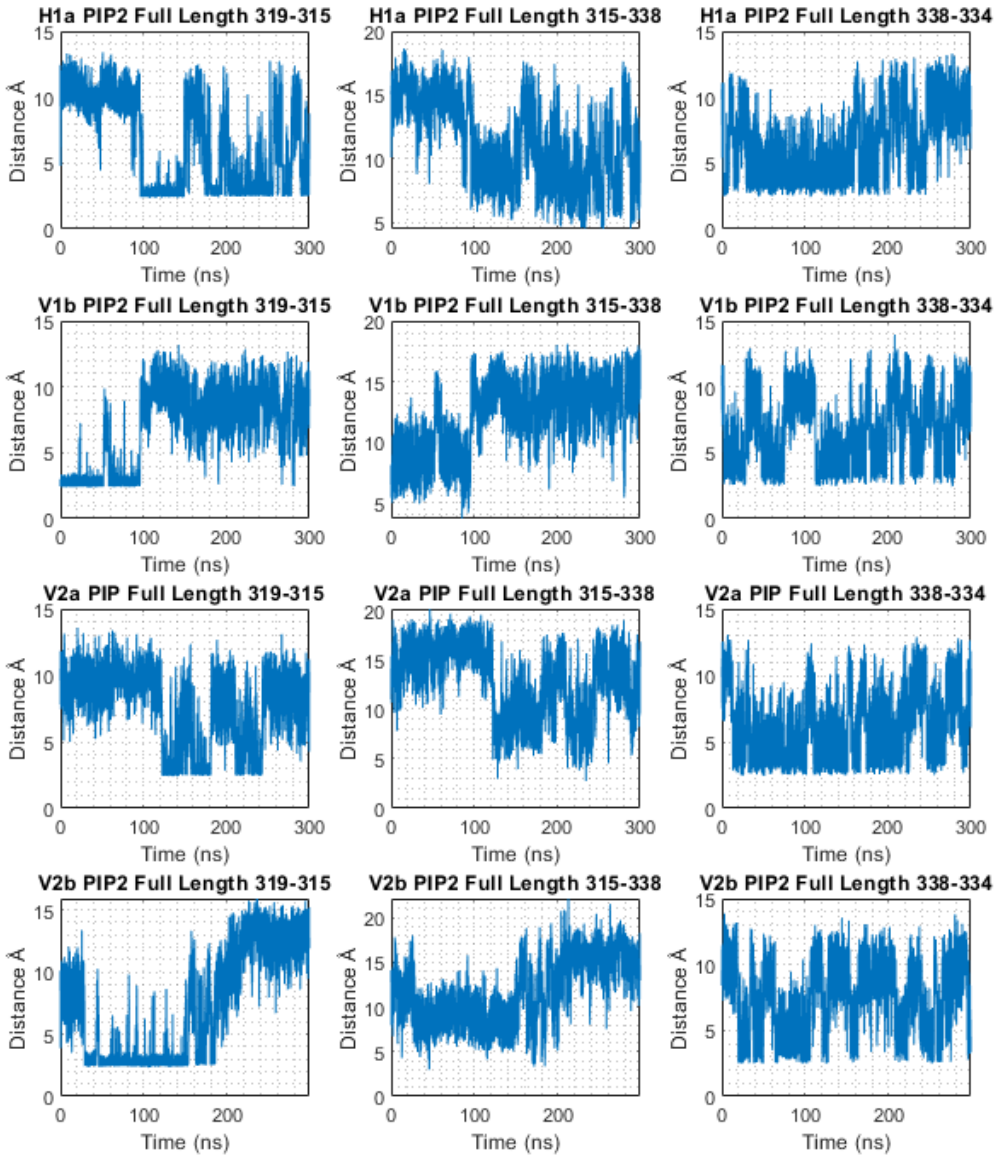


Figure A38. These full-length simulations show similar results with no interaction between the 315-338 bridge. The main point was to quantify how often the different residues were in contact, and it is clear that the other bridges are frequently close enough for interaction except for 315-338.

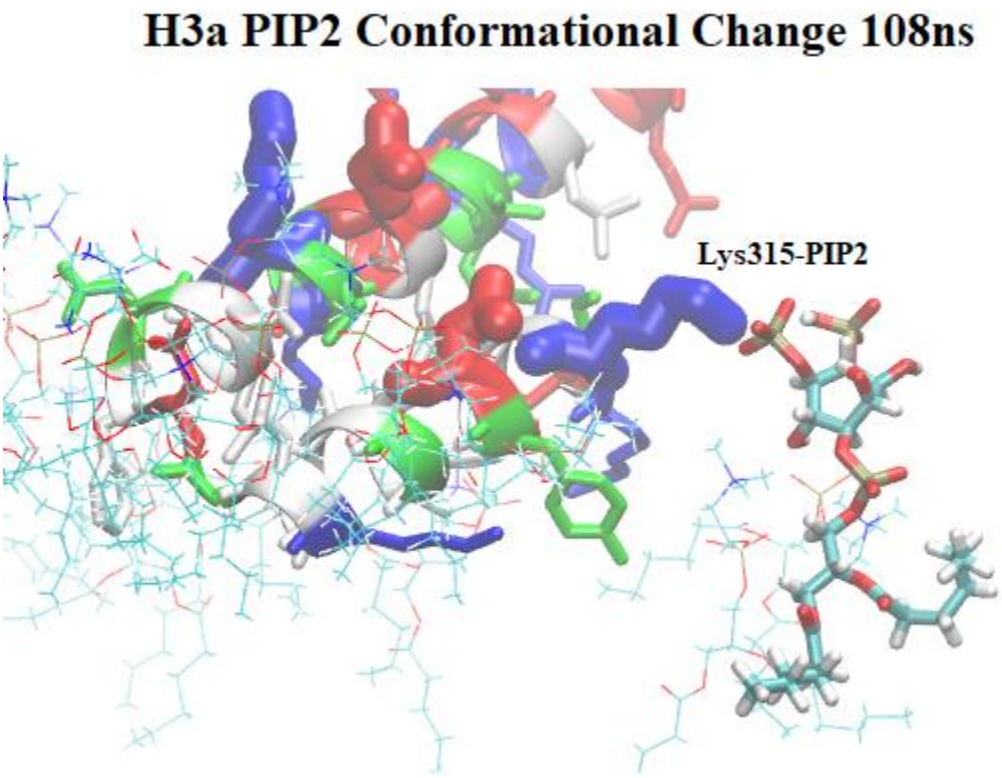


Figure A39. This is an image of the third conformation change of the side chain, showing that the PIP2 lipids have enough charge to displace the Lys315 from interacting with D319 and showing a preference for the PIP2 lipid. A similar result is expected for the other enlarged lysine residue K334.

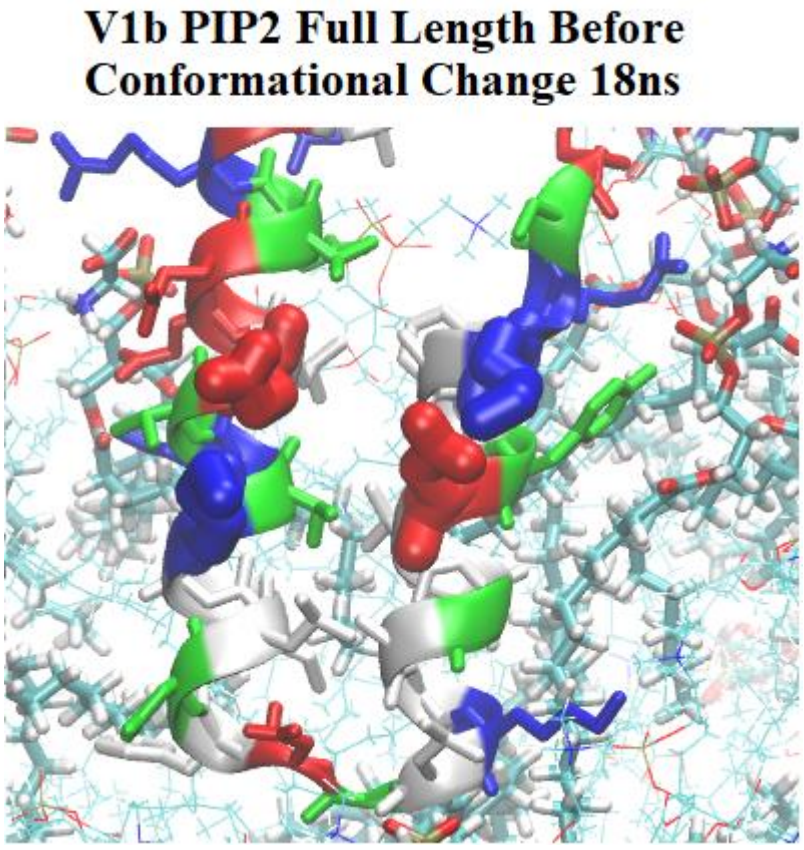


Figure A40. This is another image of the peptide before the conformational change, and the main point is that K315-D319 is close enough to interact and K334-E338 are close enough to interact even if the bridge between K315-E338 is not close enough to interact.

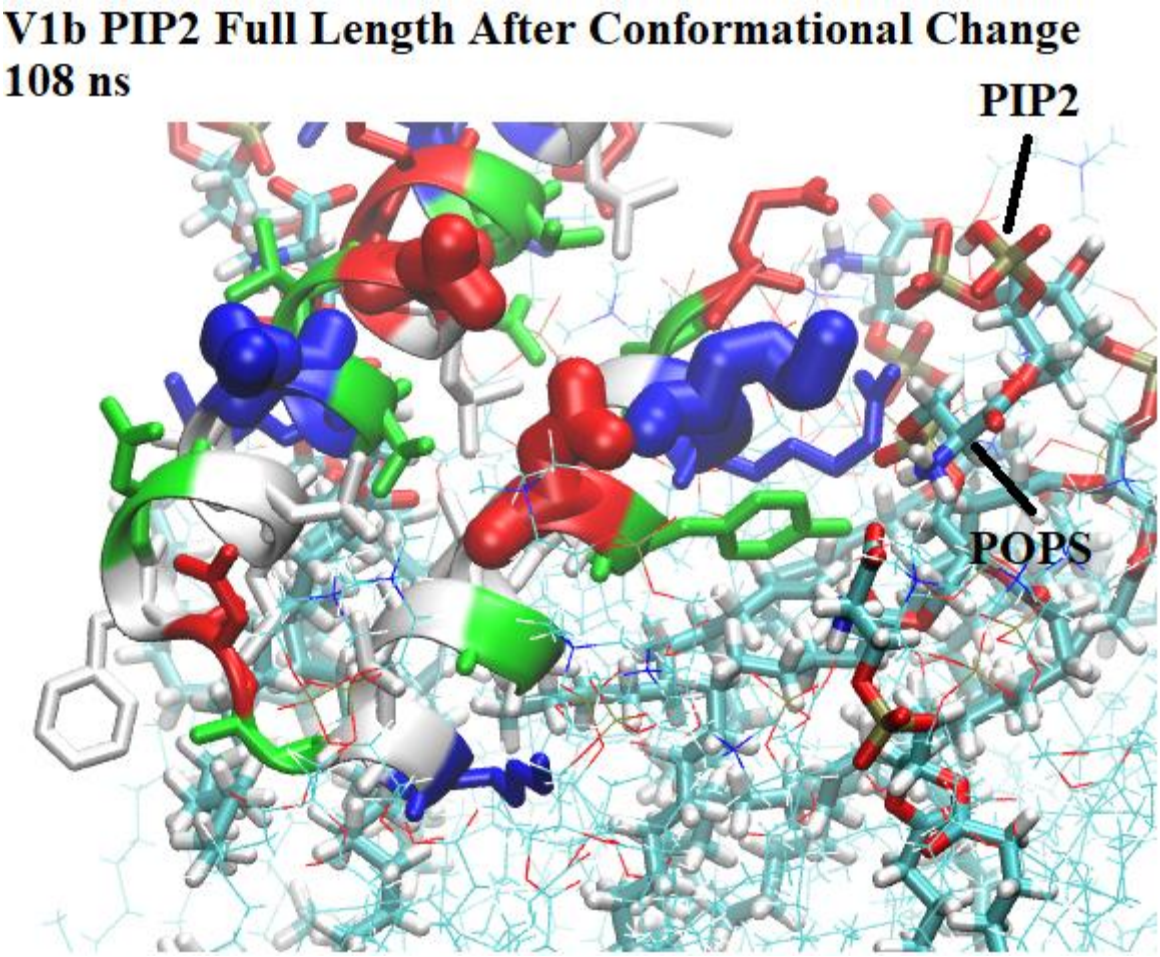


Figure A41. Here, in this case, the peptide interacts with the POPS and the PIP2 simultaneously, showing synergy between the two lipids.

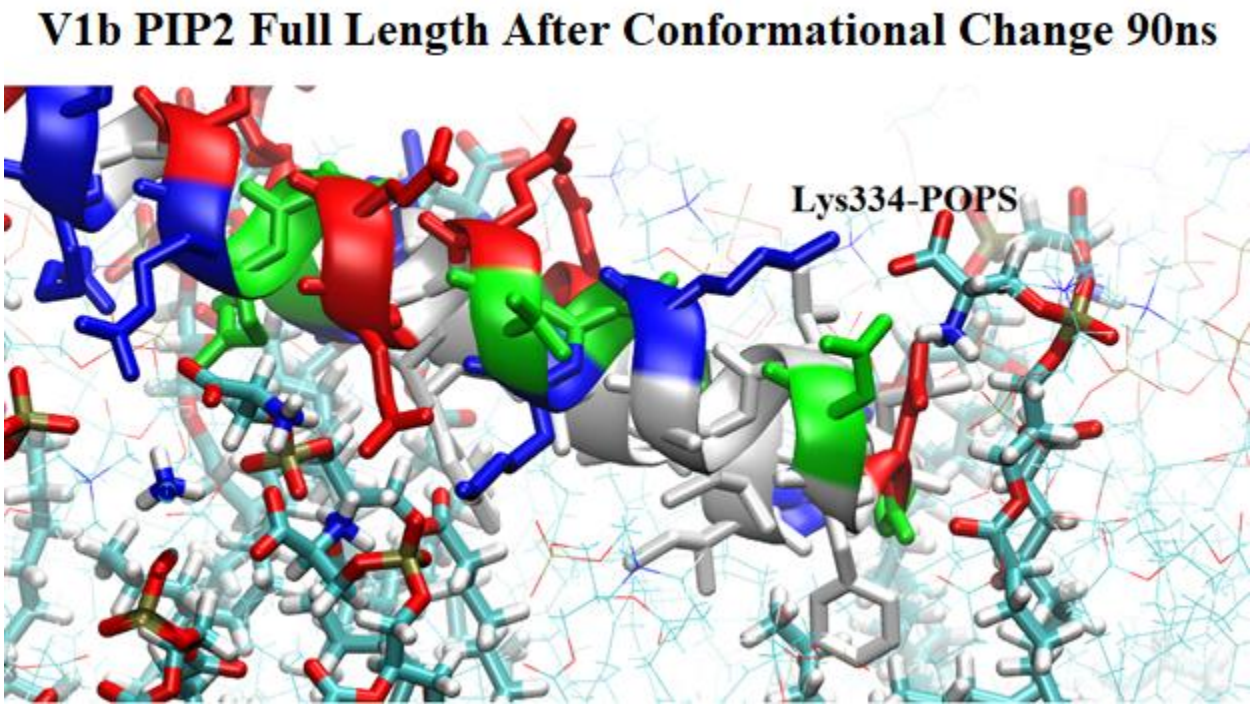


Figure A42. This image of the peptide shows the conformational change with Lys334 that was much less frequently observed. However, it is possible for the conformational change to take place without disrupting the peptide structure, and this indicates that it may be stabilized with POPS lipids.

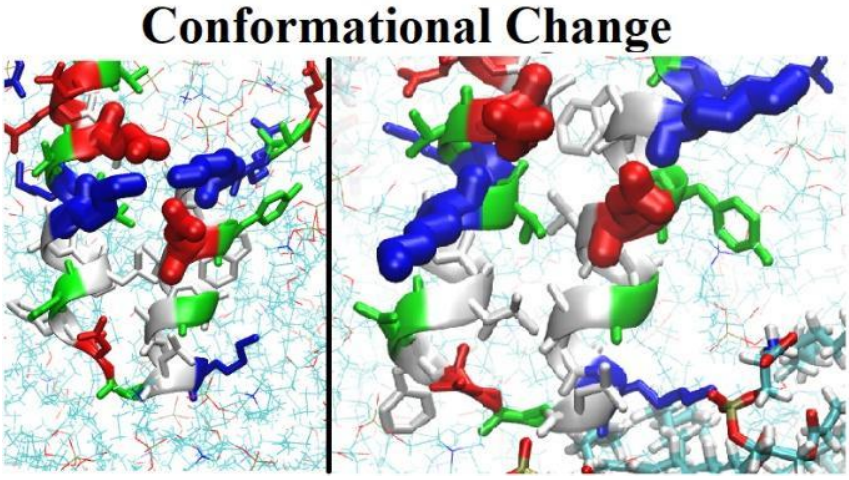


Figure A43. The image on the left is taken from simulation H1a full length at 22ns, showing the long-lived and spontaneously formed peptide-peptide salt bridges. On the right is H1a PIP2 full length, offering both amino acids in the conformational change, although a PIP2 or POPS lipid does not stabilize K344, so it quickly moves back to form peptide-peptide salt bridges.

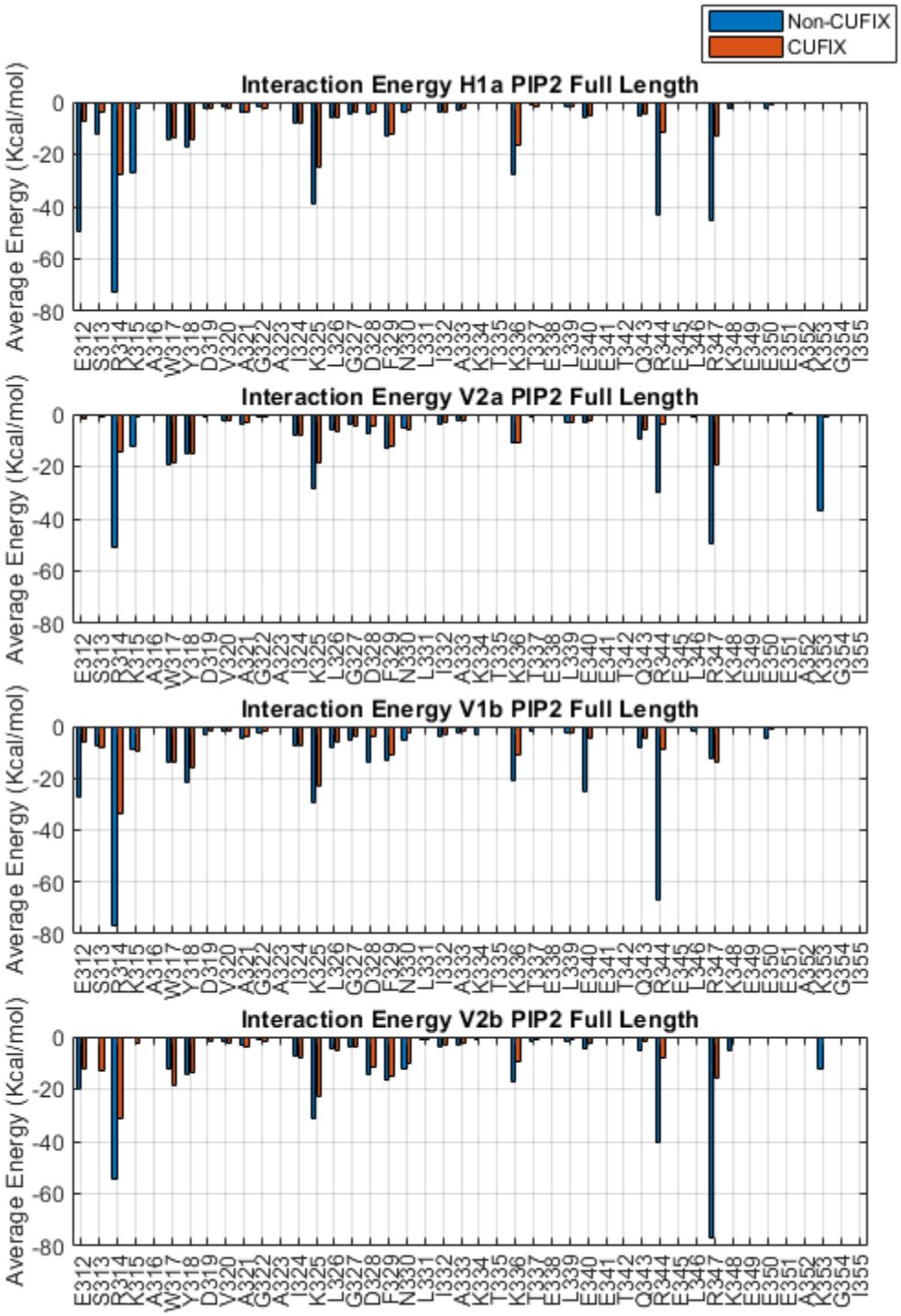


Figure A44. The per residue interaction energy of full-length simulations H1PIP and VFPIP, and the higher energy residues are on the two sides such as E312, R314, R344, and R347, and V2PIP has lower energy than H1PIP because of the PIP2 lipids being directly underneath the peptide pushing the peptide somewhat away from the membrane.

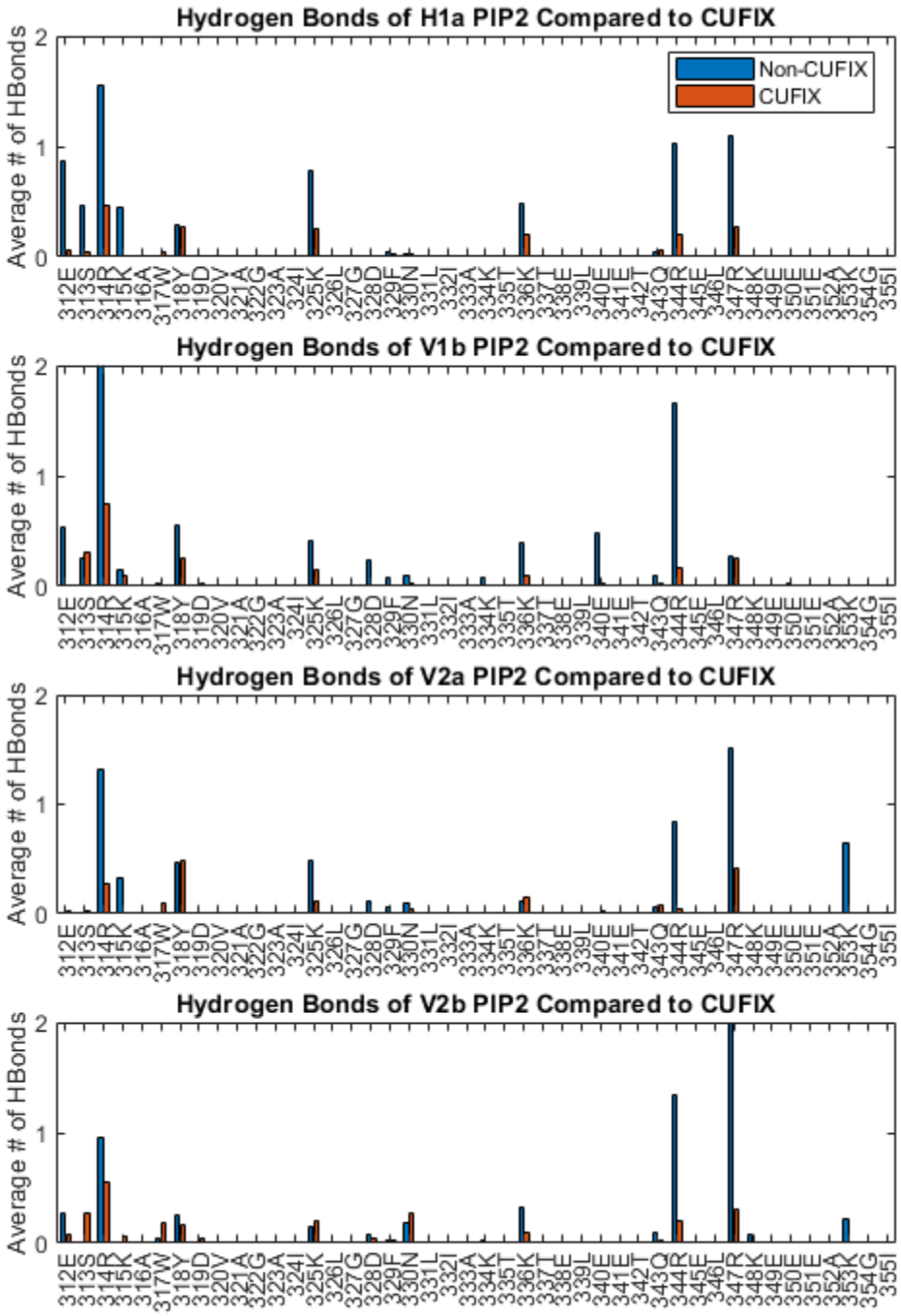


Figure A45. Shows the number of hydrogen bonds between the membrane and the peptide. It can be seen that the NBFIX parameters do drastically reduce the values of the hydrogen bonding for many residues, and there are slightly more hydrogen bonds expected in H1 PIP compared to V2 PIP.

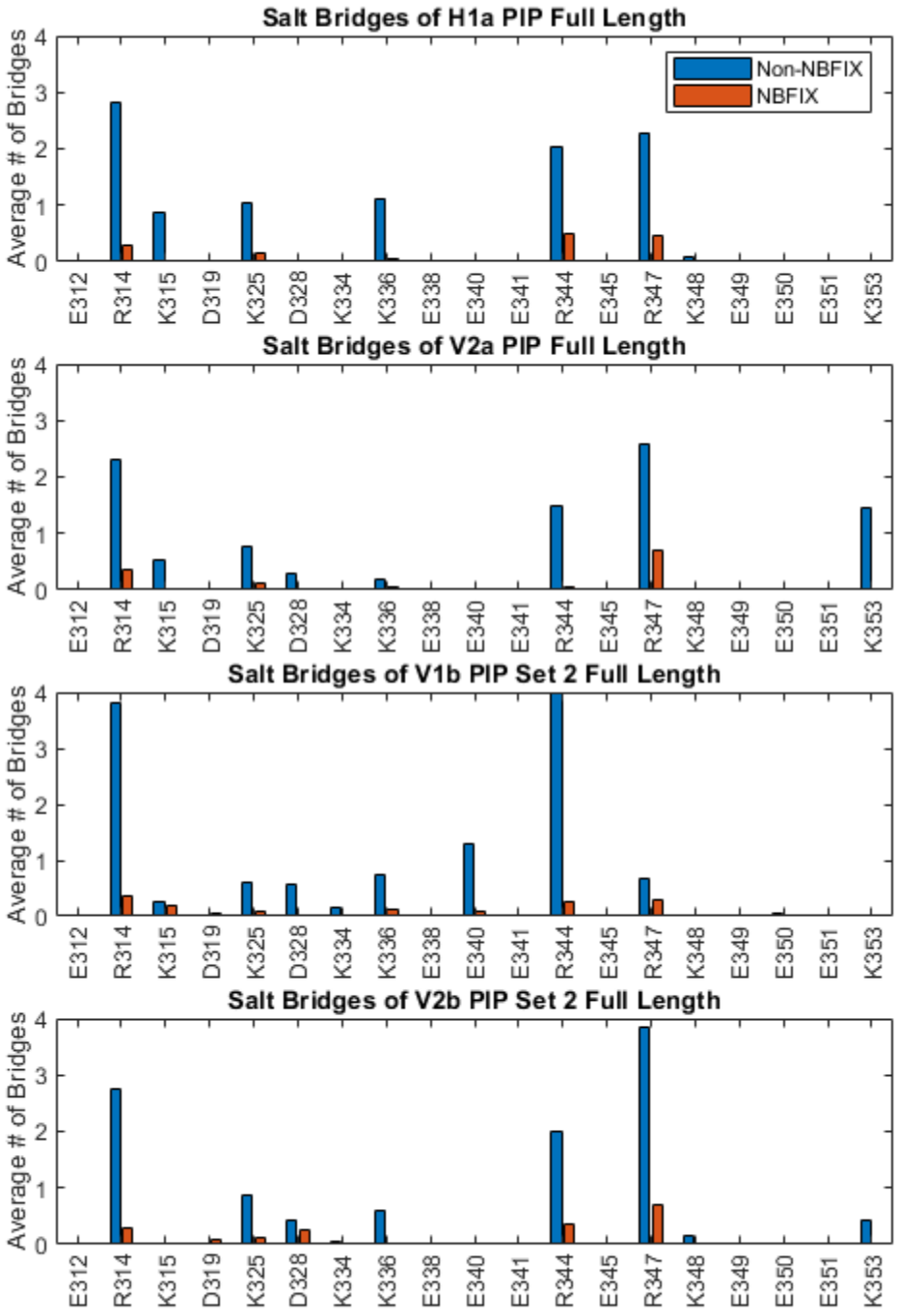


Figure A46. Displays the number of salt bridges per residue for full-length membranes, showing higher hydrogen bonds for residue R347 in both simulations. R347 is the residue that is accessible to the PIP2 lipid head groups.

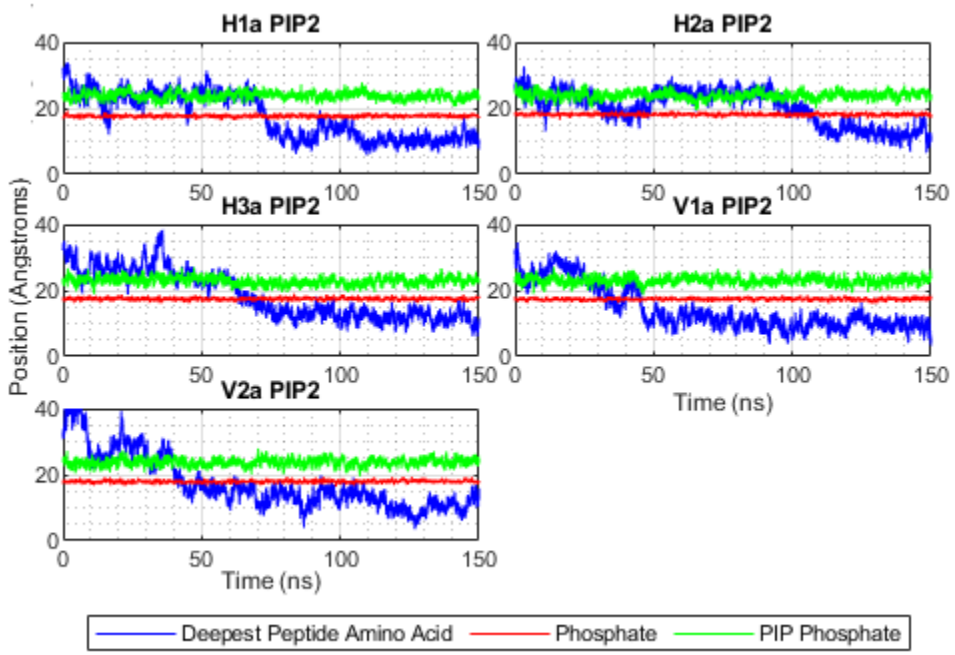


Figure A47. The image shows the position of the deepest residue for the PIP trials vs time, along with the position of the PIP phosphate and the main phosphate groups. This is similar to the other image of the position of the deepest residue vs. time, showing that the peptide starts away from the membrane and approaches the membrane, except that the PIP2 lipids head groups stick higher than the main membrane phosphates. Defined as the center of mass of the deepest residue.

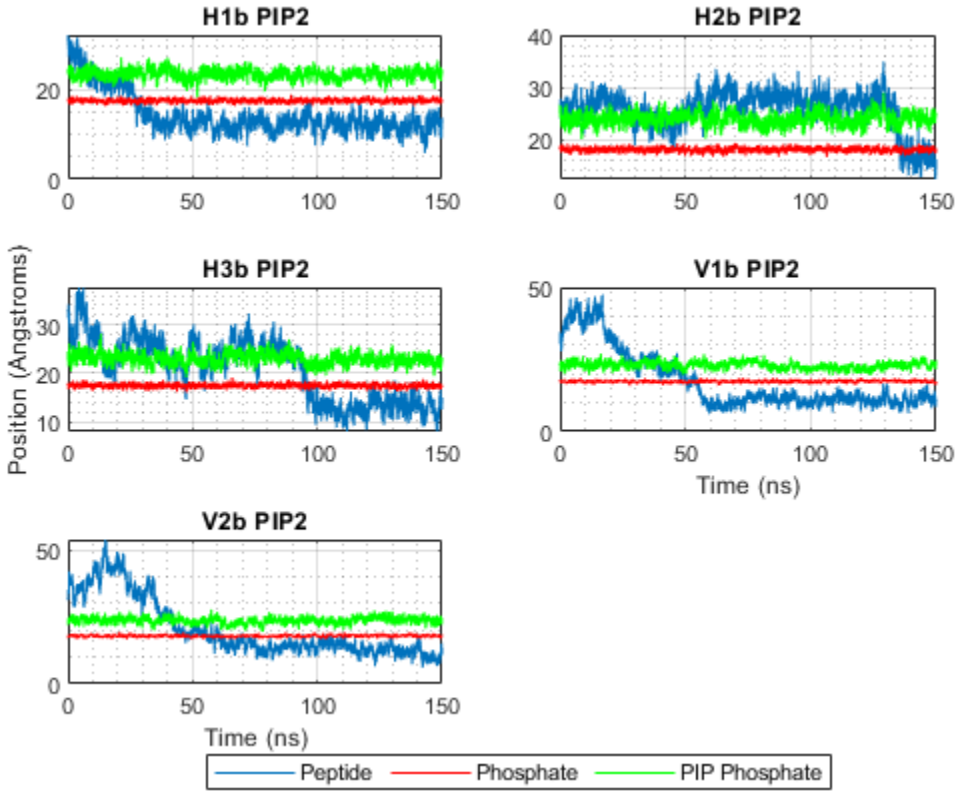


Figure A48. Image of the position of the COM of the deepest amino acid residue vs time.

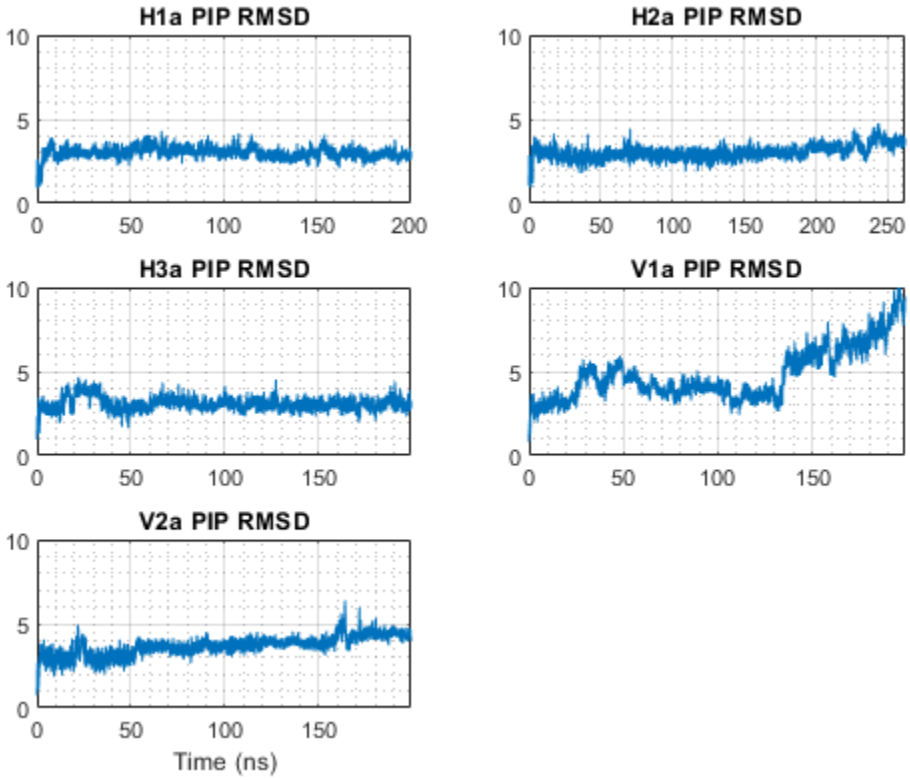


Figure A49. RMSD of the HMMM simulations, it is interesting that in simulation V1, the two helices were broken apart, causing the higher RMSD. All simulations show stability except for the V1 PIP trial, where the two helices split apart.

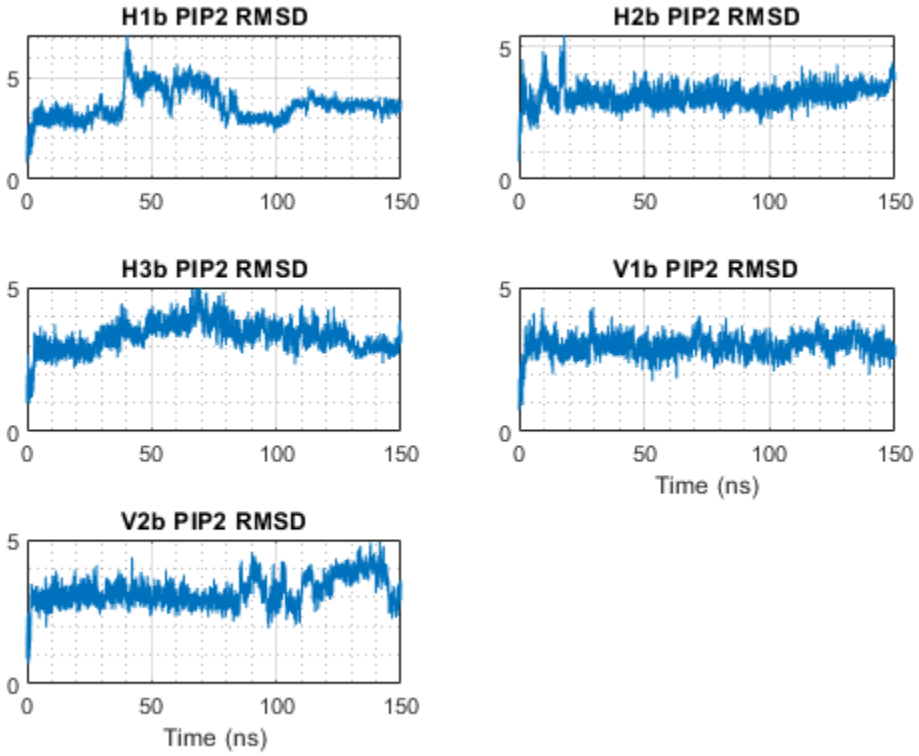


Figure A50. Image showing the RMSD of the HMMM simulations over time, the main point is that all of the simulations had stable RMSD values and that the helices never split apart.

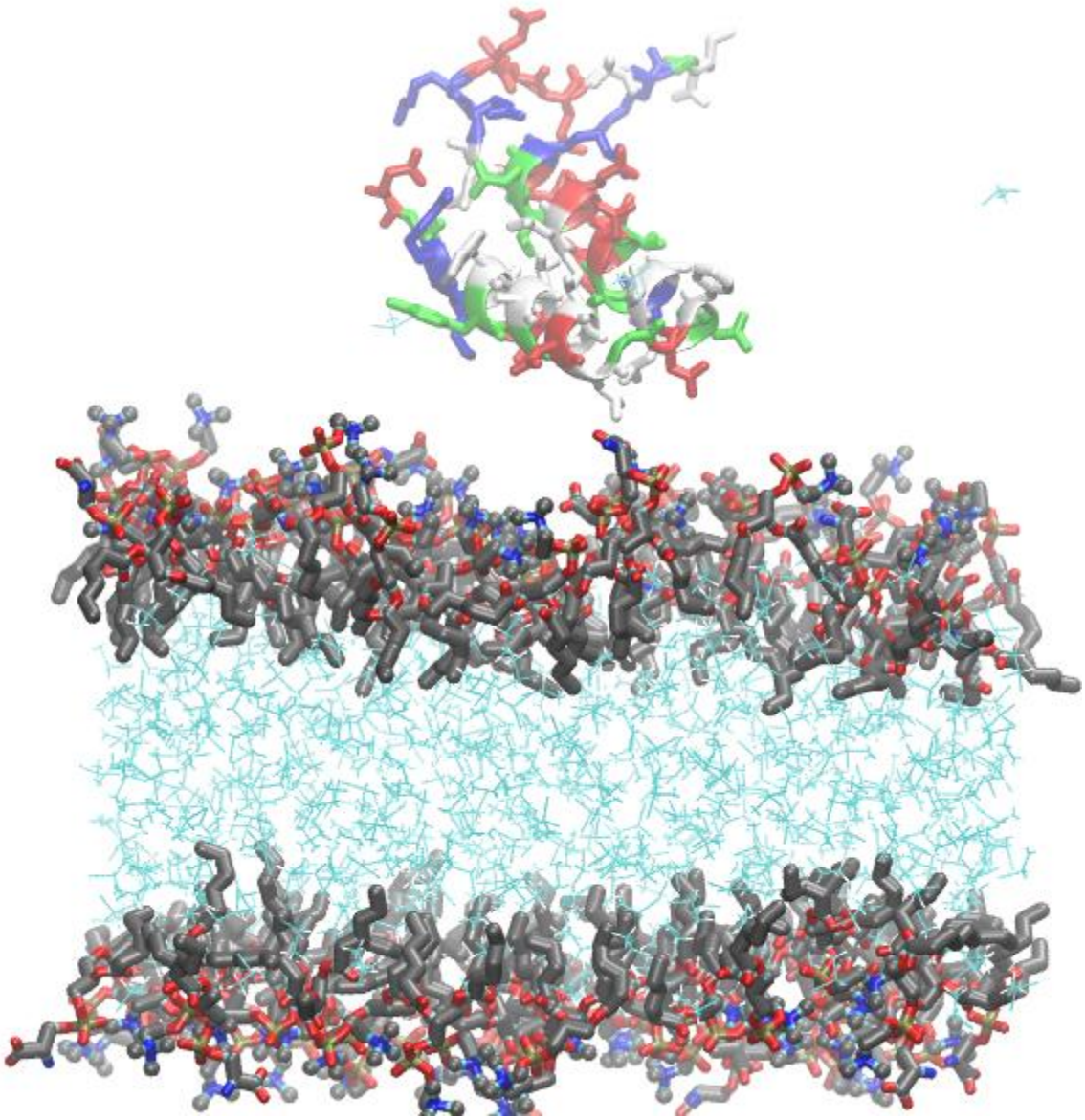


Figure A51. Image of the peptide unable to bind to the membrane without the two amino acids that were found to be essential. At the end of the simulation, 200 ns, there was no binding.

Appendix B

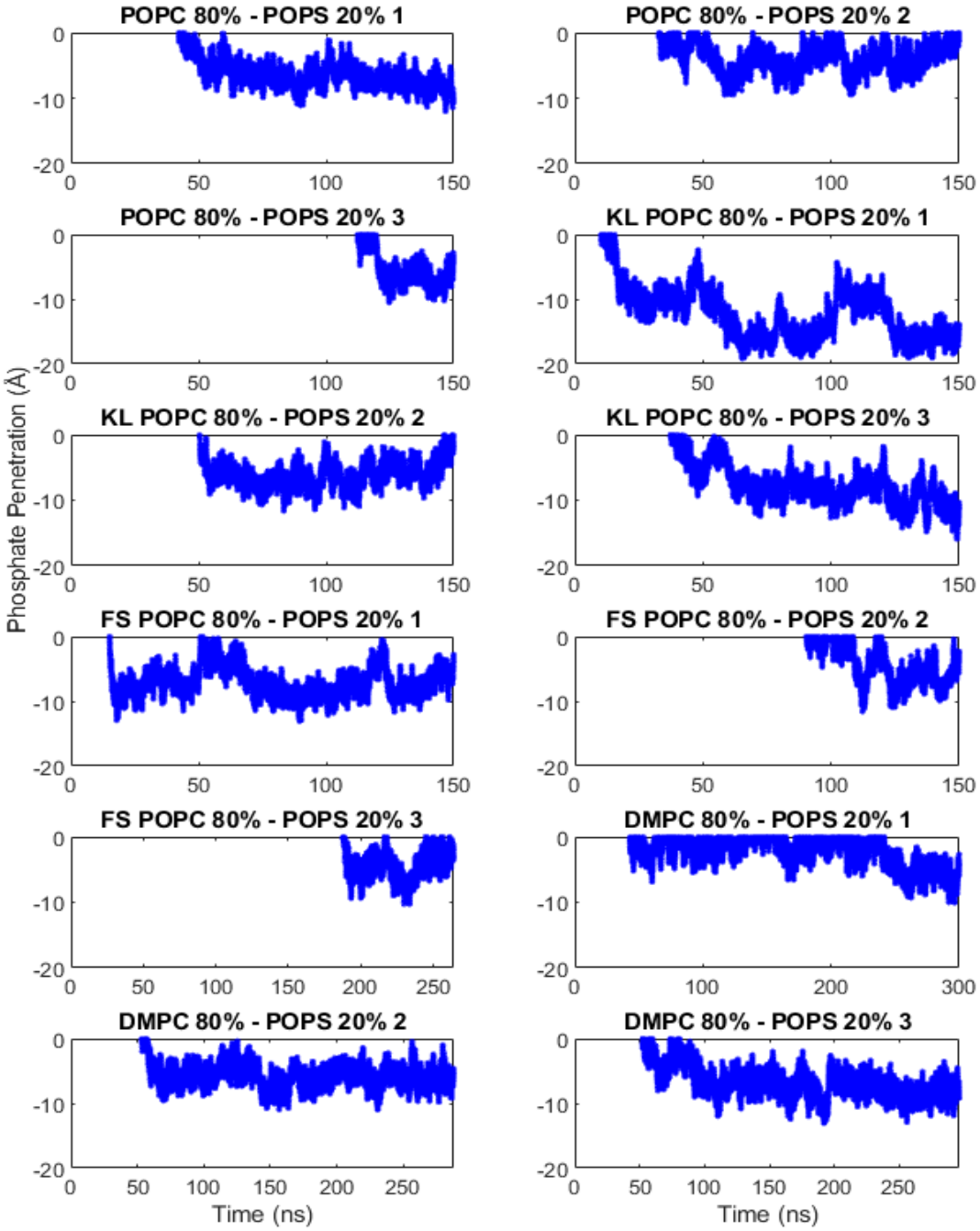


Figure B1) Shows the Osh4 HMMM Binding of some ALPS trials. The peptide comes from the aqueous phase onto the membrane, and any y values below 0 indicate the beginning of binding. This was all done using the average carbon alpha position.

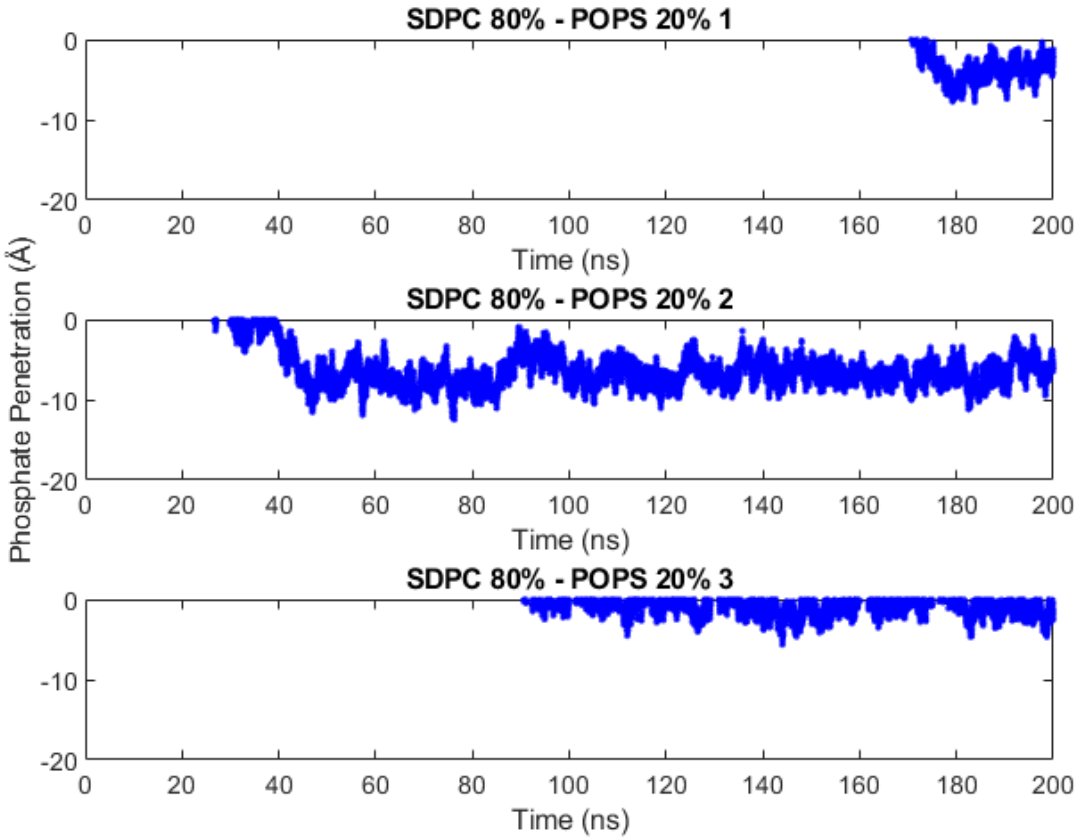


Figure B2) Shows more carbon alpha positions for the Osh4 HMMM binding sequences. There is variation between trials, but it looks like the lowest free energy simulations bind to a specific position in the membrane.

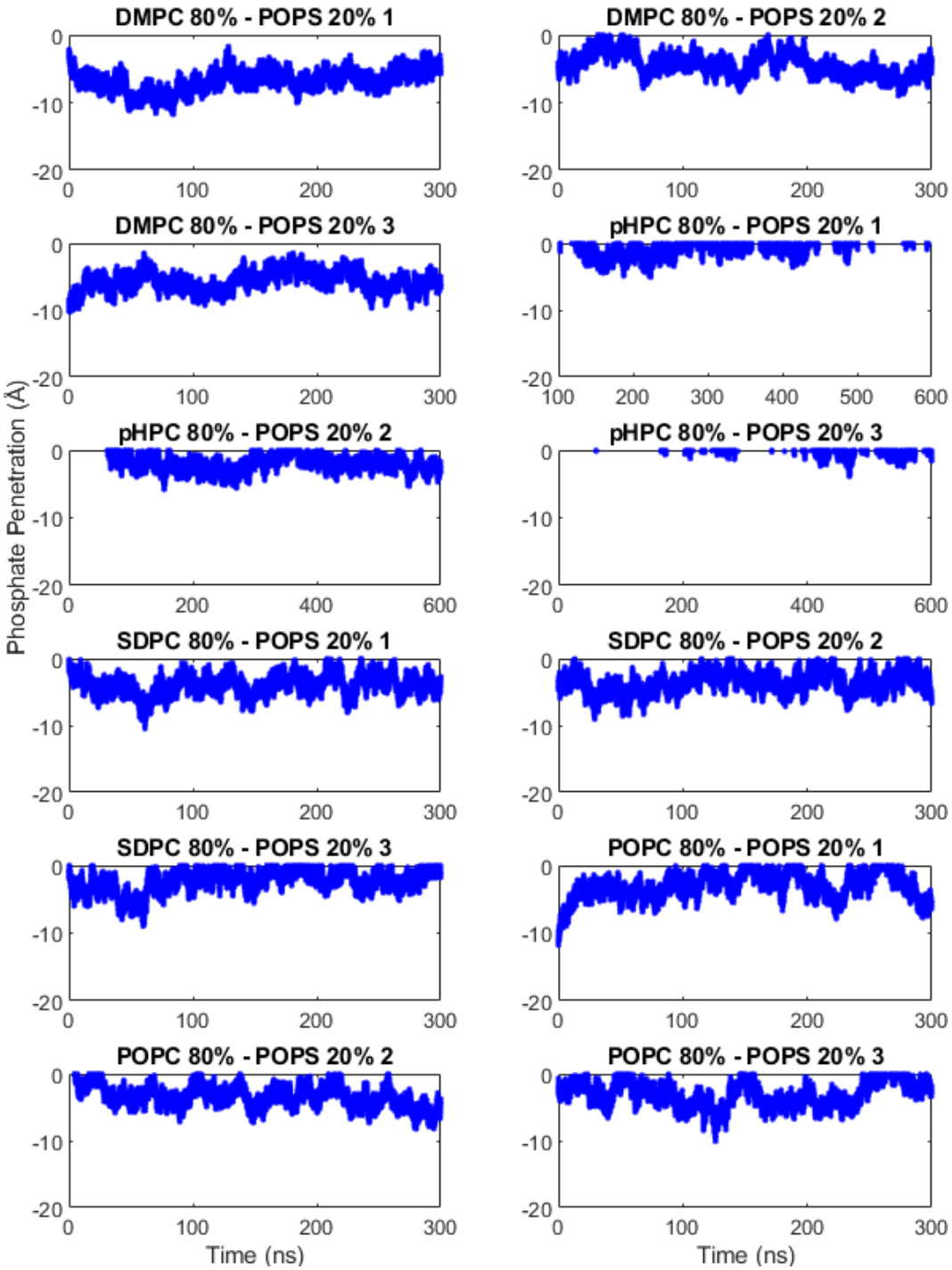


Figure B3) Carbon alpha of the full-length conversion. It should be noted that the pHPC simulations were started as full-length simulations from the beginning, so they are the only ones to bind. Many simulations were re-arranged over the 150-300 ns of simulation, although changes appeared to be less than 5 angstroms.

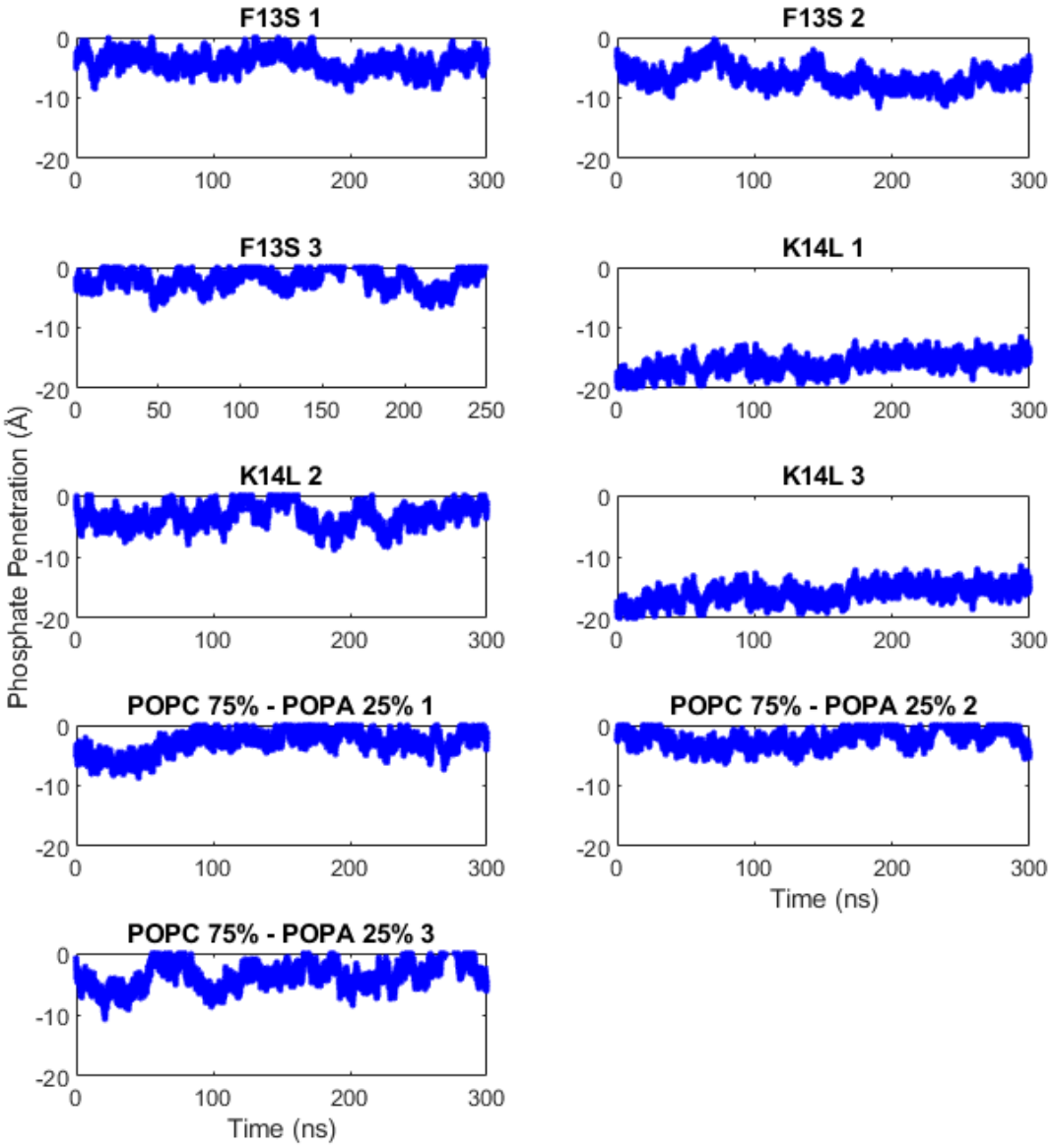


Figure B4) This figure shows the carbon alpha mutants and the POPC-POPA mixture. The KL is shown to adopt a transmembrane orientation.

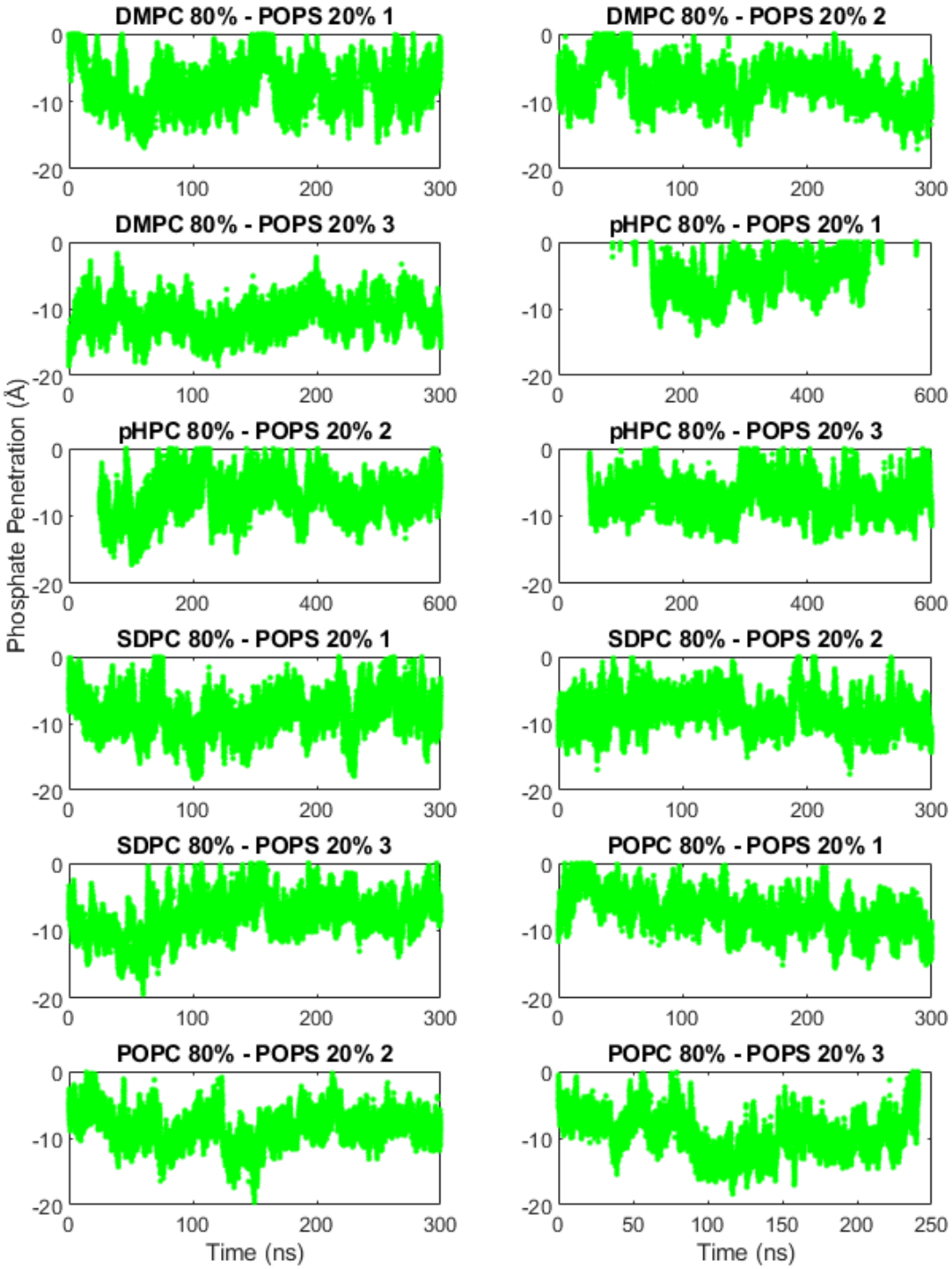


Figure B5) This measured the methionine residue 1 in the sequence and measured the HE1, HE2, and HE3 hydrogens that could be isotope labeled.

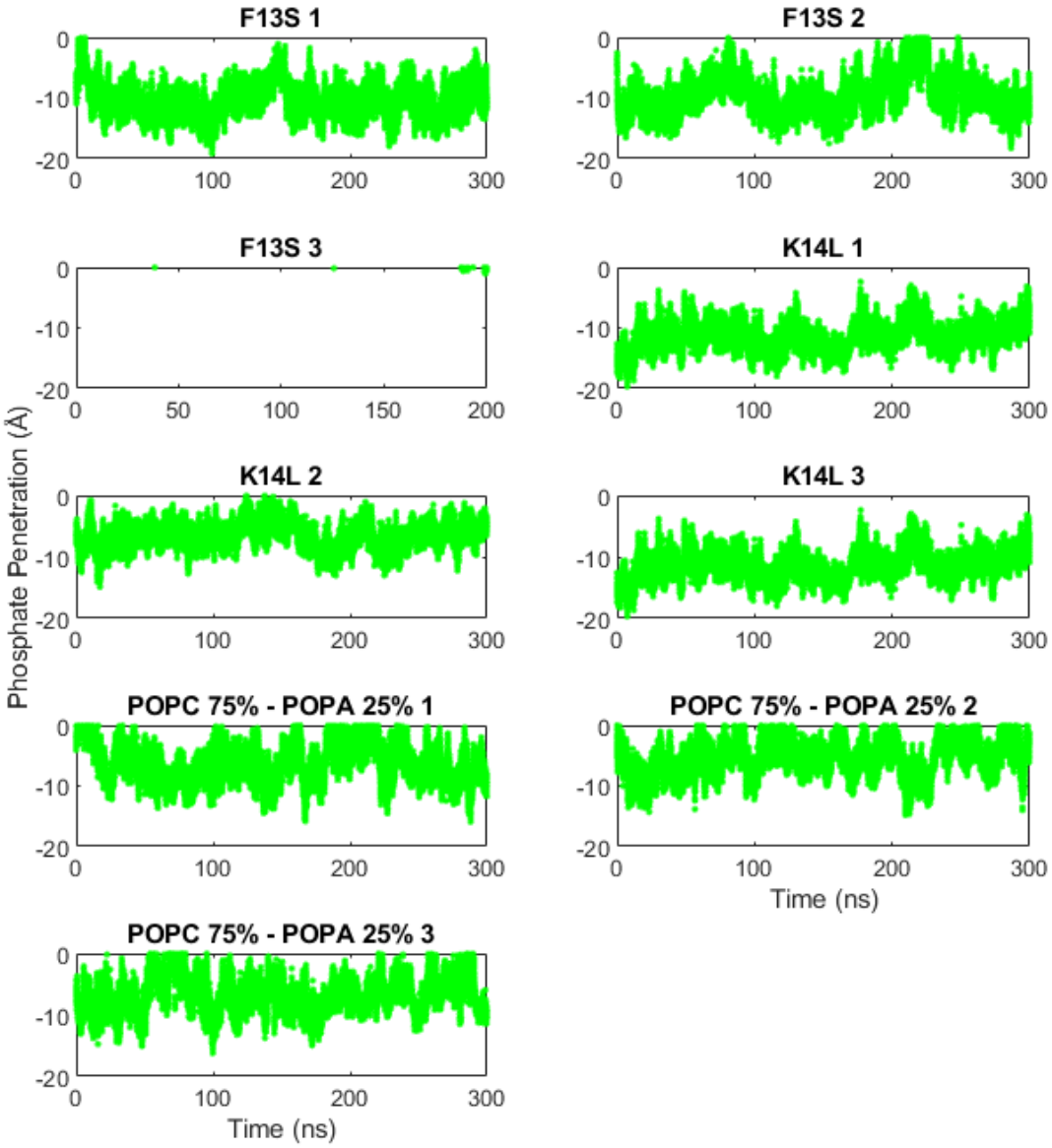


Figure B6) This shows the continued Met position data for the mutated Osh peptides and the POPC-POPA lipids. Measuring the HE1, HE2, and HE3 hydrogens that could be isotope labeled.

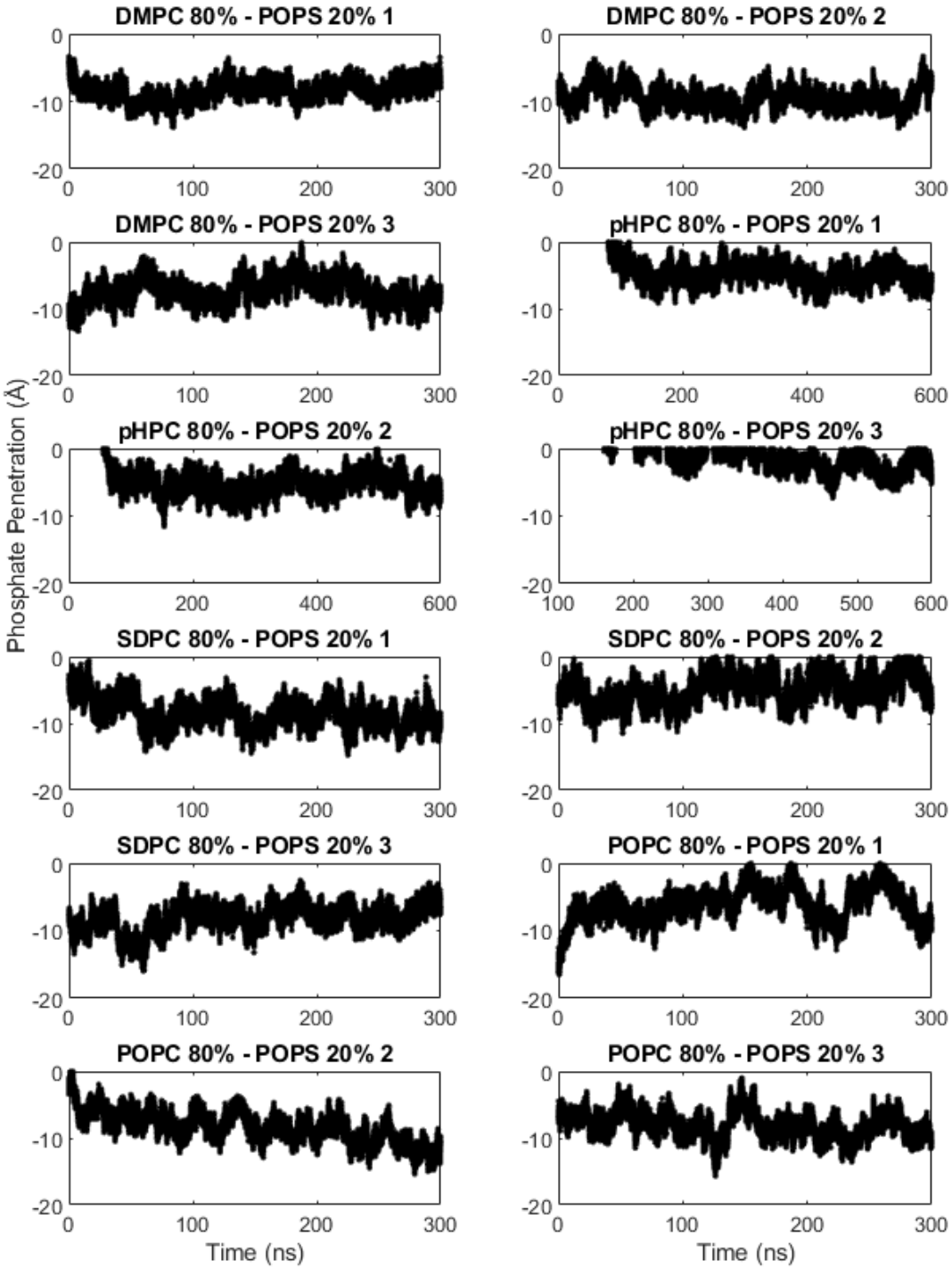


Figure B7) The PHE position and specifically the HA HB1 HB2 HD1 HE1 HZ HE2 HD2 atoms are selected in the calculation because they can be isotope labeled.

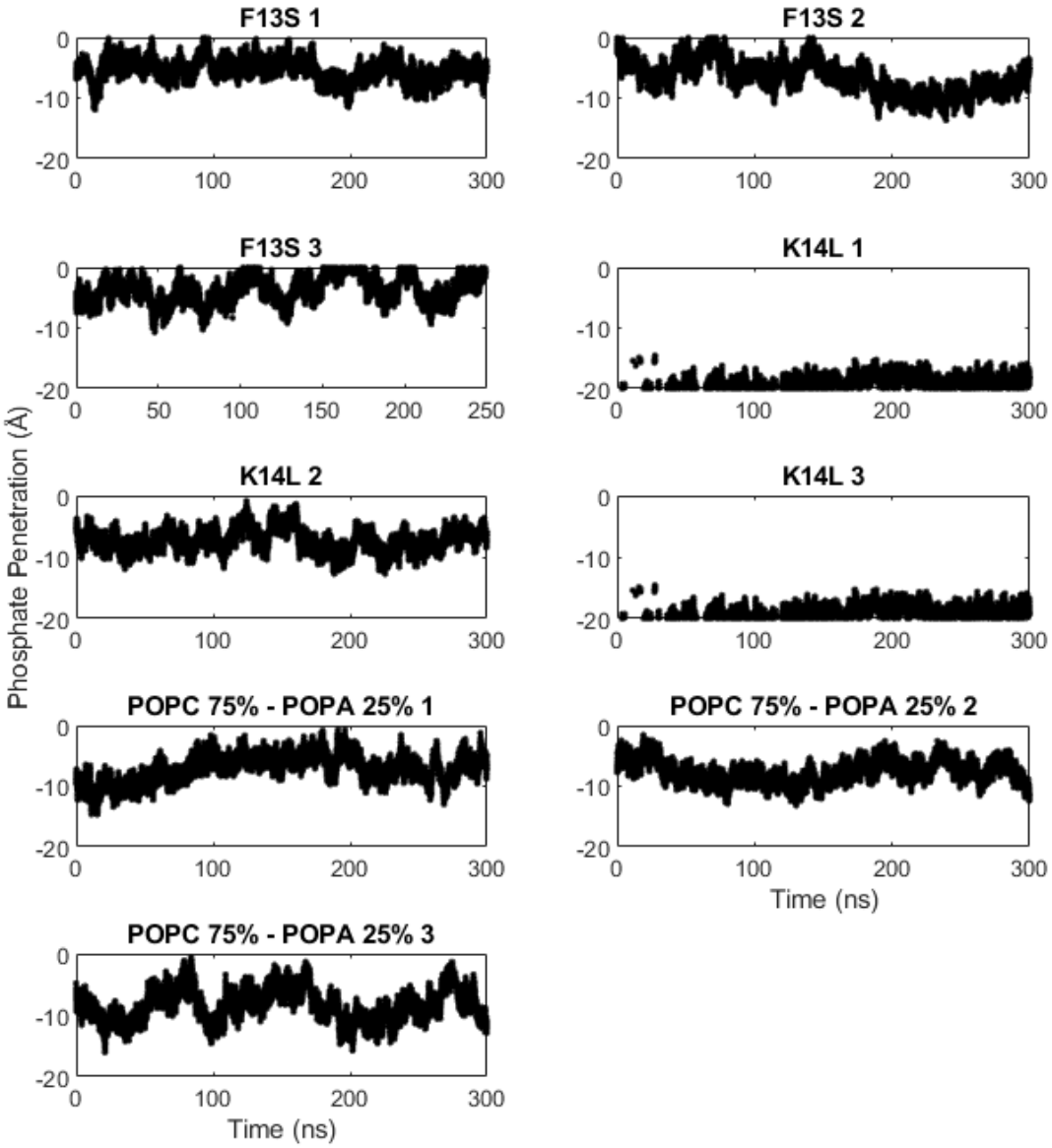


Figure B8) The PHE position for the mutated peptides and specifically the HA HB1 HB2 HD1 HE1 HZ HE2 HD2 atoms are selected in the calculation because they could be isotope labeled.

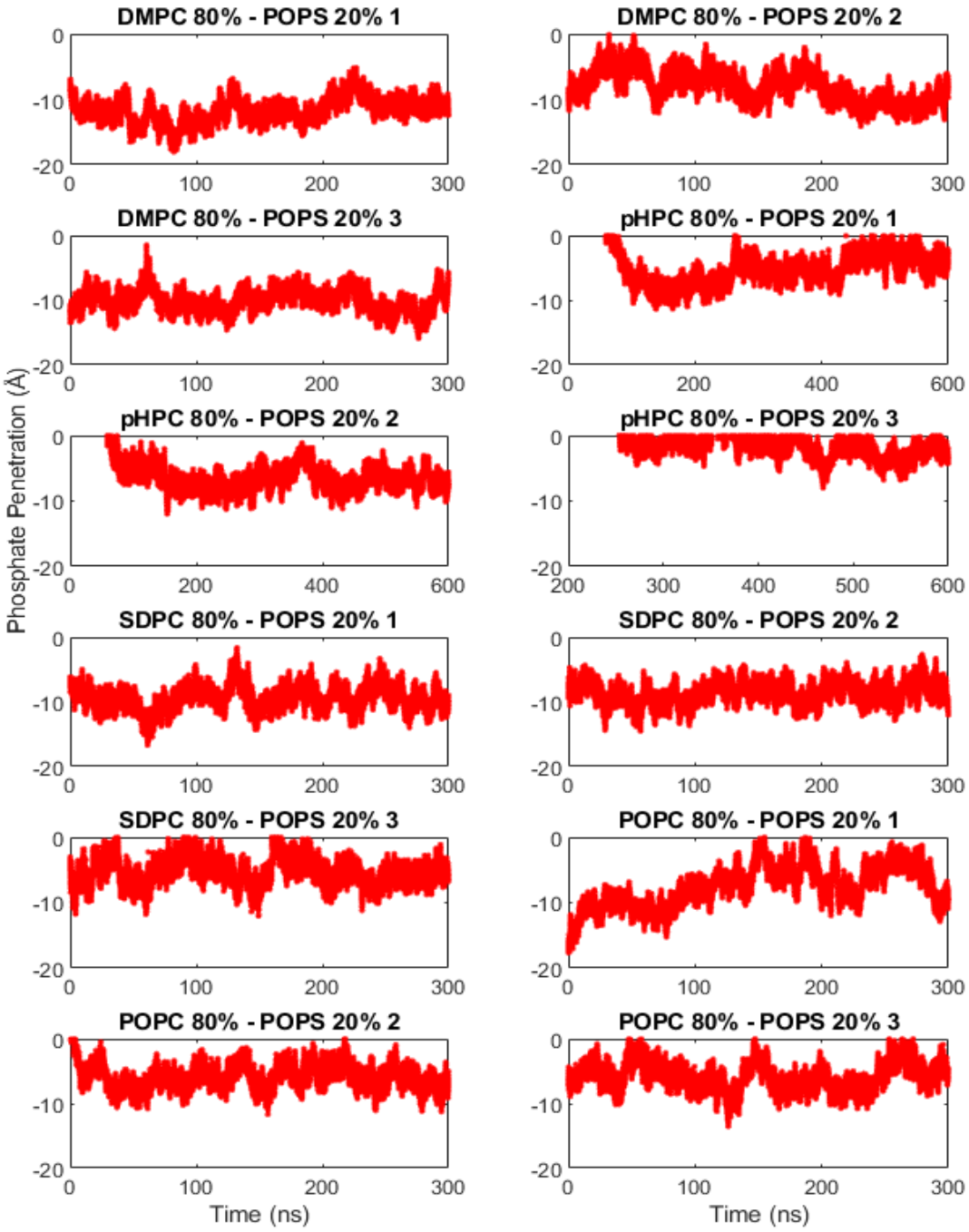


Figure B9) The LEU amino acid of the main Osh4 simulation membranes, tracked over time. Specifically measuring the HA HB1 HB2 HD11 HD12 HD13 HG HD21 HD22 HD23 that can be isotope labeled.

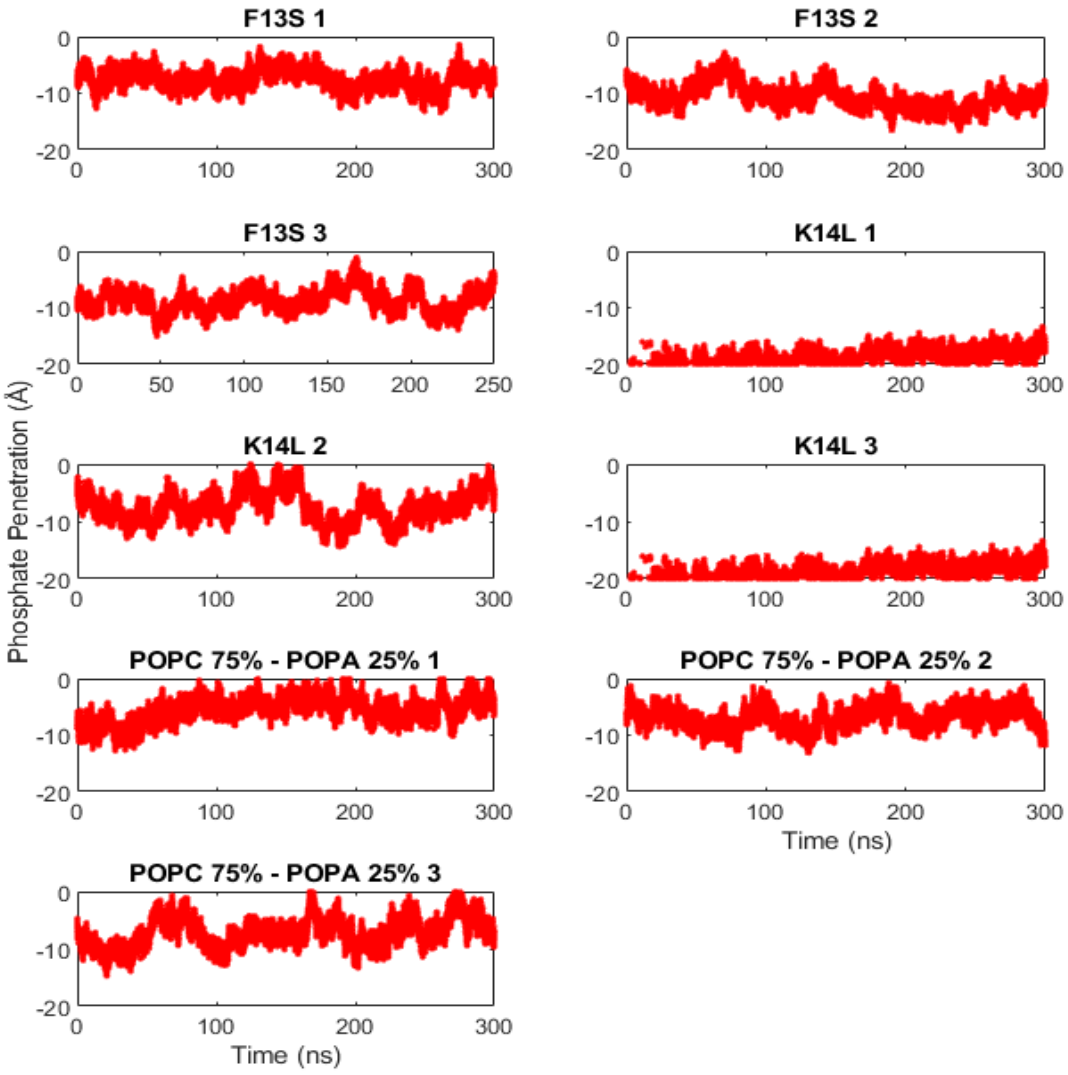


Figure B10) LEU mutant data for the Osh4 simulation membranes, tracked over time. Specifically measuring the HA HB1 HB2 HD11 HD12 HD13 HG HD21 HD22 HD23 that can be isotope labeled.

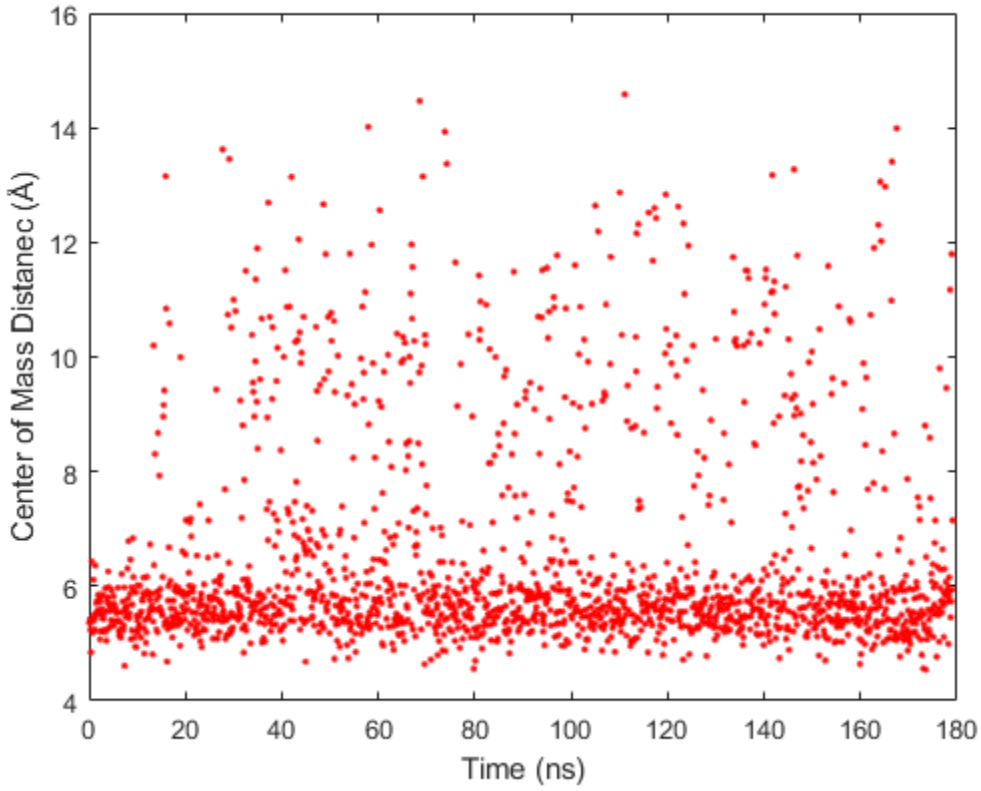


Figure B11) Shows the time series distance vs time between W10 and F13. The cutoff for π–π is not clearly documented for the Charmm force field, but it appears that the average value is between 5-6 Å

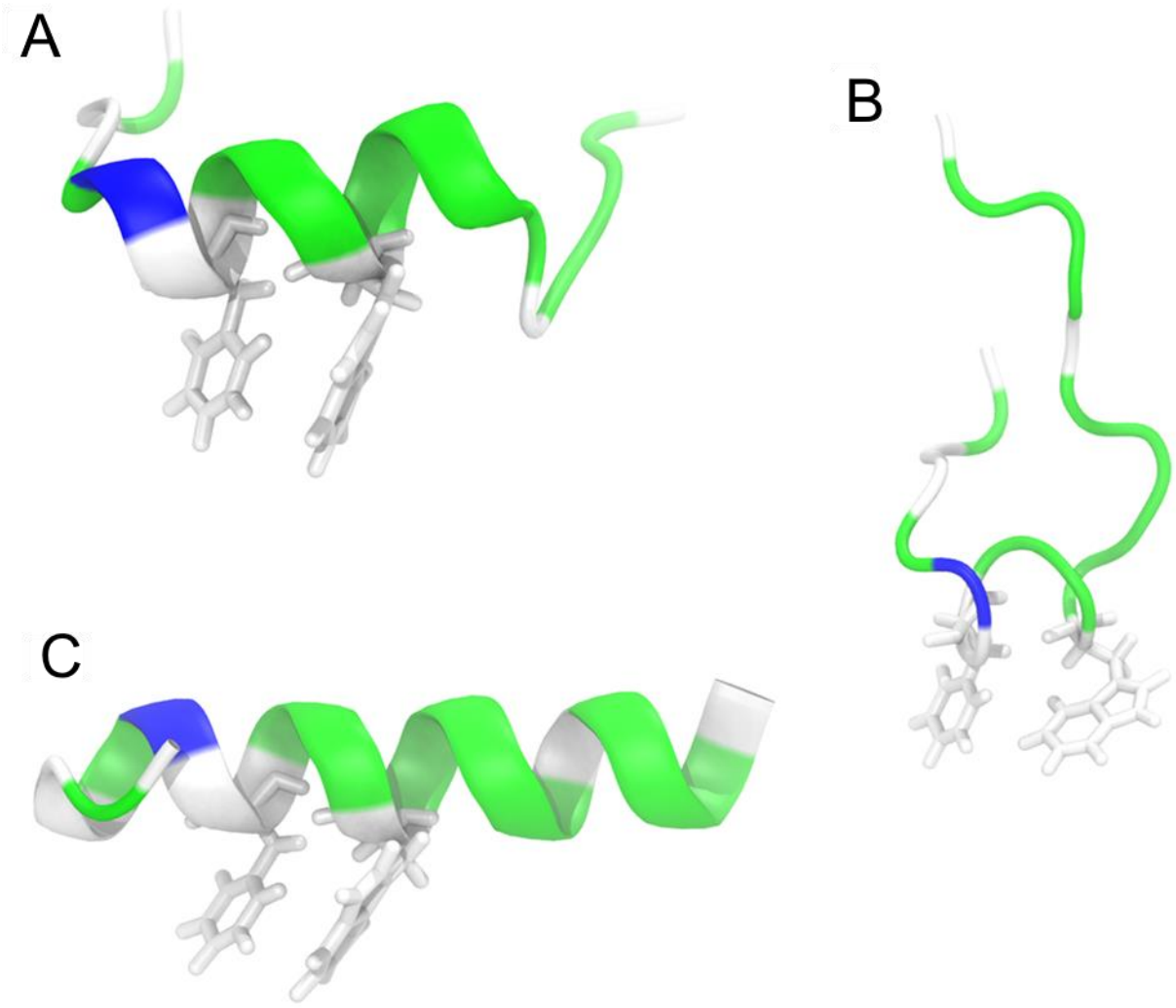


Figure B12) Examples of π–π stacking from the simulation that was studied in Figure B11. Based on these figures it seems like the majority of the interactions have π–π stacking. Panel A-B show varying levels of partially folded helix, panel C shows a mostly folded helix.

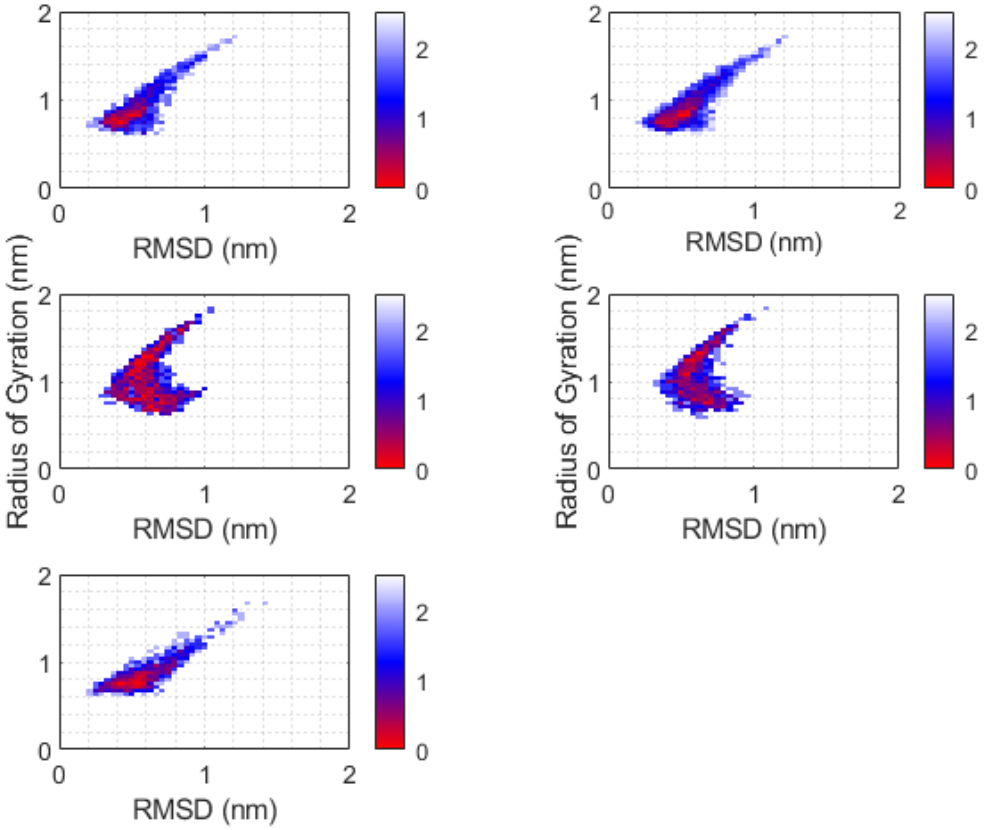


Figure B13) The free energy landscape, as defined with RMSD, is the x-axis and radius of gyration on the y-axis, both in nm.

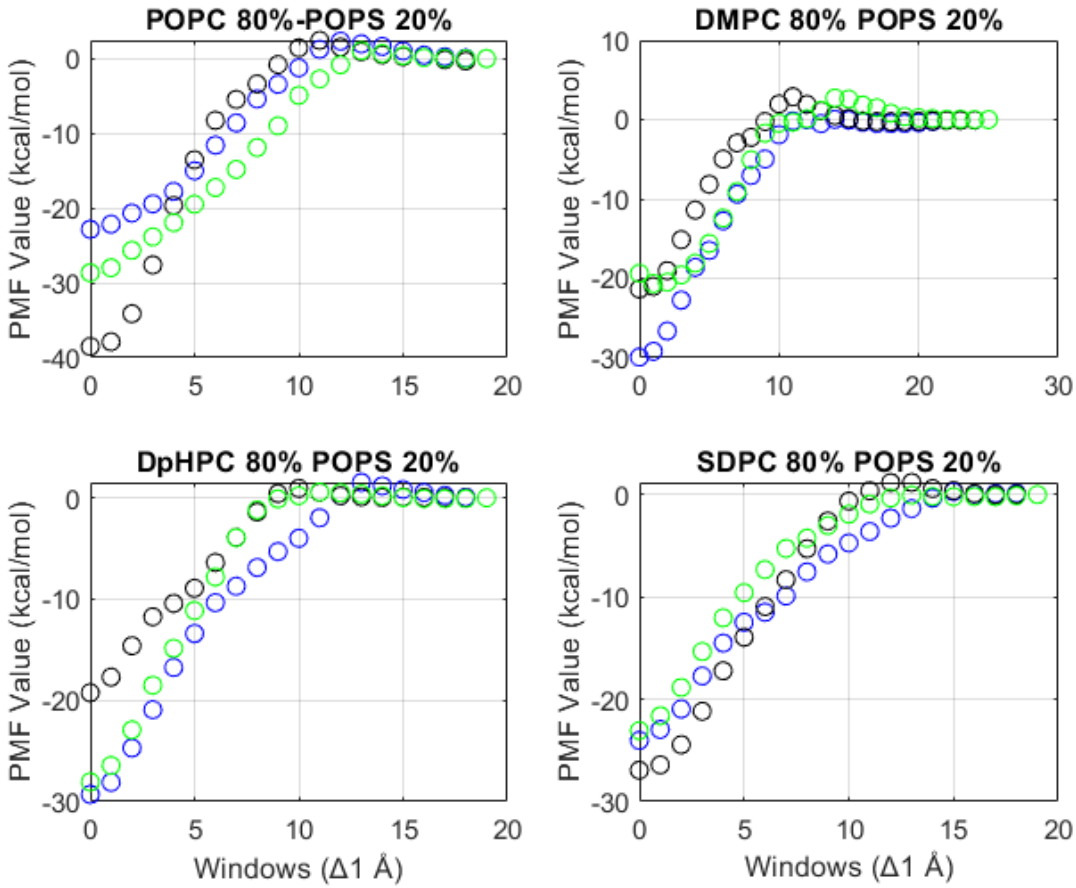


Figure B14) The PMF profiles demonstrate an adequate number of windows to achieve complete separation from the membrane interactions.

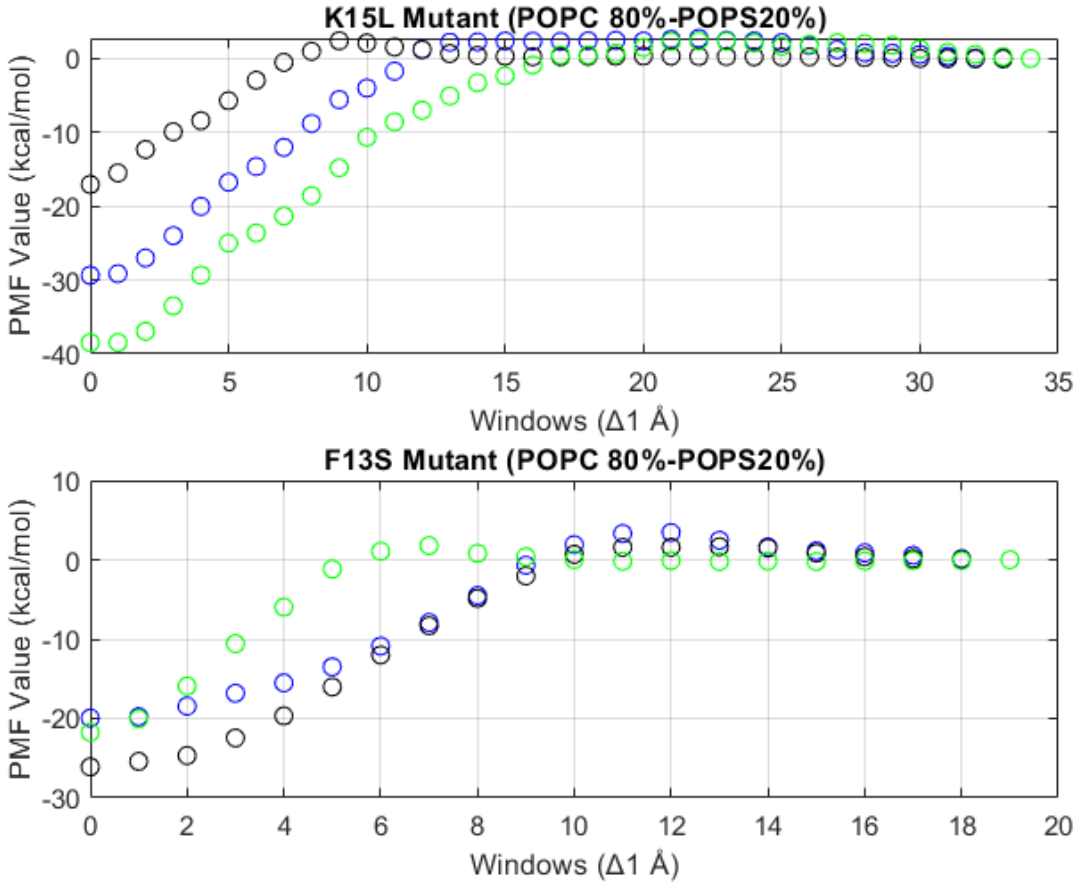


Figure B15) PMF profiles of the mutants showing that there was adequate number of windows to achieve complete separation

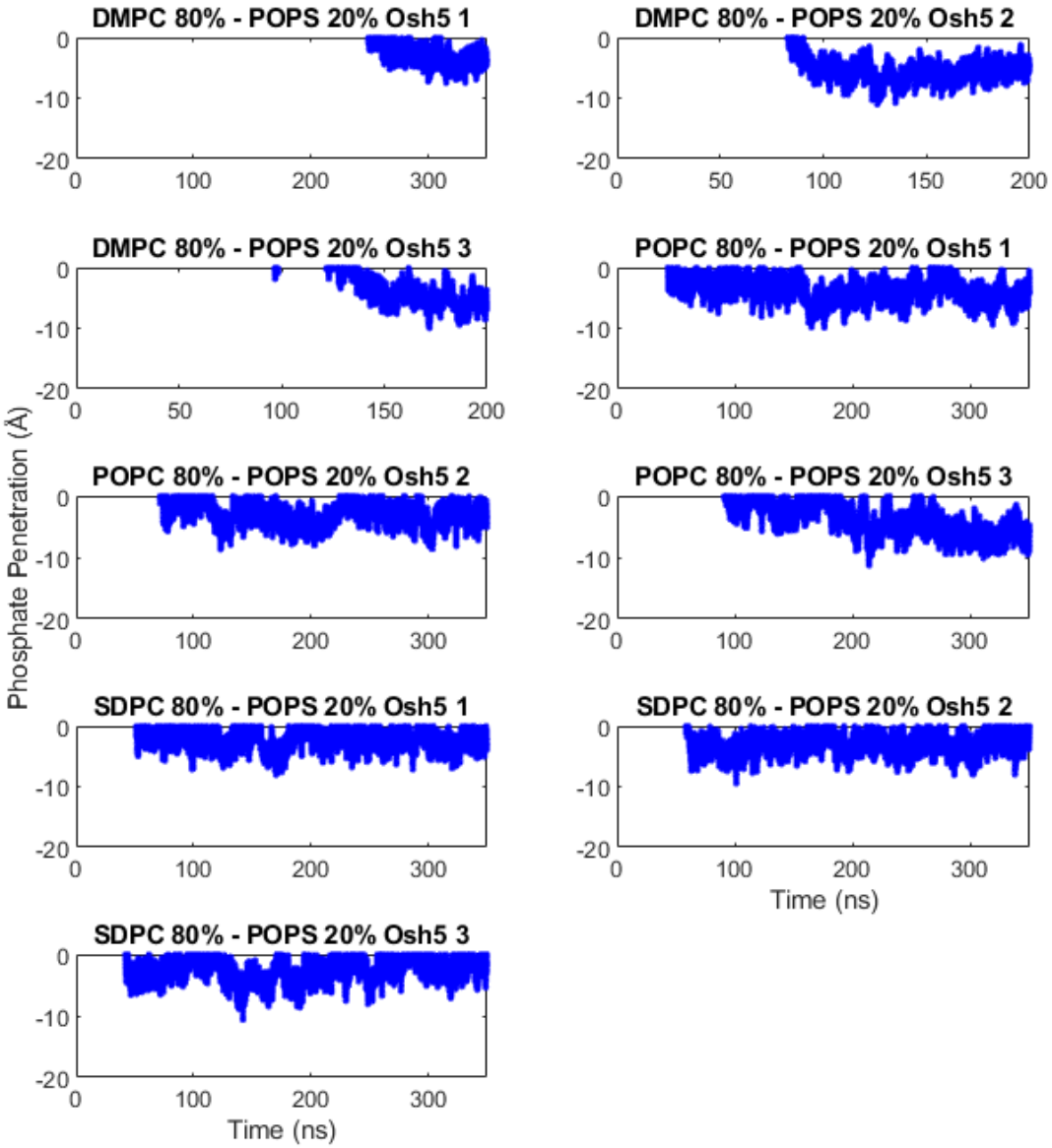


Figure B16) This is the carbon alpha of the Osh5 HMMM binding simulations. These also have some variation that increases after the conversion from HMMM to full-length lipids.

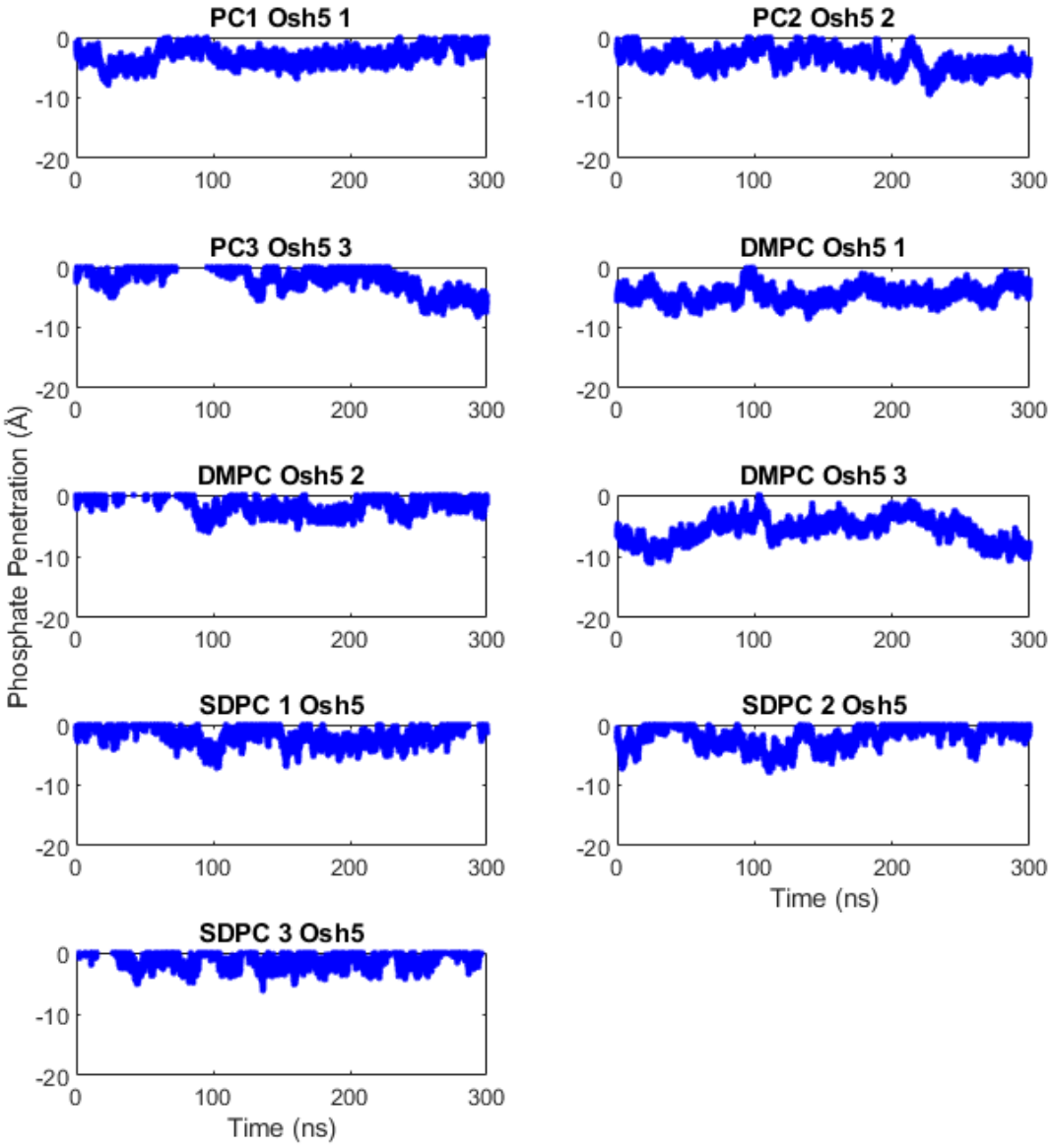


Figure B17) This is the same carbon alpha penetration metric for the osh5 simulations during the full-length simulations showing similar re-arrangements and fluctuations.

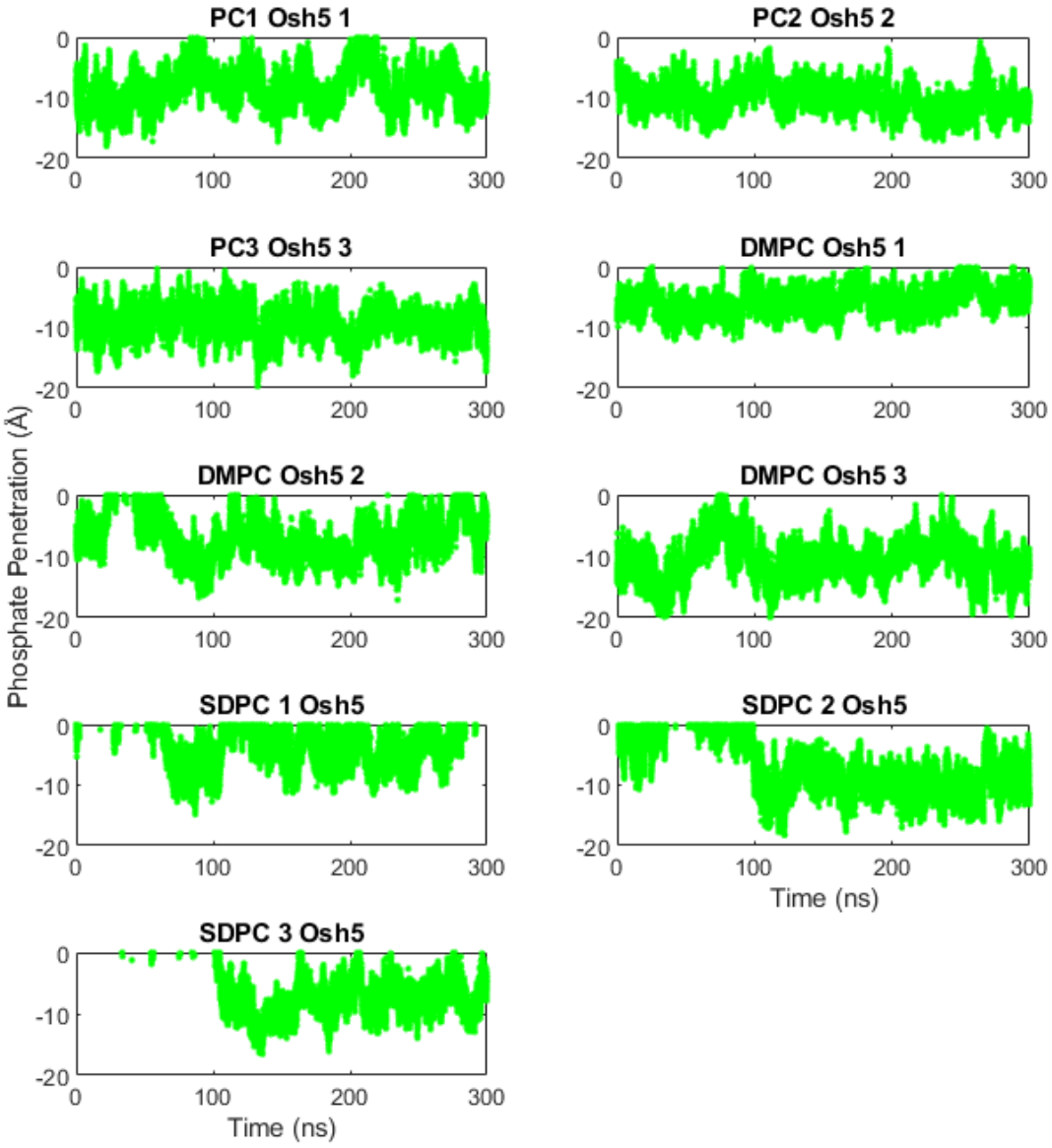


Figure B18) The Met position for the Osh5 peptide. Measuring the HE1, HE2, and HE3 hydrogens that could be isotope labeled.

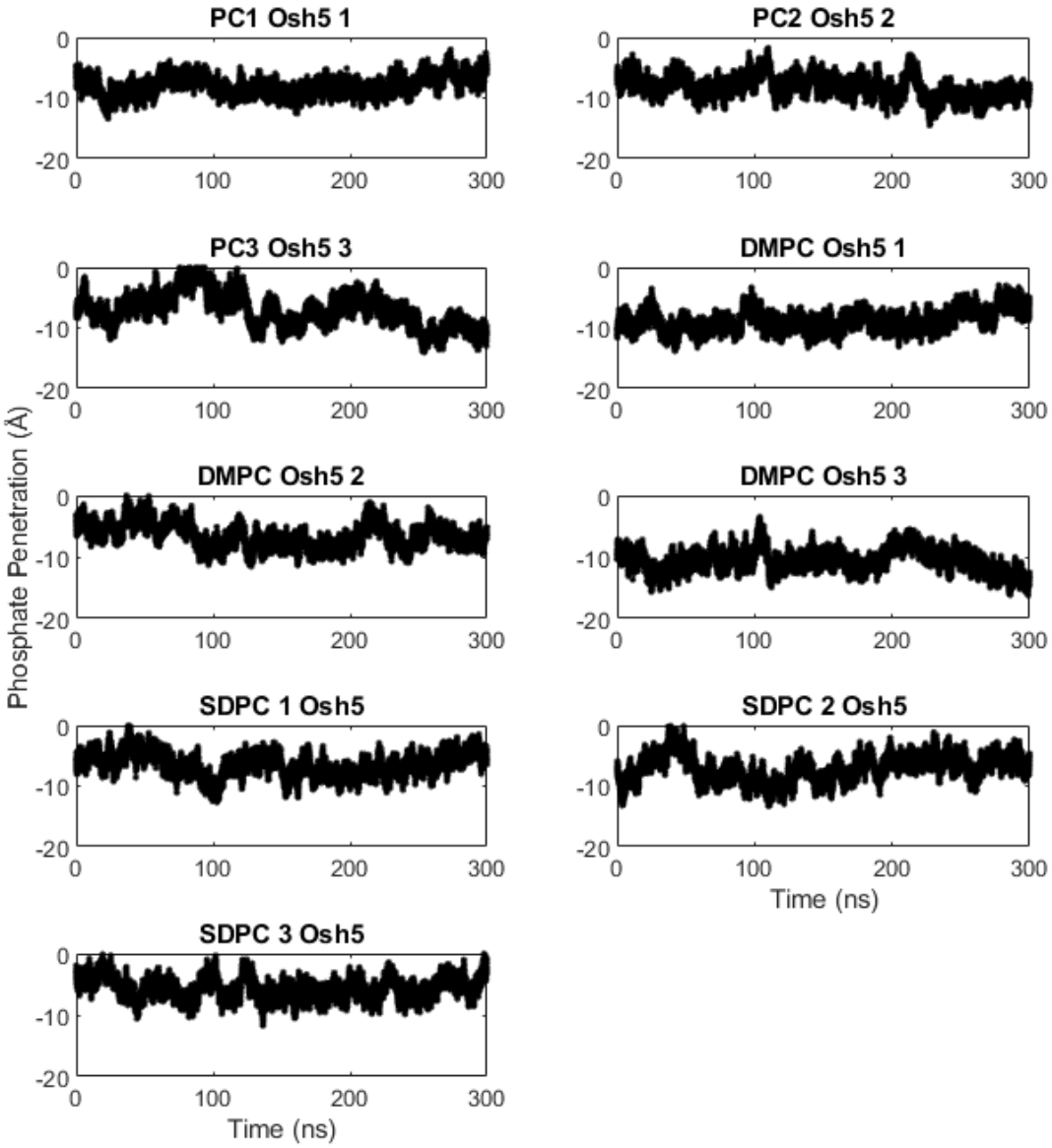


Figure B19) The PHE amino acids position over time for the Osh5 peptides in varied membrane environments. Specifically, the HA HB1 HB2 HD1 HE1 HZ HE2 HD2 atoms are selected in the calculation because they could be isotope labeled.

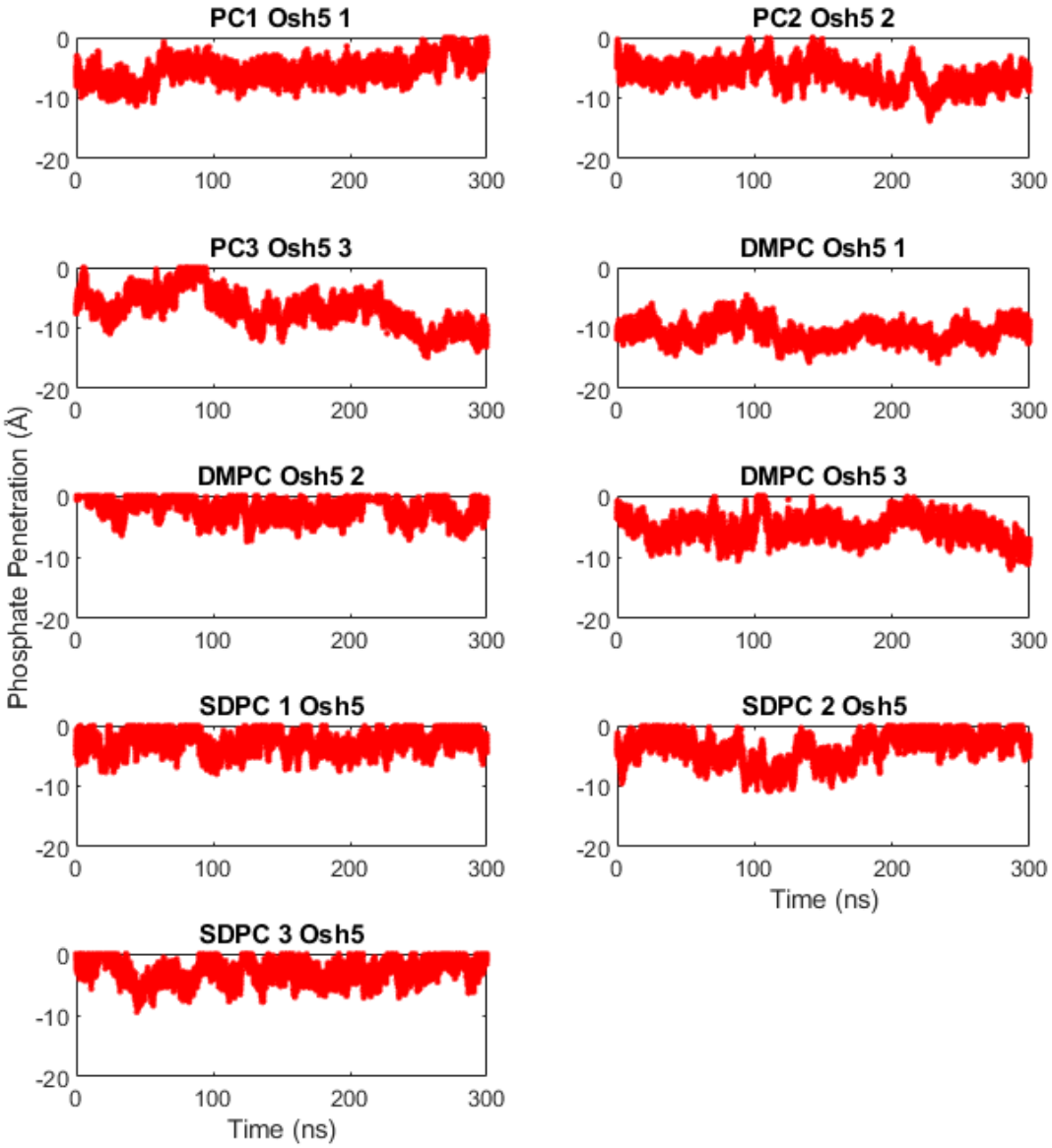


Figure B20) The LEU Osh5 amino acid position measured over time for varied Osh5 membrane types. Specifically measuring the HA HB1 HB2 HD11 HD12 HD13 HG HD21 HD22 HD23

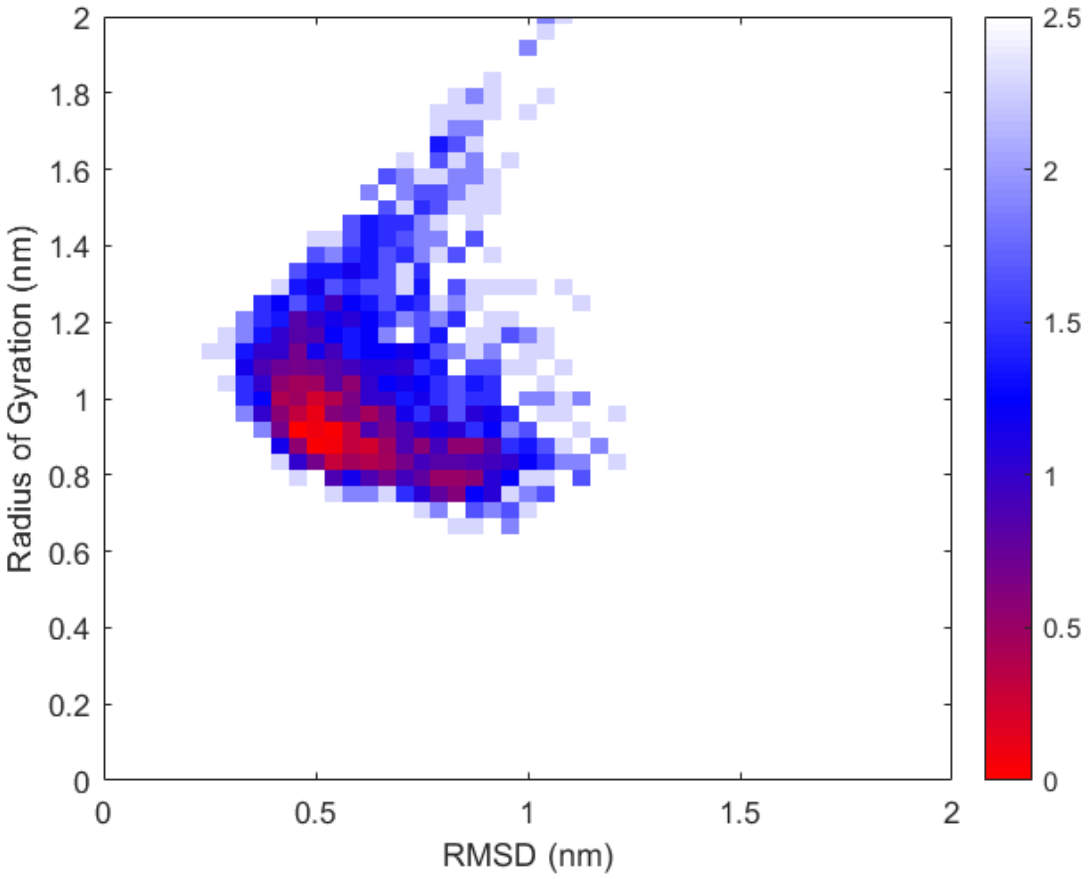


Figure B21) The free energy landscape of the Osh5 peptide as defined by RMSD and radius of gyration in nm.

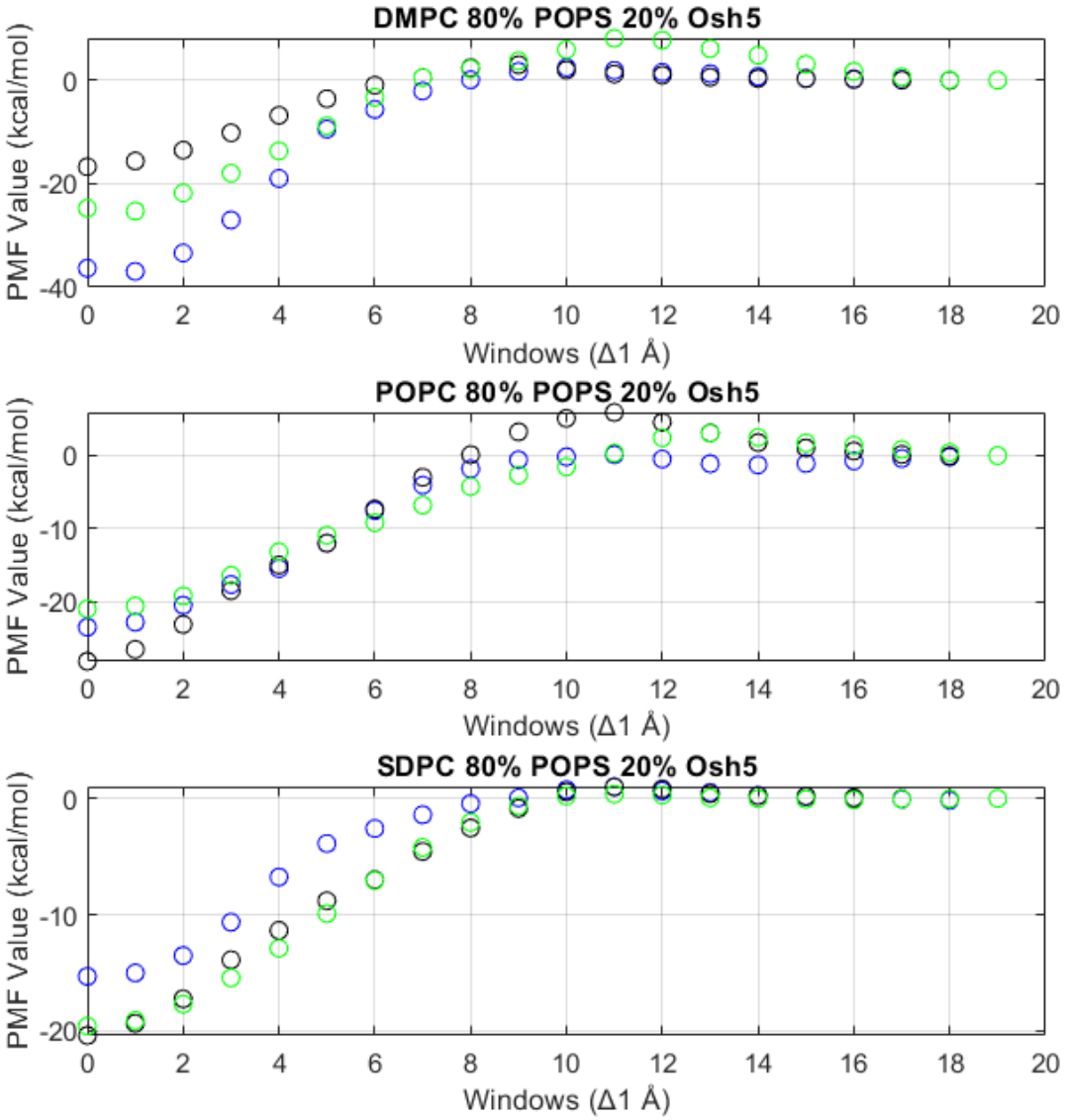


Figure B22) PMF profiles of the Osh5 mutants show an adequate number of windows to achieve complete separation for the PMF measurement.

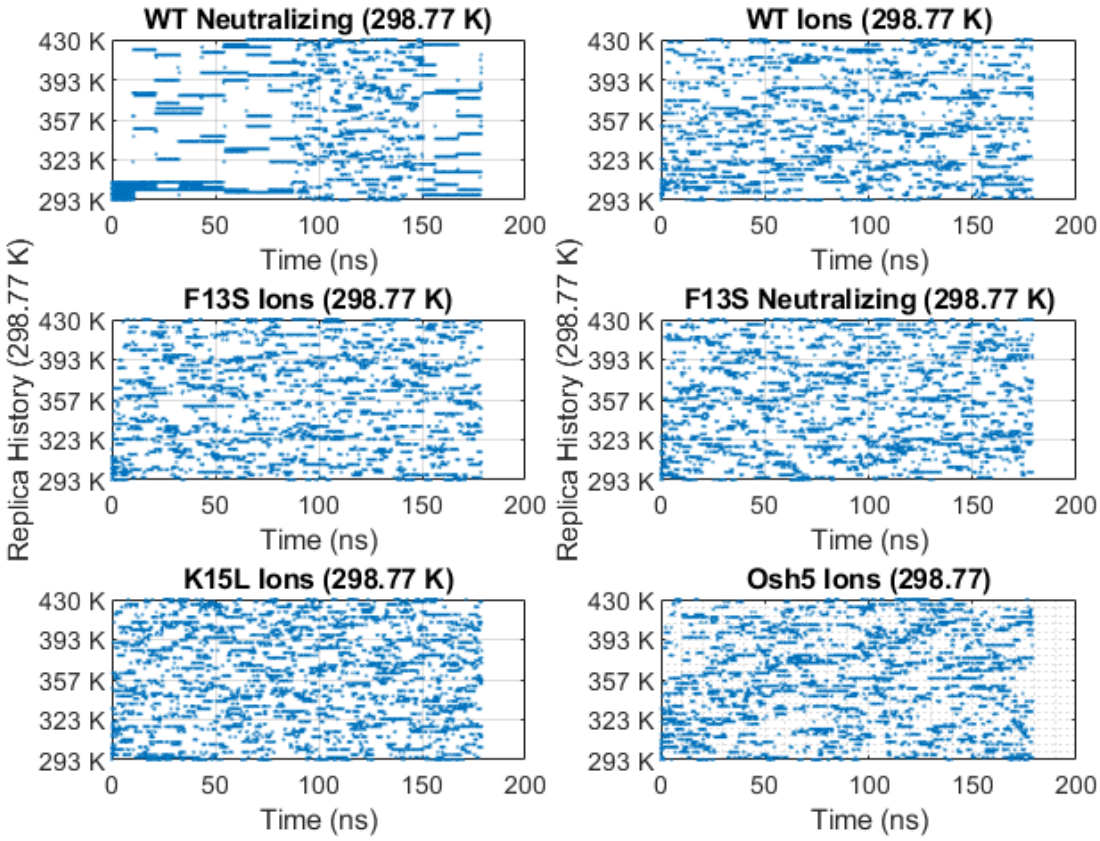


Figure B23) Window history for the 298.77 K simulations. All of these replicas had transitions except the one of the wild type in neutralizing ions.

Appendix C

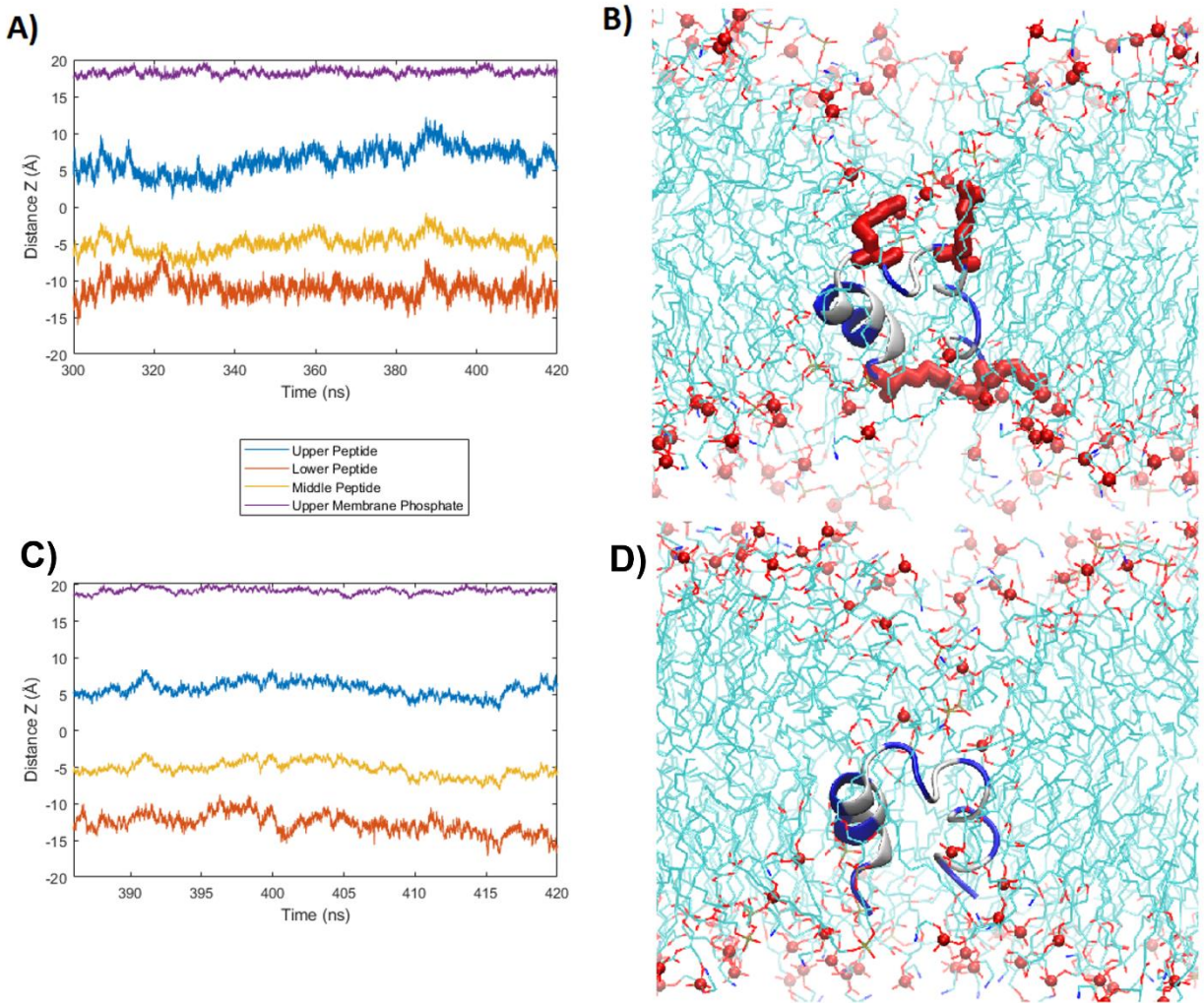


Figure C1. A and B correspond to one simulation, and C and D correspond to another. The A and C trials were used to fit the X-ray crossing points and were equilibrated for 20 ns before the data collection began. The B and D figures show a copy of the system set to 0 dyne/cm to test the extreme stability cases. The residues tracked to generate the position vs. time graph were the hydrogens on the guanidinium groups of R12 and R16 to generate the upper peptide (blue). Amino acids R1 and R24 (the termini) were used for the bottom peptide (orange). The choice of which hydrogen was the same for all AAs (HH12). Additionally, all of the residues used were always arginine for the long chain that extends to the furthest edges of the space occupied by the peptide. The middle peptide was tracked by using the entire peptide. The simulation snapshot B depicts the residues selected (red tubular shapes) and the phosphates (small red spheres) shown in all simulations.

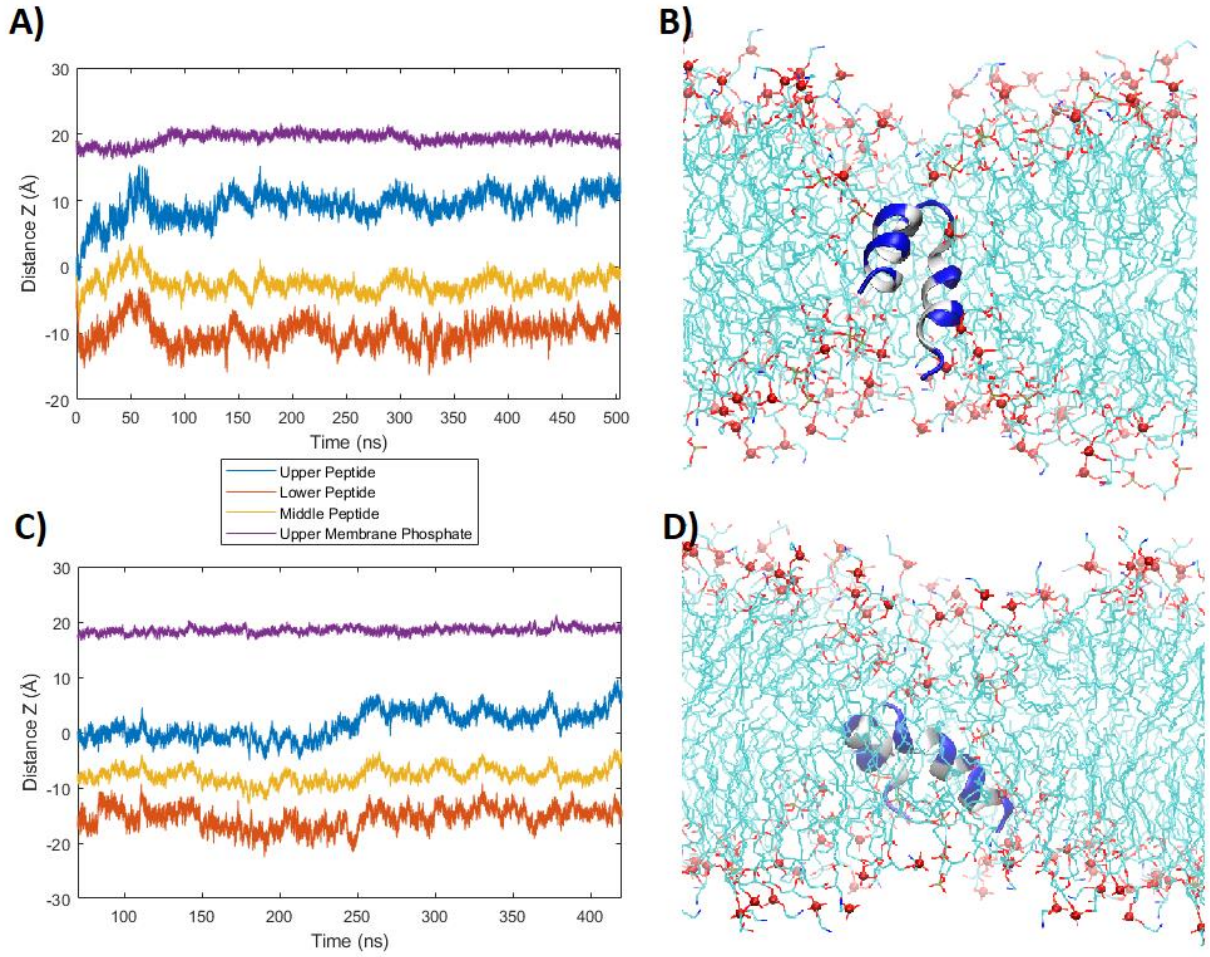


Figure C2. A and B correspond to one simulation, and C and D correspond to another. The A and B simulation was the longest simulation set at or below the target surface tension to ascertain if the peptide stayed inserted at 15 dyne/cm. It was changed to 9 dyne/cm when it was stable. The residues that were used to test stability were R12 and R16 for the upper peptide (blue), R1 and R24 for the lower peptide (orange), and the entire peptide for the middle peptide (yellow). The C and D figures show the longest simulations continuously set to 15 dyne/cm. This simulation used R8 and R12 for the upper peptide to monitor its trajectory.

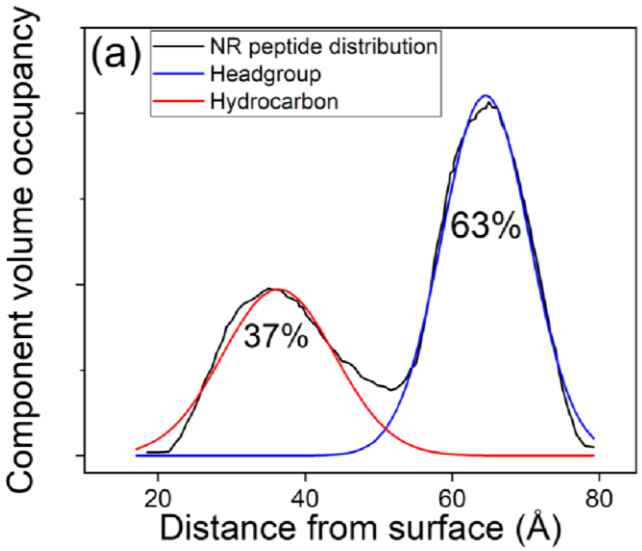


Figure C3. Method of determining WLBU2 partitioning in (a) G(-) IM, with NR peptide distribution shown as red solid circles.

Appendix D

Table D1) The original thermally engineered amino acid point mutations[217] and which sites and numbering correspond with the wild type amino acids in the report. The wild-type sequence came from Uniprot ID P09038.

Site 1	Engineered	NA	NA	NA	NA
	WT in this work	K261	R262	K267	K271
Site 1'	Engineered	H59F	R31L	V52T	E54D
	WT in this work	H192	R164	V185	E187
Site 2	Engineered	T121P	S109E		
	WT in this work	T254	S242		
Site 2'	Engineered	C96N	S94I	L92Y	
	WT in this work	C229	S227	L225	
Destabilizing Point Mutations	Engineered	Y120	W123		
	WT in this work	Y253A	W256A		

Table D2) Shows the chemical linkages of the different polysaccharides that were studied in this work.

Tri-3	α-*L-IdoA-(1->4)*-α-*D-GlcNAc-(1->4)-β-D-GlcA-1->OMe*
Tetra-1	α-*L-IdoA-(1->4)*-α-*D-GlcNAc-(1->4)-β-D-GlcA-(1->3)-β-D-Gal-1->OMe*
Heparin aFGF	α*LIdoA(1→4)*α-D*Glc(1→4)*α*LIdoA(1→4)*α-D*Glc(1→4)*α*LIdoA(1→4)*α-D*Glc* -Glucose has sulfamino at position 2 and sulfate at position 6 -Iduronic acid has sulfate at position 2

Table D3) The REMD simulation carbohydrate compositions in mM and total ng/ml of sugar. The total ng/ml is comparable to the upper sugar concentrations that were used in the work by Ornitz and co-workers. This also highlights how the fructose has less overall mass of sugar.

Trial	System Size	# of Waters	mM F-1,6-BP	mM Tri-3	mM Tetra-1	Total ng/ml Sugar
Wild Type bFGF with 4 Tri-3 and 1 Tetra-1 (High Temp Range)	80390	25923	0	8.56	2.141	6.55E+06
Wild Type bFGF with 4 Tri-3 and 1 Tetra-1 (Low Temp Range)	80390	25923	0	8.56	2.141	6.55E+06
Wild Type bFGF with 5 molecules of F-1,6-BP	77796	25133	11.04	0	0	3.71E+06
Y253A, W256A, Y257A with 5 F-1,6-BP molecules	77781	25140	11.04	0	0	3.71E+06
Control (High Temp Range)	68068	21945	0	0	0	0
Control (Low Temp Range)	68068	21945	0	0	0	0
Y253A, W256A, Y257A Control	75912	24567	0	0	0	0
Y253A, W256A Control	75884	24554	0	0	0	0
Stabilizing Mutations	68040	21933	0	0	0	0

Table D4) Showing the simulation time lengths, configuration, and number of sugars for each of the simulations with short, charged sugars that mimic heparin.

Trial	Length	Configuration	Number of Sugars
Literature bFGF	-	-	-
IP6	546 ns	HMR CUFIX	Two IP6
IP6 Replica	432 ns	HMR CUFIX	Two IP6
Single IP6	Deleted Sugar From Equilibrated Two IP6	-	One IP6
IP4	584 ns	HMR CUFIX	Two IP4
IP4 Replica	480 ns	HMR CUFIX	Two IP4
HS Strong Side	400 ns	HMR CUFIX	Heparin
HS Weak Side	400 ns	HMR CUFIX	Heparin

Table D5) There were several PMF trials conducted on aFGF protein, and these conclude that the small, charged molecules are in agreement with the literature values for IP6.

2 x Kcal/(mol IP6)	1 x Kcal/(mol IP6)	Trial	Force Field	Timestep	Unrestrained Pre-Equilibration (ns)	Equilibration (ns)	Run Time (ns)
-14.62	-7.31	Literature bFGF	-	-		-	-
-19.69	-9.85	2 x IP6	CUFIX	4fs/step	126	16	34
-21.25	-10.63	2 x IP6 replica	CUFIX	4fs/step	432	16	34
-	-16.27	1 x IP6 one	CUFIX	4fs/step	126	16	34
-17.83	-8.92	2 x IP4	CUFIX	2fs/step	584	14	26
-	-15.63	Heparan Sulfate	CUFIX	2fs/step	400	18	22

Table D6) The interaction energy analysis between either protein A or protein B and the polysaccharide CARA indicated that there was a measurable difference between the binding of the two proteins that was reproducible and could correspond to the higher and lower binding interaction supporting the -15 vs. -36 kcal/mol.

Interaction Energy Last 50ns of Relaxation (Kcal/mol)			
	Trial 1 Protein A		Trial 1 Protein B
Average:	-123.8±0.3	Average:	-102.6±0.3
VDW Average:	-28.9±0.1	VDW Average:	-23.1±0.1
ELEC Average:	-94.9±0.3	ELEC Average:	-79.5±0.3
	Trial 2 Protein A		Trial 2 Protein B
Average:	-87.4±0.2	Average:	-121.6±0.3
VDW Average:	-22.0±0.1	VDW Average:	-34.4±0.1
ELEC Average:	-65.4±0.3	ELEC Average:	-87.1±0.3

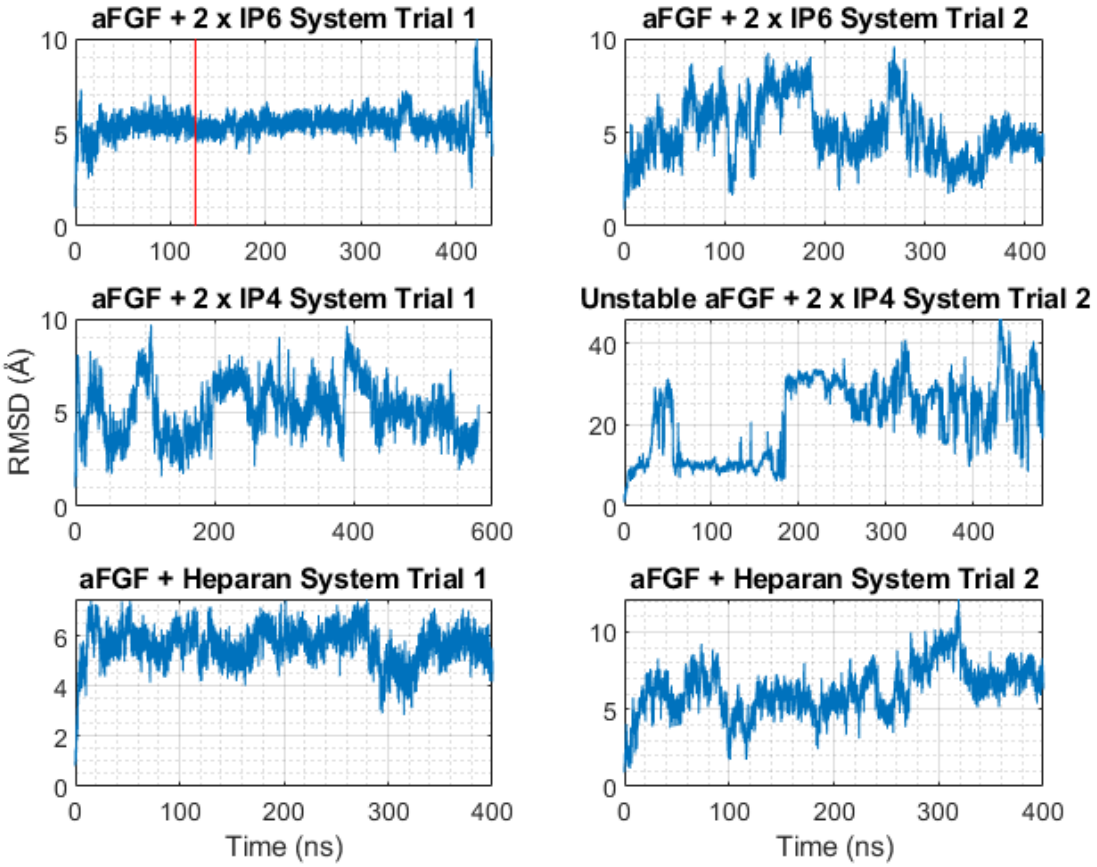


Figure D1) Showing the RMSD of trial #1 of the aFGF systems dimerized with 2 IP6 molecules. The RMSD of the global dimer RMSD relaxation from the crystal structure, rather than the individual RMSD of each protein individually. The red line indicates where the system was actually stopped to convert into the first PMF system, but the system was continued to try to get a more relaxed system, but as that happened the second IP6 molecule dislodges in the simulation.

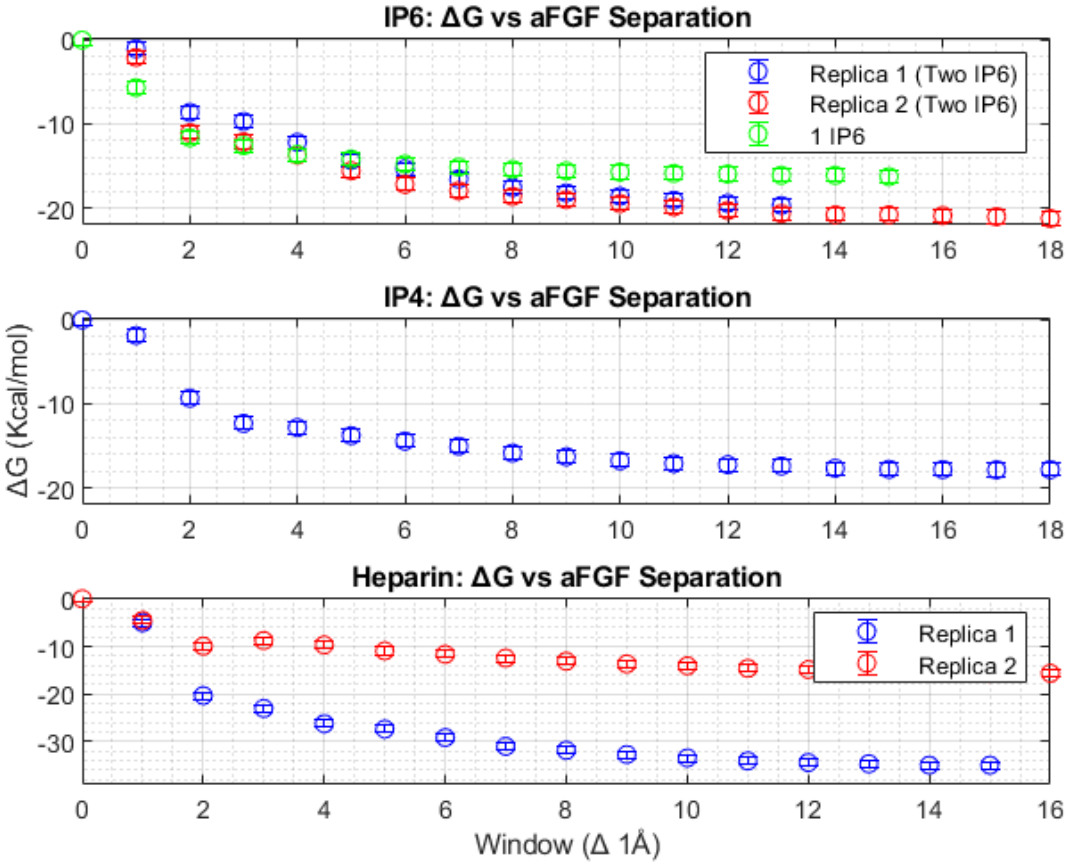


Figure D2) This figure shows the ΔG associated with pulling apart the aFGF dimers to compare with the literature Kd value and illustrate that for a system containing two IP6 sugars at the interface the energy come to ΔG -9.85 ± 0.75 kcal/mol per sugar. There were notable restraints to prevent rotation as the aFGF separated, and IP6 molecules were simulated with a lower charge (-6 simulation) vs (-8 physiological) due to the hypoxic condition of interest.

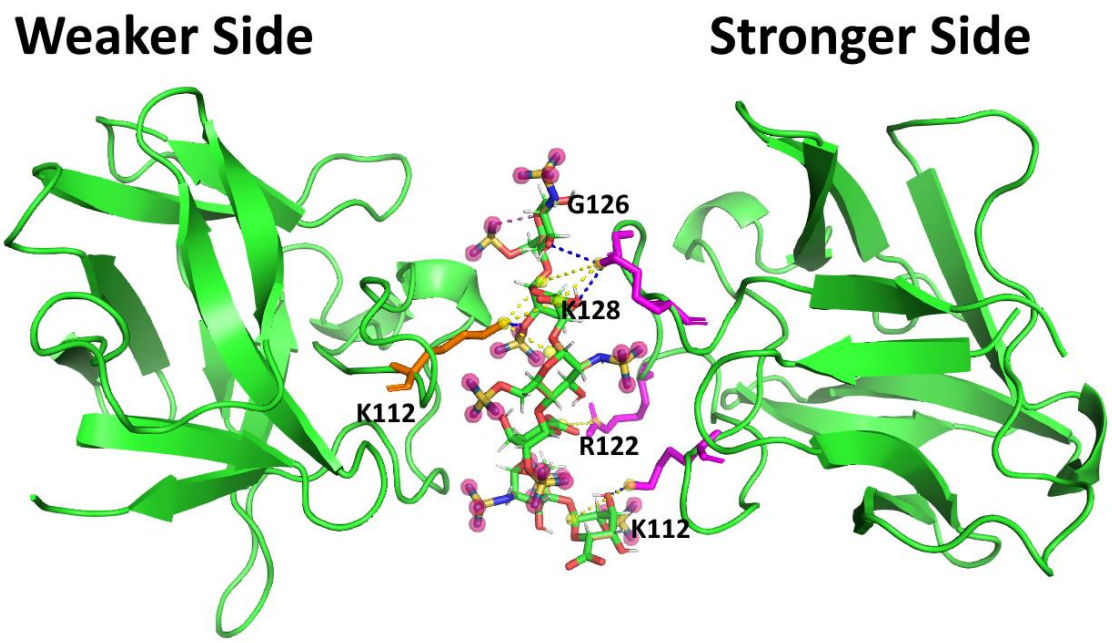


Figure D3) This figure is a snapshot of the heparin dimer. The main point is that the interactions between the two protein are uneven. The unequal interactions makes one side stronger and another side weaker. The magenta colors belong to protein A "stronger side", and the single orange color corresponds to protein B the "weaker side". The exact number of interactions varies over time. The sulfates stand out as magenta spheres, and the ionic interactions are dashed yellow and result from higher partial charges, while the hydrogen bonds are dashed blue.

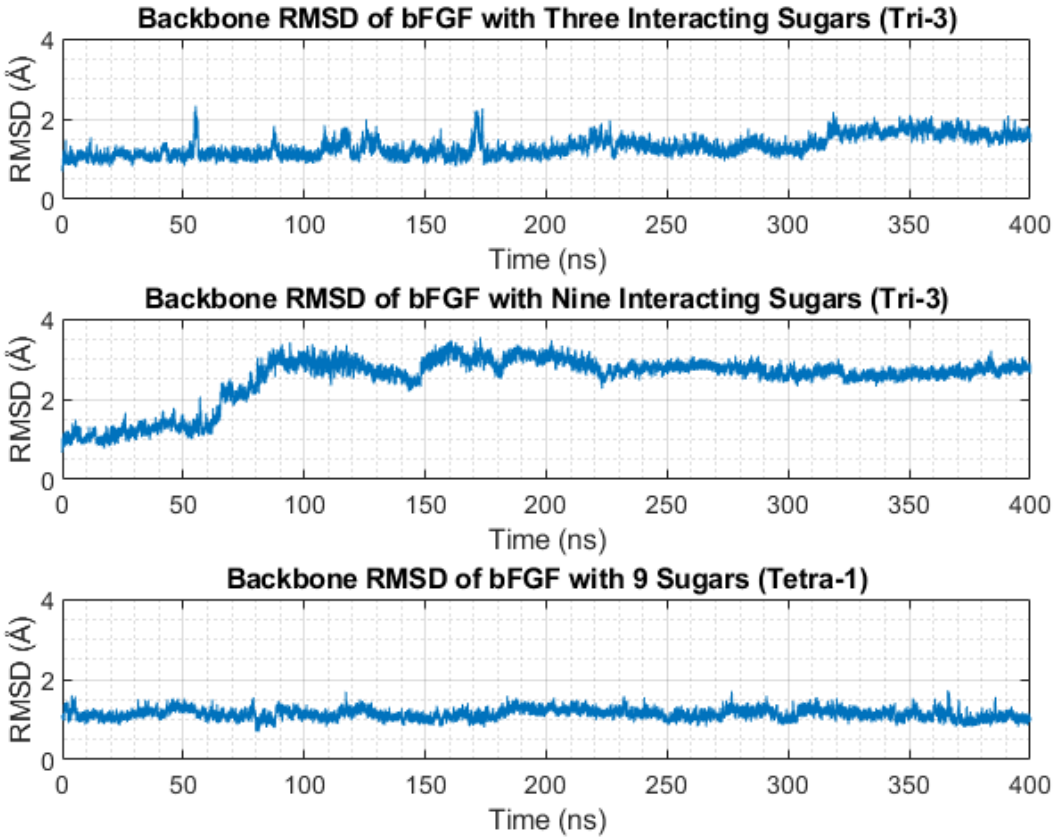


Figure D4) Showing the RMSD of the HMR simulated bFGF protein over the full time of simulation as three, or nine Tri-3 sugars, and also with nine Tetra-1 sugars each in separate simulations. The sugars would bind and unbind from the bFGF protein. This indicates little change from the backbone and so the sugars didn't cause major conformational changes from the original predicted crystal structure. The rise in RMSD at 40 ns for the nine Tri-3 polysaccharides is likely related to the slight unfolding of the protein in the primary binding site. In the Tetra-1 simulation the protein retains much of the original structure even at high sugar concentration, although it is known that this sugar has one of the weaker activities.

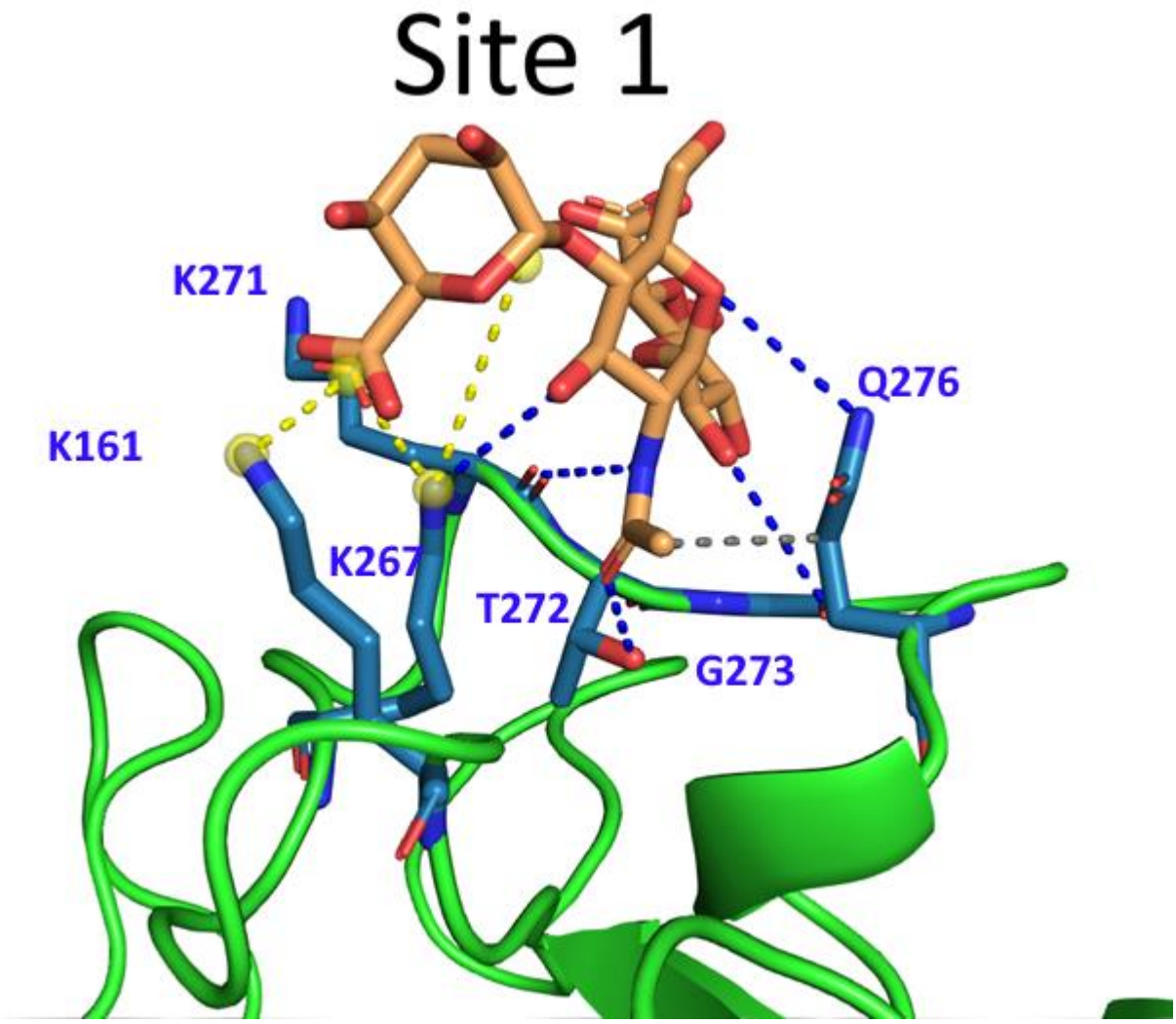


Figure D5) Showing salt bridges in yellow to the partial charges on the sugar represented by yellow spheres between two oxygens, hydrogen bonds in blue, and hydrophobic interactions in grey. The interacting amino acids are labeled blue because there were no thermodynamic point mutations. Site 1 has a strong interaction in this pocket but lacks clear interactions between this site and the other thermodynamically unstable regions. The tri-3 sugar was defined as *α-L-IdoA-(1->4)- α-D-GlcNAc-(1->4)- β-D-GlcA-1->OMe*

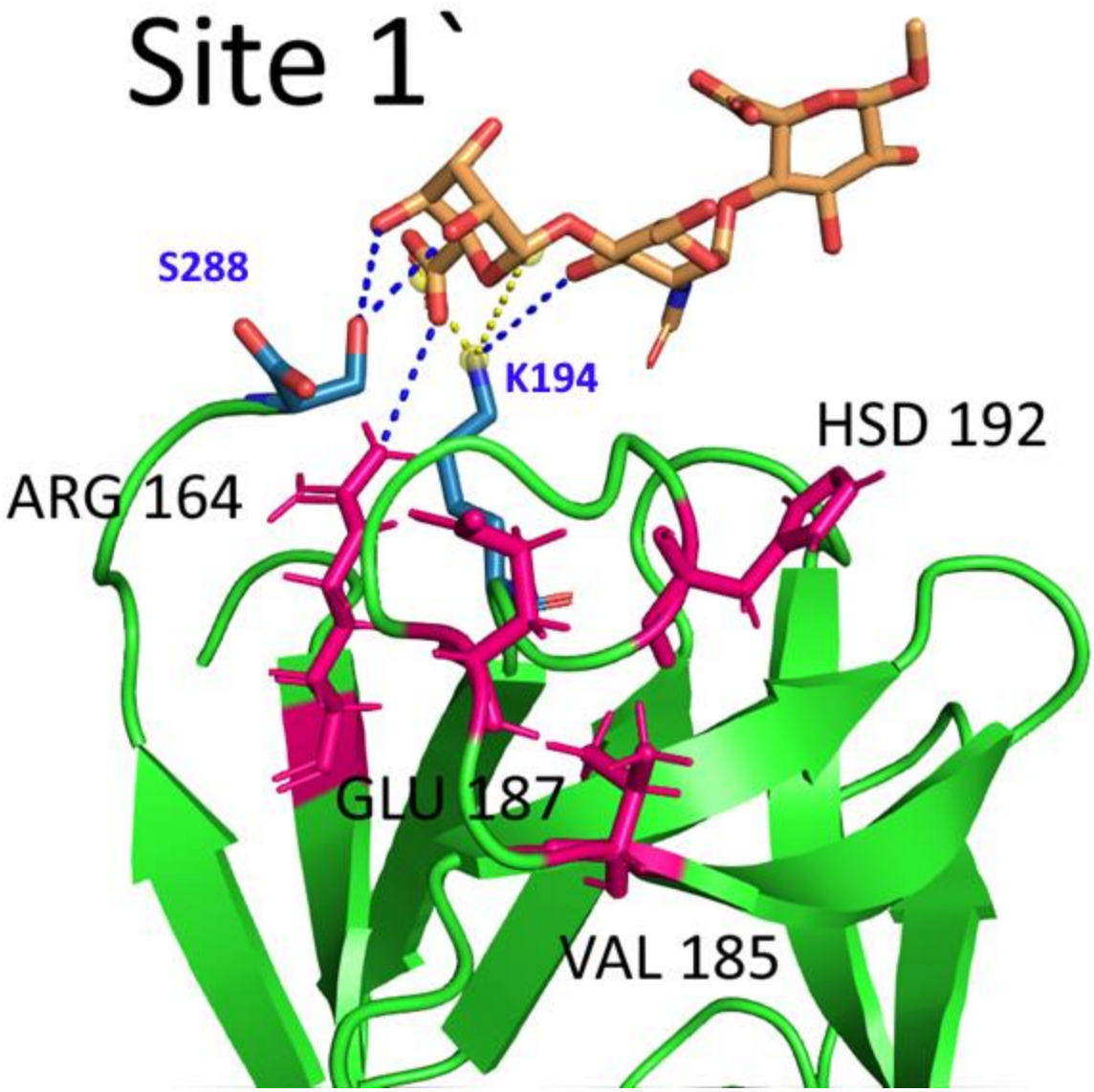


Figure D6) This binding site appears to have a slight majority of the thermally unstable amino acids. The salt bridges are shown in yellow dotted lines to the yellow spheres on the sugar between the two oxygens, and hydrogen bonds are dotted blue, the pink depicts the thermally unstable amino acids, while other amino acids had direct interactions, and were labeled with blue. The tri-3 sugar was defined as *α-L-IdoA-(1->4)- α-D-GlcNAc-(1->4)- β-D-GlcA-1->OMe*

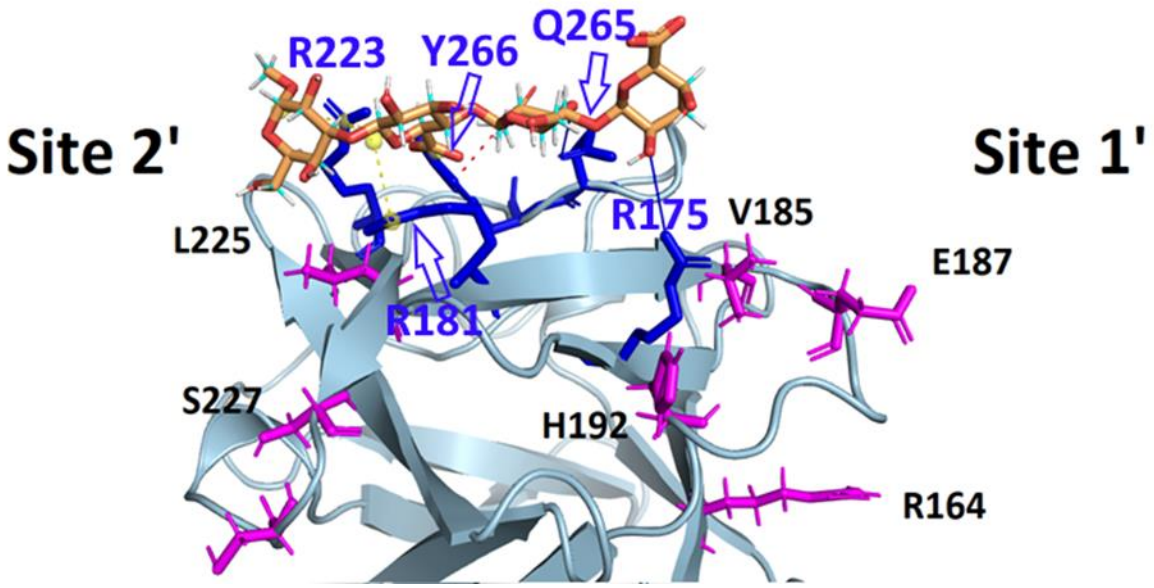


Figure D7) This demonstrates how the Tetra-1 heparan derivative can bind to and reinforce both site 2' and site 1' simultaneously. This contrasts with Figure D6 above because the sugar is not situated directly above the site 1' binding site, but rather interacting indirectly with the arginine amino acids that form the β-strands that compose site 2' and one of the strands on site 1'. There is one yellow salt bridge indicated by the dashed yellow line connecting to the partial charge between two oxygens on the sugar. The blue amino acids were found to have direct interactions, and the pink amino acids were thermally engineered amino acids. The labels were separated so that the blue ones correspond to the amino acids with direct interactions and the black text corresponds to the thermally unstable amino acids. This might stabilize as many as three amino acids by reinforcing V185 E187 and H192. The tetra-1 sugar was defined as *a-L-IdoA-(1->4)-a-D-GlcNAc-(1->4)-B-D-GlcA-(1->3)-B-D-Gal-1->OMe*

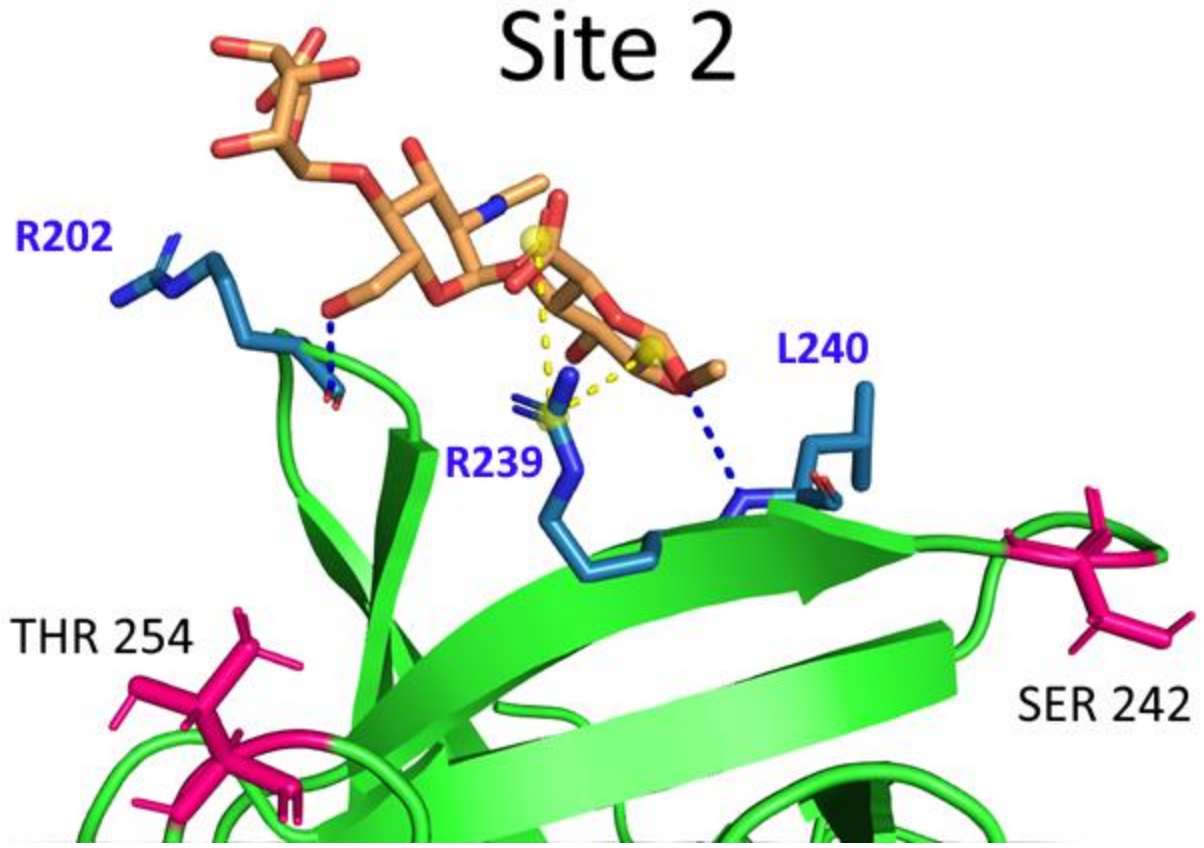


Figure D8) This demonstrates a Tri-3 sugar pose that reinforces the T254 and S242 amino acids. This is mainly ionically driven interaction with the arginine amino acids that line two separate β-strands. There are two charged regions on the sugar that are indicated by the yellow spheres with partial charge connected to R239 and the dashed blue lines correspond to the hydrogen bonds. Of the sugar binding sites that were identified this is the closest to the original crystal structure, while the site 2' and site 1'/site2' combo placements would be difficult to predict. Interestingly, even though this site would be predicted from the crystal structure it had a weaker free energy. The tri-3 sugar was defined as *α-L-IdoA-(1->4)- α-D-GlcNAc-(1->4)- β-D-GlcA-1->OMe*

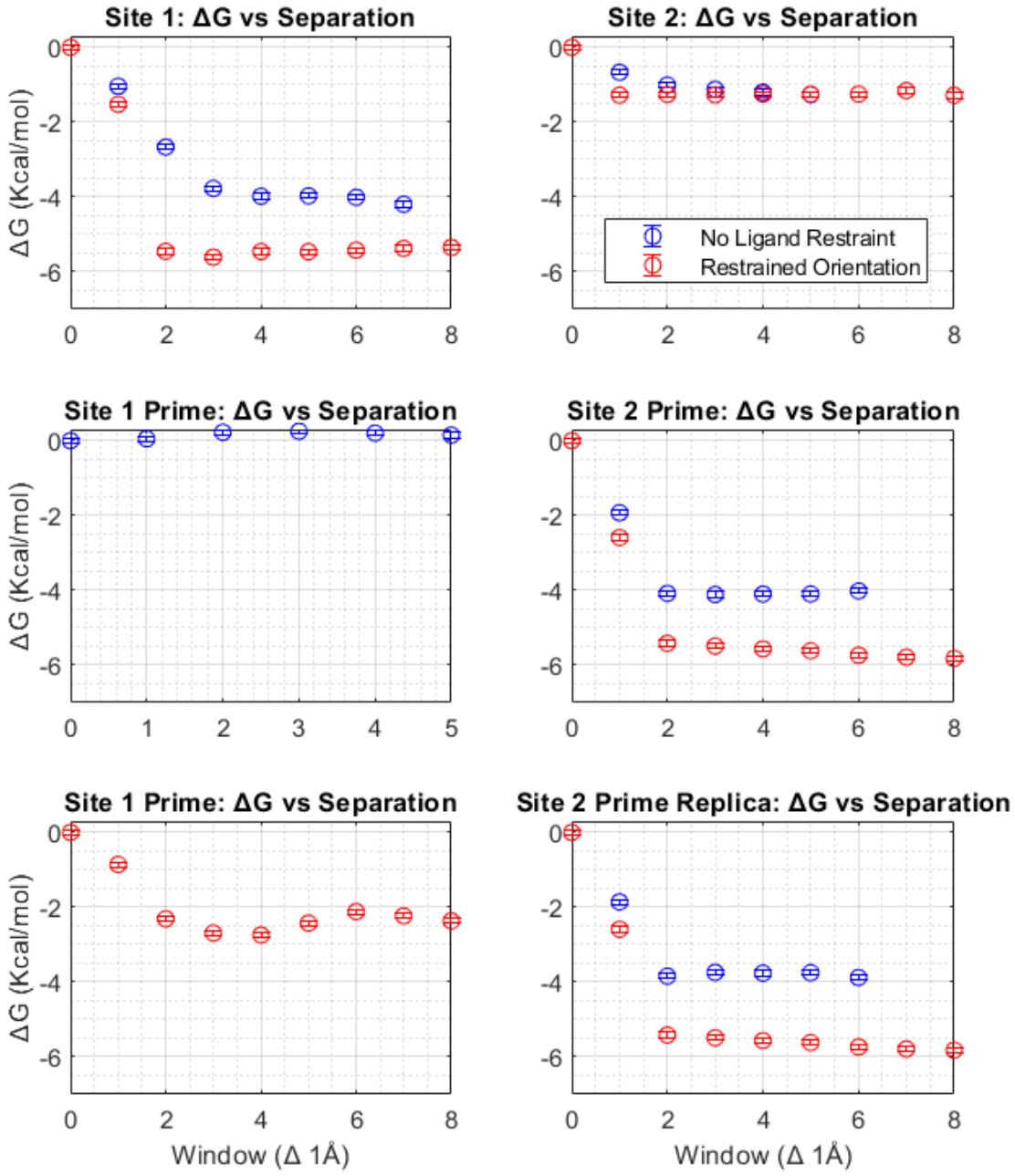


Figure D9) The figure shows the PMF of the primary binding site compared against the newly identified binding sites with different restraints on the ligand as it came off the binding site, which show minor differences. The site 2' and site 1' are the most promising binding sites that were identified. The tri-3 sugar was defined as *α-L-IdoA-(1->4)- α-D-GlcNAc-(1->4)- β-D-GlcA-1->OMe* and the Tetra-1 sugar was defined as *α-L-IdoA-(1->4)-α-D-GlcNAc-(1->4)-β-D-GlcA-(1->3)-β-D-Gal-1->OMe* and only used for the site 1' simulation that had around -2kcal/mol.

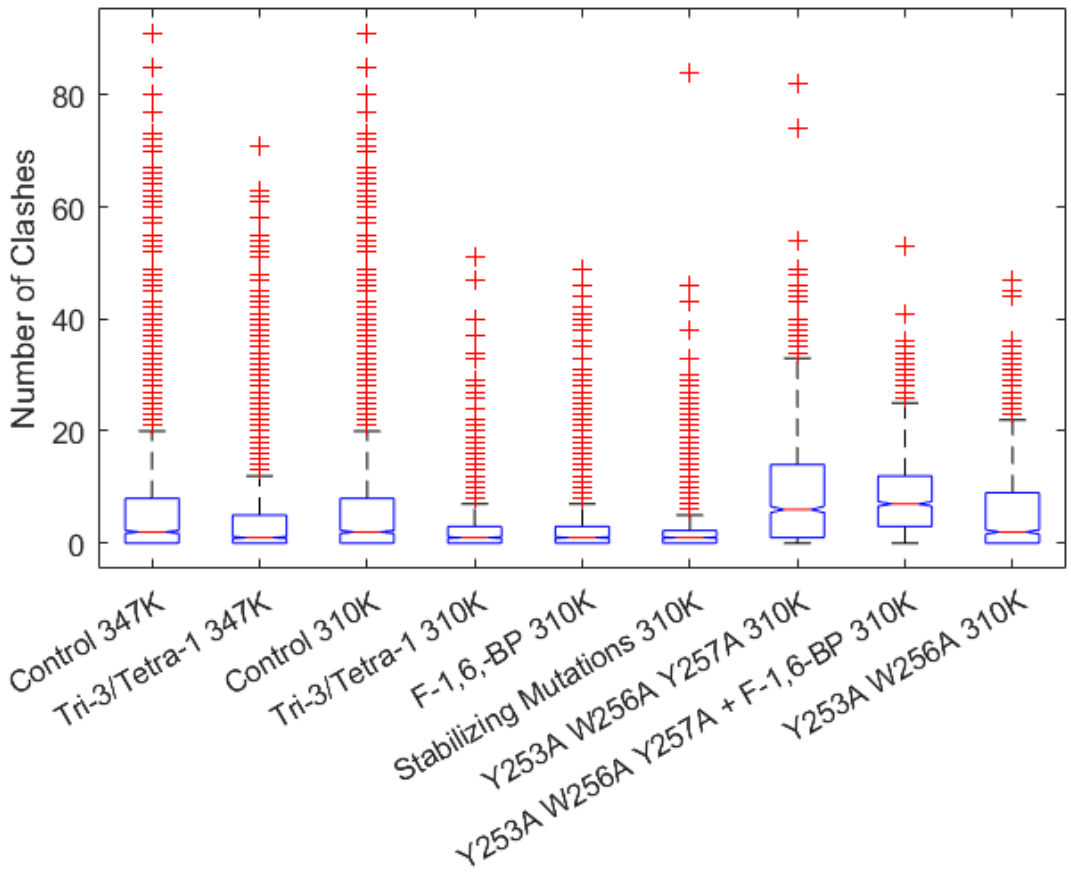


Figure D10) Same as figure 12 in the main text except the clashes that are outliers are also shown. There is a clear difference in the upper quartile of the box plots that are stabilized and those that are not stabilized.

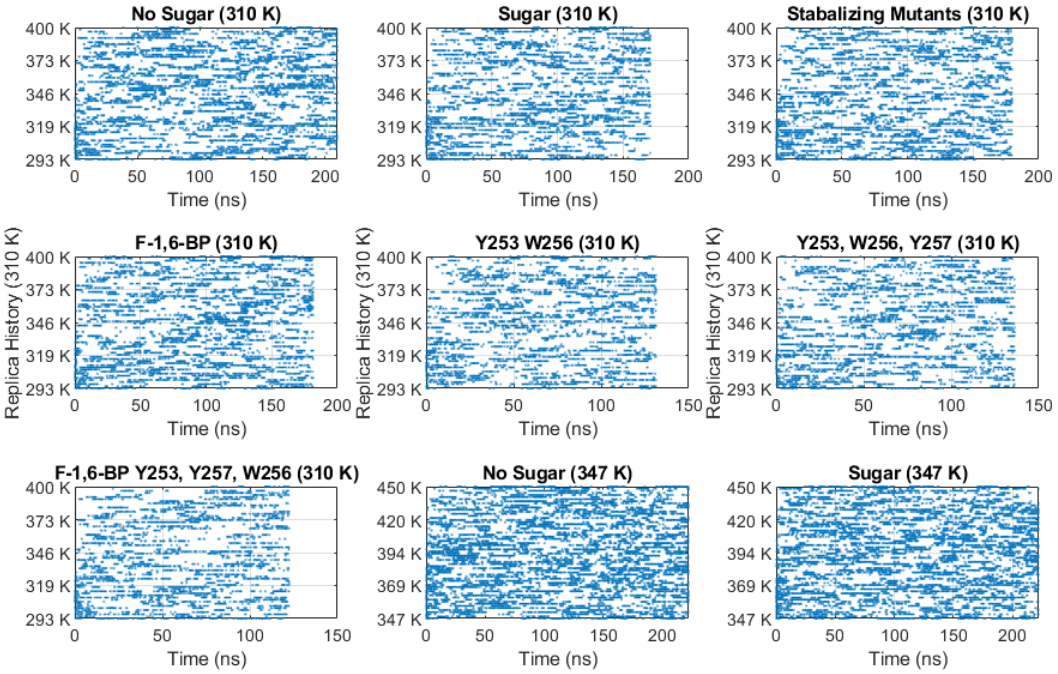


Figure D11) This figure shows the temperature history of the replicas. This indicates that all of the trials had good exchanges and sampled the full energy landscape well.

References:

(1) Fairn, G. D.; McMaster, C. R. Emerging roles of the oxysterol-binding protein family in metabolism, transport, and signaling. *Cellular and Molecular Life Sciences* **2008**, *65* (2), 228-236. DOI: 10.1007/s00018-007-7325-2.
(2) Fahlman, C.; Bickler, P.; Sullivan, B.; Gregory, G. Activation of the neuroprotective ERK signaling pathway by fructose-1, 6-bisphosphate during hypoxia involves intracellular Ca2+ and phospholipase C. *Brain research* **2002**, *958* (1), 43-51.
(3) Barnes, N. M.; Hales Tg Fau - Lummis, S. C. R.; Lummis Sc Fau - Peters, J. A.; Peters, J. A. The 5-HT3 receptor--the relationship between structure and function. (0028-3908 (Print)). From 2009 Jan.
(4) Zhang, Y.; Dijkman, P. M.; Zou, R.; Zandl-Lang, M.; Sanchez, R. M.; Eckhardt-Strelau, L.; Köfeler, H.; Vogel, H.; Yuan, S.; Kudryashev, M. Asymmetric opening of the homopentameric 5-HT3A serotonin receptor in lipid bilayers. *Nature Communications* **2021**, *12* (1), 1074. DOI: 10.1038/s41467-021-21016-7.
(5) Polovinkin, L.; Hassaine, G.; Perot, J.; Neumann, E.; Jensen, A. A.; Lefebvre, S. N.; Corringer, P.-J.; Neyton, J.; Chipot, C.; Dehez, F. Conformational transitions of the serotonin 5-HT3 receptor. *Nature* **2018**, *563* (7730), 275-279.
(6) Guros, N. B.; Balijepalli, A.; Klauda, J. B. Microsecond-timescale simulations suggest 5-HT-mediated preactivation of the 5-HT3A serotonin receptor. *Proceedings of the National Academy of Sciences of the United States of America* **2020**, *117* (1), 405-414. DOI: 10.1073/pnas.1908848117 WorldCat.org.
(7) Lina, T. T.; Alzahrani, S.; Gonzalez, J.; Pinchuk, I. V.; Beswick, E. J.; Reyes, V. E. Immune evasion strategies used by Helicobacter pylori. *World journal of gastroenterology* **2014**, *20* (36), 12753-12766. DOI: 10.3748/wjg.v20.i36.12753 PubMed.
(8) Park, B. S.; Lee, J.-O. Recognition of lipopolysaccharide pattern by TLR4 complexes. *Experimental & Molecular Medicine* **2013**, *45* (12), e66-e66. DOI: 10.1038/emm.2013.97.
(9) Armant, M. A.; Fenton, M. J. Toll-like receptors: a family of pattern-recognition receptors in mammals. (1474-760X (Electronic)). From 2002 Jul 29.
(10) Ramamoorthy, A.; Thennarasu, S.; Lee, D.-K.; Tan, A.; Maloy, L. Solid-state NMR investigation of the membrane-disrupting mechanism of antimicrobial peptides MSI-78 and MSI-594 derived from magainin 2 and melittin. *Biophysical journal* **2006**, *91* (1), 206-216.
(11) Qin, S.; Pande, A. H.; Nemec, K. N.; Tatulian, S. A. The N-terminal α-helix of pancreatic phospholipase A2 determines productive-mode orientation of the enzyme at the membrane surface. *Journal of molecular biology* **2004**, *344* (1), 71-89.
(12) Maveyraud, L.; Mourey, L. Protein X-ray crystallography and drug discovery. *Molecules* **2020**, *25* (5), 1030.
(13) Woys, A. M.; Lin, Y.-S.; Reddy, A. S.; Xiong, W.; de Pablo, J. J.; Skinner, J. L.; Zanni, M. T. 2D IR line shapes probe ovispirin peptide conformation and depth in lipid bilayers. *Journal of the American Chemical Society* **2010**, *132* (8), 2832-2838.
(14) Lotti, D.; Hamm, P.; Kraack, J. P. Surface-sensitive spectro-electrochemistry using ultrafast 2D ATR IR spectroscopy. *The Journal of Physical Chemistry C* **2016**, *120* (5), 2883-2892.
(15) Yuan, Y.; Zhu, Q.; Song, R.; Ma, J.; Dong, H. A two-ended data-driven accelerated sampling method for exploring the transition pathways between two known states of protein. *Journal of Chemical Theory and Computation* **2020**, *16* (7), 4631-4640.
(16) Bastos, M.; Abian, O.; Johnson, C. M.; Ferreira-da-Silva, F.; Vega, S.; Jimenez-Alesanco, A.; Ortega-Alarcon, D.; Velazquez-Campoy, A. Isothermal titration calorimetry. *Nature Reviews Methods Primers* **2023**, *3* (1), 17.
(17) Cino, E. A.; Choy, W.-Y.; Karttunen, M. Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations. *Journal of chemical theory and computation* **2012**, *8* (8), 2725-2740.
(18) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ1 and χ2 Dihedral Angles. *Journal of Chemical Theory and Computation* **2012**, *8* (9), 3257-3273. DOI: 10.1021/ct300400x.
(19) Matsubara, D.; Kasahara, K.; Dokainish, H. M.; Oshima, H.; Sugita, Y. Modified Protein-Water Interactions in CHARMM36m for Thermodynamics and Kinetics of Proteins in Dilute and Crowded Solutions. *Molecules* **2022**, *27* (17), 5726.
(20) Yu, Y. Parameterization of the CHARMM Lipid Force Field and Applications to Membrane Modeling. University of Maryland, College Park, 2022.

(21) Yu, Y.; Kramer, A.; Venable, R. M.; Brooks, B. R.; Klauda, J. B.; Pastor, R. W. CHARMM36 lipid force field with explicit treatment of long-range dispersion: parametrization and validation for phosphatidylethanolamine, phosphatidylglycerol, and ether lipids. *Journal of chemical theory and computation* **2021**, *17* (3), 1581-1595.
(22) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26* (16), 1781-1802. DOI: 10.1002/jcc.20289 (acccessed 2019/07/18).
(23) Phillips, J. C.; Hardy, D. J.; Maia, J. D. C.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Henin, J.; Jiang, W.; et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J Chem Phys* **2020**, *153* (4), 044130. DOI: 10.1063/5.0014475.
(24) Stevenson, P. Membrane and Membrane Protein Dynamics Studied with Time-Resolved Infrared Spectroscopy. Massachusetts Institute of Technology, Department of Chemistry, 2017.
(25) Chand, S. Investigating Membrane Fluidity in Two-dimensions. University of Sheffield, 2020.
(26) O'Brien, E. S.; Fuglestad, B.; Lessen, H. J.; Stetz, M. A.; Lin, D. W.; Marques, B. S.; Gupta, K.; Fleming, K. G.; Wand, A. J. Membrane proteins have distinct fast internal motion and residual conformational entropy. *Angewandte Chemie International Edition* **2020**, *59* (27), 11108-11114.
(27) Pott, T.; Méléard, P. The dynamics of vesicle thermal fluctuations is controlled by intermonolayer friction. *Europhysics letters* **2002**, *59* (1), 87.
(28) Ashkar, R.; Nagao, M.; Butler, P. D.; Woodka, A. C.; Sen, M. K.; Koga, T. Tuning membrane thickness fluctuations in model lipid bilayers. *Biophysical journal* **2015**, *109* (1), 106-112.
(29) Chaban, V. Computationally efficient prediction of area per lipid. *Chemical Physics Letters* **2014**, *616*, 25-29.
(30) Khakbaz, P.; Klauda, J. B. Probing the importance of lipid diversity in cell membranes via molecular simulation. *Chemistry and physics of lipids* **2015**, *192*, 12-22.
(31) Allsopp, R.; Pavlova, A.; Cline, T.; Salyapongse, A. M.; Gillilan, R. E.; Di, Y. P.; Deslouches, B.; Klauda, J. B.; Gumbart, J. C.; Tristram-Nagle, S. Antimicrobial Peptide Mechanism Studied by Scattering-Guided Molecular Dynamics Simulation. *The Journal of Physical Chemistry B* **2022**, *126* (36), 6922-6935.
(32) Elbaz, Y.; Schuldiner, M. Staying in touch: the molecular era of organelle contact sites. *Trends in biochemical sciences* **2011**, *36* (11), 616-623.
(33) Boettcher, J. M.; Davis-Harrison, R. L.; Clay, M. C.; Nieuwkoop, A. J.; Ohkubo, Y. Z.; Tajkhorshid, E.; Morrissey, J. H.; Rienstra, C. M. Atomic view of calcium-induced clustering of phosphatidylserine in mixed lipid bilayers. *Biochemistry* **2011**, *50* (12), 2264-2273.
(34) Wildermuth, K. D.; Monje-Galvan, V.; Warburton, L. M.; Klauda, J. B. Effect of Membrane Lipid Packing on Stable Binding of the ALPS Peptide. *Journal of Chemical Theory and Computation* **2019**, *15* (2), 1418-1429. DOI: 10.1021/acs.jctc.8b00945.
(35) Levitan, I.; Fang, Y.; Rosenhouse-Dantsker, A.; Romanenko, V. Cholesterol and Ion Channels. In *Cholesterol Binding and Cholesterol Transport Proteins: Structure and Function in Health and Disease*, Harris, J. R. Ed.; Springer Netherlands, 2010; pp 509-549.
(36) Barrantes, F. Structural basis for lipid modulation of nicotinic acetylcholine receptor function. *Brain Research Reviews* **2004**, *47* (1-3), 71-95.
(37) Casares, D.; Escribá, P. V.; Rosselló, C. A. Membrane lipid composition: effect on membrane and organelle structure, function and compartmentalization and therapeutic avenues. *International journal of molecular sciences* **2019**, *20* (9), 2167.
(38) Schulz, T. A.; Prinz, W. A. Sterol transport in yeast and the oxysterol binding protein homologue (OSH) family. *BBA - Molecular and Cell Biology of Lipids* **2007**, *1771* (6), 769-780. DOI: 10.1016/j.bbalip.2007.03.003 WorldCat.org.
(39) Beh, C. T.; Rine, J. A role for yeast oxysterol-binding protein homologs in endocytosis and in the maintenance of intracellular sterol-lipid distribution. *Journal of Cell Science* **2004**, *117* (14), 2983. DOI: 10.1242/jcs.01157.
(40) Im, Y. J.; Raychaudhuri, S.; Prinz, W. A.; Hurley, J. H. Structural mechanism for sterol sensing and transport by OSBP-related proteins. *Nature* **2005**, *437* (7055), 154-158. DOI: 10.1038/nature03923.
(41) Glebov, O. O. Understanding SARS-CoV-2 endocytosis for COVID-19 drug repurposing. *The FEBS journal* **2020**, *287* (17), 3664-3671.
(42) Quon, E.; Sere, Y. Y.; Chauhan, N.; Johansen, J.; Sullivan, D. P.; Dittman, J. S.; Rice, W. J.; Chan, R. B.; Di Paolo, G.; Beh, C. T.; et al. Endoplasmic reticulum-plasma membrane contact sites integrate sterol and phospholipid regulation. *PLOS Biology* **2018**, *16* (5), e2003864. DOI: 10.1371/journal.pbio.2003864.
(43) Gong, J.; Chen, Y.; Pu, F.; Sun, P.; He, F.; Zhang, L.; Li, Y.; Ma, Z.; Wang, H. Understanding membrane protein drug targets in computational perspective. *Current Drug Targets* **2019**, *20* (5), 551-564.

(44) Oyagi, A.; Morimoto, N.; Hamanaka, J.; Ishiguro, M.; Tsuruma, K.; Shimazawa, M.; Hara, H. Forebrain specific heparin-binding epidermal growth factor-like growth factor knockout mice show exacerbated ischemia and reperfusion injury. *Neuroscience* **2011**, *185*, 116-124.
(45) A single channel mutation alters agonist efficacy at 5HT3A and 5HT3AB receptors. *British Journal of Pharmacology* **2013**, *170* (2), 391-402. From http://worldcat.org /z-wcorg/.
(46) Monje-Galvan, V.; Klauda, J. B. Peripheral membrane proteins: Tying the knot between experiment and computation. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2016**, *1858* (7, Part B), 1584-1593. DOI: https://doi.org/10.1016/j.bbamem.2016.02.018.
(47) Monje-Galvan, V.; Klauda, J. B. Preferred Binding Mechanism of Osh4's Amphipathic Lipid-Packing Sensor Motif, Insights from Molecular Dynamics. *The Journal of Physical Chemistry B* **2018**, *122* (42), 9713-9723. DOI: 10.1021/acs.jpcb.8b07067.
(48) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Davila-Contreras, E. M.; Qi, Y. F.; Lee, J. M.; Monje-Galvan, V.; Venable, R. M.; et al. CHARMM-GUI membrane builder toward realistic biological membrane simulations. *Journal of computational chemistry* **2014**, *35* (27), 1997-2004. DOI: 10.1002/jcc.23702.
(49) Allsopp, R. J.; Klauda, J. B. Impact of PIP2 Lipids, Force Field Parameters, and Mutational Analysis on the Binding of the Osh4's α6–α7 Domain. *The Journal of Physical Chemistry B* **2021**, *125* (20), 5296-5308. DOI: 10.1021/acs.jpcb.0c10393.
(50) England, L. J.; Imperial, J.; Jacobsen, R.; Craig, A. G.; Gulyas, J.; Akhtar, M.; Rivier, J.; Julius, D.; Olivera, B. M. Inactivation of a Serotonin-Gated Ion Channel by a Polypeptide Toxin from Marine Snails. *Science* **1998**, *281* (5376), 575-578. WorldCat.org.
(51) Farias, L.; Smith, E.; Markov, A. Prevention of ischemic-hypoxic brain injury and death in rabbits with fructose-1, 6-diphosphate. *Stroke* **1990**, *21* (4), 606-613.
(52) Trimarchi, G. R.; Arcadi, F. A.; De Luca, R.; Imperatore, C.; Santoro, G.; Trimarchi, F.; Costa, G. Neuroprotective activity of fructose-1, 6-bisphosphate following transient forebrain ischemia in the Mongolian gerbil. *The Japanese Journal of Pharmacology* **1993**, *62* (3), 215-222.
(53) Singh, R. P.; Brooks, B. R.; Klauda, J. B. Binding and release of cholesterol in the Osh4 protein of yeast. *Proteins: Structure, Function, and Bioinformatics* **2009**, *75* (2), 468-477. DOI: 10.1002/prot.22263.
(54) Schulz, T. A.; Choi, M.-G.; Raychaudhuri, S.; Mears, J. A.; Ghirlando, R.; Hinshaw, J. E.; Prinz, W. A. Lipid-regulated sterol transfer between closely apposed membranes by oxysterol-binding protein homologues. *The Journal of Cell Biology* **2009**, *187* (6), 889-903. DOI: 10.1083/jcb.200905007.
(55) Tian, S.; Ohta, A.; Horiuchi, H.; Fukuda, R. Oxysterol-binding protein homologs mediate sterol transport from the endoplasmic reticulum to mitochondria in yeast. *Journal of Biological Chemistry* **2018**, *293* (15), 5636-5648.
(56) Funato, K.; Riezman, H. Vesicular and nonvesicular transport of ceramide from ER to the Golgi apparatus in yeast. *The Journal of Cell Biology* **2001**, *155* (6), 949-960. DOI: 10.1083/jcb.200105033 WorldCat.org.
(57) Bigay, J.; Antonny, B. Curvature, Lipid Packing, and Electrostatics of Membrane Organelles: Defining Cellular Territories in Determining Specificity. *Developmental Cell* **2012**, *23* (5), 886-895. DOI: https://doi.org/10.1016/j.devcel.2012.10.009.
(58) Drin, G.; Jf, C.; R, G.; T, B.; Tu, S.; Antonny, B. A general amphipathic alpha-helical motif for sensing membrane curvature. *Nature Structural and Molecular Biology* **2007**, (1545-9993 (Print)). From 2007 Feb.
(59) Antonny, B. Mechanisms of Membrane Curvature Sensing. *Annual Review of Biochemistry* **2011**, *80* (1), 101-123. DOI: 10.1146/annurev-biochem-052809-155121 (acccessed 2022/01/26).
(60) Raychaudhuri, S.; Im, Y. J.; Hurley, J. H.; Prinz, W. A. Nonvesicular sterol movement from plasma membrane to ER requires oxysterol-binding protein–related proteins and phosphoinositides. *The Journal of Cell Biology* **2006**, *173* (1), 107. DOI: 10.1083/jcb.200510084.
(61) Kentala, H.; Weber-Boyvat, M.; Olkkonen, V. M. OSBP-Related Protein Family: Mediators of Lipid Transport and Signaling at Membrane Contact Sites. *International review of cell and molecular biology* **2016**, *321*, 299-340. DOI: 10.1016/bs.ircmb.2015.09.006 WorldCat.org.
(62) Deslouches, B.; Islam, K.; Craigo, J. K.; Paranjape, S. M.; Montelaro, R. C.; Mietzner, T. A. Activity of the de novo engineered antimicrobial peptide WLBU2 against Pseudomonas aeruginosa in human serum and whole blood: implications for systemic applications. *Antimicrobial agents and chemotherapy* **2005**, *49* (8), 3208-3216. DOI: 10.1128/AAC.49.8.3208-3216.2005 PubMed.
(63) Van Oort, C. M.; Ferrell, J. B.; Remington, J. M.; Wshah, S.; Li, J. AMPGAN v2: Machine Learning-Guided Design of Antimicrobial Peptides. *Journal of Chemical Information and Modeling* **2021**, *61* (5), 2198-2207. DOI: 10.1021/acs.jcim.0c01441.

(64) Tencza, S. B.; Creighton, D. J.; Yuan, T.; Vogel, H. J.; Montelaro, R. C.; Mietzner, T. A. Lentivirus-derived antimicrobial peptides: increased potency by sequence engineering and dimerization. *Journal of Antimicrobial Chemotherapy* **1999**, *44* (1), 33-41. DOI: 10.1093/jac/44.1.33 (acccessed 9/20/2020).
(65) Heinrich, F.; Salyapongse, A.; Kumagai, A.; Dupuy, F. G.; Shukla, K.; Penk, A.; Huster, D.; Ernst, R. K.; Pavlova, A.; Gumbart, J. C.; et al. Synergistic Biophysical Techniques Reveal Structural Mechanisms of Engineered Cationic Antimicrobial Peptides in Lipid Model Membranes. *Chemistry (Weinheim an der Bergstrasse, Germany)* **2020**, *26* (28), 6247-6256. DOI: 10.1002/chem.202000212 PubMed.
(66) Reeves, D. C.; Lummis, S. C. The molecular basis of the structure and function of the 5-HT 3 receptor: a model ligand-gated ion channel. *Molecular membrane biology* **2002**, *19* (1), 11-26.
(67) Wu, Z.-s.; Cheng, H.; Jiang, Y.; Melcher, K.; Xu, H. E. Ion channels gated by acetylcholine and serotonin: structures, biology, and drug discovery. *Acta Pharmacologica Sinica* **2015**, *36* (8), 895-907.
(68) Kalamida, D.; Poulas, K.; Avramopoulou, V.; Fostieri, E.; Lagoumintzis, G.; Lazaridis, K.; Sideri, A.; Zouridakis, M.; Tzartos, S. J. Muscle and neuronal nicotinic acetylcholine receptors: structure, function and pathogenicity. *The FEBS journal* **2007**, *274* (15), 3799-3845.
(69) Gripp, K. W.; Smithson, S. F.; Scurr, I. J.; Baptista, J.; Majumdar, A.; Pierre, G.; Williams, M.; Henderson, L. B.; Wentzensen, I. M.; McLaughlin, H. Syndromic disorders caused by gain-of-function variants in KCNH1, KCNK4, and KCNN3—a subgroup of K+ channelopathies. *European Journal of Human Genetics* **2021**, *29* (9), 1384-1395.
(70) Guzel, T.; Mirowska-Guzel, D. The role of serotonin neurotransmission in gastrointestinal tract and pharmacotherapy. *Molecules* **2022**, *27* (5), 1680.
(71) Kaupmann, K.; Huggel, K.; Heid, J.; Flor, P. J.; Bischoff, S.; Mickel, S. J.; McMaster, G.; Angst, C.; Bittiger, H.; Froestl, W. Expression cloning of GABAB receptors uncovers similarity to metabotropic glutamate receptors. *Nature* **1997**, *386* (6622), 239-246.
(72) Terry, N.; Margolis, K. G. Serotonergic mechanisms regulating the GI tract: experimental evidence and therapeutic relevance. *Gastrointestinal Pharmacology* **2017**, 319-342.
(73) Smith, H. S.; Cox, L. R.; Smith, E. J. 5-HT3 receptor antagonists for the treatment of nausea/vomiting. *Ann Palliat Med* **2012**, *1* (2), 115-120.
(74) Kato, M.; Fukuda, T.; Wakeno, M.; Fukuda, K.; Okugawa, G.; Ikenaga, Y.; Yamashita, M.; Takekita, Y.; Nobuhara, K.; Azuma, J. Effects of the serotonin type 2A, 3A and 3B receptor and the serotonin transporter genes on paroxetine and fluvoxamine efficacy and adverse drug reactions in depressed Japanese patients. *Neuropsychobiology* **2006**, *53* (4), 186-195.
(75) Connolly, C. N.; Wafford, K. The Cys-loop superfamily of ligand-gated ion channels: the impact of receptor structure on function. *Biochemical Society Transactions* **2004**, *32* (3), 529-534.
(76) Franco, N.; Franco, R. Understanding the added value of g-protein-coupled receptor heteromers. *Scientifica* **2014**, *2014*.
(77) Morales Duque, H.; Campos Dias, S.; Franco, O. L. Structural and functional analyses of cone snail toxins. *Marine drugs* **2019**, *17* (6), 370.
(78) Olivera, B. M.; Gray, W. R.; Zeikus, R.; McIntosh, J. M.; Varga, J.; Rivier, J.; De Santos, V.; Cruz, L. J. Peptide neurotoxins from fish-hunting cone snails. *Science* **1985**, *230* (4732), 1338-1343.
(79) Lee, S.; Kim, Y.; Back, S. K.; Choi, H.-W.; Lee, J. Y.; Jung, H. H.; Ryu, J. H.; Suh, H.-W.; Na, H. S.; Kim, H. J. Analgesic effect of highly reversible ω-conotoxin FVIA on N type Ca2+ channels. *Molecular pain* **2010**, *6*, 1744-8069-1746-1797.
(80) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Tunyasuvunakool, K.; Ronneberger, O.; Bates, R.; Žídek, A.; Bridgland, A. AlphaFold 2. *In Fourteenth Critical Assessment of Techniques for Protein Structure Prediction (Abstract Book* **2020**.
(81) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583-589. DOI: 10.1038/s41586-021-03819-2.
(82) Zhang, H.; Wang, L.; Yang, X.; Lian, Z.; Qiu, Y.; Dong, Z.; Wu, X.; Pan, X. Identification of Novel Conopeptides and Distinct Gene Superfamilies in the Marine Cone Snail Conus quercinus. *Frontiers in Marine Science* **2021**, *8*, Original Research. DOI: 10.3389/fmars.2021.766792.
(83) Papke, R. L.; Brunzell, D. H.; De Biasi, M. Cholinergic Receptors and Addiction. (1866-3370 (Print)). From 2020.
(84) Abraham, N.; Healy, M.; Ragnarsson, L.; Brust, A.; Alewood, P. F.; Lewis, R. J. Structural mechanisms for α-conotoxin activity at the human α3β4 nicotinic acetylcholine receptor. *Scientific Reports* **2017**, *7* (1), 45466. DOI: 10.1038/srep45466.

(85) Platt, J. L.; Vercellotti, G. M.; Lindman, B. J.; Oegema Jr, T.; Bach, F. H.; Dalmasso, A. P. Release of heparan sulfate from endothelial cells. Implications for pathogenesis of hyperacute rejection. *The Journal of experimental medicine* **1990**, *171* (4), 1363-1368.
(86) Gesuete, R.; Kohama, S. G.; Stenzel-Poore, M. P. Toll-like receptors and ischemic brain injury. *Journal of Neuropathology & Experimental Neurology* **2014**, *73* (5), 378-386.
(87) Bae, M.; Cassilly, C. D.; Liu, X.; Park, S.-M.; Tusi, B. K.; Chen, X.; Kwon, J.; Filipčík, P.; Bolze, A. S.; Liu, Z. Akkermansia muciniphila phospholipid induces homeostatic immune responses. *Nature* **2022**, *608* (7921), 168-173.
(88) Wang, Q.; McLoughlin, R. M.; Cobb, B. A.; Charrel-Dennis, M.; Zaleski, K. J.; Golenbock, D.; Tzianabos, A. O.; Kasper, D. L. A bacterial carbohydrate links innate and adaptive responses through Toll-like receptor 2. *The Journal of experimental medicine* **2006**, *203* (13), 2853-2863.
(89) Matsumoto, M.; Seya, T. TLR3: interferon induction by double-stranded RNA including poly (I: C). *Advanced drug delivery reviews* **2008**, *60* (7), 805-812.
(90) Li, J.; Shworak, N. W.; Simons, M. Increased responsiveness of hypoxic endothelial cells to FGF2 is mediated by HIF-1α-dependent regulation of enzymes involved in synthesis of heparan sulfate FGF2-binding sites. *Journal of cell science* **2002**, *115* (9), 1951-1959.
(91) Tang, S.-C.; Arumugam, T. V.; Xu, X.; Cheng, A.; Mughal, M. R.; Jo, D. G.; Lathia, J. D.; Siler, D. A.; Chigurupati, S.; Ouyang, X. Pivotal role for neuronal Toll-like receptors in ischemic brain injury and functional deficits. *Proceedings of the National Academy of Sciences* **2007**, *104* (34), 13798-13803.
(92) Wang, X.; Stridh, L.; Li, W.; Dean, J.; Elmgren, A.; Gan, L.; Eriksson, K.; Hagberg, H.; Mallard, C. Lipopolysaccharide sensitizes neonatal hypoxic-ischemic brain injury in a MyD88-dependent manner. *The Journal of Immunology* **2009**, *183* (11), 7471-7477.
(93) Li, S.; Marks, J. A.; Eisenstadt, R.; Kumasaka, K.; Samadi, D.; Johnson, V. E.; Holena, D. N.; Allen, S. R.; Browne, K. D.; Smith, D. H. Enoxaparin ameliorates post–traumatic brain injury edema and neurologic recovery, reducing cerebral leukocyte endothelial interactions and vessel permeability in vivo. *The journal of trauma and acute care surgery* **2015**, *79* (1), 78.
(94) Stutzmann, J. M.; Mary, V.; Wahl, F.; Grosjean-Piot, O.; Uzan, A.; Pratt, J. Neuroprotective profile of enoxaparin, a low molecular weight heparin, in in vivo models of cerebral ischemia or traumatic brain injury in rats: a review. *CNS drug reviews* **2002**, *8* (1), 1-30.
(95) Bickler, P.; Kelleher, J. A. Fructose-1, 6-bisphosphate stabilizes brain intracellular calcium during hypoxia in rats. *Stroke* **1992**, *23* (11), 1617-1622.
(96) Donohoe, P.; Fahlman, C.; Bickler, P.; Vexler, Z.; Gregory, G. Neuroprotection and intracellular Ca2+ modulation with fructose-1, 6-bisphosphate during in vitro hypoxia–ischemia involves phospholipase C-dependent signaling. *Brain research* **2001**, *917* (2), 158-166.
(97) de la Torre, J. C. Treatment of head injury in mice, using a fructose 1, 6-diphosphate and dimethyl sulfoxide combination. *Neurosurgery* **1995**, *37* (2), 273-279.
(98) Rønn, L. C.; Doherty, P.; Holm, A.; Berezin, V.; Bock, E. Neurite outgrowth induced by a synthetic peptide ligand of neural cell adhesion molecule requires fibroblast growth factor receptor activation. *Journal of neurochemistry* **2000**, *75* (2), 665-671.
(99) Ardizzone, A.; Scuderi, S. A.; Giuffrida, D.; Colarossi, C.; Puglisi, C.; Campolo, M.; Cuzzocrea, S.; Esposito, E.; Paterniti, I. Role of fibroblast growth factors receptors (FGFRs) in brain tumors, focus on astrocytoma and glioblastoma. *Cancers* **2020**, *12* (12), 3825.
(100) Poelje, P. D. v.; Potter, S. C.; Erion, M. D. Fructose-1, 6-bisphosphatase inhibitors for reducing excessive endogenous glucose production in type 2 diabetes. *Diabetes-Perspectives in Drug Therapy* **2011**, 279-301.
(101) Gawarammana, I.; Mohamed, F.; Bowe, S. J.; Rathnathilake, A.; Narangoda, S. K.; Azher, S.; Dawson, A. H.; Buckley, N. A. Fructose-1, 6-diphosphate (FDP) as a novel antidote for yellow oleander-induced cardiac toxicity: a randomized controlled double blind study. *BMC Emergency Medicine* **2010**, *10* (1), 1-6.
(102) Kouider, S.; Kolb, F.; Müller, I.; Pfüller, K. Behavior of various blood constituents (glucose, fructose, insulin, lactate, pyruvate, free fatty acids, inorganic phosphate) and half-life of monosaccharides in the plasma after iv infusion of glucose, fructose, galactose and invert sugar solutions in ruminants. 2. Studies in cattle. *Archiv fur Experimentelle Veterinarmedizin* **1978**, *32* (5), 699-714.
(103) Kunath, T.; Saba-El-Leil, M. K.; Almousailleakh, M.; Wray, J.; Meloche, S.; Smith, A. FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. **2007**.
(104) Mishra, O. P.; Delivoria-Papadopoulos, M. Inositol tetrakisphosphate (IP4)-and inositol triphosphate (IP3)-dependent Ca2+ influx in cortical neuronal nuclei of newborn piglets following graded hypoxia. *Neurochemical research* **2004**, *29* (2), 391-396.

(105) Ain, Q. u.; Batool, M.; Choi, S. TLR4-targeting therapeutics: structural basis and computer-aided drug discovery approaches. *Molecules* **2020**, *25* (3), 627.
(106) Ma, Q.; Dudas, B.; Hejna, M.; Cornelli, U.; Lee, J. M.; Lorens, S.; Mervis, R.; Hanin, I.; Fareed, J. The blood–brain barrier accessibility of a heparin-derived oligosaccharides C3. *Thrombosis research* **2002**, *105* (5), 447-453.
(107) Weinberg, S. E.; Sun, L. Y.; Yang, A. L.; Liao, J.; Yang, G. Y. Overview of inositol and inositol phosphates on chemoprevention of colitis-induced carcinogenesis. *Molecules* **2020**, *26* (1), 31.
(108) Morrison, R. S.; Shi, E.; Kan, M.; Yamaguchi, F.; McKeehan, W.; Rudnicka-Nawrot, M.; Palczewski, K. Inositolhexakisphosphate (InsP6): An antagonist of fibroblast growth factor receptor binding and activity. *In Vitro Cellular & Developmental Biology - Animal* **1994**, *30* (11), 783-789. DOI: 10.1007/BF02631302 WorldCat.org.
(109) Del Tatto, V. A Fully Anisotropic Formulation of Stochastic Cell Rescaling. *arXiv preprint arXiv:2111.06402* **2021**.
(110) Rosenbluth, M. N.; Rosenbluth, A. W. Further results on Monte Carlo equations of state. *The Journal of Chemical Physics* **1954**, *22* (5), 881-884.
(111) MacKerell Jr, A. D.; Bashford, D.; Bellott, M.; Dunbrack Jr, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B* **1998**, *102* (18), 3586-3616.
(112) Allen, M. P.; Tildesley, D. J. *Computer simulation of liquids*; Oxford university press, 2017.
(113) Yoo, J.; Aksimentiev, A. New tricks for old dogs: improving the accuracy of biomolecular force fields by pair-specific corrections to non-bonded interactions. *Physical Chemistry Chemical Physics* **2018**, *20* (13), 8432-8449, 10.1039/C7CP08185E. DOI: 10.1039/C7CP08185E.
(114) Zhang, Y.; Feller, S. E.; Brooks, B. R.; Pastor, R. W. Computer simulation of liquid/liquid interfaces. I. Theory and application to octane/water. *The Journal of chemical physics* **1995**, *103* (23), 10252-10266.
(115) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *The Journal of Chemical Physics* **1995**, *103* (11), 4613-4621. DOI: 10.1063/1.470648 (acccessed 2019/07/18).
(116) Martyna, G. J.; Tobias, D. J.; Klein, M. L. Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics* **1994**, *101* (5), 4177-4189. DOI: 10.1063/1.467468 (acccessed 2019/07/18).
(117) Parrinello, M.; Rahman, A. Crystal structure and pair potentials: A molecular-dynamics study. *Physical review letters* **1980**, *45* (14), 1196.
(118) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied physics* **1981**, *52* (12), 7182-7190.
(119) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet algorithm for efficient molecular dynamics simulations with long-range interactions. *Molecular Simulation* **1991**, *6* (1-3), 121-142.
(120) Hut, P.; Makino, J.; McMillan, S. Building a better leapfrog. *Astrophysical Journal, Part 2-Letters (ISSN 0004-637X), vol. 443, no. 2, p. L93-L96* **1995**, *443*, L93-L96.
(121) Yeh, I.-C.; Hummer, G. System-size dependence of diffusion coefficients and viscosities from molecular dynamics simulations with periodic boundary conditions. *The Journal of Physical Chemistry B* **2004**, *108* (40), 15873-15879.
(122) Piggot, T. J.; Pineiro, A.; Khalid, S. Molecular dynamics simulations of phosphatidylcholine membranes: a comparative force field study. *Journal of chemical theory and computation* **2012**, *8* (11), 4593-4609.
(123) Patra, M.; Karttunen, M.; Hyvönen, M. T.; Falck, E.; Lindqvist, P.; Vattulainen, I. Molecular dynamics simulations of lipid bilayers: major artifacts due to truncating electrostatic interactions. *Biophysical journal* **2003**, *84* (6), 3636-3645.
(124) Leonard, A. N.; Simmonett, A. C.; Pickard IV, F. C.; Huang, J.; Venable, R. M.; Klauda, J. B.; Brooks, B. R.; Pastor, R. W. Comparison of additive and polarizable models with explicit treatment of long-range Lennard-Jones interactions using alkane simulations. *Journal of chemical theory and computation* **2018**, *14* (2), 948-958.
(125) Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of chemical physics* **1980**, *72* (4), 2384-2393.
(126) Berendsen, H. J.; Postma, J. v.; Van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics* **1984**, *81* (8), 3684-3690.
(127) Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics* **1984**, *52* (2), 255-268.
(128) Nosé, S. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics* **1984**, *81* (1), 511-519.
(129) Hoover, W. G. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A* **1985**, *31* (3), 1695-1697. DOI: 10.1103/PhysRevA.31.1695.

(130) Rogge, S.; Vanduyfhuys, L.; Ghysels, A.; Waroquier, M.; Verstraelen, T.; Maurin, G.; Van Speybroeck, V. A comparison of barostats for the mechanical characterization of metal–organic frameworks. *Journal of chemical theory and computation* **2015**, *11* (12), 5583-5597.
(131) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmuller, H.; MacKerell, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* **2017**, *14* (1), 71-73. DOI: 10.1038/Nmeth.4067.
(132) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM all-atom additive force field for lipids: Validation on six lipid types. *Journal of Physical Chemistry B* **2010**, *114* (23), 7830-7843. DOI: 10.1021/jp101759q.
(133) Venable, R. M.; Sodt, A. J.; Rogaski, B.; Rui, H.; Hatcher, E.; MacKerell, A. D.; Pastor, R. W.; Klauda, J. B. CHARMM all-atom additive force field for sphingomyelin: elucidation of hydrogen bonding and of positive curvature. *Biophysical journal* **2014**, *107* (1), 134-145.
(134) Guvench, O.; Hatcher, E.; Venable, R. M.; Pastor, R. W.; MacKerell Jr, A. D. CHARMM additive all-atom force field for glycosidic linkages between hexopyranoses. *Journal of chemical theory and computation* **2009**, *5* (9), 2353-2370.
(135) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics* **1977**, *23* (3), 327-341. DOI: https://doi.org/10.1016/0021-9991(77)90098-5.
(136) Andersen, H. C. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of computational Physics* **1983**, *52* (1), 24-34.
(137) Vanommeslaeghe, K.; MacKerell Jr, A. D. Automation of the CHARMM General Force Field (CGenFF) I: bond perception and atom typing. *Journal of chemical information and modeling* **2012**, *52* (12), 3144-3154.
(138) Vanommeslaeghe, K.; Raman, E. P.; MacKerell Jr, A. D. Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *Journal of chemical information and modeling* **2012**, *52* (12), 3155-3168.
(139) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98* (12), 10089-10092. DOI: 10.1063/1.464397 (acccessed 2019/07/18).
(140) Steinbach, P. J.; Brooks, B. R. New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry* **1994**, *15* (7), 667-683. DOI: 10.1002/jcc.540150702 (acccessed 2019/07/18).
(141) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14* (1), 33-38.
(142) Lee, J.; Cheng, X.; Swails, J. M.; Yeom, M. S.; Eastman, P. K.; Lemkul, J. A.; Wei, S.; Buckner, J.; Jeong, J. C.; Qi, Y.; et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation* **2016**, *12* (1), 405-413. DOI: 10.1021/acs.jctc.5b00935.
(143) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79* (2), 926-935. DOI: 10.1063/1.445869 (acccessed 2019/07/18).
(144) Jo, S.; Kim, T.; Im, W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. *PLOS ONE* **2007**, *2* (9), e880. DOI: 10.1371/journal.pone.0000880.
(145) Qi, Y.; Cheng, X.; Lee, J.; Vermaas, J. V.; Pogorelov, T. V.; Tajkhorshid, E.; Park, S.; Klauda, J. B.; Im, W. CHARMM-GUI HMMM Builder for Membrane Simulations with the Highly Mobile Membrane-Mimetic Model. (1542-0086 (Electronic)). From 2015 Nov 17.
(146) Lee, J.; Patel, D. S.; Ståhle, J.; Park, S.-J.; Kern, N. R.; Kim, S.; Lee, J.; Cheng, X.; Valvano, M. A.; Holst, O.; et al. CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans. *Journal of Chemical Theory and Computation* **2019**, *15* (1), 775-786. DOI: 10.1021/acs.jctc.8b01066.
(147) Jo, S.; Lim, J. B.; Klauda, J. B.; Im, W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. *Biophysical Journal* **2009**, *97* (1), 50-58. DOI: https://doi.org/10.1016/j.bpj.2009.04.013.
(148) Jo, S.; Kim, T.; Iyer, V. G.; Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *Journal of Computational Chemistry* **2008**, *29* (11), 1859-1865. DOI: 10.1002/jcc.20945.
(149) Ohkubo, Y. Z.; Tv, P.; Mj, A.; Ga, C.; Tajkhorshid, E. Accelerating membrane insertion of peripheral proteins with a novel membrane mimetic model. *Biophys, J.* (1542-0086 (Electronic)). From 2012 May 2.

(150) Grossfield, A. An Implementation of WHAM: The Weighted Histogram Analysis Method, version 2.0. 9. 18. *Google Scholar There is no corresponding record for this reference* **2014**.
(151) Lev, S. Non-vesicular lipid transport by lipid-transfer proteins and beyond. *Nature Reviews Molecular Cell Biology* **2010**, *11* (10), 739-750. DOI: 10.1038/nrm2971.
(152) De Matteis, M. A.; Di Campli, A.; D'Angelo, G. Lipid-transfer proteins in membrane trafficking at the Golgi complex. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **2007**, *1771* (6), 761-768. DOI: https://doi.org/10.1016/j.bbalip.2007.04.004.
(153) Cockcroft, S. Mammalian phosphatidylinositol transfer proteins: emerging roles in signal transduction and vesicular traffic. *Chemistry and Physics of Lipids* **1999**, *98* (1), 23-33. DOI: https://doi.org/10.1016/S0009-3084(99)00015-8.
(154) Holthuis, J. C. M.; Levine, T. P. Lipid traffic: floppy drives and a superhighway. *Nature Reviews Molecular Cell Biology* **2005**, *6* (3), 209-220. DOI: 10.1038/nrm1591.
(155) Levine, T. Short-range intracellular trafficking of small molecules across endoplasmic reticulum junctions. *Trends in Cell Biology* **2004**, *14* (9), 483-490. DOI: https://doi.org/10.1016/j.tcb.2004.07.017.
(156) Kawano, M.; Kumagai, K.; Nishijima, M.; Hanada, K. Efficient Trafficking of Ceramide from the Endoplasmic Reticulum to the Golgi Apparatus Requires a VAMP-associated Protein-interacting FFAT Motif of CERT*. *Journal of Biological Chemistry* **2006**, *281* (40), 30279-30288. DOI: https://doi.org/10.1074/jbc.M605032200.
(157) Hanada, K. Discovery of the molecular machinery CERT for endoplasmic reticulum-to-Golgi trafficking of ceramide. *Molecular and Cellular Biochemistry* **2006**, *286* (1), 23-31. DOI: 10.1007/s11010-005-9044-z.
(158) Lemmon, M. A. Pleckstrin homology (PH) domains and phosphoinositides. *Biochemical Society Symposium* **2007**, *74*, 81-93. DOI: 10.1042/BSS2007c08.
(159) Li, X.; Rivas, M. P.; Fang, M.; Marchena, J.; Mehrotra, B.; Chaudhary, A.; Feng, L.; Prestwich, G. D.; Bankaitis, V. A. Analysis of oxysterol binding protein homologue Kes1p function in regulation of Sec14p-dependent protein transport from the yeast Golgi complex. *The Journal of Cell Biology* **2002**, *157* (1), 63. DOI: 10.1083/jcb.200201037.
(160) de Saint-Jean, M.; Delfosse, V.; Douguet, D.; Chicanne, G.; Payrastre, B.; Bourguet, W.; Antonny, B.; Drin, G. Osh4p exchanges sterols for phosphatidylinositol 4-phosphate between lipid bilayers. *The Journal of Cell Biology* **2011**, *195* (6), 965. DOI: 10.1083/jcb.201104062.
(161) MarvinSketch. https://chemaxon.com/: 2020.
(162) Wu, E. L.; Cheng, X.; Jo, S.; Rui, H.; Song, K. C.; Dávila-Contreras, E. M.; Qi, Y.; Lee, J.; Monje-Galvan, V.; Venable, R. M.; et al. CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *J Comput Chem* **2014**, *35* (27), 1997-2004. DOI: 10.1002/jcc.23702 From NLM.
(163) Brooks, B. R.; Brooks Iii, C. L.; Mackerell Jr, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry* **2009**, *30* (10), 1545-1614. DOI: 10.1002/jcc.21287 (acccessed 2019/07/18).
(164) Li, Z.; Venable, R. M.; Rogers, L. A.; Murray, D.; Pastor, R. W. Molecular Dynamics Simulations of PIP2 and PIP3 in Lipid Bilayers: Determination of Ring Orientation, and the Effects of Surface Roughness on a Poisson-Boltzmann Description. *Biophysical Journal* **2009**, *97* (1), 155-163.
(165) Stone, J. An Efficient Library for Parallel Ray Tracing and Animation. Computer Science Department, University of Missouri-Rolla: 1998.
(166) Harayama, T.; Riezman, H. Understanding the diversity of membrane lipid composition. *Nature reviews Molecular cell biology* **2018**, *19* (5), 281-296.
(167) Zinser, E.; Sperka-Gottlieb, C.; Fasch, E.-V.; Kohlwein, S. D.; Paltauf, F.; Daum, G. Phospholipid synthesis and lipid composition of subcellular membranes in the unicellular eukaryote Saccharomyces cerevisiae. *Journal of bacteriology* **1991**, *173* (6), 2026-2034.
(168) Rogaski, B.; Klauda, J. B. Membrane-Binding Mechanism of a Peripheral Membrane Protein through Microsecond Molecular Dynamics Simulations. *Journal of Molecular Biology* **2012**, *423* (5), 847-861. DOI: https://doi.org/10.1016/j.jmb.2012.08.015.
(169) Beh, C. T.; Cool, L.; Phillips, J.; Rine, J. Overlapping functions of the yeast oxysterol-binding protein homologues. *Genetics* **2001**, *157* (3), 1117-1140.
(170) Gautier, R.; Douguet, D.; Antonny, B.; Drin, G. HELIQUEST: a web server to screen sequences with specific α-helical properties. *Bioinformatics* **2008**, *24* (18), 2101-2102.
(171) Vanni, S.; Vamparys, L.; Gautier, R.; Drin, G.; Etchebest, C.; Fuchs, Patrick F. J.; Antonny, B. Amphipathic Lipid Packing Sensor Motifs: Probing Bilayer Defects with Hydrophobic Residues. *Biophysical Journal* **2013**, *104* (3), 575-584. DOI: https://doi.org/10.1016/j.bpj.2012.11.3837.

(172) Klose, D.; Wallace, B. A.; Janes, R. W. 2Struc: the secondary structure server. *Bioinformatics* **2010**, *26* (20), 2624-2625.
(173) O'Niell, J. *Tackling drug-resistant infections globally: Final report and recommendations*; United Kingdom, 2016.
(174) Durr, U. H. N.; Sudheendra, U. S.; Ramamoorthy, A. LL-37, the only human member of the cathelicidin family of antimicrobial peptides. *Bba-Biomembranes* **2006**, *1758* (9), 1408-1425. DOI: 10.1016/j.bbamem.2006.03.030.
(175) Lohner, K. The role of membrane lipid composition in cell targeting of antimicrobial peptides. In *Development of Novel Antimicrobial Agents: Emerging Strategies*, Lohner, K. Ed.; Horizon Scientific Press, 2001; pp 149-165.
(176) Deslouches, B.; Gonzalez, I. A.; DeAlmeida, D.; Islam, K.; Steele, C.; Montelaro, R. C.; Mietzner, T. A. De novo-derived cationic antimicrobial peptide activity in a murine model of Pseudomonas aeruginosa bacteraemia. *The Journal of antimicrobial chemotherapy* **2007**, *60* (3), 669-672, Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't. DOI: 10.1093/jac/dkm253.
(177) Chen, C.; Deslouches, B.; Montelaro, R. C.; Di, Y. P. Enhanced efficacy of the engineered antimicrobial peptide WLBU2 via direct airway delivery in a murine model of Pseudomonas aeruginosa pneumonia. *Clin Microbiol Infec* **2018**, *24* (5), 547(e541-e548).
(178) Deslouches, B.; Steckbeck, J. D.; Craigo, J. K.; Doi, Y.; Mietzner, T. A.; Montelaro, R. C. Rational design of engineered cationic antimicrobial peptides consisting exclusively of arginine and tryptophan, and their activity against multidrug-resistant pathogens. *Antimicrob Agents Chemother* **2013**, *57* (6), 2511-2521. DOI: AAC.02218-12 [pii]

10.1128/AAC.02218-12.
(179) Steckbeck, J. D.; Deslouches, B.; Montelaro, R. C. Antimicrobial peptides: new drugs for bad bugs? *Expert opinion on biological therapy* **2014**, *14* (1), 11-14, Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't

Review. DOI: 10.1517/14712598.2013.844227.
(180) Lashua, L. P.; Melvin, J. A.; Deslouches, B.; Pilewski, J. M.; Montelaro, R. C.; Bomberger, J. M. Engineered cationic antimicrobial peptide (eCAP) prevents Pseudomonas aeruginosa biofilm growth on airway epithelial cells. *J Antimicrob Chemoth* **2016**, *71* (8), 2200-2207.
(181) Deslouches, B.; Steckbeck, J. D.; Craigo, J. K.; Doi, Y.; Burns, J. L.; Montelaro, R. C. Engineered cationic antimicrobial peptides to overcome multidrug resistance by ESKAPE pathogens. *Antimicrob Agents Chemother* **2015**, *59* (2), 1329-1333, Research Support, N.I.H., Extramural

Research Support, Non-U.S. Gov't. DOI: 10.1128/AAC.03937-14.
(182) Deslouches, B.; Phadke, S. M.; Lazarevic, V.; Cascio, M.; Islam, K.; Montelaro, R. C.; Mietzner, T. A. De novo generation of cationic antimicrobial peptides: influence of length and tryptophan substitution on antimicrobial activity. *Antimicrob Agents Chemother* **2005**, *49* (1), 316-322, Research Support, Non-U.S. Gov't

Research Support, U.S. Gov't, P.H.S. DOI: 10.1128/AAC.49.1.316-322.2005.
(183) Di, Y. P.; Lin, Q.; Chen, C.; Montelaro, R. C.; Doi, Y.; Deslouches, B. Enhanced therapeutic index of an antimicrobial peptide in mice by increasing safety and activity against multidrug-resistant bacteria. *Science Advances* **2020**, *6* (18). DOI: ARTN eaay6817

10.1126/sciadv.aay6817.
(184) Huang, D., Dobbins, D., Ghahramani, P., Friedland, I., Steckbeck, J. A phase 1 study of safety, tolerability and pharmacokinetics of single ascending doses of a first in human engineered cationic peptides, PLG0206, intravenously administered in healthy subjects. *Antimicrob Agents Chemother* **2021**, [Preprint], doi: https://doi.org/10.1128/AAC.01441-01421.
(185) Durell, S. R.; Brooks, B. R.; Bennaim, A. Solvent-induced forces between two hydrophilic groups. *J Phys Chem-Us* **1994**, *98* (8), 2198-2202.
(186) Klauda, J. B.; Kucerka, N.; Brooks, B. R.; Pastor, R. W.; Nagle, J. F. Simulation-based methods for interpreting X-ray data from lipid bilayers. *Biophysical journal* **2006**, *90* (8), 2796-2807. DOI: 10.1529/biophysj.105.075697.
(187) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald - an N.log(N) method for Ewald sums in large systems. *J Chem Phys* **1993**, *98* (12), 10089-10092. DOI: Doi 10.1063/1.464397.

(188) Kucerka, N.; Katsaras, J.; Nagle, J. F. Comparing membrane simulations to scattering experiments: Introducing the SIMtoEXP software. *J Membrane Biol* **2010**, *235* (1), 43-50. DOI: 10.1007/s00232-010-9254-5.
(189) Raman, S.; Vernon, R.; Thompson, J.; Tyka, M.; Sadreyev, R.; Pei, J. M.; Kim, D.; Kellogg, E.; DiMaio, F.; Lange, O.; et al. Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins* **2009**, *77*, 89-99. DOI: 10.1002/prot.22540.
(190) Song, Y. F.; DiMaio, F.; Wang, R. Y. R.; Kim, D.; Miles, C.; Brunette, T. J.; Thompson, J.; Baker, D. High-resolution comparative modeling with RosettaCM. *Structure* **2013**, *21* (10), 1735-1742. DOI: 10.1016/j.str.2013.08.005.
(191) Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S. J.; Patel, D. S.; Beaven, A. H.; Lee, K. I.; Rui, H.; Park, S.; et al. CHARMM-GUI 10 years for biomolecular modeling and simulation. *Journal of computational chemistry* **2017**, *38* (15), 1114-1124. DOI: 10.1002/jcc.24660.
(192) Loney, R. W.; Panzuela, S.; Chen, J.; Yang, Z. M.; Fritz, J. R.; Dell, Z.; Corradi, V.; Kumar, K.; Tieleman, D. P.; Hall, S. B.; et al. Location of the hydrophobic surfactant proteins, SP-B and SP-C, in fluid-phase bilayers. *Journal of Physical Chemistry B* **2020**, *124* (31), 6763-6774.
(193) Petrache, H. I.; Feller, S. E.; Nagle, J. F. Determination of component volumes of lipid bilayers from simulations. *Biophysical journal* **1997**, *72* (5), 2237-2242. DOI: Doi 10.1016/S0006-3495(97)78867-2.
(194) Wimley, W. C.; White, S. H. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat Struct Biol* **1996**, *3* (10), 842-848. DOI: Doi 10.1038/Nsb1096-842.
(195) Eglen, R.; Swank, S.; Walsh, L.; Whiting, R. Characterization of 5-HT3 and 'atypical'5-HT receptors mediating guinea-pig ileal contractions in vitro. *British journal of pharmacology* **1990**, *101* (3), 513-520.
(196) Abdel-Salam, O. M.; El-Shamarka, M. E.-S.; Salem, N. A.; El-Mosallamy, A. E.; Sleem, A. A. Amelioration of the haloperidol-induced memory impairment and brain oxidative stress by cinnarizine. *EXCLI journal* **2012**, *11*, 517.
(197) Gupta, D.; Radhakrishnan, M.; Kurhe, Y.; Thangaraj, D.; Prabhakar, V.; Kanade, P. Antidepressant-like effects of a novel 5-HT3 receptor antagonist 6z in acute and chronic murine models of depression. *Acta pharmacologica Sinica* **2014**, *35* (12), 1493-1503. DOI: 10.1038/aps.2014.89 WorldCat.org.
(198) Robillard, R.; Saad, M.; Ray, L. B.; BuJáki, B.; Douglass, A.; Lee, E. K.; Soucy, L.; Spitale, N.; De Koninck, J.; Kendzerska, T. Selective serotonin reuptake inhibitor use is associated with worse sleep-related breathing disturbances in individuals with depressive disorders and sleep complaints: a retrospective study. (1550-9397 (Electronic)). From 2021 Mar 1.
(199) Johnson, B. A.; Ait-Daoud, N.; Ma, J. Z.; Wang, Y. Ondansetron reduces mood disturbance among biologically predisposed, alcohol-dependent individuals. *Alcoholism: Clinical and Experimental Research* **2003**, *27* (11), 1773-1779.
(200) Hu, H.; Ataka, K.; Menny, A.; Fourati, Z.; Sauguet, L.; Corringer, P.-J.; Koehl, P.; Heberle, J.; Delarue, M. Electrostatics, proton sensor, and networks governing the gating transition in GLIC, a proton-gated pentameric ion channel. *Proceedings of the National Academy of Sciences* **2018**, *115* (52), E12172-E12181.
(201) Sauguet, L.; Shahsavar, A.; Poitevin, F.; Huon, C.; Menny, A.; Nemecz, À.; Haouz, A.; Changeux, J.-P.; Corringer, P.-J.; Delarue, M. Crystal structures of a pentameric ligand-gated ion channel provide a mechanism for activation. *Proceedings of the National Academy of Sciences* **2014**, *111* (3), 966-971.
(202) Bergh, C.; Heusser, S. A.; Howard, R.; Lindahl, E. Markov state models of proton-and pore-dependent activation in a pentameric ligand-gated ion channel. *Elife* **2021**, *10*, e68369.
(203) Li, Z.; Chan, K. C.; Nickels, J. D.; Cheng, X. Molecular Dynamics Refinement of Open State Serotonin 5-HT3A Receptor Structures. *Journal of Chemical Information and Modeling* **2023**, *63* (4), 1196-1207.
(204) Peters, J. A.; Cooper, M. A.; Carland, J. E.; Livesey, M. R.; Hales, T. G.; Lambert, J. J. Novel structural determinants of single channel conductance and ion selectivity in 5-hydroxytryptamine type 3 and nicotinic acetylcholine receptors. *Journal of Physiology* **2010**, *588* (4), 587-596. DOI: 10.1113/jphysiol.2009.183137 WorldCat.org.
(205) Kim, D. E.; Chivian, D.; Baker, D. Protein structure prediction and analysis using the Robetta server. *Nucleic acids research* **2004**, *32* (Web Server issue), W526-W531. DOI: 10.1093/nar/gkh468 PubMed.
(206) Nanni, L.; Lumini, A. Ensemblator: An ensemble of classifiers for reliable classification of biological data. *Pattern Recognition Letters* **2007**, *28* (5), 622-630.
(207) Trott, O.; Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. (1096-987X (Electronic)). From 2010 Jan 30.
(208) Comeau, S. R.; Gatchell, D. W.; Vajda, S.; Camacho, C. J. ClusPro: a fully automated algorithm for protein–protein docking. *Nucleic Acids Research* **2004**, *32* (suppl_2), W96-W99. DOI: 10.1093/nar/gkh354 (acccessed 12/2/2021).
(209) Grossfield, A. WHAM, Ver. 2.05: an implementation of the weighted histogram analysis method. 2011.

(210) Hon, K. L.; Leung, A. K.; Torres, A. R. Concussion: a global perspective. In *Seminars in pediatric neurology*, 2019; Elsevier: Vol. 30, pp 117-127.
(211) Kerr, Z. Y.; Chandran, A.; Nedimyer, A. K.; Arakkal, A.; Pierpoint, L. A.; Zuckerman, S. L. Concussion incidence and trends in 20 high school sports. *Pediatrics* **2019**, *144* (5).
(212) Simard, J. M.; Kent, T. A.; Chen, M.; Tarasov, K. V.; Gerzanich, V. Brain oedema in focal ischaemia: molecular pathophysiology and theoretical implications. *The Lancet Neurology* **2007**, *6* (3), 258-268.
(213) Hussain, R.; Tithof, J.; Wang, W.; Cheetham-West, A.; Song, W.; Peng, W.; Sigurdsson, B.; Kim, D.; Sun, Q.; Peng, S. Potentiating glymphatic drainage minimizes post-traumatic cerebral oedema. *Nature* **2023**, 1-9.
(214) Hu, Y.-P.; Zhong, Y.-Q.; Chen, Z.-G.; Chen, C.-Y.; Shi, Z.; Zulueta, M. M. L.; Ku, C.-C.; Lee, P.-Y.; Wang, C.-C.; Hung, S.-C. Divergent Synthesis of 48 Heparan Sulfate-Based Disaccharides and Probing the Specific Sugar–Fibroblast Growth Factor-1 Interaction. *Journal of the American Chemical Society* **2012**, *134* (51), 20722-20727. DOI: 10.1021/ja3090065.
(215) DiGabriele, A. D.; Lax, I.; Chen, D. I.; Svahn, C. M.; Jaye, M.; Schlessinger, J.; Hendrickson, W. A. Structure of a heparin-linked biologically active dimer of fibroblast growth factor. *Nature* **1998**, *393* (6687), 812-817. DOI: 10.1038/31741.
(216) Faham, S.; Hileman, R. E.; Fromm, J. R.; Linhardt, R. J.; Rees, D. C. Heparin Structure and Interactions with Basic Fibroblast Growth Factor. *Science* **1996**, *271* (5252), 1116-1120. DOI: doi:10.1126/science.271.5252.1116.
(217) Dvorak, P.; Bednar, D.; Vanacek, P.; Balek, L.; Eiselleova, L.; Stepankova, V.; Sebestova, E.; Kunova Bosakova, M.; Konecna, Z.; Mazurenko, S. Computer-assisted engineering of hyperstable fibroblast growth factor 2. *Biotechnology and bioengineering* **2018**, *115* (4), 850-862.
(218) Koledova, Z.; Sumbal, J.; Rabata, A.; de La Bourdonnaye, G.; Chaloupkova, R.; Hrdlickova, B.; Damborsky, J.; Stepankova, V. Fibroblast growth factor 2 protein stability provides decreased dependence on heparin for induction of FGFR signaling and alters ERK signaling dynamics. *Frontiers in cell and developmental biology* **2019**, 331.
(219) Sher, I.; Weizman, A.; Lubinsky-Mink, S.; Lang, T.; Adir, N.; Schomburg, D.; Ron, D. Mutations uncouple human fibroblast growth factor (FGF)-7 biological activity and receptor binding and support broad specificity in the secondary receptor binding site of FGFs. *Journal of Biological Chemistry* **1999**, *274* (49), 35016-35022.
(220) Springer, B. A.; Pantoliano, M. W.; Barbera, F. A.; Gunyuzlu, P. L.; Thompson, L. D.; Herblin, W. F.; Rosenfeld, S. A.; Book, G. W. Identification and concerted function of two receptor binding surfaces on basic fibroblast growth factor required for mitogenesis. *Journal of Biological Chemistry* **1994**, *269* (43), 26879-26884.
(221) Govind Kumar, V.; Agrawal, S.; Kumar, T. K. S.; Moradi, M. Mechanistic picture for monomeric human fibroblast growth factor 1 stabilization by heparin binding. *The Journal of Physical Chemistry B* **2021**, *125* (46), 12690-12697.
(222) Vemuri, S.; Beylin, I.; Sluzky, V.; Stratton, P.; Eberlein, G.; Wang, Y. J. The stability of bFGF against thermal denaturation. *Journal of pharmacy and pharmacology* **1994**, *46* (6), 481-486.
(223) Tobey, S. L.; Anslyn, E. V. Determination of inorganic phosphate in serum and saliva using a synthetic receptor. *Organic letters* **2003**, *5* (12), 2029-2031.
(224) Lacerda-Abreu, M. A.; Russo-Abrahão, T.; Meyer-Fernandes, J. R. The roles of sodium-independent inorganic phosphate transporters in inorganic phosphate homeostasis and in cancer and other diseases. *International Journal of Molecular Sciences* **2020**, *21* (23), 9298.
(225) Pavani, P.; Kumar, K.; Rani, A.; Venkatesu, P.; Lee, M.-J. The influence of sodium phosphate buffer on the stability of various proteins: Insights into protein-buffer interactions. *Journal of Molecular Liquids* **2021**, *331*, 115753.
(226) Ornitz, D. M.; Herr, A. B.; Nilsson, M.; Westman, J.; Svahn, C.-M.; Waksman, G. FGF binding and FGF receptor activation by synthetic heparan-derived di-and trisaccharides. *Science* **1995**, *268* (5209), 432-436.
(227) Morrison, R. S.; Giordano, S.; Yamaguchi, F.; Hendrickson, S.; Berger, M.; Palczewski, K. Basic fibroblast growth factor expression is required for clonogenic growth of human glioma cells. *Journal of neuroscience research* **1993**, *34* (5), 502-509.
(228) Maffucci, T.; Piccolo, E.; Cumashi, A.; Iezzi, M.; Riley, A. M.; Saiardi, A.; Godage, H. Y.; Rossi, C.; Broggini, M.; Iacobelli, S. Inhibition of the phosphatidylinositol 3-kinase/Akt pathway by inositol pentakisphosphate results in antiangiogenic and antitumor effects. *Cancer Research* **2005**, *65* (18), 8339-8349.
(229) Berendsen, H. J.; van der Spoel, D.; van Drunen, R. GROMACS: A message-passing parallel molecular dynamics implementation. *Computer physics communications* **1995**, *91* (1-3), 43-56.
(230) Lindahl, E.; Hess, B.; Van Der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Molecular modeling annual* **2001**, *7*, 306-317.

(231) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. GROMACS: fast, flexible, and free. *Journal of computational chemistry* **2005**, *26* (16), 1701-1718.
(232) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation* **2008**, *4* (3), 435-447. DOI: doi:10.1021/ct700301q.
(233) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29* (7), 845-854.
(234) 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*, 19-25.
(235) Park, S.-J.; Lee, J.; Qi, Y.; Kern, N. R.; Lee, H. S.; Jo, S.; Joung, I.; Joo, K.; Lee, J.; Im, W. CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates. *Glycobiology* **2019**, *29* (4), 320-331.
(236) Lee, J.; Hitzenberger, M.; Rieger, M.; Kern, N. R.; Zacharias, M.; Im, W. CHARMM-GUI supports the Amber force fields. *The Journal of chemical physics* **2020**, *153* (3), 035103.
(237) Jo, S.; Song, K. C.; Desaire, H.; MacKerell Jr, A. D.; Im, W. Glycan Reader: automated sugar identification and simulation preparation for carbohydrates and glycoproteins. *Journal of computational chemistry* **2011**, *32* (14), 3135-3141.
(238) Park, S.-J.; Lee, J.; Patel, D. S.; Ma, H.; Lee, H. S.; Jo, S.; Im, W. Glycan Reader is improved to recognize most sugar types and chemical modifications in the Protein Data Bank. *Bioinformatics* **2017**, *33* (19), 3051-3057.
(239) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of computational chemistry* **2010**, *31* (4), 671-690.
(240) Yu, W.; He, X.; Vanommeslaeghe, K.; MacKerell Jr, A. D. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *Journal of computational chemistry* **2012**, *33* (31), 2451-2468.
(241) Kim, S.; Lee, J.; Jo, S.; Brooks III, C. L.; Lee, H. S.; Im, W. CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. Wiley Online Library: 2017.
(242) Gao, Y.; Lee, J.; Smith, I. P. S.; Lee, H.; Kim, S.; Qi, Y.; Klauda, J. B.; Widmalm, G.; Khalid, S.; Im, W. Charmm-gui supports hydrogen mass repartitioning and different protonation states of phosphates in lipopolysaccharides. *Journal of chemical information and modeling* **2021**, *61* (2), 831-839.
(243) Guvench, O.; Greene, S. N.; Kamath, G.; Brady, J. W.; Venable, R. M.; Pastor, R. W.; Mackerell Jr, A. D. Additive empirical force field for hexopyranose monosaccharides. *Journal of computational chemistry* **2008**, *29* (15), 2543-2564.
(244) Hatcher, E.; Guvench, O.; MacKerell Jr, A. D. CHARMM additive all-atom force field for aldopentofuranoses, methyl-aldopentofuranosides, and fructofuranose. *The Journal of Physical Chemistry B* **2009**, *113* (37), 12466-12476.
(245) Raman, E. P.; Guvench, O.; MacKerell Jr, A. D. CHARMM additive all-atom force field for glycosidic linkages in carbohydrates involving furanoses. *The Journal of Physical Chemistry B* **2010**, *114* (40), 12981-12994.
(246) Adasme, M. F.; Linnemann, K. L.; Bolz, S. N.; Kaiser, F.; Salentin, S.; Haupt, V. J.; Schroeder, M. PLIP 2021: expanding the scope of the protein–ligand interaction profiler to DNA and RNA. *Nucleic acids research* **2021**, *49* (W1), W530-W534.
(247) LLC, S. The PyMOL Molecular Graphics System, Version 2.5.3. 2015.
(248) Costello, A. J.; Glonek, T.; Myers, T. C. 31P Nuclear magnetic resonance pH titrations of myo-inositol hexaphosphate. *Carbohydrate Research* **1976**, *46* (2), 159-171.
(249) Govind Kumar, V.; Polasa, A.; Agrawal, S.; Kumar, T. K. S.; Moradi, M. Binding affinity estimation from restrained umbrella sampling simulations. *Nature Computational Science* **2023**, *3* (1), 59-70.
(250) Agrawal, S.; Govind Kumar, V.; Gundampati, R. K.; Moradi, M.; Kumar, T. K. S. Characterization of the structural forces governing the reversibility of the thermal unfolding of the human acidic fibroblast growth factor. *Scientific Reports* **2021**, *11* (1), 1-13.
(251) Hajnóczky, G.; Thomas, A. P. Minimal requirements for calcium oscillations driven by the IP3 receptor. *The EMBO journal* **1997**, *16* (12), 3533-3543.
(252) Foskett, J. K.; White, C.; Cheung, K.-H.; Mak, D.-O. D. Inositol trisphosphate receptor Ca2+ release channels. *Physiological reviews* **2007**, *87* (2), 593-658.

(253) Burdakov, D.; Petersen, O. H.; Verkhratsky, A. Intraluminal calcium as a primary regulator of endoplasmic reticulum function. *Cell calcium* **2005**, *38* (3-4), 303-310.
(254) Fu, S.; Yang, L.; Li, P.; Hofmann, O.; Dicker, L.; Hide, W.; Lin, X.; Watkins, S. M.; Ivanov, A. R.; Hotamisligil, G. S. Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity. *Nature* **2011**, *473* (7348), 528-531.
(255) Micaroni, M. The role of calcium in intracellular trafficking. *Current molecular medicine* **2010**, *10* (8), 763-773.
(256) Oancea, E.; Meyer, T. Reversible desensitization of inositol trisphosphate-induced calcium release provides a mechanism for repetitive calcium spikes. *Journal of Biological Chemistry* **1996**, *271* (29), 17253-17260.
(257) Ilyin, V.; Parker, I. Role of cytosolic Ca2+ in inhibition of InsP3-evoked Ca2+ release in Xenopus oocytes. *The Journal of Physiology* **1994**, *477* (3), 503-509.

Refereed Journal Articles and Conference Proceedings

1. Robert Allsopp and Jeffery Klauda "Impact of PIP2 Lipids, Force Field Parameters, and Mutational Analysis on the Binding of the Osh4's α6–α7 Domain." 2021, ***The Journal of Physical Chemistry B,*** Volume 125, Issue 20, Pages 5296-5308

2. Robert Allsopp, Anna Pavlova, Tyler Cline, Aria Salyapongse, Richard Gillilan, Y. Peter Di, Berthony Deslouches, Jeffery Klauda, James Gumbart, and Stephanie Tristram-Nagle, "Antimicrobial Peptide Mechanism Studied by Scattering-Guided Molecular Dynamics Simulation." 2022, ***The Journal of Physical Chemistry B,*** Volume 126, Issue 36, Pages 6922-6935

3. Robert Allsopp, and Jeffery Klauda. "Conformational Dynamics of bFGF Interacting with Polysaccharides and Point Mutations." **(Under review in the Journal of Biophysical Chemistry)**

4. Robert Allsopp, Dinara Konakbayeva, Amy Karlsson, and Jeffery Klauda, "Osh4 Peptide Folding and Binding Free Energies as Studied from Simulations and Wet Lab Techniques." **(In Preparation)**

5. Robert Allsopp and Jeffery Klauda, "Comparing the Effects of Varied Lipid Packing on the Binding Free Energies of Osh4 and Osh5 ALPS Peptides." **(In Preparation)**

6. Robert Allsopp and Jeffery Klauda, "Microsecond Deactivation Studies of Conotoxin σ-GVIIIA on 5-HT_{3A} Ion Channel and Partial Water Filling of the Apo Crystal Structure." **(In Preparation)**

7. Dylan Mendes, Tyla Holoman, Robert Allsopp, and Jeffery Klauda, "A Study of Maganin's Binding to Model Stratum Corneum Membrane." **(In Preparation)**

Presentations

1. Robert Allsopp, "Peripheral Membrane Binding Characteristics of Osh α6-α7 Helices and the Effect of the Presence of PIP2 Lipids and Force Field Parameters." Poster session at the Biophysical Society Annual Meeting, 2021.

2. Robert Allsopp, "Computational Study of Membrane Binding and Free Energy of Amphipathic Helix (ALPS) on Osh4 to Varied Lipid Bilayers." Poster session at the Biophysical Society Annual Meeting, 2022.

3. Robert Allsopp, "Computational Study of Membrane Binding and Free Energy of Amphipathic Helix (ALPS) on Osh4 to Varied Lipid Bilayers." Poster session at the University of Maryland Chemical Engineering Showcase, 2022.

4. Robert Allsopp, “MD Simulations of the 5-HT_{3A} Ion Channel.” Presentation at the Formulations Division weekly meeting at Abbvie, 2022.

5. Robert Allsopp, “Controlled Release Properties of a Novel Minimally Processed Formulation Consisting Primarily of Ionizable Polymers.” End of the internship presentation for the Formulations Division at Abbvie, 2022.

6. Robert Allsopp, “Sigma Conotoxin GVIIIA: Binding to 5HT3A with Microsecond Simulations and Docking of Diverse Structures.” Poster session at the Biophysical Society Annual Meeting, 2023.

7. Robert Allsopp, “Sigma Conotoxin GVIIIA: Binding to 5HT3A with Microsecond Simulations and Docking of Diverse Structures.” Poster session at the University of Maryland Chemical Engineering Showcase, 2023.

8. Robert Allsopp, “Examples of Protein Ligand Binding Sites with a Focus On Heparan Sulfate Induced Protein Folding of the basic Fibroblast Growth Factor.*”* Research presentation at the Institute for Bioscience and Biotechnology Research/NIST, 2023.