2022

Analyzing the Free Energy of Ions Sampling a Voltage Gated Sodium Ion Channel

Isabel Varghese
Colby College

Analyzing the Free Energy of Ions Sampling a Voltage Gated Sodium Ion Channel.

By Iz Varghese

A Thesis Presented to the Department of Chemistry,

Colby College, Waterville, ME

In Partial Fulfillment of the Requirements for Graduation

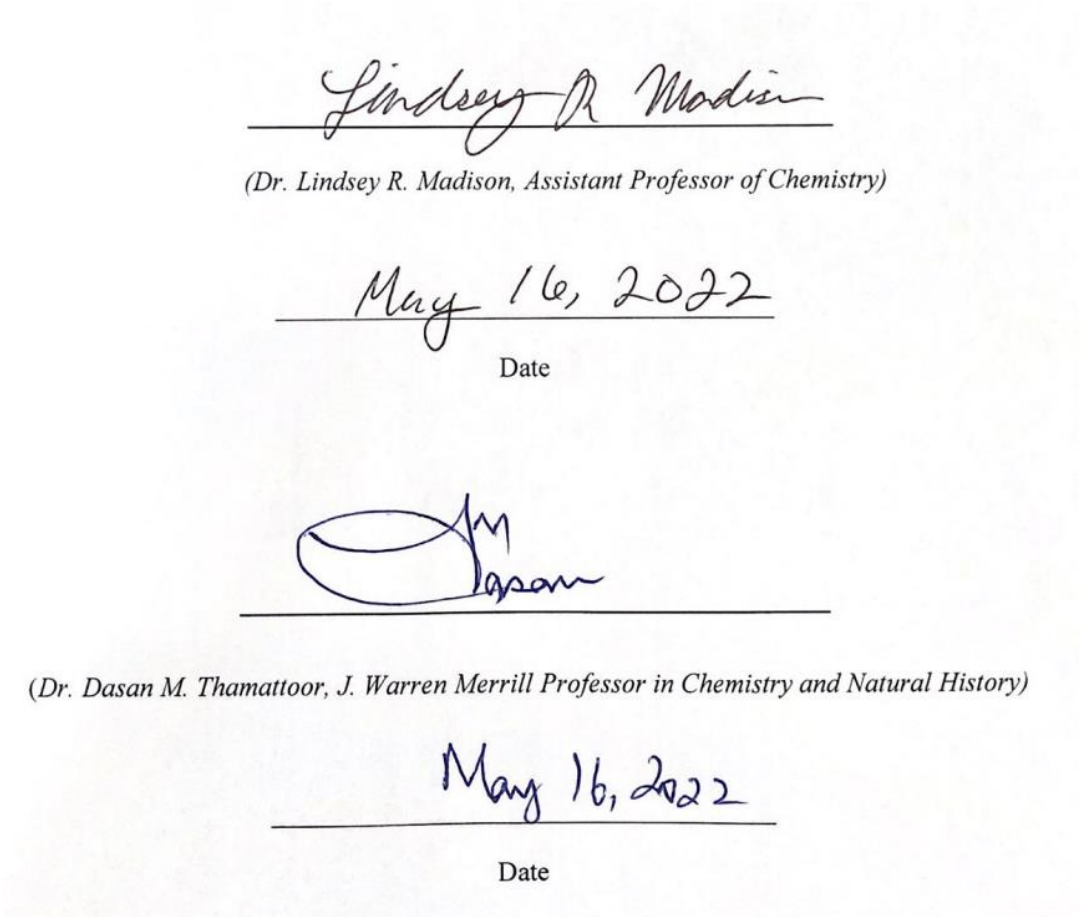
With Honors in Chemistry

Submitted May, 2022

Analyzing the Free Energy of Ions Sampling a Voltage Gated Sodium Ion Channel.

By Iz Varghese

Approved:



Lindsey R Madison

(Dr. Lindsey R. Madison, Assistant Professor of Chemistry)

May 16, 2022

Date

Dasan

(Dr. Dasan M. Thamattoor, J. Warren Merrill Professor in Chemistry and Natural History)

May 16, 2022

Date

Vitae

Isabel ‘Iz’ Varghese was born on August 25th, 2000, to Smitha Varghese in Troy, Michigan. They graduated from Forest Hills Central High School in Grand Rapids, Michigan, and joined Colby College as a Ralph J. Bunche Scholar. After much fluctuation, they finally reached an equilibration stage where they decided to major in Chemistry with a concentration in Biochemistry and a minor in English. Starting in the summer after their first year, they worked in the Madison lab conducting computational research about diffusion of guest molecules in clathrate hydrates. They also had the opportunity to participate in an REU at the University of Colorado-Anschutz Medical Campus where they studied the role of fibronectin in BRAF-inhibitor resistance in thyroid cancer. They balanced their scientific research with research in the Women, Gender, and Sexuality Department under Dr. Jay Sibara, and have completed an independent study with Dr. Aaron Hanlon in the English department. They have been accepted into the NIH Post-baccalaureate Research Education Program (PREP) at the University of Michigan where they hope to continue pursuing their passion for physical chemistry and biophysics.

Acknowledgements

There are so many people I want to thank. First, my mother and my grandparents who moved across an ocean in search of better opportunities for me and my siblings. I would not be here without you. Especially my dear Appachan: you fought and worked so hard for us to have the best life possible. You spent every breath reminding me how much me and my siblings are loved and supported. You were my strength, light, and joy. I hope I am making you proud.

I am also grateful for Anabel and Meribell who have been with me through thick and thin, and who both give the best hugs. Thank you to John for making me food whenever I am hungry, and to Bix for dipping his paws in all my unattended glasses of water.

Thank you to everyone in Apartment 419. I will cherish our conversations and will miss the weird noises you all make, karaoke sessions, and our study sessions (that inevitably became karaoke sessions) at the island kitchen table.

Professor Lindsey Madison, I am so grateful for all your guidance and support. You took my seedling of an idea and helped me to grow it to the 70+ page gargantuan reality that is this thesis. I have learned a lot about being a scientist; especially how to gracefully handle challenges and advocate for myself. Thank you also to Ginny for patiently listening to me talk about my work. I am very humbled and proud to be graduating as a member from the Madison lab.

Professor Das Thamattoor, thank you for agreeing to be my second reader, for all the enthusiastic “Iz!”’s and our exchanges in Malayalam. It made Colby a little more homely.

Thank you to all the other Madison lab members: Jordyn, Bob, Shailin, Emerson, Coy, and Muna. As theoretical chemists, we rarely saw each other, but it was always a blast when we did. Thank you to all members of the chemistry department who have guided me and supported me in my journey as a scientist. I truly could not have done it without you all.

Table of Contents

Analyzing the Free Energy of Ions Sampling A Voltage Gated Ion Channel. 1

Vitae **Error! Bookmark not defined.**

Acknowledgements 4

Table of Contents 5

Abstract 6

Background 7

Selectivity and Mechanisms of Diffusion in Voltage Gated Sodium Ion Channels. 9

Research Questions 10

Molecular Dynamics 13

Theoretical Justification for MD 14

CHARMM Forcefield 15

Adaptive Bias Force (ABF). 19

Computing Free Energy from Positions 19

Methods 21

Molecular Dynamics of Biological Systems. 21

Creating PDB Files 22

Capturing Interactions of Particles 25

Running NAMD Simulations 26

Minimization 28

Equilibration 29

Production Runs 30

Results and Discussion 32

Structure of the Protein 32

Free energy profiles of ions in the protein channel 34

Hydration Shells Of Ions 42

Effects Of Protonation States Of EEEE Ring On Ion Binding And Free Energy 49

Differences In Homogeneous And Heterogeneous Systems 53

Limitations of this Study 55

Future Directions 58

Conclusions 59

References 60

Appendix 67

Abstract

Voltage gated sodium ion channels are implicated in cardiac diseases, seizures, etc., and they play a role in maintaining ionic homeostasis in cells. Computational studies use prokaryotic model because they are simpler but function similarly to human voltage gated sodium ion channels. This study uses molecular dynamics (MD) to study three specific questions regarding voltage gated sodium ion channels of *Magnetococcus Marinus.* The first question in this study is how the free energy of sodium diffusion compares to that of calcium ion diffusion. We were not able to find any physically significant information due to poor sampling and a lack of ion diffusion. We also examined the role of the hydration shell and protonation state in conferring selectivity to the voltage gated sodium ion channel. This study does not find evidence that the size of the hydration shell or protonation state of the glutamate ring in the selectivity filter contributes to selectivity against calcium ions. Additionally, the water itself has no significant interactions with the amino acid residues. A key conclusion from this study is that approximating the location of the uncrystallized residues using the GalaxyFill algorithm does not accurately represent experimental voltage gated ion channels because it occasionally predicts non-physical conformations that result in the interlocking of rings of the residues with other residues and may be biased towards the formation of loops in the system.

Background

The cell membrane is selectively permeable to substances. A 'selectively permeable' membrane means that the membrane only allows certain substances to diffuse through it. Small, nonpolar substances such as oxygen and carbon dioxide can pass through the membrane without any intervention in a process called passive diffusion.[1] Charged particles such as sodium, calcium and potassium ions need to undergo facilitated diffusion.[2] Ion channels that facilitate the diffusion of sodium are known as sodium ion channels. These channels are dependent on the membrane potential of the cell, and, therefore, are called voltage gated ion channels. Voltage gated channels are regulated by the membrane potential.[3] The membrane potential is represented by a modified version of the Nernst Equation:

$$V_{eq} = \frac{RT}{zF} ln \frac{[Na^+_{extracellular}]}{[Na^+_{intracellular}]} \qquad (1)$$

where V_{eq} is the membrane potential of the cell, *R* is the gas constant, *T* is the temperature, *z* is the ion charge which is +1 in the case of sodium ions, and *F* is the Faraday's constant. As seen in this equation, the membrane potential is logarithmically related to the concentration of the sodium ions inside and outside the cell membrane.

Depending on the membrane potential, the ion channels can either be in an open or closed configuration. Traditional models suggest that in an open configuration, the ions can pass through the channel, and in a closed configuration, they cannot. However, some studies suggest that even in a closed configuration, ions are capable of entering the protein.[4] Thus, the conventional categorization of proteins as either open or closed can be misleading and does not capture the complexity of ion transport through the membrane.

Prokaryotic cells are used as models to study human voltage gated sodium ion channels because of similarities in their structure and function.[13] In prokaryotic cells, the voltage gated sodium channel has two regions: the pore domain and voltage sensor domain. The pore domain is a transmembrane region of the protein through which the ions traverse while the voltage sensor domain is a transmembrane region of the protein that mediates opening and closing of the ion channel to allow for ion passage through the protein. A functional voltage gated sodium channel would fluctuate between an open and closed configuration dependent on the movement of the voltage sensor domains. Prokaryotic voltage gated channels are homotetrameric, with each monomer contributing to both the pore domain (S1-S4) and the voltage sensing domain (S5 and S6, and the P1 and P2 helices).

The membrane potential that triggers the open and closed conformations is species dependent. For example, in a human cell, when the ion channel is deactivated and closed, the membrane potential is -70 mV.[4] When the ion channel is open, the membrane potential is between -55 mV and +40 mV. When the ion channel is in its refractory period (when it is closed and cannot open for a while), the membrane potential is between +40 mV and -70 mV.[5] For *E. coli*, the resting membrane potential is around -135 mV, and alters to about -24 mV, depending on the study.[6]

Selectivity is the property of ion channels that allow for certain ions to pass through and not others. Ion channels have a region in the protein structure, called a selectivity filter, that have amino acid residues contributing to the protein's selectivity.[7] The amino acid residues are species dependent; selectivity is conferred through a combination of the amino acid residues involved and the size of the selectivity filter. For example, potassium ion channels are known for their highly conserved selectivity filter sequence, TXXTXGYG, and a small selectivity filter.[8] Ions

are solvated in water when they enter the ion channel; the water structure in which the ion is solvated is called a hydration shell. In a potassium ion channel, the ions "shed" their hydration shell in the selectivity filter.[9] When sodium ions enter a voltage gated potassium ion channel, they have a higher energy barrier for dehydration in comparison to potassium ions.[10] Consequently, it is less energetically favorable for sodium ions to traverse the protein channel. In other words, the potassium ion channel is selective against sodium ions. As detailed in the next section, the converse is also true: voltage gated sodium ion channels select for sodium, and against potassium.

Selectivity and Mechanisms of Diffusion in Voltage Gated Sodium Ion Channels.

Voltage gated sodium ion channels, as their name suggests, are selective for sodium ion diffusion. In general, the pore of the selectivity filter of voltage gated sodium ion channels is larger than that of the potassium voltage gated ion channel.[11] Corresponding to this increase in size, is the presence of a hydration shell as the sodium ion traverses through the ion channel.[9,11,12] Depending on the species, there are different recognized amino acid residues that the hydration shell and/or the sodium ion interacts with, known as binding sites. There is limited understanding of the structures of the voltage gated ion channel in its open configuration across species. In computational studies of ion channels, bacterial voltage gated sodium ion channels have generally been used as a model to further understand the role of voltage gated sodium ion channels. Bacterial sodium channels are structurally simpler to model than human sodium ion channels, and thus, require less computational power and time.[13] This study will focus on the ion diffusion and permeation of sodium ions in *Magnetococcus marinus* (Na_vMs). Specifically, this study will focus on the free energy of ion diffusion and permeation, as it traverses through these channels.

Bacteria do not have sodium-potassium pumps, and therefore, any transmission of ions through the channels would not be able to return to the resting membrane potential. Thus, the function of voltage gated ion channels in prokaryotes is less certain.[14] Some studies suggest that sodium ion channels play a more protective role in bacteria, where ion channels prevent osmotic shock when placed in extreme environments.[15] Five binding sites have been identified in the selectivity filter of Na_v Ms, called S0-S4.[16] Of the five binding sites, S1 is the dominant binding site with a negatively charged amino acid residue (Glu) in the EEEE ring.[16] The positively charged sodium ion has attractive interactions with this binding site; as a result, in this binding site, the ion sheds its hydration shell partially.[16] The S0, S2-S4, are binding sites that interact with the whole hydration shell.[16]

The selectivity of Na_vMs can be understood thermodynamically using free energy. The free energy changes of the system as the particle traverses the channel indicates the relative stability of the particle as a function of its location and gives insight into why sodium ions can pass through the ion channel, but other ions cannot. Thus, free energy profiles are useful for identifying energy barriers and regions of favorable interactions such as binding sites.

Research Questions

This study will focus on the energetics of diffusion of different particles in Na_vMs. There are four unknown factors that this study will focus on. Prior to answering any questions regarding energetics of ion diffusion through the channel, it is imperative to construct a system that contains an experimentally relevant model of the protein. For this study, we are using an open configuration of the Na_vMs channel (PDB: 4F4L). As mentioned earlier, prokaryotic cells have two primary regions: the pore and the voltage sensing region. The crystal structure that this study uses has been engineered to have a pore only region (i.e., a truncated voltage sensing

region). This means that the protein will not be sensitive to the membrane potential, and thus, the charge of the ions in the system should not affect whether the protein is open or closed. Na_vMs is stable without the voltage sensing region and continues to conduct ions. This structure also contains a disordered region near the C-terminus that has been hypothesized to play a role in stabilizing the tetrameric Na_vMs structure. The 4F4L structure does not have high resolution locations for the disordered region due to stacking in the region; however, the residues present in this location are known and documented in the Protein Data Bank. When studying this protein to answer questions related to energetics and selectivity, prior groups have excluded the disordered regions from their studies and only included residues 8-94. In this study, we approximate the location of the residues, and include them in the study. In doing so, a question that arose was whether these approximations were reasonable, and how these approximations affect ion diffusion. Additionally, Ulmschneider et al. applied harmonic restraints to the protein to force the channel to remain open.[16] In this study, we are interested in how the presence of the tails affect ion diffusion even without the harmonic restraints.

Under these conditions, this study will focus on the selectivity of the ion channel; specifically, we will examine how the protein confers selectivity against calcium in comparison to sodium. The free energy profiles of sodium ions in Na_vMs show that at either entrance of the channel, Na^+ and K^+ have a change in free energy of zero.[16] However, at locations within the proteins, specifically corresponding to the binding site, the sodium ions have lower energy barriers while traveling through the ion channel. Both sodium and potassium are ions that cross the membrane using voltage gated ion channels.[16,17] Voltage gated ion channels specific to calcium ions (Ca^{2+}) also exist. Furthermore, bacterial voltage gated sodium ion channels are evolutionarily closer to calcium ion channels.[13,18] Accordingly, this study aims to address how

the energetics of calcium ions compare to that of sodium ions. Given that the calcium ion is more positive in charge, we expect the calcium ion to have a larger hydration shell. Consequently, we hypothesize that the large hydration shell plays a role in conferring selectivity against the calcium ions.

To gain a clearer picture of ion diffusion, it would also be useful to examine ion hydration during traversal within the protein. While it is known that water hydrates the sodium ion to facilitate transportation, there are open questions about the free energy of water as it moves through the sodium ion channel.[12,16] Water has a net diffusion rate of zero in the ion channel.[16] Yet, the hydration shell is hypothesized to interact with amino acid residues in the channel. We are interested in determining a free energy profile of water to see if there are any such interactions. By examining the free energy of the water as it moves through the ion channel, we can examine the role that water plays in selectivity and ion transport.

Another factor that needs to be considered is the protonation state of ionizable amino acid residues in Na_vMs. For Na_vMs, the selectivity filter has an EEEE ring whose protonation state is of interest because of its effect on ion conductance. This ring, as the name suggests, has one glutamate residue on each of the four chains that face inwards towards each other when none of the residues are protonated (Figure 1). At a physiological pH of 7.4, the glutamate residue can exist in 3 different states: singly protonated, doubly protonated, and fully deprotonated. However, depending on the protonation state, the conductance of ions through the protein varies.[19] In other words, the rate of diffusion of ions is related to the protonation states of the EEEE ring structure. A fully deprotonated state is correlated to faster diffusion with a decrease in the rate of diffusion with each additional proton. Thus, this study will also examine the effects

of protonation state on the selectivity of the protein against other ions, and the role of water in this process.

Finally, in a biological environment such as a cell or organism the protein is also exposed to other ions and substances. However, many computational studies assume a homogeneous environment with only the ion of interest. We are interested in seeing how having both sodium and potassium in a system affects the free energy profile of the ions as they diffuse through the channel.

Molecular Dynamics

Computational models are necessary to further understand the energetics of ion transport in a channel because an analysis of the first two questions require an atomistic interaction. To answer any of the proposed questions will require an analysis of the free energy as a function of location. Location over time can be determined by analyzing the trajectory of a particle in a system. Molecular Dynamics (MD) is a computational algorithm that allows for the simulation of molecular and atomic motion.

MD is useful in simulating large systems of atoms, such as proteins. MD has been used to simulate human glucose transporter 3 (GLUT 3), a central protein in nervous system metabolism, and tyrosine kinase receptors.[20,21] When simulating these proteins, MD is used to study the dynamics of the proteins and can answer various questions regarding how amino acid residues move, what the free energy of a particle is, the Michaelis-Menten constant, etc. Visual Molecular Dynamics (VMD) is a computer program that allows users to visualize and analyze the results of a MD simulation.[22] Molecular visualization is a valuable tool in gaining qualitative observations regarding how amino acid residues and ions move within a channel. This data can be used in

tandem with any quantitative data to glean a comprehensive understanding of the protein in question.

This study will use Nanoscale Molecular Dynamics (NAMD) which is a computer program incorporating MD that can be used to study millions of atoms. NAMD uses a Charm++ parallel programming that can be linked to VMD.[23] By simulating voltage gated sodium ion channels in NAMD, not only will this study be able to simulate a system as large as a protein, but we will also be able to visualize what happens as different ions traverse through the voltage gated sodium ion channel.

Theoretical Justification for MD

For MD simulations, the user sets up a simulation box within which all particles reside; in this case, the box would have to be large enough to accommodate the proteins and the ion, but small enough to minimize computational costs in term of time and memory.[24] Setting up a simulation box is the equivalent of setting up an ensemble. An ensemble is a set up when a system has certain variables of state set to a constant value.[25] There are various categories of ensembles: microcanonical, canonical, isothermal-isobaric, and grand canonical ensemble. This study will be using the isothermal-isobaric which is dependent on the following constant thermodynamic variables of states: number of particles, pressure and temperature.[25] Ensemble systems are a simplified statistical model that can be used to derive thermodynamic properties of systems from quantum mechanical properties.[26]

In MD, each particle is simulated as a point charge. Point charges are used to find intermolecular Coulombic potentials. By taking the derivative of the intermolecular Coulombic potential as a function of its position, the force acting on each particle can be determined. This relationship will further be detailed below in the ‘CHARMM Forcefield’ section. MD uses

classical mechanics to model the motion of particles. Thus, the force (*F*) on each particle can be computed using the following equation:

$$F = ma \qquad (2)$$

where *m* is the mass of the particle, and *a* is the acceleration of the particle. The motions are studied as a function of time, in femtoseconds (fs). Since the increments of time (Δt) are so small, the acceleration of each particle can be regarded as constant over that time step.[27] Consequently, the kinematic equations can be used to determine the position and velocity of each particle (Equation 3 and 4).

$$x_1 = x_0 + v_0\Delta t + \frac{1}{2}a(\Delta t)^2 \qquad (3)$$

$$v_1 = v_0 + a\Delta t \qquad (4)$$

where x_0 and v_0 is the initial position and velocity. Once these positions have been determined, they act as the new initial position and velocity of the particle. Since acceleration is constant, at an infinitesimal time increment, the new positions of the particles can be calculated. Each time a position is calculated, based on increments of time, it is called an iteration. Each increment of time is known as a timestep. To calculate the motion of particles, called trajectories, over an extended period, we repeat these iterations for many timesteps.

CHARMM Forcefield

To calculate acceleration, we need a model that can calculate the forces that each particle experiences. Forcefields are a set of parameters and equations that are used to model the geometry and forces of a molecule. The Chemistry at Harvard Macromolecular Mechanics (CHARMM) is a commonly used forcefield in molecular dynamics and simulations. CHARMM accounts both for intermolecular and intramolecular interactions.[28] Intramolecular interactions

that are accounted for include the forces and parameters that comprise the structure of the molecule such as bond lengths, bond angles, dihedral angles, etc. The intramolecular vibrational motion of atoms is modeled using the classical spring model, or the harmonic model. The harmonic model has the general equation:

$$U = \frac{1}{2}kx^2 \qquad (5)$$

where *U* is the potential energy, *k* is the spring constant, and *x* is the location of the particle. The CHARMM equation for potential energy has the form:

$$V = \sum_{bonds} k_b(b - b_o)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{dihedrals} k_\phi[1 + cos(n\phi - \delta)]$$
$$+ \sum_{impropers} k_\omega(\omega - \omega_0)^2 + \sum_{Urey-Bradley} k_u(u - u_0)^2$$
$$+ \sum_{nonbonded}(\epsilon[(\frac{R_{min,ij}}{r_{ij}})^{12} - (\frac{R_{min,ij}}{r_{ij}})^6] + \frac{q_i q_j}{\epsilon r_{ij}}) \qquad (6)$$

where k_b, k_θ, k_ϕ, k_ω, and k_u are the bonds, angle, dihedral angle, improper dihedral angle force, and Urey-Bradley constants, respectively; b, θ, ϕ, ω, and u are the bond length, angle, dihedral angle, improper torsion angle, and Urey-Bradley 1,3-distance, respectively.[29] Here, a zero in the subscript indicates an equilibrium value for the individual terms. The Urey-Bradley cross term accounts for 1,3-nonbonded interactions. The CHARMM forcefield accounts for the harmonic potential model in the bonds, angles, improper torsion angle, and Urey-Bradley term, as seen in the general form of the equations.

The CHARMM forcefield accounts for external nonbonded interactions such as van der Waals forces. It does so using the Lennard-Jones model (12-6 model) which emulates the attractive and repulsive forces of particles. The Lennard-Jones model for potential forces (V_{LJ}) follows the form:

$$V_{LJ}(r) = 4\varepsilon[(\frac{\sigma}{r})^{12} - (\frac{\sigma}{r})^6] \qquad (7)$$

where *r* is the distance between the two interacting particles; ε is the depth of the well potential, and a measure how strongly two particles are attracted to each other; and σ is the distance at which the intermolecular forces between the two particles is zero. The Lennard-Jones model shows potential energy as a function of the distance between two particles. The model suggests that as *r* approaches infinity, i.e., as the particles are infinitely far apart, the potential energy approaches zero because it is unlikely that they will interact with each other. As they come closer to each other, the particles start to decrease in potential energy because they are attracted to each other. At a certain distance, the potential energy reaches a minimum, after which any decrease in distance between the two particles increases the potential energy of the system.

When σ and ε values are reported in the literature, the values indicate interactions of a pure substance. For example, the σ of oxygen in the TIP3P forcefield of water, which is used in CHARMM (described below), is 3.151 Å, and ε of water is 0.152 kcal/mol.[30] This indicates that at 3.151 Å, the intermolecular forces between two oxygen molecules is zero, and that the energy well's depth is 0.152 kcal/mol. However, proteins are not a pure system, and therefore, require a method to determine Lennard-Jones parameters between two different atoms.

Combining rules are equations that calculate the interactions between two different atoms. There are different models of combination rules; Lorentz-Berthelot is used in molecular dynamics to analyze hard sphere systems. Hard-sphere molecules are those that follow the following rule:

$$V(r_1, r_2) = \begin{cases} 0 \; if \; |r_1 - r_2| \geq \sigma \\ \infty \; if \; |r_1 - r_2| < \sigma \end{cases} \qquad (8)$$

where *V* is the potential energy as a function of location, r_1 and r_2 are the position of the two particles, and σ is the diameter of the particles. The Lorentz-Berthelot equations for the Lennard-Jones parameters are as follows:

$$\sigma_{ij} = \frac{\sigma_{ii}+\sigma_{jj}}{2} \quad (9)$$

$$\epsilon_{ij} = \sqrt{\epsilon_{ij}\epsilon_{jj}} \quad (10)$$

Where epsilon or sigma for *ii* and *jj* represent the parameters of two different particles, *i* and *j*, that interact with itself. While it does have its shortcomings, the Lorentz-Berthelot rules are still used widely in computational simulations to calculate the parameters for the Lennard-Jones model for interactions of different molecules.[31,32]

CHARMM also accounts for external interactions from electrostatic interactions. This takes the form:

$$U = -\frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \quad (11)$$

Where *U* is electrostatic potential energy, ε_0 is the vacuum electric permittivity (or electric constant), q_i is the charge of a particle, *i*, and q_j is the charge of particle *j*, and r_{ij} is the distance between the two particles. The electrostatic model mathematically shows how particles attract each other when they are opposite in charges and repel each other when the particles have the same charge.

The CHARMM equation yields potential energy which can be related to force using the following equation:

$$F(x) = -\frac{dU}{dx} \quad (12)$$

where *F(x)* is force as a function of location, and $\frac{dU}{dx}$ is the derivative of potential energy as a function of location.

Adaptive Bias Force (ABF)

One of the objectives of this study is to determine what happens when ions enter a voltage gated sodium ion channel. However, given the constraints of time, in many of the systems the simulation may not be long enough to allow for an ion to sample the entire length of the channel. Alternatively, in some cases such as the calcium ion, it is also plausible that the ion does not pass the EEEE binding site, because there is not enough force to surmount the energy barrier of calcium ions and amino acid residue interactions. ABF is a force method that allows the system to escape any "kinetic traps" that the simulation would otherwise be stuck in till the time steps run out.[33] In doing so, ABF is a powerful and versatile tool that allows users to increase sampling even in a limited timescale. It does so by applying a specified force along a transition coordinate, thereby reducing the overall energy needed for calcium to enter the voltage gated sodium ion channels.

Computing Free Energy from Positions

The simulation will yield the location of the particle as a function of time. The CHARMM forcefield, through Lennard Jones and other force models, account for the enthalpic properties of the system. To analyze free energy, there also needs to be an entropic component analyzed. The output files can be used to convert position as function of time into probability as a function of time. The probability accounts for the entropy, or disorder, of the system. The equation for free energy is

$$G = -RT\ ln\ (p(x)) \qquad (13)$$

where G is Gibbs free energy, R is the gas constant, T is the temperature, and p is the probability that a particle will be at a certain location, x.[27] Considering that the output files yield position as

a function of time, the probability can be derived by determining the ratio frequency of a particle being at a location for the iterations of the simulation. Thus, by plotting the free energy of the system, as a function of location, this study will be able to gain further insight into the selectivity of ion diffusion in voltage gated sodium ion channels.

Methods

Molecular Dynamics of Biological Systems.

The purpose of this study is to further understand the energetics of diffusion in voltage gated sodium ion channels in *Magnetococcus Marinus* (Na_vMs). Specifically, this study aims to glean further information about the protein's selectivity against calcium ions, and the role of water as it traverses through the ion channel. To answer these questions, we will use molecular dynamics (MD). MD is a computational algorithm that can model the motion of particles in a system based on predetermined attractive and repulsive forces and Newtonian mechanics. By using MD, we can attain an atomistic and dynamic understanding of diffusion through Na_vMs.

This study will be using Nanoscale Molecular Dynamics (NAMD), an MD software that allows users to study larger systems of particles such as proteins. The two crucial pieces of information necessary to use NAMD is knowing where the particles are located, and how those particles interact with each other. These aspects are accounted for in the Protein Data Bank (PDB) and Protein Structure files (PSF) which will be examined in the section titled "Creating PDB Files" and "Capturing Interactions of Particles." These sections will further expand on how the location of each particle in a Na_vMs system is recorded, and how attractive and repulsive forces are modeled in a system. Furthermore, these sections will provide additional details about how the protein is placed in a bilayer lipid membrane, justifications for the lipid type, and the protonation state of ionizable residues.

When the protein is crystalized, the lipid bilayer membrane is not in the system. Instead, the membrane was added using CHARMM-GUI as described in the 'Creating PDB Files' section.[34] This can lead to the lipids and amino acids in the protein starting in unnatural positions. As a result, the initial protein-membrane system is not necessarily the most

energetically stable configuration of the system. To push the system into an energy minimum, the system needs to be minimized and equilibrated before running simulations that produce analyzable data. The process of minimization, equilibration and collecting data runs through a process called production runs will be detailed in the respective subsections of the section titled "Running a NAMD Simulation." Given that the ions do not always traverse through the channel during the simulation, additional force was applied to allow this to occur, as explained in the section titled "Adaptive Biasing Force (ABF)."

Creating PDB Files

NAMD is a computational software that implements MD with capabilities to handle memory and processing requirements for larger systems of molecules, thus, making it useful for simulating proteins. A set of coordinates of the system in question is necessary to run a NAMD simulation. For an MD simulation, these coordinates are stored in a PDB file. In the context of this project, PDB refers to files downloaded from the Protein Data Bank in a specific format.[35] Other than the coordinates, the PDB file contains information about the name of the protein, scientists who determined the structure, species from which the protein came, amino acid sequence, how the protein structure was determined, the location of secondary structures, and so on. The PDB file of proteins are based on experimental results from NMR or X-ray crystallography. Thus, the input file contains the coordinates of any particle that has also been crystallized including ions and water. Since X-ray crystallography cannot resolve hydrogen atom structures, any hydrogen atoms present in the PDB file are added from theoretical models and NMR analysis. The files can also include heterogeneous atoms that are used in the process of crystallization or any drugs that are bound to the protein. The PDB file also includes a list of "missing" residues; this means that the identity of the amino acid residue in the protein structure

could not be resolved during X-ray diffraction. They can be approximated in the system, however, as explained later. The coordinate files used in this study are the crystallized open conformation of the voltage gated sodium ion channels in *Magnetococcus marinus* [PDB code: 4F4L]. The crystallized Na_vMs had extra water molecules that were excluded in this study using CHARMM GUI, a graphical user interface that allows users to build complex systems.[34] Using the CHARMM GUI input generator, the location of the missing residues that are listed in the PDB file are also approximated using an algorithm known as GalaxyFill.[34,36]

The proteins are placed in a 1-palmitoyl-2-oleoyl-glycero-3-phosphocholine (POPC) membrane. POPC has been used to model the membrane environment in which voltage gated sodium ions exist, yielding computational results comparable to experimental results.[16] This is a simplistic model of the cell membrane that does not account for any interactions by glycolipids, cholesterols, etc. While there are some studies that suggest that the lipid constituents and elasticity affect proteins, these are not accounted for in this study and can be focused on in future studies by using CHARMM-GUI.[37] To account for appropriate orientation of the protein in the membrane, this study uses crystallized protein structures from the Orientation of Proteins in Membranes (OPM) database. Proteins in this database are aligned spatially with the membrane normal, which is defined as the z-axis. The POPC membrane is added using the CHARMM GUI membrane bilayer input generator using the replacement technique. The replacement technique packs proteins with lipid-like spheres which are then replaced with POPC. There is a mismatch in the number of lipids in the outer and inner leaflet of the bilayer due to the shape of the protein (Table 1).

Table 1: Number of lipids in the outer and inner leaflets. There is a mismatch between the upper and lower leaflets of a maximum of 14 lipids due to the shape of the protein.

Protonation State	Upper Leaflets	Lower Leaflets
Calcium Ions		
Deprotonated	60	71
Doubly Protonated	60	71
Sodium Ions		
Deprotonated	68	74
Doubly Protonated	60	67

The protonation states and any disulfide bonds can also be modified. There are no cysteine residues in the Na_vMs structure that are close enough to form a disulfide bond. However, the protonation states needed to be modified. The input generator assumes the neutral charge of the residues unless stated otherwise. As aforementioned in Na_vMs, the selectivity pore has an EEEE ring whose protonation state can change. This ring has one glutamate residue on each of the four chains that face inwards towards each other when none of the residues are protonated (Figure 1).[19,38] At a physiological pH of 7.4, the glutamate residue can exist in 3 different states: singly protonated, doubly protonated and fully deprotonated. The conductance of the ion channel depends on the protonation state.[19] In other words, the rate of diffusion of ions is related to the protonation states of the EEEE ring structure. A fully deprotonated state is correlated to faster diffusion with a decrease in the rate of diffusion with each additional proton. This study aims to further understand the effects of the protonation states on the free energy of ion diffusion by modeling the doubly protonated and deprotonated states at a concentration of 0.15 M Na^+ or Ca^{2+}.

Figure 1: Looking down the Na_vMs channel (green), the EEEE ring faces inwards towards the pore opening (shown using a VDW configuration) and is emphasized with atomistic detail. The image was generated using VMD and represents the equilibrated structure.[22]

Capturing Interactions of Particles

In addition to a file documenting the location of each atom in the system, NAMD also requires an input file that contains information about the force field that is to be applied on each particle in the system; this file is called a protein structure file, or a PSF file. The PSF file is generated from the PDB file and a force field topology file. In this study, we use the CHARMM36 force field for proteins and lipids, and TIP3P water model (Equation 6).[39]

In a computational simulation, the CHARMM force field is integrated through two files: parameter files and topology files. The parameter files contain information about bonded and

nonbonded interactions between the atoms described in the topology files. Specifically, they contain the parameters, which are numerical constants, necessary for the CHARMM force field. These include the Lennard-Jones parameters, k_b, k_θ, k_ϕ, k_ω, and k_u. By using Equation 6, these constants can be used to calculate the forces and energies acting on the particles. The topology files contain information regarding the name, type, mass, and partial charge about ions and every atom in amino acid residues, lipids, and water. The topology files also identify which atoms are connected to each other, the angles of the atoms, and so on. In essence, the topology files contain all the connectivity information about bonds which complements the potential energy functions in the parameter file. By identifying the residues unique to Na_vMs and Nav in the PDB file, a PSF file is generated to describe the partial charges, mass, and connectivity of each atom in the system.

It is important to note that while the CHARMM force field typically uses the Lennard-Jones model to simulate nonbonded interactions, some non-bonded interactions are inaccurate when modeled using the Lennard-Jones model and can exaggerate the attractive forces. For example, the attraction between charged and hydrophobic groups are exaggerated in comparison to experimental models.[40] The force fields of these interactions have been corrected using the Non-Bonded Fix (NBFIX) parameters to account for the interaction between the ions in the system and the carbonyl atoms. This is done to better capture the bulk properties of ion interactions which is essential for this study in due to the focus on ion diffusion.

Running NAMD Simulations

Once the system has been set up with all necessary components to run the NAMD simulations, a final file called the configuration needs to be created. The configuration file is written using Tcl scripting language, and it contains all information needed to run an MD

simulation. This includes information such as the location of the PDB files, the PSF files, the location of the parameter files, and any other files that it needs to refer to. To answer the questions that pertain to this study, twelve different configuration files have been created, each of which is a different combination of species (Na_vMs), Glu53 protonation state (singly, doubly, or deprotonated), and ion type (calcium or sodium ions).

In addition to the files that are read into the system, the configuration files also detail any information that is needed to model a system that approximates thermodynamically relevant conditions. In this study, we use periodic boundary conditions which means that the system that is created acts as a unit cell that is a part of an infinitely large system extending in the x, y and z direction. In this system, the boundaries are rectangular in shape. The center and size of each side of the box was determined from VMD using the TkConsole system. In this study, the system had the coordinates for all particles, including water, wrapped around the periodic boundary conditions. This means that if a particle crosses the periodic boundary in any vector direction, the particle is translated to the other side of the cell. This ensures that the number of the particles in the system stays constant.

The pressure and temperature of the system also stay constant in this system. The temperature is set to a constant of 310 K using Langevin dynamics. Langevin dynamics is an approach to modeling macromolecular systems that accounts for a frictional and stochastic term, thereby modeling particle dynamics more accurately.[41,42] This temperature is chosen because it is around physiological temperature, and therefore has biological relevance. The pressure is held constant by turning on the Langevin Piston at a constant pressure of 1.01325 bar. The pressure corresponds to atmospheric pressure. Keeping the number of particles in the system, the pressure

and the temperature of the system constant means that the system is set to align with the isothermal-isobaric ensemble (also called the NPT ensemble).

The vibrational energy of water is not accounted for in this study. The bonds of the water are kept rigid rather than allowing for harmonic vibrational motion. Thus, this study does not account for any energetic fluctuation caused by the vibrational motion of water; however, as the magnitude of this energy is significantly smaller than that caused by translational motion, we do not anticipate any significant deviations. Furthermore, the CHARMM force field has parameterized water in its rigid conformation.

The configuration file also contains information about the distance that forces can be felt. The cutoff value is 12 Å, beyond which the electrostatic and van der Waals forces between particles are neglected. The switching function is turned on at 10 Å for this study. This function ensures that the van der Waals and electrostatic forces are decreased gradually, rather than an abrupt truncation of the forces at 12 Å.

Once these conditions have been set, four simulation types are run sequentially using the simulation parameters described above: minimization, equilibration, production, and adaptive bias forcing.

Minimization

The initial structure read in from the PDB file does not necessarily have the most stable configuration of the protein. To find a stable configuration of the system, it needs to undergo a process known as minimization. This process pushes each particle to a local energy minimum by calculating the energy associated with various positions of each atom. Equilibrating a system without minimizing it might lead to structures that do not align with reality. In this study, the system was minimized for 6 picoseconds in a series of 2 minimization steps. First, the structure

is minimized by holding everything, but the tails fixed without motion (5 picosecond). Then, everything, but the protein was released and minimized for (1 picosecond), thus accounting for the solvent which is water.

Equilibration

Once a local minimum is discovered, the system is equilibrated. Equilibration is different from minimization because it is an attempt to attain equilibrium using molecular dynamics. As mentioned in the 'Introduction,' the CHARMM potential can be related to force using the following equation:

$$F(x) = -\frac{dU}{dx} \qquad (12)$$

where *F(x)* is force as a function of location, and $\frac{dU}{dx}$ is the derivative of potential energy as a function of location. Given force, the acceleration of each particle can be solved using Newton's law of motion:

$$F = ma, \qquad (2)$$

where *m* is the mass of the particle, and *a* is the acceleration of the particle. Now, solving the kinematic equations (Equations 3 and 4), the distance traveled by each atom can be calculated. Equilibration is conducted in an NPT ensemble in 3 steps, as detailed below.

For this study, the system is equilibrated sequentially in three parts under NPT conditions. The equilibration is done in three parts so that the simulation does not crash. First, everything else but the lipid tails are held constant, and equilibrated for 25,000 fs. Then, everything else but the protein is equilibrated for 2,500 fs. Here, the first time step is the last time step of the first equilibration procedure. Finally, everything in the system is released, and equilibrated for 250,000 fs. The timesteps for every equilibration is 1fs/step. The achievement of

equilibration is measured by ensuring that energy of the system as a function of time has no significant slope. A slope of zero indicates that the energy of the system has no significant fluctuations.

Production Runs

Once the system has been equilibrated, the system is ready for production runs. When the simulation is run to produce analyzable results, it is called a production run. This differs from the equilibration process in two ways. First, the equilibration process is shorter than the production run. Second, the equilibration process is run to reach an energy equilibrium, so that the system that is being analyzed does not have widely fluctuating energies. The production run is set for 2.5 million fs under the same conditions that the equilibration is conducted. A summary of the simulation conditions is listed in Table 2. Simulation run times for all systems are listed in Table 3.

Table 2: Simulation conditions for Equilibration and Production Run Conditions.

Ensemble	Isothermal-Isobaric (NPT)
Pressure (bar)	1.01325
Temperature (K)	310
Cutoff (Å)	12
Timestep (fs/step)	1

Table 3: Simulation run times for the production and equilibration of different systems.

	Timestep (ns)	
Protonation State	Equilibration	Production
Sodium Ions		
Deprotonated	0.3	15
Doubly Protonated	0.3	25

Calcium Ions		
Deprotonated	0.3	5
Doubly Protonated	0.3	5

Results and Discussion

Structure of the Protein

The voltage gated sodium ion channel ion consists of two regions: the pore domain and voltage sensor domain. The pore domain is the region that allows for ion passage while the voltage sensor domain is the region of the protein that allows for the voltage-gated mediation of ion transfer through the protein. A functional voltage gated sodium channel would fluctuate between an open and closed configuration dependent on the movement of the voltage sensor domains. This study is modeling voltage gated sodium channels in *Magnetococcus Marinus*, a prokaryotic organism. In prokaryotic organisms, the voltage gated channels are homotetrameric, with each monomer contributing to both the pore domain (S1-S4) and the voltage sensing domain (S5 and S6, and the P1 and P2 helices).[19]

McCusker et al., created and crystallized a structure without the voltage-sensing region.[43] The C-terminal domain of their structure has a disordered region whose residues were identified but the residual locations have not been due to stacking. Consequently, most studies that use the 4F4L structure exclude this region in their studies. However, this region is hypothesized to play a stabilizing role for the tetrameric structure.[43] We included this region in our study using GalaxyFill, an algorithm in CHARMM-GUI, that can approximate the location of missing residues. Our structure was placed in a homogenous POPC membrane bilayer, as detailed in the 'Methods' section. The resulting structure has long tail-like ends at the end of the protein, which should open up to allow for ion diffusion (Figure 2; Table 4). However, the algorithm seems to bias the formation of loop structures, which resulted in one of the disordered regions looping back on itself into the protein. Thus, ions were blocked and did not diffuse through the channel

in the course of this study. The extracellular region of the protein is above Region A and the intracellular side of the protein is below Region D, as shown in Figure 2.

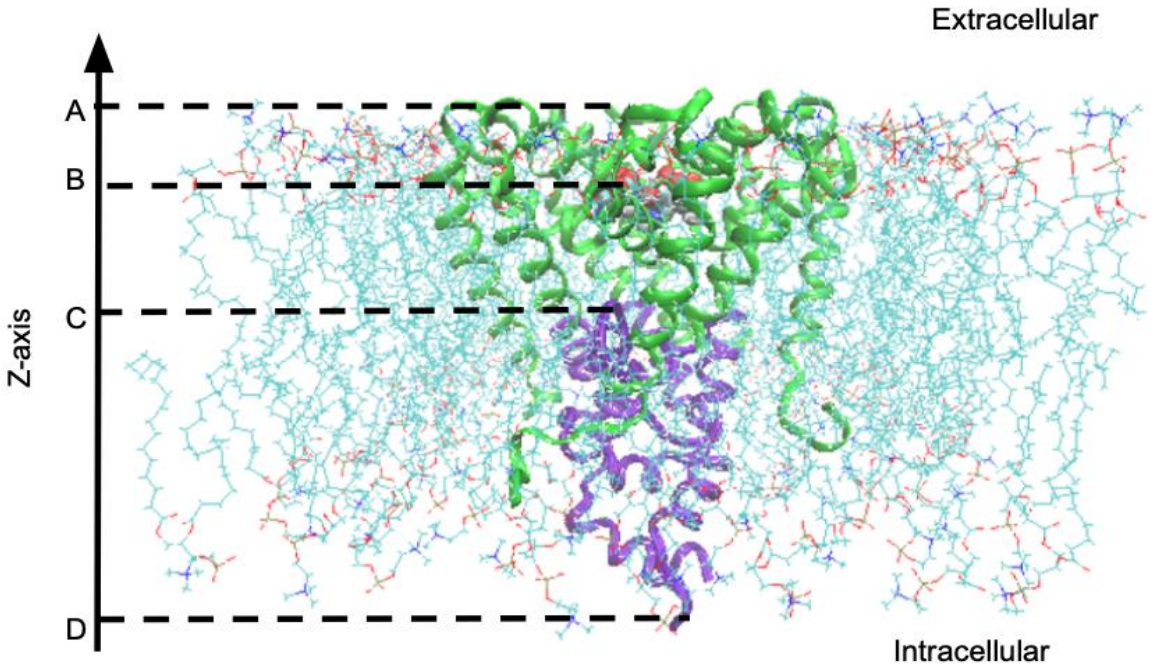


Figure 2. A voltage gated sodium ion channel (new ribbon configuration in green) placed in a POPC membrane (CPK configuration with polar heads shown in red and dark blue) after equilibration. The purple region in the protein corresponds to the disordered ends of the C-terminal. Locations are dependent on the systems and are recorded in Table 4. Region A corresponds to the top of the protein on the extracellular side of the membrane. Region B corresponds to the center of mass of the EEEE ring. Region C corresponds to the start of the disordered region, and Region D corresponds to the end of the disordered region which is also the end of the protein on the intracellular side.

Table 4: The z-coordinates of various regions in the system shown in Figure 2.

	Location (Å)			
System	Region A	Region B	Region C	Region D
Sodium Ions				
Deprotonated	26.9	11.9	4.14	-54.6
Doubly Protonated	25.4	11.5	-5.0	-38.0
Calcium Ions				
Deprotonated	26.4	13.2	-4.6	-51.2
Doubly Protonated	27.1	12.4	-6.1	-39.4
Sodium and Potassium				
Deprotonated	27.1	13.5	-0.6	-41.2

Free energy profiles of ions in the protein channel

The second area of focus in this study is how the free energy of sodium ions and calcium ions compare to each other. Given that the EEEE ring in Na_vMs can exist in three different protonation states, this study focuses on the two most extreme protonation states when it comes to ion conductance: deprotonated and doubly protonated.[19] Since significant sampling was not observed from just a production run, an adaptive biasing force (ABF) of 20 kcal/Å was applied to select ions in all systems along the z-axis in a radial sampling fashion (Supplemental Information).

Sodium and Calcium Ions in Deprotonated Na_vMs

The free energy of sodium and calcium ions was plotted as a function of where they are on the z-coordinate (Figure 3). The profile displays ions as they move from right to left, with the regions between 50 and 20 Å being the extracellular region, and -20 to -60 Å being the extracellular region. For both calcium and sodium, these regions are relatively flat with minor

fluctuations. As described in the “Hydration shells of ions” section, this is due to the energetic favorability of the hydration shell and solvation of ions in water. When the ions move closer to the protein, we see a dip in the free energy profile for both calcium and sodium at around 18 Å. Visually, this location corresponds to the EEEE binding site in the Na_vMs. The calcium ions do not sample the channel in the given production run time (5.3 ns) whereas the sodium ions do (15.3 ns). However, given the difference in simulation time, it is unclear as to whether this is a result of the sodium ions having a longer production time. The sampling events (and lack thereof) were confirmed using VMD. This is why we see a non-physical plateau for the calcium ions within the channel (Figure 3). In the free energy profile, it appears that there is a lower free energy of binding by the calcium ions to the EEEE ring. However, when this system is visually examined, the calcium ions never bind to the ring, and instead interact with the polar heads of the membrane at the same z-coordinate (Figure 4). This interaction points towards a limitation of using the z-coordinate as a coordinate for ion diffusion that will further be expanded on in the section titled “Limitations of this Study.”

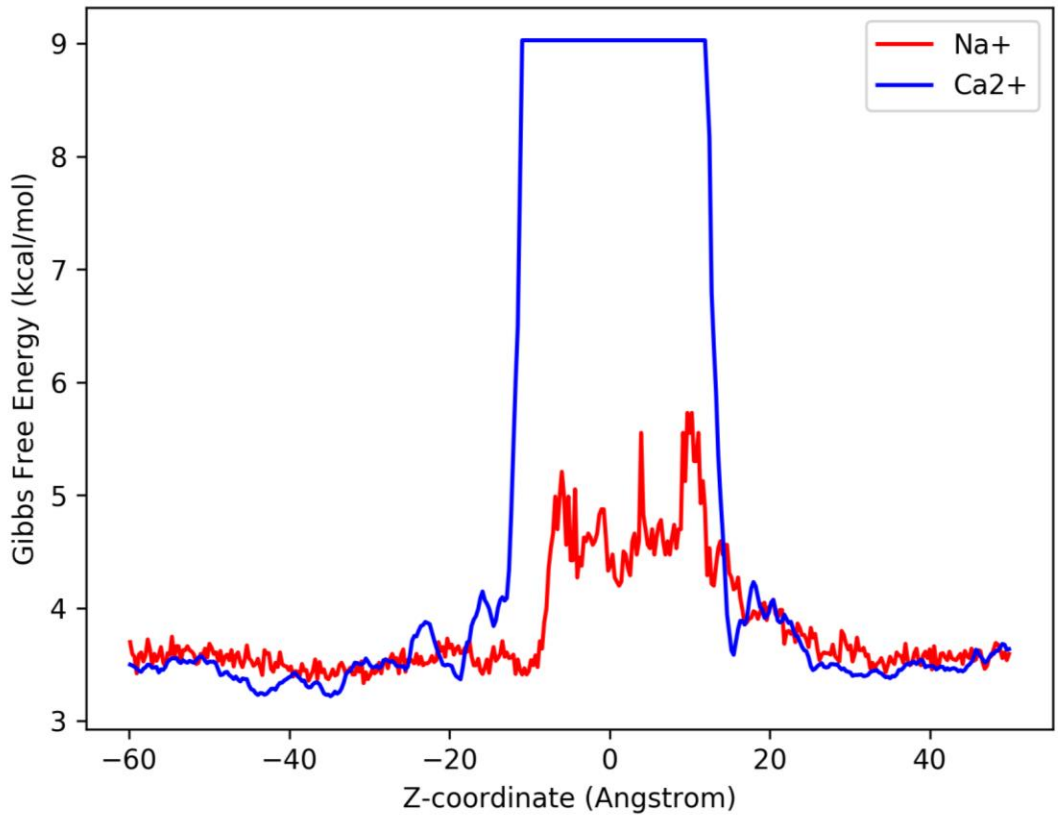


Figure 3. Free energy profile of production runs of deprotonated EEEE ring of homogenous 0.15 M Ca^{2+} and Na^{+} systems.

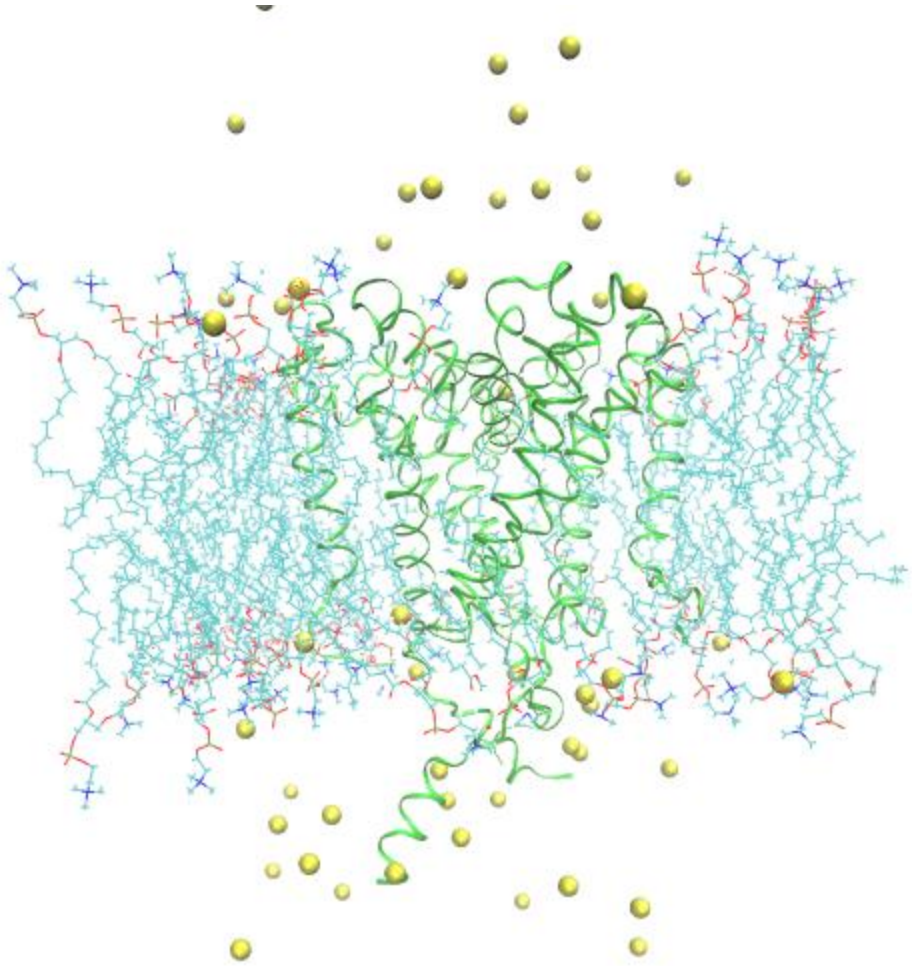


Figure 4. Calcium ions (yellow) interact with the polar heads of the membrane on the extracellular and intracellular side (shown in a Lines configuration with red and blue).

When it comes to the sampling of sodium ions in the channel, there are multiple dips in the free energy within the channel at around 0 Å and 10 Å. VMD is used to examine the events in the channel visually that may correspond to these dips and peaks. The dip at around 4 Å corresponds to the ion beginning of the disordered region of the C-terminal. When the system is examined visually using VMD, the disordered region seems to act as a plug that prevents the ions from diffusing to the other side of the protein. We see that the tails have a stabilizing effect on the ions and impact the hydration shell as delineated in the section titled "Hydration Shells of Ions." Additionally, that section will also show why the deprotonated system contains non-physical bonds that mar the experimental predictability of this model.

There is also an observable peak at around 10 Å in the sodium free energy curve. The center of mass of the EEEE ring is at 11.9 Å. Thus, the peak corresponds to the energy barrier that the sodium ion surmounts to pass the selectivity filter (SF) into the pore. Of note is that the free energy of binding for sodium ions is larger than that of the calcium's. However, this might be a result of the lack of sampling within the channel that biases the region to look like it has a lower free energy. When the channel was visually examined, it was interesting to note that there were three ions in the channel at once (Figure 5).

The presence of another ion in the channel does not seem to affect ions, and it is plausible that the presence of an ion plays a role in pushing the ion out of the channel. During the simulation, it appeared that even as the ion closest to the disordered region moved away from this region, it was repelled by the sodium ion right above it. However, this hypothesis requires further testing as the probability graph shows low sampling in regions where the ions move closer to each other (Figure 6). The simulation needs to be run longer to get more sampling in these locations.

We also see a dip in the free energy profile at around -10 Å. This region corresponds to the intracellular polar heads of the bilayer membrane and some parts of the disordered C-terminal region. The dip in the free energy profile indicates a stabilization by the polar heads of the POPC membrane, and the tails. This is confirmed visually using VMD where we are seeing the ion interact with the membrane and residues on the disordered region of the protein as opposed to inside the channel. The tails and the POPC bilayers act as a more stabilizing force in the deprotonated system with sodium ions than the EEEE binding site, as the free energy is lower in this region.

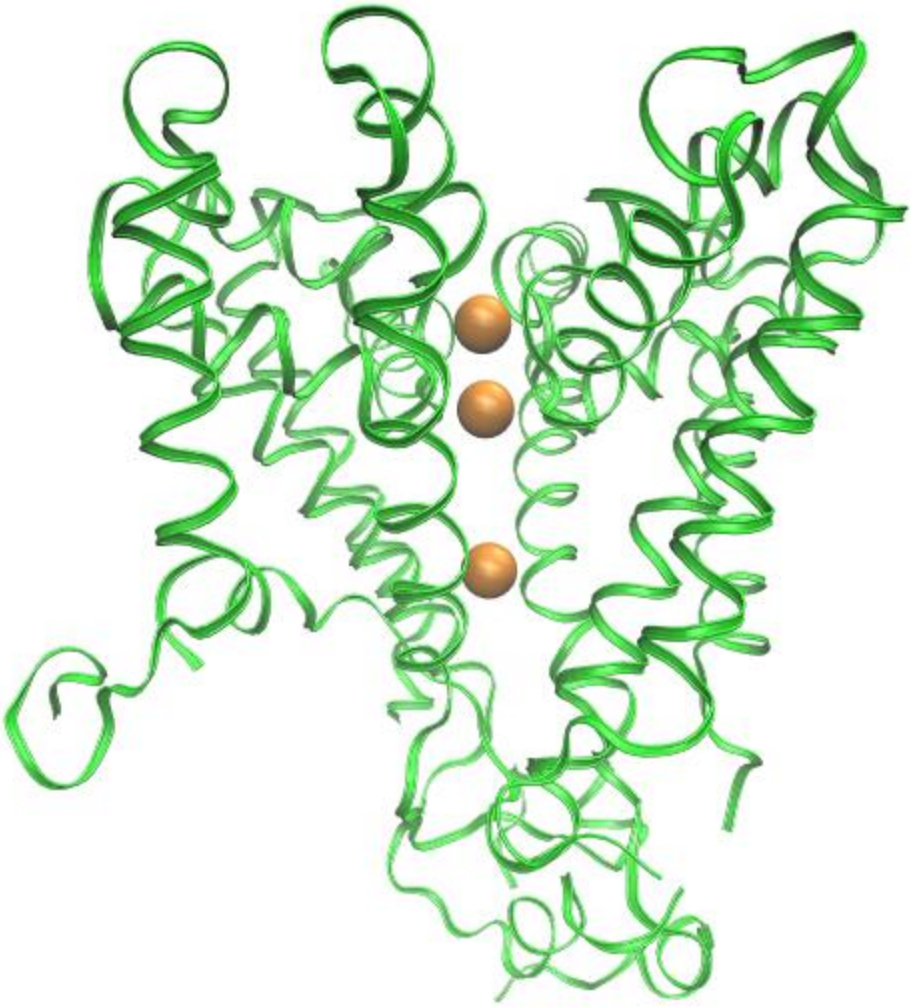


Figure 5. A VMD generated image of three sodium ions (orange) within a deprotonated Na_vMs (green) during the production run.

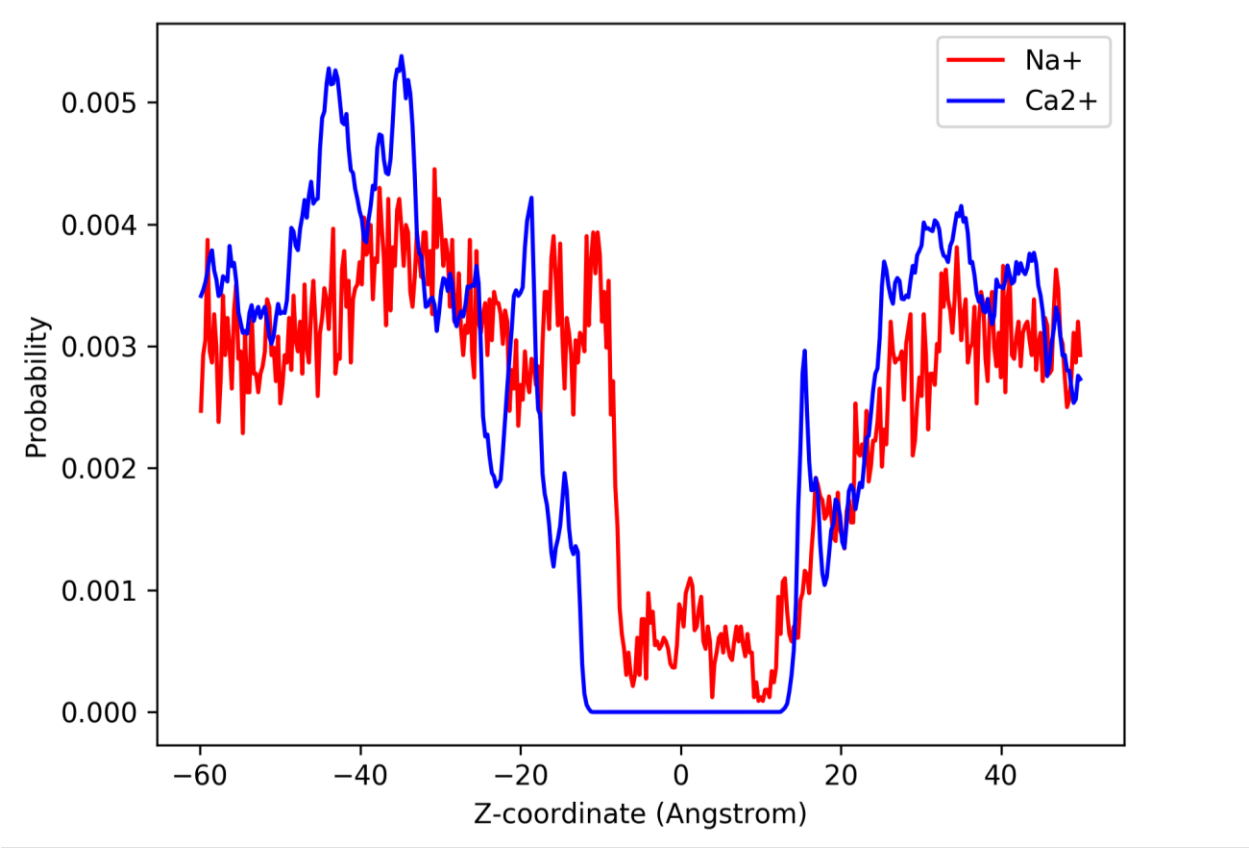


Figure 6. Probability of calcium and sodium ions being present at various locations along the z-coordinate.

Sodium and Calcium Ions in Doubly Protonated Na_vMs

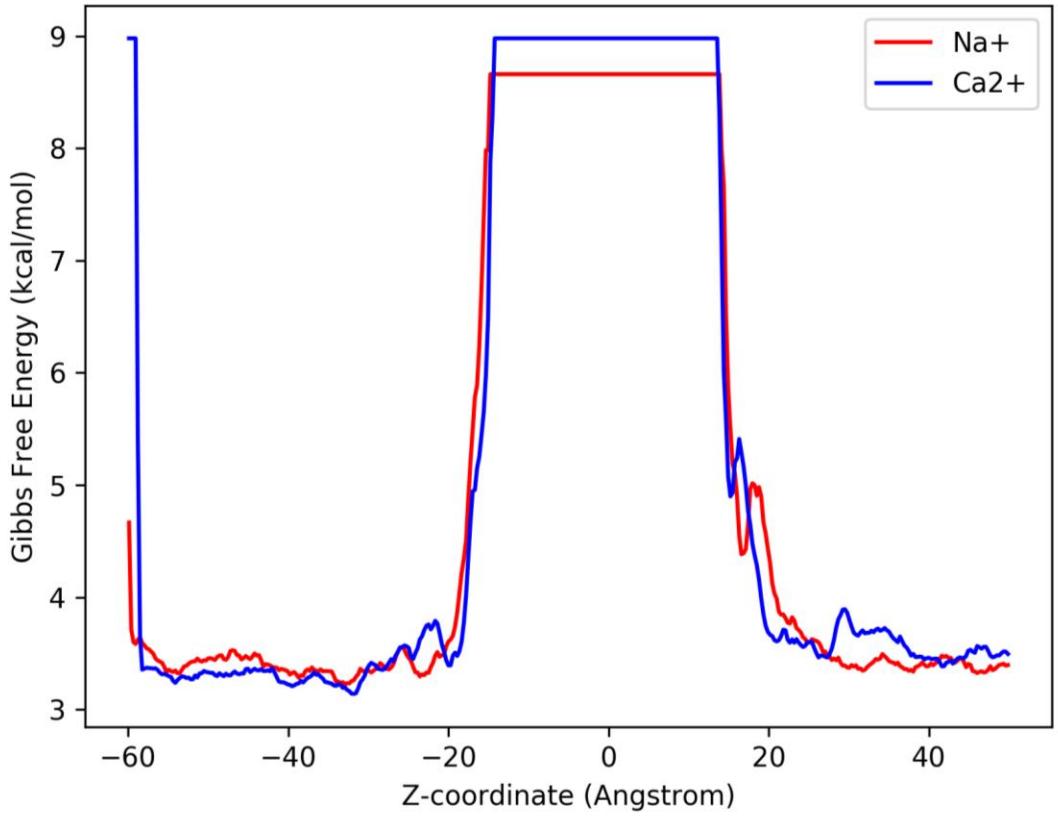


Figure 7. Free energy profile of production runs when the EEEE is doubly protonated in either a homogenous 0.15 M Ca^{2+} or Na^{+} system.

In a doubly protonated system, the ions don't sample the channel (Figure 7). The sodium and calcium ions both have a plateau in the region of the channel which is a non-physical value/profile of the free energy in the channel that look like each other. The similarity in the free energy profile points to the role of the EEEE ring as a funnel for the sodium ion.

Hydration Shells of Ions

As ions traverse the voltage gated ion channel, they are solvated in a layer of water known as the hydration shell. We used VMD to visualize the interaction between the ion and the water shell, and the water shell and amino-acid residues. In this study, we set a cutoff of 4.0 Å away from the sodium ion as a cutoff for the hydration shell. This value was chosen because prior studies have

determined that the radius of the hydration shell is 2.9 Å, and that there are 5-6 water molecules in the first layer of the hydration shell.[44,45] Using VMD, we were able to count about 5 to 7 molecules of water surrounding the sodium ions at the EEEE binding site at 4.0Å (Figure 8). For standardization purposes, we used the 4.0 Å cutoff for the calcium ions, as well. In a deprotonated EEEE ring, there is no visible shedding of the hydration shell when the sodium ion is in the binding site. The hydration shell seems to be larger in the EEEE ring than it does in the channel, especially during the partial shedding event near the tails.

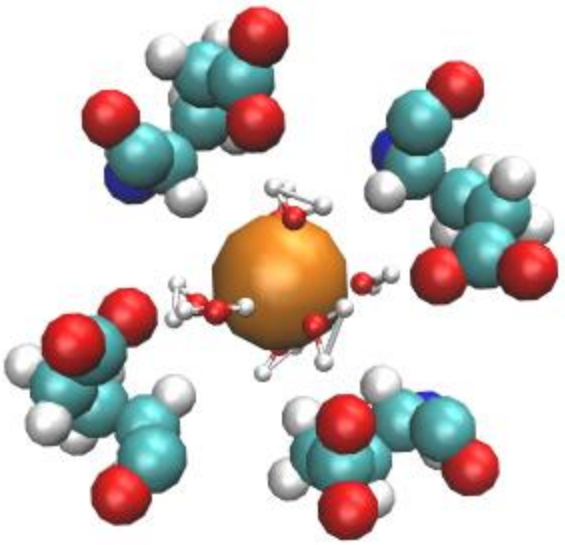


Figure 8. Sodium ion (orange) surrounded by 5 water molecules at 4.0 Å away from the ion. The ion is in the EEEE ring (shown with a VDW configuration) at the entrance of the channel. The protein is not shown here for visualization purposes.

We notice that there is a partial shedding event that occurs when the ion gets to the intracellular end of the protein near the disordered regions (Figure 9).[16,46] During the production runs, we see that there are more interactions with leucine and tryptophan, an aromatic amino acid when the ions move near the disordered region (Figure 10). There is an amino-acid ring penetration occurring (a GLU residue through a PHE ring) meaning that the result from this

system is non-physical. Prior computational calculation of phenylalanine interactions with water show that the carboxylic and amine region of the amino acid residue serve as sites for favorable interactions; however, given that the rings are surrounding the ion, it seems more likely that the steric hindrance is preventing the presence of many particles in the region.[47]

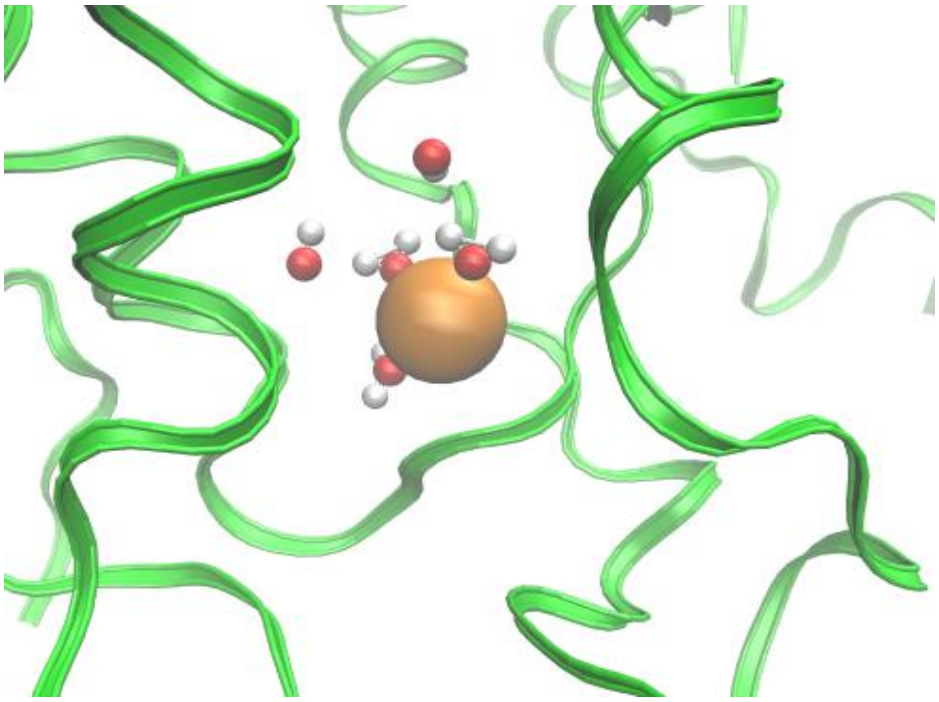


Figure 9. Sodium ion (orange) surrounded by waters within 4.0Å of the ion in a deprotonated ion channel during a production run. One side of the ion does not have water. The proteins are shown in green, in a ribbon style,

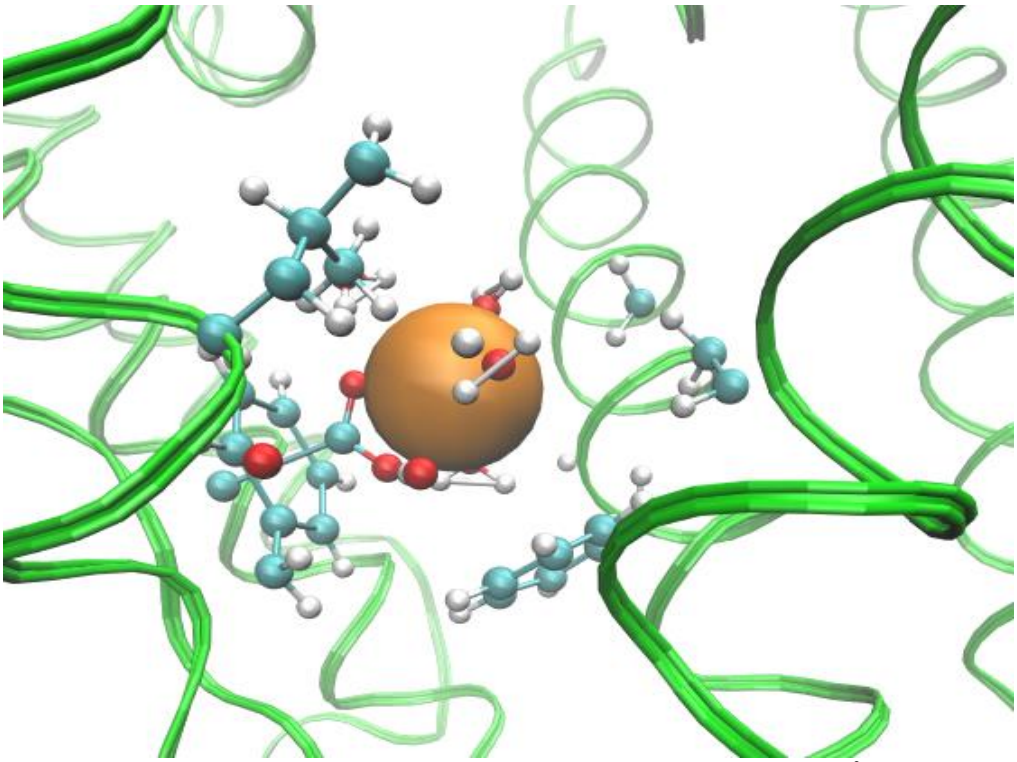


Figure 10: Sodium ion (orange) surrounded by waters within 4.0 Å of the ion in a deprotonated ion channel during a production run without ABF applied. One side of the ion does not have water. The amino acid residues that are less than 6.0 Å away from the ion are shown in a CPK style. These include tryptophan and leucine. Ring penetration occurs here.

We computed the free energy profile of all the sodium in the system and all the water molecules in the system from a production run (15 ns) with no forces applied to the system and plotted it as a function of z-coordinate (Figure 11). The energy in the extracellular region and the intracellular region plateaus. This makes sense given that in the extracellular region and intracellular region is filled with water. The water is modeled with a rigid TIP3P model which places a partial negative charge on the oxygen and partial positive charges on the hydrogen, to mimic the charge distribution in real water. The Coulombic interactions between the partial negative charge on the oxygen with the positive charge of the sodium cation are stabilizing, and consequently, are relatively lower in free energy. As sodium moves closer to the protein, we see an increase in the free energy. The noise in the figure is likely a result of poor sampling at those locations. We see a dip in the sodium's free energy profile at the binding EEEE binding site

which has a center of mass at approximately 18 Å. The Na_vMs protein spans from -37 to 34 Å. There is about a 1.5 kcal/mol free energy barrier for the sodium to move beyond the binding site into the center of the protein.

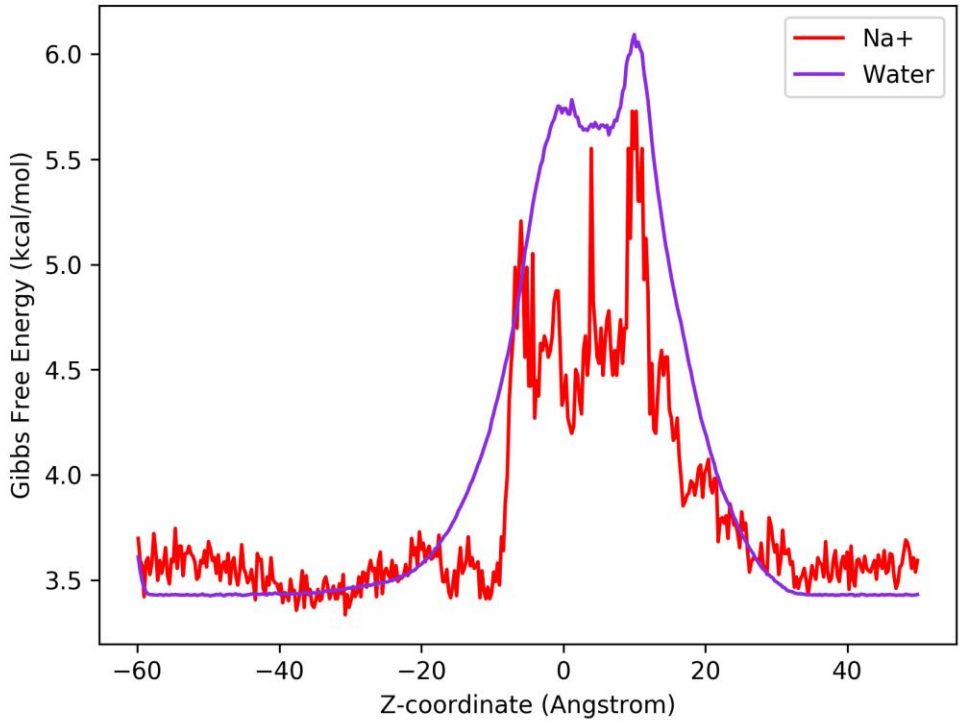


Figure 11. Free energy profile of sodium ions (Na+) and water as they move along the z-axis in a deprotonated Na_vMs system during a production run.

The free energy curve of water, however, does not display similar dips even at regions such as the binding spots suggesting that those locations are binding spots for the Na^+ rather than sodium and water. Prior studies have shown that the net diffusion of water is 0, and that the protein pore is water-lined.[16,46] Thus, the gradual increase and decrease reflect that the water is not engaging in binding events with the various amino acids. Rather, it seems that the narrowness of the protein channel acts as a funnel to reduce the amount of water present in the region due to steric hindrance, thereby, increasing the free energy. The dip that is observed at

around 0 Å is closer to the beginning of the tail blockage at the extracellular end of the protein. This is where we see a partial shedding event occur where one face of the ion is stripped of the waters. The tails exist in the z-coordinate range of 6 to -37.2 Å. There are two stabilizing forces that correspond to the dip at approximately 10 Å: the polar heads of the POPC membrane bilayer and the tails of the Na_vMs. At around 10 Å in the z-axis, we see that the sodium ions seem to be pulled towards the membrane, and that the ions are interacting with the amino acid residues in the tail.

There are no observable traversal events in the channels with calcium. However, using VMD, we can confirm that the calcium does bind to the EEEE ring. Thus, the question became what causes traversal in calcium to be an energetically unfavorable event? Prior studies have shown the hydration shell of calcium ions consist of 9-10 water molecules. To maintain a standard frame of comparison to sodium, this study only observed water molecules less than 4.0Å away from the calcium ion. Even at this distance, it is possible to see that the hydration shell for the calcium ions is larger than that of the sodium ion with 7 water molecules in comparison to the 5 water molecules, respectively (Figure 12). This is because calcium ions have a higher positive charge than the sodium ions, thus, having a higher attractive pull on water. The size of the hydration shell, thus, plays an important role in conferring selectivity to the Na_vMs against calcium ions. This can be examined quantitatively using the free energy profile of water in this system (Figure 13). Given that there is not a traversal event, it is not possible to glean accurate insight into the free energy profile of the calcium ions within the channel, as will be discussed in the section “Free energy profiles of ions in the protein channel.” However, there is a significant energy barrier that seems to inhibit the ion from sampling the channel after the EEEE

binding site at around 18 Å. The shape of the water's free energy profile curve remains the same even without the calcium in the channel.

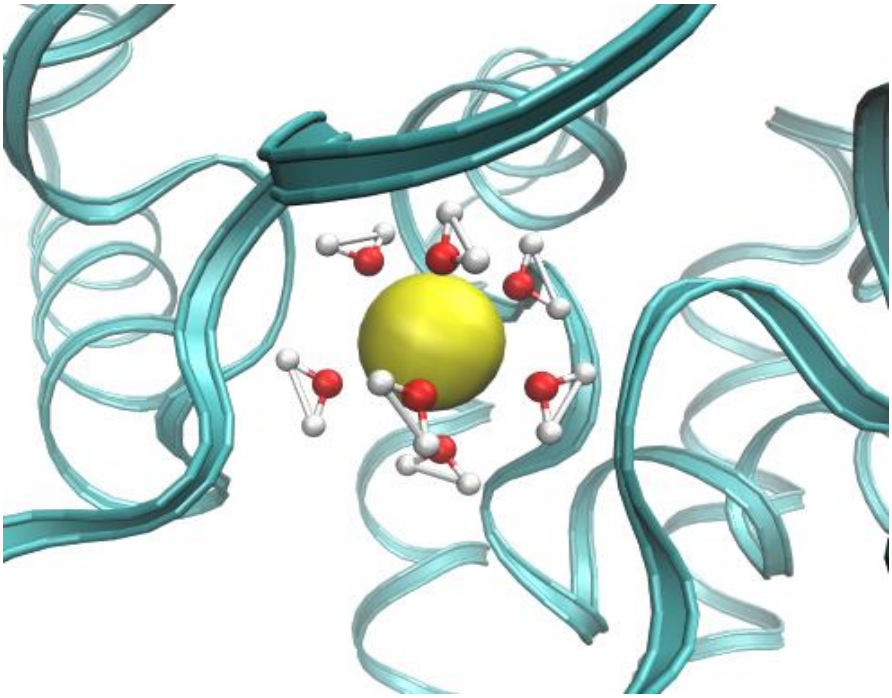


Figure 12. Calcium ions are surrounded by 7 water ions in the EEEE binding site of Na_vMs. The calcium ion is yellow; the water molecules are shown in a CPK configuration, and the protein is shown in a teal ribbon style configuration.

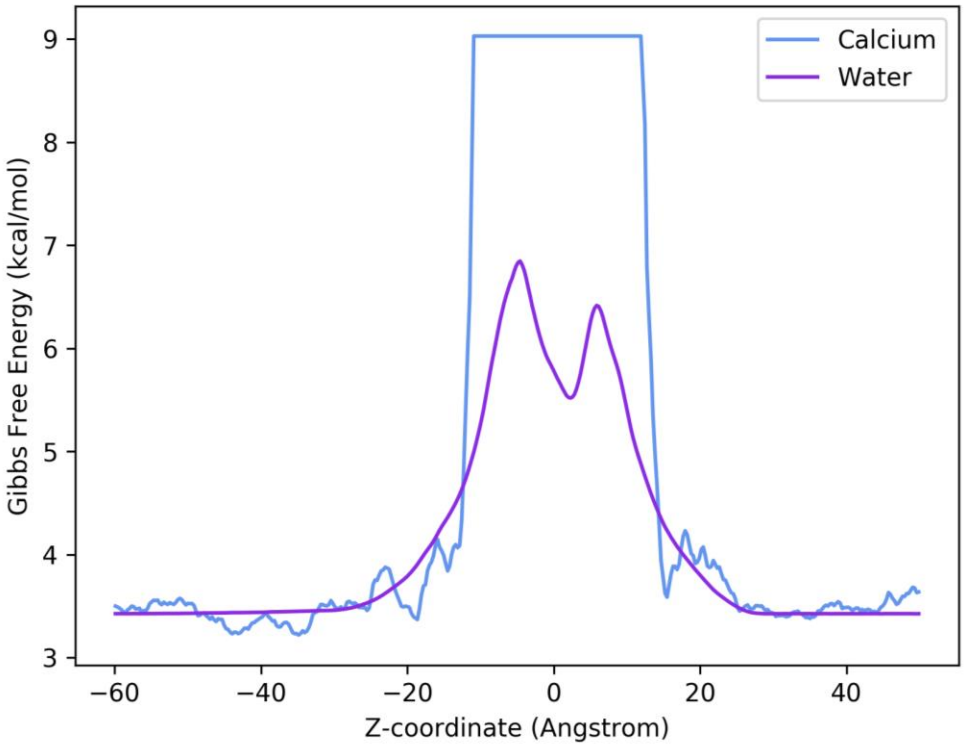


Figure 13. Free energy profile of the calcium ions and water as they move along the z-axis in a deprotonated Na_vMs system during a production run.

Effects of Protonation States of EEEE Ring on Ion Binding and Free Energy

Prior studies have examined the effects of the protonation states of the Glu residues in the EEEE ring ion conductance.[19] Chen et al. used a structure of Na_vMs, which was fully open, containing the VSD (PDB code: 5HVX). Results from this study suggest that there is a difference in the size of the selectivity filter, and the way that the ions bind at various sites. We expand on this by also examining whether the protonation state confers selectivity against other ions, in this case, calcium ions. In this study, we use the 4F4L fully open configuration of Na_vMs. The 4F4L does not have the voltage sensing domains nor the C-terminal domain but it does have residues for the disordered region of the C-terminal domain, but without

crystallographic information. We attempted to simulate the missing residues of the voltage sensing region within the 4F4L structures using GalaxyFill in order to assess the role of these residues in ion diffusion. However, given that the disordered region's residues block ions from diffusing, we specifically examined how the ion sampled near the binding site. We compared how the ions interact with the binding site in a protonated and deprotonated protein with sodium (Figure 14) and calcium ions (Figure 15).

When examining the sodium ions with the Na_vMs, we noticed that there is lower sampling within the protein pore for the doubly protonated structure than there is for the deprotonated one. Given that the deprotonated system has 15 ns of simulation time, and the doubly protonated protein has 2.5 ns of simulation time. Thus, given the varying simulation times, it is unclear whether the low sampling is a result of lower conductance or if conductance of ions is lower.

At the EEEE binding site, the free energy for the deprotonated ring is lower than that of the doubly protonated rings. However, due to the discrepancy in production run times and the low sampling, this question requires further examination to establish that the protonation and deprotonation is what results in the lower free energy. For the deprotonated structure, we see a stabilization event occurring at -10 Å which corresponds to the partial shedding of the hydration shell, and the interactions with the amino acid residues mentioned in 'Hydration Shells of Ions.'

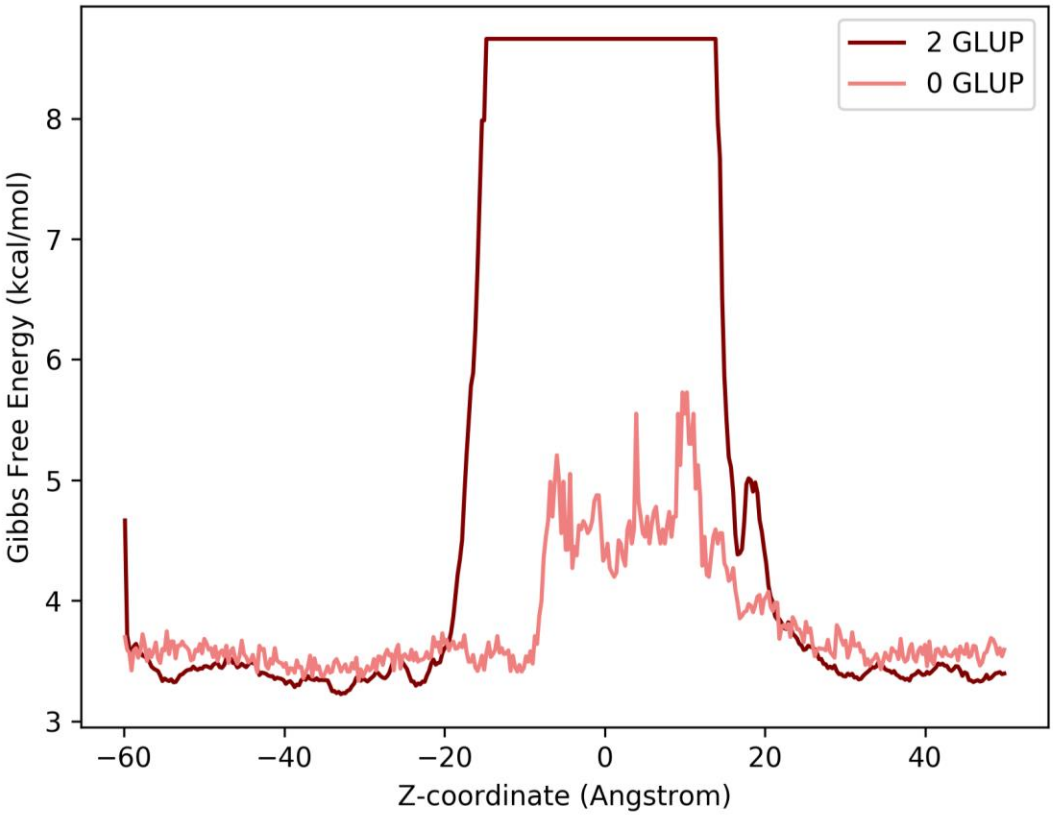


Figure 14. Free energy profile of Na^+ ions in a singly (0 GLUP) and doubly (2 GLUP) protonated EEEE ring.

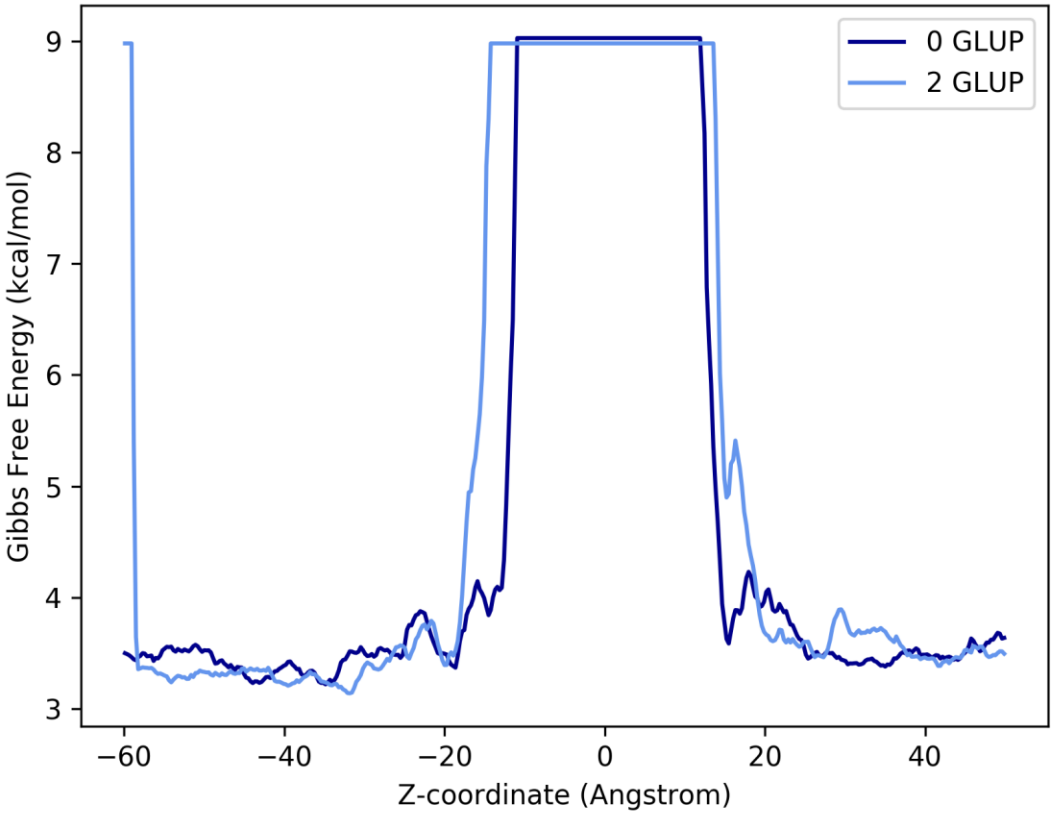


Figure 15. Free energy profile of Ca^{2+} ions in a singly (0 GLUP) and doubly (2 GLUP) protonated EEEE ring.

There are no traversal events of calcium ions in proteins with either a protonated or deprotonated EEEE ring. Thus, it is not possible to determine whether there are any interactions of the calcium ions once it is within the channel. The disordered region of the C-terminal has the same change in free energy when it comes to stabilizing the ions. At the EEEE binding site, the calcium ions seem to have a similar free energy in both the protonated and deprotonated structures. The similarity in free energy suggests that the protonation state does not affect the selectivity against calcium; however, due to the short sampling time (5 ns) further sampling and simulation runs are necessary to confirm this.

Differences in Homogeneous and Heterogeneous Systems

In most simulations that were conducted, we examined ions in the system homogeneously. In a biological environment such as a cell or organism, however, the protein is also exposed to other ions and substances. We simulated 0.15 M of KCl and 0.15 M of NaCl in this system and observed that the sodium ions never sample the channel. There are two potassium ions that sample the channel and bind to the EEEE binding site (Figure 16). However, the free energy of this sampling is significantly higher than the potassium ion in a homogeneous system (Figure 17). This suggests two possibilities: either it is energetically unfavorable for the potassium to sample the channel, or the heterogeneity contributes to the high free energy of the system. Since the sodium does not sample the channel in this simulation, we do not have a metric to compare the data to.

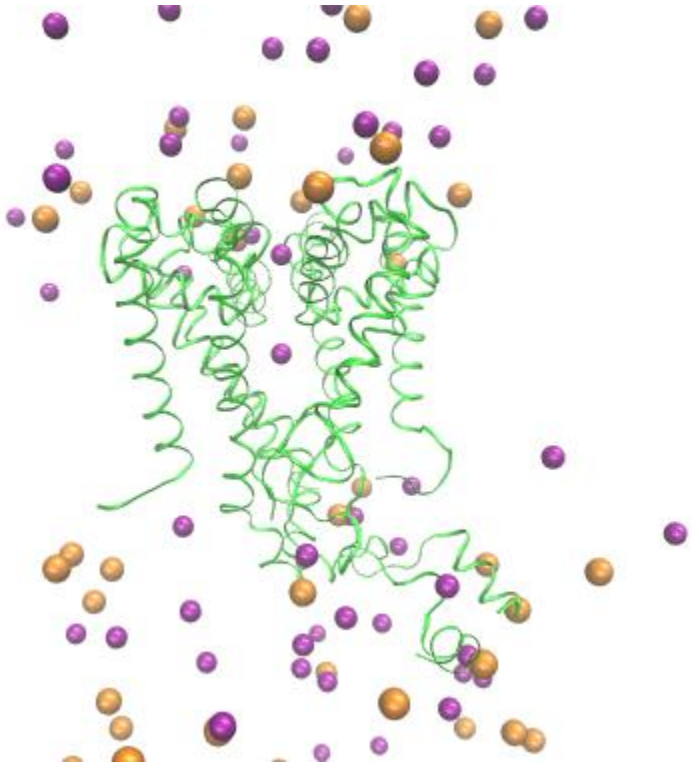


Figure 16. Potassium ions (purple) sampling the voltage gated ion channel (green) in a heterogeneous system of sodium (orange) and potassium ions.

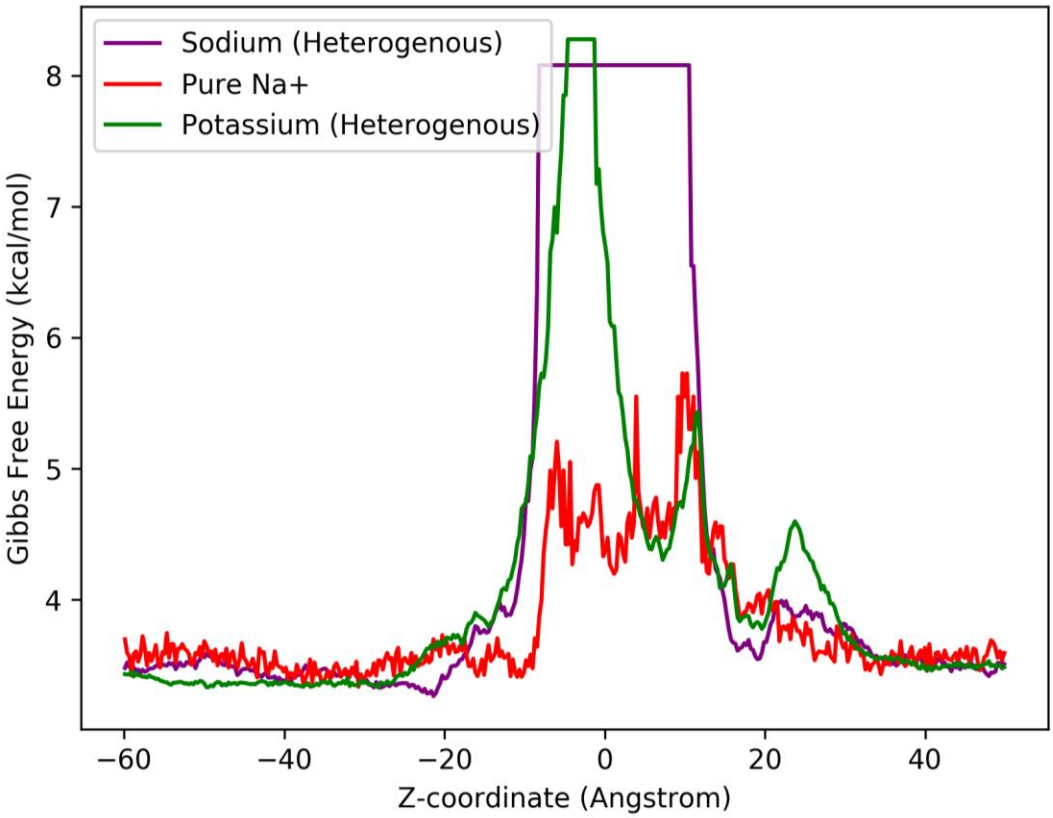


Figure 17. Free energy profile of a heterogeneous system of potassium ions (“Potassium (Heterogenous)”) and sodium ions (“Sodium (Heterogenous)”) in comparison to sodium ions from a homogenous system (“Pure Na+”).

Limitations of this Study

This study used GalaxyFill to approximate the location of the missing residues in the disordered region of the C-terminal. The resulting structure prevented ions from diffusing through the channel. Consequently, this study could not truly capture what occurs during a traversal event. It seems that this is largely due to the steric hindrance caused by the tails. McCusker et al. predicts that as the monomers (S6 helix, specifically), the disordered region should move to accommodate ion diffusion.[43] We do not see this event occurring, but this is most likely due to the approximations made by GalaxyFill. While GalaxyFill has built-in checks to prevent this, in the deprotonated system with sodium, we see a ring penetration event occurring which is not physically possible (Figure 10). Thus, GalaxyFill does not accurately capture an experimental reality.

We attempted to apply an adaptive biasing force (ABF) on all the systems to facilitate the diffusion event. An ABF acts in the opposite direction of the estimated mean force, thus requiring less energy by the ion to surmount any energy barriers in order to diffuse through the channel.[33] We defined the z-axis as the collective variable along which the force would be applied using the center of mass of the protein as the reference point. Ions are selected based on their proximity to the entrance of the protein. A lower boundary of -50 Å from the center of mass, an upper of boundary of 50 Å, and a force constant of 20 kcal/Å is implemented. Table A1 in the ‘Appendix’ provides information about the simulation run times.

When the ABF is applied to a select calcium ion in a deprotonated protein system, we see that the free energy is over 600 kcal/mol to move the ion along the z-coordinate (Figure A1 in Appendix). In a doubly protonated system, both the sodium and calcium ion also have a high free energy at over 125 kcal/mol (Figure A2 in Appendix). When examined visually, we noticed that

the ions are moving through the bilayer membrane (Figure A3). Despite the fact that an ion traversal through the membrane is energetically improbable, the ion was taking this path rather than traversing through the channel. This reveals two important limitations.

First, the z-coordinate might not be the most accurate collective variable along which traversal is measured. If we had used this definition without looking at the VMD simulation, it might have seemed like the energetics of diffusion is in the order of magnitude of 100 kcal/mol, which is not experimentally relevant. Thus, if particles have another path to move along the z-axis, then the profile does not solely capture ion sampling. A potential way to address this is by creating a constrained z-variable. The collective z-variable used in this study allows the selected ion to sample radially along the z-axis. If there is a way to constrain the variable so that the ion moves through a single axis without sampling radially (or making the radius significantly smaller), then it is possible that the ion might not move towards the membrane.

Second, the presence of the tails in the location assumed by the GalaxyFill algorithm constructs a disordered region that is more energetically unfavorable for the ion to traverse than a lipid bilayer. This points to the limitation of using this algorithm in our study. The values generated from this simulation are non-physical as ions have been shown to diffuse through channels. An alternative solution has been proposed in the 'Future Directions' section.

A second limitation to this study is the approximation of the hydration shell to 4.0 Å. In this study, we set a cutoff of 4.0 Å away from the sodium ion as a cutoff for the hydration shell. This value was chosen because prior studies have determined that the radius of the hydration shell is 2.9 Å, and that there are 5-6 water molecules in the first layer of the hydration shell.[44,45] Based on this, we used a standard value of 4.0 Å for both the sodium and calcium ions. Instead, in the future it might make more sense to establish a radial distribution function (RDF) to define

the hydration shell. An RDF curve measures the probability that one atom exists at a certain radius from another atom through the course of the entire simulation. By establishing an RDF curve, we can find the most probable location of waters from the ion, thus allowing us to glean a more comprehensive picture of how the hydration shell changes at various binding sites.

Future Directions

One potential future direction for this study is to use the 5HVX structure which contains a fully crystallized structure of the voltage gated ion channel with both the pore region and the voltage sensing region. Doing so would ensure that the location of residues would not need to be approximated and could provide an understanding of selectivity of the voltage gated sodium ion channel that has experimental relevance.

Another potential direction is to conduct a comparative analysis with human voltage gated sodium ion channels. The crystal structure of a human voltage gated sodium channel in an open configuration has not yet been determined. In humans, the Na_v is necessary for initiating action potentials, in electrically excitable cells.[48] Consequently, they have been implicated in diseases such as arrhythmia, epilepsy, and chronic pain.[49,50] These structures are more complicated than the bacterial voltage gated ion channel. Thus, they are more difficult to computationally model, and require more time. Determining a crystal structure of the open configuration of Na_v and performing a comparative analysis with the bacterial voltage gated ion channel using MD could provide insight into how accurate our models are in comparison to human voltage gated ion channels.

Conclusions

While the relative change in free energy within the channel for sodium ions is comparable to that from prior studies, the energetics of ion interaction with the disordered region does not accurately represent experimental results about voltage gated ions. The GalaxyFill software approximates residues in a manner that does not reflect reality possibly due to the fact this is an algorithm that seems to bias loop closure. Additionally, there is no voltage sensing region to regulate the opening and closing of this channel. Under these circumstances, it is not possible to glean any experimental data from this study. However, understanding the limitations of GalaxyFill is valuable for future research.

We also did not see any calcium enter the channel and could not gain any insight into what confers selectivity against calcium ions. We were able to determine that at 4.0 Å both calcium ions and sodium ions have a similar number of waters in the hydration shell. Using the definition that a hydration shell is 4.0 Å away from the ion, we have evidence to suggest that the hydration shell does not play a role in conferring selectivity. However, given that the charge of the calcium is more positive (+2 in comparison to +1), we would expect a larger hydration shell than that of sodium, and thus, we would need to redefine a hydration shell for it to have experimental relevance. We also notice that water does not have any significant binding events within the channel past the EEEE ring. It is interesting to note that the hydration shell is shed partially when the ions interact with the disordered region of the protein.

References

(1) Cooper, G. M. Transport of Small Molecules. *Cell Mol. Approach 2nd Ed.* **2000**.

(2) Chen, R.; Buyan, A.; Corry, B. Voltage-Gated Sodium Channel Pharmacology. In *Advances in Pharmacology*; Elsevier, 2017; Vol. 79, pp 255–285. https://doi.org/10.1016/bs.apha.2017.02.002.

(3) Yu, F. H.; Catterall, W. A. Overview of the Voltage-Gated Sodium Channel Family. *Genome Biol.* **2003**, *4* (3), 207. https://doi.org/10.1186/gb-2003-4-3-207.

(4) Ahmed, M.; Jalily Hasani, H.; Ganesan, A.; Houghton, M.; Barakat, K. Modeling the Human Nav1.5 Sodium Channel: Structural and Mechanistic Insights of Ion Permeation and Drug Blockade. *Drug Des. Devel. Ther.* **2017**, *11*, 2301–2324. https://doi.org/10.2147/DDDT.S133944.

(5) Splinter, R. Ultrashort Phenomena in Biochemistry and Biological Signaling. *J. Phys. Gen. Phys.* **2014**, *548*. https://doi.org/10.1088/1742-6596/548/1/012008.

(6) Lo, C.-J.; Leake, M. C.; Pilizota, T.; Berry, R. M. Nonequivalence of Membrane Voltage and Ion-Gradient as Driving Forces for the Bacterial Flagellar Motor at Low Load. *Biophys. J.* **2007**, *93* (1), 294–302. https://doi.org/10.1529/biophysj.106.095265.

(7) de Lera Ruiz, M.; Kraus, R. L. Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. *J. Med. Chem.* **2015**, *58* (18), 7093–7118. https://doi.org/10.1021/jm501981g.

(8) Heginbotham, L.; Lu, Z.; Abramson, T.; MacKinnon, R. Mutations in the K+ Channel Signature Sequence. *Biophys. J.* **1994**, *66* (4), 1061–1067.

(9) Kuang, Q.; Purhonen, P.; Hebert, H. Structure of Potassium Channels. *Cell. Mol. Life Sci.* **2015**, *72*, 3677–3693. https://doi.org/10.1007/s00018-015-1948-5.

(10) Biggin, P. C.; Smith, G. R.; Shrivastava, I.; Choe, S.; Sansom, M. S. P. Potassium and Sodium Ions in a Potassium Channel Studied by Molecular Dynamics Simulations. *Biochim. Biophys. Acta BBA - Biomembr.* **2001**, *1510* (1), 1–9. https://doi.org/10.1016/S0005-2736(00)00345-X.

(11) Carnevale, V.; Delemotte, L.; Howard, R. J. Molecular Dynamics Simulations of Ion Channels. *Trends Biochem. Sci.* **2021**, *46* (7), 621–622. https://doi.org/10.1016/j.tibs.2021.04.005.

(12) Qiu, H.; Shen, R.; Guo, W. Ion Solvation and Structural Stability in a Sodium Channel Investigated by Molecular Dynamics Calculations. *Biochim. Biophys. Acta BBA - Biomembr.* **2012**, *1818* (11), 2529–2535. https://doi.org/10.1016/j.bbamem.2012.06.003.

(13) Payandeh, J.; Minor, D. L. Bacterial Voltage-Gated Sodium Channels (BacNaVs) from the Soil, Sea, and Salt Lakes Enlighten Molecular Mechanisms of Electrical Signaling and Pharmacology in the Brain and Heart. *J. Mol. Biol.* **2015**, *427* (1), 3–30. https://doi.org/10.1016/j.jmb.2014.08.010.

(14) Ren, D.; Navarro, B.; Xu, H.; Yue, L.; Shi, Q.; Clapham, D. E. A Prokaryotic Voltage-Gated Sodium Channel. *Science* **2001**, *294* (5550), 2372–2375. https://doi.org/10.1126/science.1065635.

(15) Compton, E. L. R.; Mindell, J. A. Bacterial Ion Channels. *EcoSal Plus* **2010**, *4* (1). https://doi.org/10.1128/ecosalplus.3.3.2.

(16) Ulmschneider, M. B.; Bagnéris, C.; McCusker, E. C.; DeCaen, P. G.; Delling, M.; Clapham, D. E.; Ulmschneider, J. P.; Wallace, B. A. Molecular Dynamics of Ion Transport through the Open Conformation of a Bacterial Voltage-Gated Sodium Channel. *Proc. Natl. Acad. Sci.* **2013**, *110* (16), 6364–6369. https://doi.org/10.1073/pnas.1214667110.

(17) Furini, S.; Domene, C. Ion-Triggered Selectivity in Bacterial Sodium Channels. *Proc. Natl. Acad. Sci.* **2018**, *115* (21), 5450–5455. https://doi.org/10.1073/pnas.1722516115.

(18) Liebeskind, B. J.; Hillis, D. M.; Zakon, H. H. Independent Acquisition of Sodium Selectivity in Bacterial and Animal Sodium Channels. *Curr. Biol.* **2013**, *23* (21), R948–R949. https://doi.org/10.1016/j.cub.2013.09.025.

(19) Chen, A. Y.; Brooks, B. R.; Damjanovic, A. Determinants of Conductance of a Bacterial Voltage-Gated Sodium Channel. *Biophys. J.* **2021**, *120* (15), 3050–3069. https://doi.org/10.1016/j.bpj.2021.06.013.

(20) Liang, H.; Bourdon, A. K.; Chen, L. Y.; Phelix, C. F.; Perry, G. Gibbs Free-Energy Gradient along the Path of Glucose Transport through Human Glucose Transporter 3. *ACS Chem. Neurosci.* **2018**, *9* (11), 2815–2823. https://doi.org/10.1021/acschemneuro.8b00223.

(21) Lesgidou, N.; Eliopoulos, E.; Goulielmos, G. N.; Vlassi, M. Insights on the Alteration of Functionality of a Tyrosine Kinase 2 Variant: A Molecular Dynamics Study. *Bioinformatics* **2018**, *34* (17), i781–i786. https://doi.org/10.1093/bioinformatics/bty556.

(22) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14* (1), 33–38. https://doi.org/10.1016/0263-7855(96)00018-5.

(23) Nelson, M. T.; Humphrey, W.; Gursoy, A.; Dalke, A.; Kale, L. V.; Skeel, R. D.; Schulten, K. NAMD: A Parallel, Object-Oriented Molecular Dynamics Program. *Int. J. High Perform. Comput. Appl.* **1996**, *10* (4), 251–268. https://doi.org/10.1177/109434209601000401.

(24) Gapsys, V.; de Groot, B. L. On the Importance of Statistics in Molecular Simulations for Thermodynamics, Kinetics and Simulation Box Size. *eLife* **2020**, *9*, e57589. https://doi.org/10.7554/eLife.57589.

(25) Atkins, P. W.; De Paula, J.; Keeler, J. *Atkins' Physical Chemistry*, Eleventh edition.; Oxford University Press: Oxford, United Kingdom ; New York, NY, 2018.

(26) Landau, L. D.; Lifshits, E. M.; Pitaevskiĭ, L. P.; Landau, L. D.; Landau, L. D. *Statistical Physics*; Pergamon international library of science, technology, engineering, and social studies; Pergamon Press: Oxford ; New York, 1980.

(27) Spitznagel, B.; Pritchett, P. R.; Messina, T. C.; Goadrich, M.; Rodriguez, J. An Undergraduate Laboratory Activity on Molecular Dynamics Simulations. *Biochem. Mol. Biol. Educ. Bimon. Publ. Int. Union Biochem. Mol. Biol.* **2016**, *44* (2), 130–139. https://doi.org/10.1002/bmb.20939.

(28) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30* (10), 1545–1614. https://doi.org/10.1002/jcc.21287.

(29) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102* (18), 3586–3616. https://doi.org/10.1021/jp973084f.

(30) Tran, K. N.; Tan, M.-L.; Ichiye, T. A Single-Site Multipole Model for Liquid Water. *J. Chem. Phys.* **2016**, *145* (3), 034501. https://doi.org/10.1063/1.4958621.

(31) Delhommelle, J.; Millié, P. Inadequacy of the Lorentz-Berthelot Combining Rules for Accurate Predictions of Equilibrium Properties by Molecular Simulation. *Mol. Phys.* **2001**, *99* (8), 619–625. https://doi.org/10.1080/00268970010020041.

(32) Wennberg, C. L.; Murtola, T.; Hess, B.; Lindahl, E. Lennard-Jones Lattice Summation in Bilayer Simulations Has Critical Effects on Surface Tension and Lipid Properties. *J. Chem. Theory Comput.* **2013**, *9* (8), 3527–3537. https://doi.org/10.1021/ct400140n.

(33) Comer, J.; Gumbart, J. C.; Hénin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C. The Adaptive Biasing Force Method: Everything You Always Wanted To Know but Were Afraid To Ask. *J. Phys. Chem. B* **2015**, *119* (3), 1129–1151. https://doi.org/10.1021/jp506633n.

(34) CHARMM-GUI: A web-based graphical user interface for CHARMM - Jo - 2008 - Journal of Computational Chemistry - Wiley Online Library https://onlinelibrary.wiley.com/doi/10.1002/jcc.20945 (accessed 2022 -01 -25).

(35) PDB Files https://www.ks.uiuc.edu/Training/Tutorials/namd/namd-tutorial-unix-html/node22.html (accessed 2022 -01 -25).

(36) Coutsias, E. A.; Seok, C.; Jacobson, M. P.; Dill, K. A. A Kinematic View of Loop Closure. *J. Comput. Chem.* **2004**, *25* (4), 510–528. https://doi.org/10.1002/jcc.10416.

(37) Lundbæk, J. A.; Birn, P.; Hansen, A. J.; Søgaard, R.; Nielsen, C.; Girshman, J.; Bruno, M. J.; Tape, S. E.; Egebjerg, J.; Greathouse, D. V.; Mattice, G. L.; Koeppe, R. E.; Andersen, O. S. Regulation of Sodium Channel Function by Bilayer Elasticity. *J. Gen. Physiol.* **2004**, *123* (5), 599–621. https://doi.org/10.1085/jgp.200308996.

(38) Buyan, A.; Sun, D.; Corry, B. Protonation State of Inhibitors Determines Interaction Sites

within Voltage-Gated Sodium Channels. *Proc. Natl. Acad. Sci.* **2018**, *115* (14), E3135–E3144. https://doi.org/10.1073/pnas.1714131115.

(39) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell, A. D. CHARMM36m: An Improved Force Field for Folded and Intrinsically Disordered Proteins. *Nat. Methods* **2017**, *14* (1), 71–73. https://doi.org/10.1038/nmeth.4067.

(40) Yoo, J.; Aksimentiev, A. New Tricks for Old Dogs: Improving the Accuracy of Biomolecular Force Fields by Pair-Specific Corrections to Non-Bonded Interactions. *Phys. Chem. Chem. Phys.* **2018**, *20* (13), 8432–8449. https://doi.org/10.1039/C7CP08185E.

(41) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103* (11), 4613–4621. https://doi.org/10.1063/1.470648.

(42) Paquet, E.; Viktor, H. L. Molecular Dynamics, Monte Carlo Simulations, and Langevin Dynamics: A Computational Review. *BioMed Res. Int.* **2015**, *2015*, e183918. https://doi.org/10.1155/2015/183918.

(43) McCusker, E. C.; Bagnéris, C.; Naylor, C. E.; Cole, A. R.; D'Avanzo, N.; Nichols, C. G.; Wallace, B. A. Structure of a Bacterial Voltage-Gated Sodium Channel Pore Reveals Mechanisms of Opening and Closing. *Nat. Commun.* **2012**, *3* (1), 1102. https://doi.org/10.1038/ncomms2077.

(44) Carrillo-Tripp, M.; Saint-Martin, H.; Ortega-Blake, I. A Comparative Study of the Hydration of Na+ and K+ with Refined Polarizable Model Potentials. *J. Chem. Phys.* **2003**, *118* (15), 7062–7073. https://doi.org/10.1063/1.1559673.

(45) Yang, K.-L.; Yiacoumi, S.; Tsouris, C. Monte Carlo Simulations of Electrical Double-

Layer Formation in Nanopores. *J. Chem. Phys.* **2002**, *117* (18), 8499–8507. https://doi.org/10.1063/1.1511726.

(46) Payandeh, J.; Scheuer, T.; Zheng, N.; Catterall, W. A. The Crystal Structure of a Voltage-Gated Sodium Channel. *Nature* **2011**, *475* (7356), 353–358. https://doi.org/10.1038/nature10238.

(47) Gao, B.; Wyttenbach, T.; Bowers, M. T. Hydration of Protonated Aromatic Amino Acids: Phenylalanine, Tryptophan, and Tyrosine. *J. Am. Chem. Soc.* **2009**, *131* (13), 4695–4701. https://doi.org/10.1021/ja8085017.

(48) Glass, W. G.; Duncan, A. L.; Biggin, P. C. Computational Investigation of Voltage-Gated Sodium Channel B3 Subunit Dynamics. *Front. Mol. Biosci.* **2020**, *7*, 40. https://doi.org/10.3389/fmolb.2020.00040.

(49) Watanabe, H.; Koopmann, T. T.; Le Scouarnec, S.; Yang, T.; Ingram, C. R.; Schott, J.-J.; Demolombe, S.; Probst, V.; Anselme, F.; Escande, D.; Wiesfeld, A. C. P.; Pfeufer, A.; Kääb, S.; Wichmann, H.-E.; Hasdemir, C.; Aizawa, Y.; Wilde, A. A. M.; Roden, D. M.; Bezzina, C. R. Sodium Channel B1 Subunit Mutations Associated with Brugada Syndrome and Cardiac Conduction Disease in Humans. *J. Clin. Invest.* **2008**, *118* (6), 2260–2268. https://doi.org/10.1172/JCI33891.

(50) Bouza, A. A.; Isom, L. L. Chapter 14: Voltage-Gated Sodium Channel β Subunits and Their Related Diseases. *Handb. Exp. Pharmacol.* **2018**, *246*, 423–450. https://doi.org/10.1007/164_2017_48.

(51) Li, Y.; Liu, J.; Gumbart, J. C. Membrane Proteins Tutorial (Introductory). 28.

(52) Aksimentiev, A.; Sotomayor, M.; Wells, D. Membrane Proteins Tutorial. 43.

Appendix

Table A1: Simulation run times for the ABF runs of different systems.

Protonation State	Timestep (ns)
Sodium Ions	
Deprotonated	2.57
Doubly Protonated	22.50
Calcium Ions	
Deprotonated	16.48
Doubly Protonated	15.00

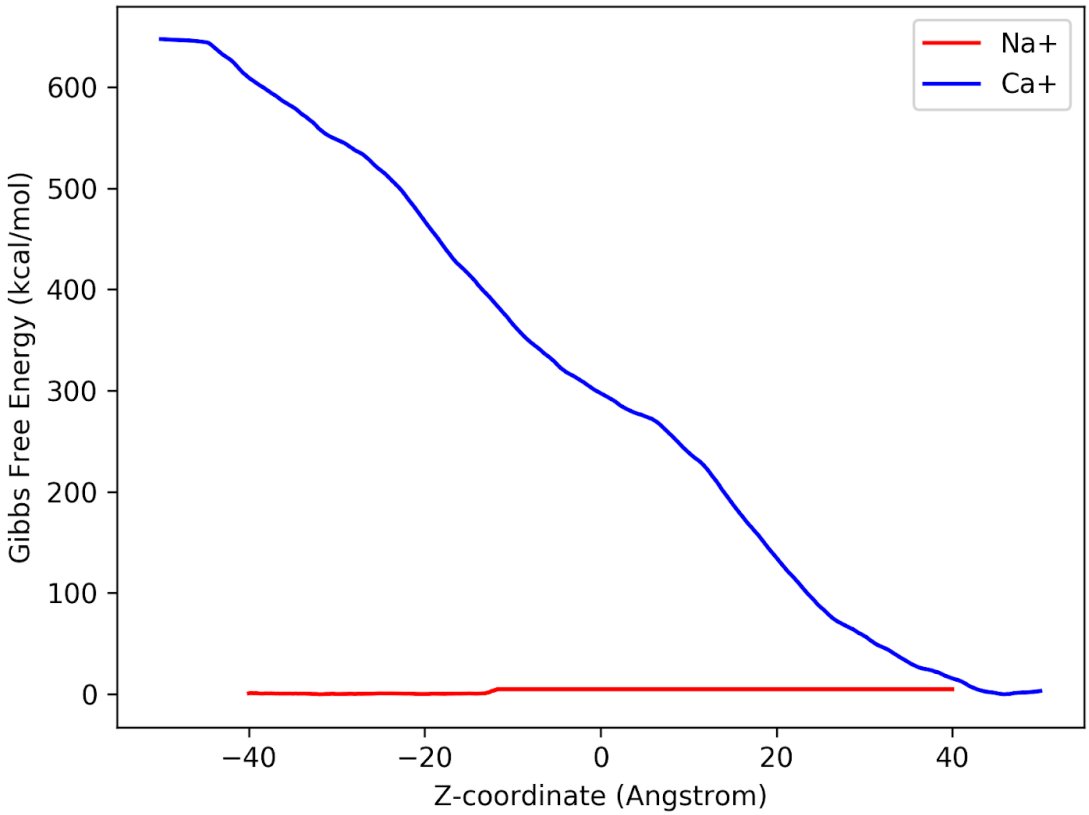


Figure A1. PMF profile of ABF runs of deprotonated EEEE ring of homogenous 0.15 M Ca^{2+} and Na^{+} systems.

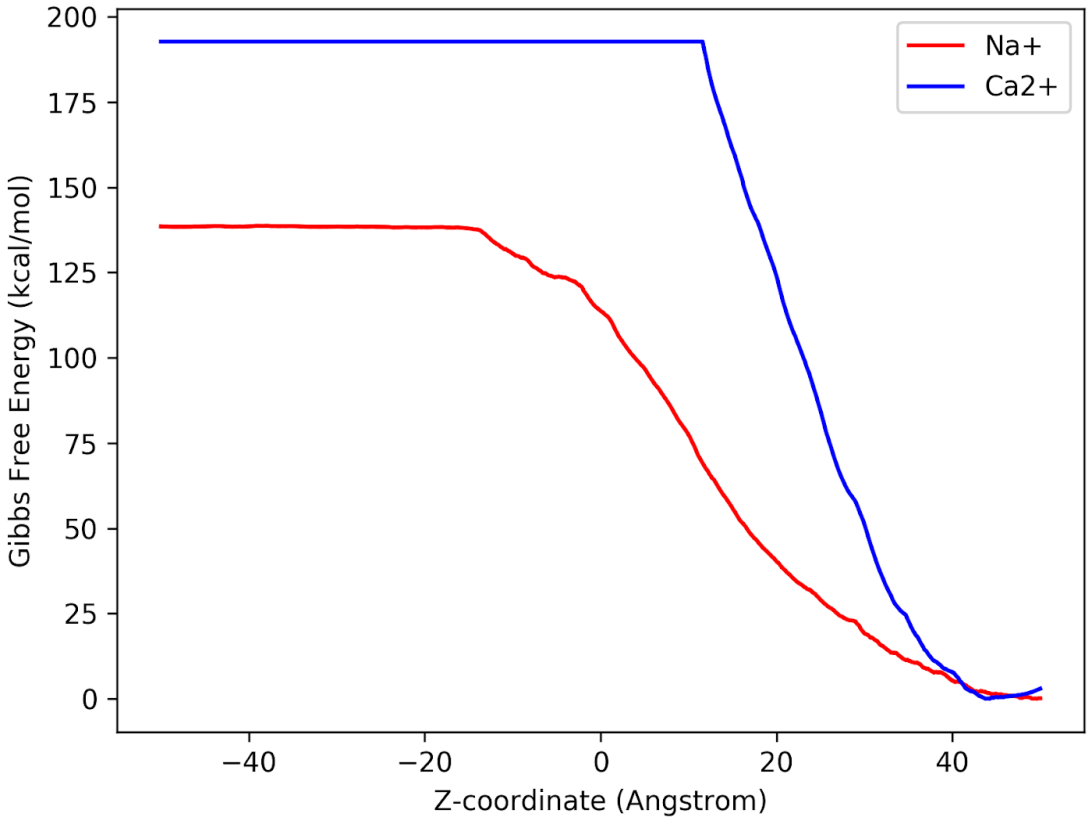


Figure A2. Free energy profile of ABF runs of a doubly protonated EEEE ring of homogenous 0.15 M Ca^{2+} and Na^{+} systems.

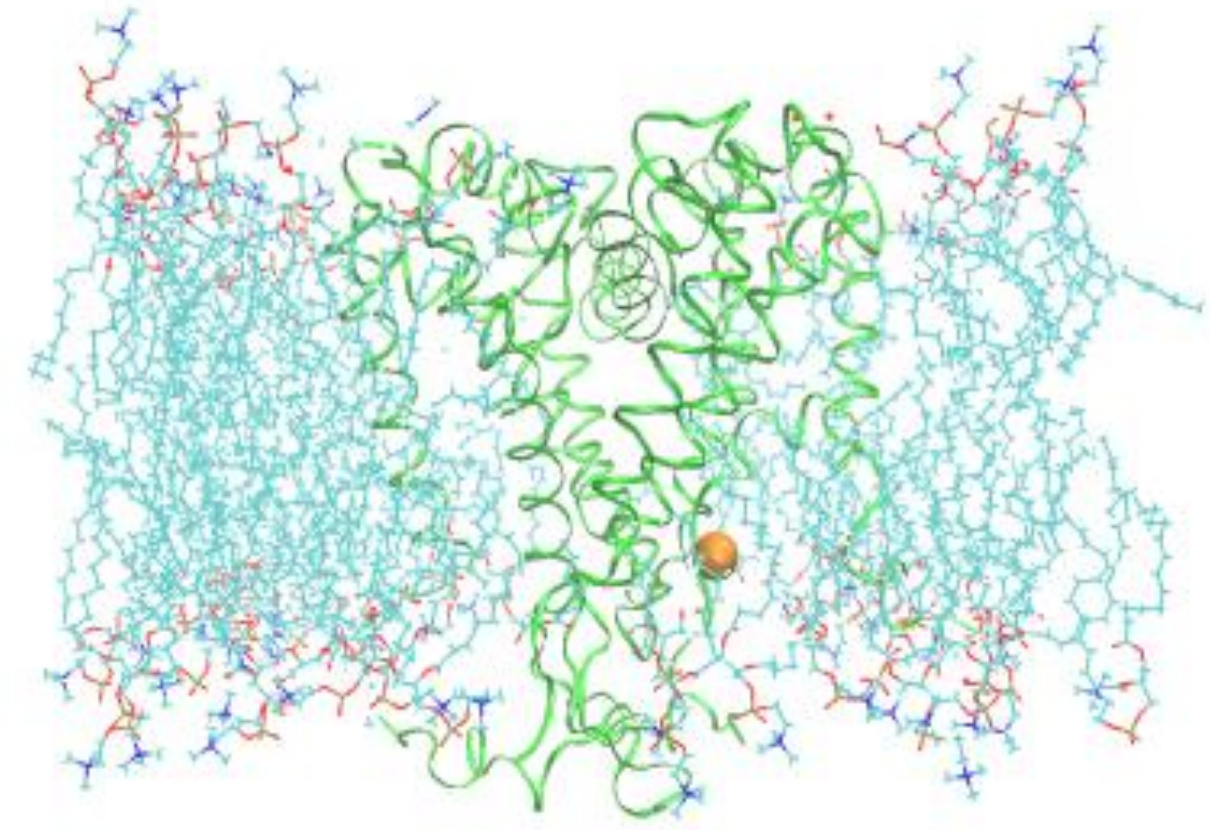


Figure A3. Sodium ion (orange) penetrates the lipid bilayer (teal in a line configuration) after having an ABF applied to it.

Location of Files

In nscc, the file that contains all the tutorials and files that created is in the following path:

\\filer\Research\lmadison\ivargh22\Thesis.

Within the 'Thesis' file, there are three files:

"DataAnalysis," "Fall," snd "JanPlan_Spring,"

"DataAnalysis" contains all the python scripts that I've used to trace atoms along a z-coordinate and create the free energy profiles. Additionally, it contains three files:

"Analysis": Contains scripts from prior research in the Madison lab, that served as templates for the data analysis for this study.

"Images_Thesis": Contains all the images that was generated for this work and a talk based on this work.

"Test_Images": Contains all the images that was created while testing the scripts I created.

"Fall" contains tutorials from Dr. Madison's' computational chemistry class that I attempted in the Fall.

"JanPlan_Spring" contains work conducted over Jan Plan and Spring of 2022. It contains 4 files:

"Bacterial": Contains all the simulations run by Iz Varghese to use in this study.

"NavMs": Contains simulations with an electric field of 250 mV and 750 mV applied.

"Tutorials": Contains all the tutorials followed and practiced during JanPlan. They served as templates and inspiration for the simulations run in the spring in the "Bacterial" file. [51,52]

"UnusedFiles": Contains files that were extra and not used for this study.

The "Bacterial" file contains data from this study, and is broken down in the following way:

"Tails" – contains the simulations with the approximated disordered region.

This section is broken up into 8 files for each simulation system. The file names are named with the concentration of ions, the type of ion studied, and the protonation state of the GLU ring. For example, for a 0.15 M concentration of sodium and potassium with a deprotonated EEEE ring, the name would be: "015MNaK_0GLUP." Within each of these files, the PDB and PSF files are named "NavMs.pdb" and "NavMs.psf." The "MeltTails" files contains all the simulation runs, and the "toppar" files contain all the parameters.

Note that for all the files with sodium ions in it (**except the deprotonated homogenous system)**, when opened has a nested file with the same name. That is the file with all the updated information. So, for example, the "015MNa_2GLUP" file contains the most up-to date simulations in the following directory:

\\filer\Research\lmadison\ivargh22\Thesis\JanPlan_Spring\Bacterial\Tails\015MNa_2GLUP\015MNa_2GLUP.

Also note that in the deprotonated pure sodium file ("015MNa_0GLUP"), the structure is a little different. The file that contains all the information used in this study is "NaVMs." This file contains simulations run by Dr. Lindsey Madison to solve for the interlocking of rings of the residues in a deprotonated protein structure with sodium ions in the system. This study used files from the following directory:

Production runs:

\\filer\Research\lmadison\ivargh22\Thesis\JanPlan_Spring\Bacterial\Tails\015MNa_0GLUP\NaVMs\Na0-slide\production

ABF:

\\filer\Research\lmadison\ivargh22\Thesis\JanPlan_Spring\Bacterial\Tails\015MNa_0GLUP\NaVMs\Na0-slide\abf

“NoTails” – contain just simulation runs of the pore domain. No data from this file was used in this study.