THE UNIVERSITY OF CHICAGO

PROTON TRANSPORT MECHANISMS IN THE INFLUENZA A M2 CHANNEL AND *DE NOVO* DESIGNED PROTEINS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY

BY
LAURA WATKINS

CHICAGO, ILLINOIS
DECEMBER 2021

To my parents, for their unwavering love and support.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Graduate school has been a stimulating, at times infuriating, and overall life-changing time of growth for me, professionally and personally. In my time at UChicago I have been surrounded by incredible people who have awed me not only with their intellect, but also their kindness and sense of responsibility for changing this world for the better. I am indebted to many, many people for making this journey meaningful, fun, and survivable.

First of all, a huge thank you to my advisor, Professor Greg Voth, for his mentorship and the access to incredible resources as a member of his group that made this research possible. I also acknowledge the Department of Energy for four years of support as a Computational Sciences Graduate Fellow, which pushed me to take classes I otherwise would not have and shaped my career choice.

Thank you to the Chemistry Department and the opportunities it provided me. I am particularly grateful for my experiences in student leadership positions, and to Professor Andrei Tokmakoff for being a thoughtful and proactive department chair and for allowing me to form the Culture Committee and run with it.

Thank you to Dr. Vera Dragisich, an incredible mentor, friend, and student ally, without whom the department would surely fall apart. I am thankful for her support and the many conversations. I will miss having an office by you and running into you in the hallway.

Thank you to the many Voth group members who were mentors and friends. Though far too many to name, Dr. Xinyou Ma and Dr. Chenghan Li were always available to answer questions and help with research problems. Thank you to Dr. Paul Calio for being a supportive friend, always available to listen to whatever I needed to get off my chest. Your friendship definitely helped me push through the toughest times.

One of the best parts of grad school has been the community of other students. I started with a great cohort of students who made classes and teaching in first year fun. Thank you especially to Bodhi Vani (my fellow WiTCh) and Sarah Brown for being inspiring women in science and even more so for being great friends. Thank you to Jon Keim for understanding and meeting me in all the philosophical questions and laments on the struggles of life, and for being a dependable and encouraging friend. A special shout out to the Baseliners for being the most fun softball team, and thank you to Coach Tim Grabnic for letting me be team manager with him, for always bringing the fun, and for making everyone feel included. Thank you to all the regulars at cookie hour (looking at you, Dan Micheroni, Charlie Matthews, Elizabeth Bain, Jake Higgins, Adam Antoszewski and more) for providing a wonderful respite from work throughout the week. Again, there are far too many people to name them all, but I am so grateful for my fellow students and friends who made grad school fun.

Outside of the university, I am thankful for my church community at Christ Church Chicago and especially to Pastor Bing Nieh for being an empathetic mentor and support. Thank you to Drew Donaldson for first being one of my closest friends, and then for being the best roommate. Thank you to Lisa Vongchingtrong and Alyssa Wilson for always caring about me, my life, and for making time for phone calls.

Thank you to Julian Nicks for being the best pandemic buddy and an incredible boyfriend. I love our trips together and your inspiring passion for social impact. Thank you, thank you, thank you for your support in the most stressful times, for kicking me when I needed a kick, and for always encouraging me.

Finally, thank you to my family for being the best family I could have asked for. Thank you to my older brothers Matthew and Stephen for always setting the bar high for me—I'm grateful that we have grown into such good friends. To my parents, Mark and Sally Watkins, none of this would have happened without all your love, guidance, and encouragement. Dad, your constant support and confidence in me, when the world says it's hard to be a woman in science, made me believe I could do it. Mom, your unwavering love and being my biggest cheerleader encouraged me to keep going. I am so blessed to have you as my parents and don't know how anyone else does it without you.

Abstract

Proton transport (PT) in biomolecular systems is the controlled movement of hydrogen ions across a cellular membrane, frequently through a channel or transporter, and is essential for sustaining life. PT is necessary for a variety of functions including maintaining pH gradients, driving ATP synthesis, and facilitating the co- or anti-transport of other small molecules, among many others. Because of their crucial role, protein channels and transporters are often chosen as drug targets in bacteria or viruses that cause disease. Thus, beyond elucidating PT mechanisms to understand how a specific channel or transporter works, studying the detailed interactions that facilitate PT is critical as it can inform drug design efforts to help treat disease, and it can reveal fundamental principles of PT that can then be used to control this function in designed systems.

In this thesis, I use computer simulation to study processes of PT in two systems, the influenza A M2 channel and two *de novo* designed proteins. Through three different projects focused on M2, I provide a detailed view on PT in M2, deducing important interactions and providing insight into inhibitor binding. First, using extensive Multiscale Reactive Molecular Dynamics (MS-RMD) simulations with explicit Grotthuss-shuttling hydrated excess protons, I show how a hydrated excess proton strongly influences both the protein and water hydrogen-bonding network throughout the channel, providing further insight into the channel's acid-activation mechanism and rectification behavior. Second, I extend this work to focus on inhibitor binding. In this work, I illuminate a dynamic understanding of the mechanism of drug inhibition in M2, grounded in the fundamental properties that enable the channel to transport and stabilize excess protons, with critical implications for future drug design efforts. Third, I use combined MS-RMD and quantum mechanics/molecular mechanics simulations to calculate the potential of mean force (PMF) of PT

in a prominent mutant of M2, showing how the D44N mutation alters the free energy and mechanism of PT.

Finally, I use these methods to study two *de novo* designed proteins in collaboration with experimental experts in protein design. In this work, our collaborators find that introducing a single polar residue into an otherwise hydrophobic channel is sufficient to induce proton conduction. Using MS-RMD simulations, I calculate 2D PMFs of PT for two systems, one with the mutation and one without, to understand how the mutation facilitates PT. This work reveals fundamental characteristics of PT and provides insight into how it can be harnessed in future protein design.

Altogether, the work in this thesis presents a dynamic picture of proton transport in proteins and reveals key characteristics that are critical for understanding PT in biological systems overall, with implications for future drug design efforts and *de novo* protein design.

Chapter 1

Introduction

1.1 Proton Transport

Proton transport (PT), the movement of hydrogen ions across membranes, is a ubiquitous process in biomolecular systems and a functional component of many channels and transporters.[4-7] It is necessary, for example, to maintain pH gradients,[8, 9] to drive ATP synthesis,[10] and to facilitate the co- or anti-transport of other small molecules.[5, 6, 11] Because of their essential role in such systems, channels and transporters with PT functionality are often targets for drug design to inhibit or control PT—in the case of viruses and bacteria, to slow or prevent infection, but there are myriad other disease applications.[12-14] Drug design is notoriously challenging, as both thermodynamic and kinetic factors must be considered but are difficult to predict and control, and its success depends on high quality structures, an understanding of structural dynamics, and a knowledge of the protein's function and its mechanism. Beyond biomolecular systems, PT is also relevant to hydrogen fuel storage, the development of biomimetic devices, and any systems with acidic or aqueous conditions. Thus, beyond elucidating PT mechanisms to understand how a specific channel or transporter works, studying the detailed interactions that facilitate PT is critical for two main reasons: 1) it can inform drug design efforts to help treat disease, and 2) it can reveal fundamental principles of PT that can then be used to control this function in designed systems.

1.2 Influenza A M2

The influenza A M2 proton channel is considered an archetype for biological proton channels and has been studied extensively due to its relevance in viral infection and its curious PT mechanism. One of the primary functions of the M2 channel is to transport protons across the viral membrane after it has been encapsulated in a cellular endosome, acidifying the interior in order to escape and infect the cell. M2 has a precisely tuned activation process to turn on proton conductions only under the low exterior pH conditions in the endosome, leading to unidirectional conductance, and it selectively transports protons.

It is the target of several antivirals available for influenza A, but prevalent mutations have rendered these drugs ineffective, and they are no longer prescribed.

1.3 Computer Simulations of Proton Transport

Proton transport is an inherently quantum mechanical process—a proton in water does not exist as an isolated ion or hydronium molecule, but rather exists in a complex hydrogen-bonded network that rearranges dynamically as bonds break and form according to the Grotthuss shuttling mechanism.[15-17] This constant bond rearrangement presents a difficulty for computer modeling, as this requires a reactive method with an accurate representation of bonds breaking and forming. Moreover, the PT process can be coupled to other complex processes such as hydration changes, electron transfer, and protein conformational changes. These couplings of PT to other processes also present further complications for computer modeling, due to the difference in timescales at play: proton shuttling between any two water molecules can occur on the femto- to picosecond timescales, whereas other processes influencing the overall PT may take nano- to milliseconds or longer. Thus, classical molecular dynamics (MD) with fixed bonding topology cannot be used to study PT; additionally, *ab initio* methods are not efficient enough to reach the many nanosecond timescales necessary to obtain sufficient sampling in biomolecular systems that may have important degrees of freedom several orders of magnitude slower than proton shuttling. Multiscale Reactive Molecular Dynamics (MS-RMD)[18-21] was developed to efficiently and accurately capture the solvation and delocalization of an excess proton in water, such that the quantum-chemical nature of the hydrated proton can be studied in the context of membrane proteins over the long timescales needed for accurate simulation of such systems. MS-RMD has been successfully applied in several protein systems to predict and explain mechanisms of PT.[3, 22-30]

1.4 Thesis Aims and Chapter Overview

The aim of this thesis is to derive fundamental principles undergirding PT in biomolecular systems by studying PT in two proteins, the influenza A M2 channel and *de novo* designed proteins. Most of this work was done using MS-RMD and enhanced sampling techniques to capture the delocalized proton and its behavior during PT. All projects were done in collaboration with William DeGrado of University of California, San Francisco.

In chapter 2, I discuss my work mining insight from extensive MS-RMD simulations of wildtype M2. Using simulations with an excess proton placed throughout the channel, I examine the molecular-level interactions that accompany PT and find that the hydrated excess proton dynamically influences the protein structure and water hydrogen-bonding network. These results further our understanding of the M2 PT mechanism.

In chapter 3, I continue my study of M2 by focusing on drug binding. Specifically, I investigate how the adamantyl-amine class of inhibitors take advantage of key features of M2 and its PT mechanism to bind in the channel and block PT.

In chapter 4, I conclude my work on M2 by studying the prevalent D44N mutant, which exhibits increased conduction and an altered pH-dependence. By calculating the free energy profile associated with PT and comparing with wildtype M2, I explain the role of Asp44 and the effect of its mutation.

Chapter 5 presents the PT mechanism for the first *de novo* designed protein with PT functionality, designed by our collaborators. This study reveals fundamental characteristics of PT and provides insight into how it can be harnessed.

Chapter 2

Proton Induced Conformational and Hydration Dynamics in the Influenza A M2 Channel

Reprinted with permission from:

Laura C. Watkins, Ruibin Liang, Jessica M. J. Swanson, William F. DeGrado, and Gregory A. Voth. Proton Induced Conformational and Hydration Dynamics in the Influenza A M2 Channel. *Journal of the American Chemical Society* **2019** *141* (29), 11667-11676.

2.1 Introduction

Proton transport (PT) is a ubiquitous process in biomolecular systems and a functional component of many channels and transporters.[4-7] However, accurately modeling an excess proton in large, complex biomolecular systems is difficult due to the quantum-chemical nature of the solvated proton – it does not exist as a hydronium ion, but rather in a complex hydrogen bonded network that constantly rearranges as bonds break and form dynamically,[31] allowing the protons involved in the structure to move according to the Grotthuss shuttling mechanism.[16, 17] (Note that the phrase "excess proton" in this work refers to the hydrated proton structure, which at its core is H_3O^+ but generally is a more complicated structure including additional water molecules.[18] It is helpful to think of the solvated proton not as a fixed particle that hops from water to water, but rather as a positive charge spread across a water complex that dynamically alters its structure and dynamic properties in response.) Moreover, the PT process can be coupled to other complex processes such as hydration changes, electron transfer, and protein conformational changes. One critical implication of these potential couplings is that the mechanism of PT cannot be easily inferred from standard molecular dynamics (MD) simulations without an explicit excess proton actually being in the system, even if "water-wires" are otherwise present (or not). In fact, such inferences may be wrong. These couplings of PT to other processes also present further complications for computer modeling, due to the difference in timescales at play: proton shuttling between any two water molecules can occur on the femto- to picosecond timescales, whereas other processes influencing the overall PT may take nano- to milliseconds or longer. To overcome these challenges, Multiscale Reactive Molecular Dynamics (MS-RMD)[18-21] was developed (and Multi-State Empirical Bond, MS-EVB, before it; see, e.g., Ref. [32]) to efficiently and accurately capture the dynamic solvation and delocalization of the excess proton in complex aqueous phase and biomolecular systems. Since MS-RMD evolves the system according to deterministic Newtonian dynamics, it is capable of accessing the longer timescales needed to realize couplings with slower degrees of freedom and to describe the rare event PT processes that are often biologically relevant. This method (and MS-EVB before it) has successfully been applied to several protein systems to predict and explain mechanisms of PT.[1, 22-29, 33]

The M2 channel is a homo-tetrameric protein in influenza A viral capsids that is responsible for the acidification of the viral interior, a critical step in the viral life cycle.[34-36] It is the target of the antiviral drugs amantadine and rimantadine,[37] but the prevalence of drug-resistant mutants has necessitated a continued effort to understand the PT mechanism and channel properties in order to better inform drug-design efforts. A critical histidine residue (His37) located at the center of the transmembrane portion of each helix acts as a "pH-sensor" and is responsible for enabling activation as the pH is lowered.[38] Each His37 can bind one additional proton, such that the tetrad can have a total charge ranging from +0 to +4, referred to as the His37 tetrad charge state. As the charge increases and approaches the activated state, electrostatic repulsion between these residues causes the channel to widen, opening the Trp41 gate located just below the His37 tetrad[39-43] and enabling proton flux to the viral interior.[1, 23, 44-48] It is hypothesized that the channel becomes activated in the +2 state, such that conduction primarily occurs as a cycle through the +2 and +3 states,[49] based on experimentally determined pKa's, pH-dependent conduction, and computer simulations.[1, 23, 50-53] Another important characteristic of the channel is its rectification behavior—channel activation and inward proton flux occurs when the exterior pH is lowered, but there is no outward proton flux observed when the interior pH is lowered.[54, 55]

Recently, we carried out multiscale simulations to investigate the PT mechanism and pH-dependent activation behavior in the influenza A M2 proton channel.[1, 23] In contrast to other prior simulation studies, our work employed multiscale simulations (including MS-RMD and quantum mechanics/molecular mechanics) with explicit, reactive excess protons to calculate the free energy profiles of PT through the full transmembrane M2 channel in the +0, +1, and +2 protonation states, thereby providing definitive insight into both the acid activation mechanism of M2 and the critical role of the His37 tetrad. This effort characterized the physical basis of PT through extensive sampling with an explicit proton and explained the rectification behavior of the channel through the asymmetry of the PT free energy profiles and estimated rate calculations. Importantly, the calculated rates of deprotonation and conductance quantitatively agreed with proton exchange rates measured by NMR[56] and experimentally determined conductance values.[44, 55, 57] While this previous study revealed some molecular understanding of the acid activation and rectification mechanisms, the detailed interactions between the excess proton and the channel that contribute to these behaviors were not explored. In this work, our previously published simulations are further analyzed to resolve the channel's dynamic response as an excess charge moves through the channel and how these dynamic changes enable acid activation.

Water-mediated PT in proteins occurs through three dominant regimes: 1) water is excluded from the lowest energy structures, but fluctuations allow short water networks to transiently form; 2) water forms a single-file hydrogen bonded path in narrow regions; and 3) water exists in clusters in larger pore regions. The M2 channel exhibits both cases 1 and 3, and is thus an excellent system for studying the dynamic collaboration between the protein, pore-lining residues, and water that enables PT.

We note that a recent computational study from Chen et al. used constant pH MD (CpHMD) to describe the conformational activation of the M2 channel.[58] By simulating the channel at several different pH's, rather than in fixed protonation states, they elucidate the pH-dependence of equilibrium conformational changes that occur as part of the activation mechanism. Our work here, in contrast, is focused on the changes coupled to an explicit proton's movement through the channel in order to determine the molecular-level details of the PT mechanism, which should be viewed in tandem with CpHMD results. While they sample the configuration space at a given pH, we sample the configuration space of the dominant protonation state in several pH ranges as a function of the excess proton's transport through the channel.

To determine the channel's dynamic response during explicit PT, we focus on the water structure in different regions of the channel and how the number and orientation of the water molecules fluctuate based on where the excess proton is in the channel. Importantly, we find that the channel structure is transiently affected in the presence of the proton, especially for the +0 state, which exhibits distinct conformations when the excess proton is in different regions of the channel. Additionally, the three-dimensional spatial distribution of the excess proton is explored for the first time, revealing an asymmetric path due to asymmetric water distributions and interactions with pore-lining residues. This analysis further emphasizes that having an explicit, reactive excess proton MD description is crucial for understanding such a PT mechanism. It is difficult, if not impossible, to predict or observe the changes and fluctuations that are coupled with PT from non-reactive empirical MD simulations that do not explicitly include a dynamically evolving excess proton in them.

2.2 Methods

The analysis performed here was done using previously published simulations; readers are referred to ref[1] for more details on the simulation set-up and potential of mean force (PMF) free energy profile calculations. Briefly, the simulations were initiated from a crystal structure of the transmembrane portion of the M2 channel (this construct is referred to as M2TM) resolved at room-temperature and high pH (PDB: 4QKL[59]) embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer solvated with water. We have shown that the presence of amphipathic helices, included in the full-length M2 protein, do not significantly influence the PT mechanism,[1] and experiments have indicated proton conduction is similar for M2TM with or without the amphipathic helices.[60] Following classical equilibration of the +0, +1, and +2 His37 tetrad charge states, MS-RMD simulations were run using the replica-exchange umbrella sampling method.[61] The collective variable (CV) for umbrella sampling was defined as the z-coordinate difference between the excess proton center of excess charge (CEC) and the center of mass of the 4 Gly34 alpha carbons, referred to as $\mathrm{CEC_Z}$. The excess proton CEC was defined as:[62]

$$\vec{r}_{CEC} = \sum_{i}^{N} c_i^2 \vec{r}_{COC}^{\,i}$$

where $\vec{r}_{COC}^{\,i}$ is the center-of-charge of the i^{th} diabatic state, and c_i^2 is the EVB amplitude of that state. The sum is over all N states. The CV was restrained over 75-80 windows every 0.5 Å for 1-2 ns each after equilibration of the system, including any hydration changes. Frames were saved every 10 ps. The PMF calculations were performed with both MS-RMD and quantum mechanics/molecular mechanics (QM/MM) simulations. While the MS-RMD methodology allows for reactive amino acids, there is currently no reliable procedure for fitting the MS-RMD forcefield for four highly coupled residues in a complex environment (i.e., the His37 tetrad), so QM/MM was used for this region of the channel (CEC_Z = 1.5 to 11.5 Å). Due to the significantly improved sampling enabled by the efficiency of MS-RMD simulations, only the MS-RMD trajectories are used in the following analysis. Thus, proton-dependent properties are only calculated for CEC_Z = [-22.0, 1.0] and CEC_Z = [12.0, 22.0] Å. Simulations were run until the PMF was converged. Simulations were additionally extended, following the same protocol, as needed for increased sampling in this study.

The analyses in this paper were performed using all of the MS-RMD umbrella sampling trajectories to delineate changes as the proton moves through the channel. Since we did not track unbiased dynamics, our conclusions are based on the assumption that PT through each part of the channel is slower than the observed hydration and conformational changes. As partial, though not conclusive, support of this assumption, we note that the reported changes occur and equilibrate on the sub-nanosecond timescale, while the rate of proton flux is multiple orders of magnitude slower (100-1000 sec^{-1} at pH 6).[57] In the water and radial distribution analyses, channel properties are calculated with respect to the protein's principal axis (PA) rather than the z-axis, which we label Z' (see **Figure 2.9** in the Supporting Information for visualization of this difference). The z-coordinate of the CEC in this new frame of reference is defined as $CEC_{Z'}$. This is the z-component of the CEC position along the protein's PA, rather than the z-axis, with respect to the center of mass of the 4 Gly34 alpha carbons such that $CEC_{Z'} \approx CEC_Z$. This was done due to the channel's slightly tilted axis (see Results and Discussion). Thus, Z' refers to the channel's principal axis, and

CEC$_{Z'}$ refers to the CEC position along that axis. Detailed descriptions of all analysis calculations are in the supporting information (SI).

2.3 Results and Discussion

2.3.1 Acid-Activation Mechanism

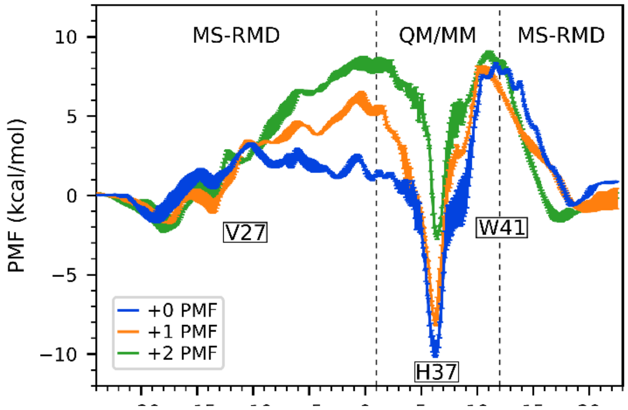

Figure 2.1. Previously calculated and published PMFs for proton transport through M2 in the +0, +1, and +2 states. Reproduced from reference [1].

In our previous work, multiscale simulation methods were used to calculate PMFs for the movement of an excess proton through the channel in the +0, +1, and +2 His37 protonation states to gain insight into the acid-activation mechanism, shown in **Figure 2.1**. These PMF profiles show deep local free energy minima for the excess proton to bind to the His37 tetrad in the +0 and +1 states, with large barriers for deprotonation towards the viral interior. The deprotonation barrier in the +2 state, however, is lower and the calculated conductance matches the experimental value, indicating how the channel reaches an activated state at low external pH.

2.3.2 Water Structure and Hydrogen Bonding Networks are Sensitive to Proton Position

It has previously been shown that a hydrated excess proton can change the water dynamics in confined areas, including increasing the level of hydration and rearranging hydrogen bonds.[32, 63, 64] Here, we evaluate and show the extent to which an excess proton affects the water structure in different regions of the M2 channel. To quantify these effects, the direction of each water-water hydrogen bond was calculated as the cosine of the angle between the hydrogen bond vector and the z-axis, taking on values from -1.0 to 1.0, as defined in **Figure 2.2b**. This value was calculated for all hydrogen bonds and averaged over all hydrogen bonds in each region to determine the average direction of water hydrogen bonding in a given configuration. This was done for all simulations and proton positions and then plotted as a function of the proton position to quantify the change as the proton moves through the channel. See the Supporting Information (SI) for more details.

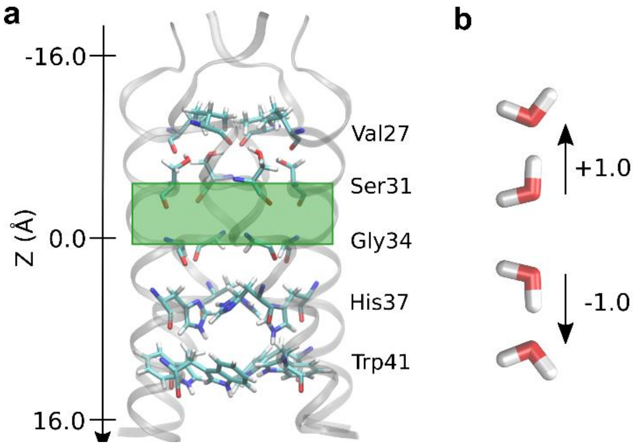

Figure 2.2. (a) The M2 channel with pore-lining residues shown as sticks. The green box highlights the region of water analyzed in **Figure 2.3b**. The direction of the z-axis indicates the direction of proton flux from the exterior to the viral interior. **(b)** Definition of hydrogen bond directions, with respect to the channel as shown, used throughout.

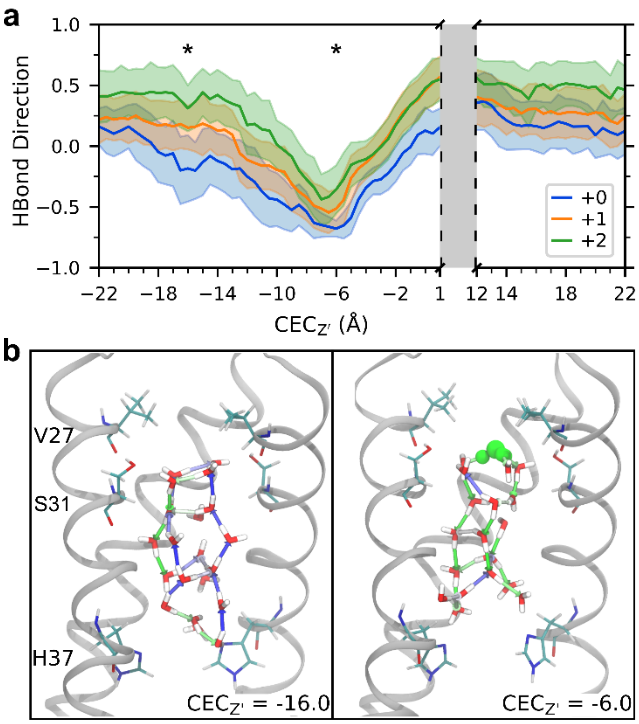

Figure 2.3. Hydrogen bond rearrangement as proton moves through the channel. (a) Average and standard deviation of the direction of water-water hydrogen bonds in the central region of M2 (see green box in **Figure 2.2a**. as a function of CEC position. (b) Example water structures in the middle region of the +1 channel for CEC$_{Z'}$ = -16.0 (the proton is near the entrance of the channel, not in view) and -6.0 Å (above this central region – see stars in 3a). Hydrogen bonds are shown as arrows, colored as a gradient based on direction where blue is +1.0 and green is -1.0. The most hydronium-like water is shown in green. Front helix and residues on back helix are removed for clarity.

Figure 2.3 shows how a cluster of water in the middle of the channel (comprised of all water contained in the region indicated by a green box in **Figure 2.2a**, from roughly Z' = -5.0 A to Z' = 0.0 A) interacts with and responds to the hydrated excess proton charge for three His37 charge states as a function of the CEC position along the Z' axis. The average direction of the hydrogen bonds in this middle region is positive before the proton enters and declining as it moves down the channel. The difference is maximized when the excess proton CEC is directly above this middle region (CEC$_{Z'}$ = -6.0 Å), at which point the waters are completely reoriented. However, this effect initiates when the excess proton is several layers of water away as the rearrangement "trickles down" from the excess charge. The water follows this trend in all sections of the channel – the average hydrogen bond direction reaches its minimum value when the proton is situated just above that region, and its maximum value when the proton is just below that region. As the excess proton moves *through* the region, there is a transition as the hydrogen bonds change polarity due to the excess charge passing through.

The backbone and sidechain atoms also play an important, but dynamic, role in stabilizing water structures through hydrogen bonding. As the hydrated excess proton moves through the channel, these interactions transiently shift both in strength and orientation. The detailed results and discussion of Ser31 hydrogen bonding dependency on the excess proton position are in the SI (see **Figure 2.11** and corresponding text).

Altogether, these changes in hydrogen bonding clearly demonstrate that PT in M2 is a highly dynamic and variable process, with a range of water reorganization trends in response to both the presence of an excess proton charge and the protein structural changes associated with the PT. This analysis also provides insight into water interactions that enable PT through water clusters (regime 3). Considering a hydrated proton approaching from the virus exterior, the hydrogen bonds of connecting water generally have a negative orientation (as defined in **Figure 2.2b**) for facile PT in accordance with the Grotthuss shuttling mechanism. The rearrangement of water as shown above indicates how the proton can, in a sense, "pave its own path." Before the proton enters the M2 channel, the water is not arranged optimally for PT, but the water structure is flexible and rearranges in response to the passage of the excess proton. Additionally, the changes in water-protein hydrogen bonding shows the flexibility of such interactions – the presence of the excess proton can result in both increased and decreased frequencies of these interactions, as well as shift the positions of involved water molecules. These types of proton-induced changes are expected to play a critical role in many biomolecular (and other) systems involving PT.

2.3.3 The Water Structure in +2 State is "primed" for PT

The water and hydrogen bonding analysis can be used to understand aspects of the acid activation mechanism in terms of the free energy of PT through the channel. The barrier for the proton to deprotonate the His37 tetrad and move past the Trp41 tetrad is large (~18, 16 kcal/mol) in the +0 and +1 states, but only ~10 kcal/mol in the +2 state (**Figure 2.1**). Comparison with the relatively lower barriers of deprotonation towards the viral *exterior* indicates that these high inward-deprotonation barriers, particularly in the +0 state, are not just due to the favorable His-binding. To better understand the role of the water structure in PT from the His37 tetrad to the viral interior, and how it may contribute to the high deprotonation barrier, we calculated the number of waters and the average direction of hydrogen bonding (as in the previous section) of water in the region just below the His37 tetrad and above the Trp41 tetrad, which approximately corresponds to the range $6 < Z' < 11$ in **Figure 2.2a**. At $CEC_{Z'} = 12.0$ Å, the excess proton is sufficiently beyond the His37 tetrad such that no His37-bound states significantly contribute to its charge delocalization, as previously determined.[1, 23] This is also the region where PMF calculations from MS-RMD, which does not account for His37 contributions to delocalization, and QM/MM, which does, overlap. Thus, the water structure below the His37 tetrad, when the proton is at that position, represents the average water structure needed to facilitate proton unbinding from the His37 tetrad towards the interior. As shown in **Figure 2.4a**, the number of waters in this region increases by 3 to 5 water molecules in the +0 and +1 states as the proton passes through (comparing $CEC_{Z'} = 1.0$ with $CEC_{Z'} = 12.0$ Å). However, the number of waters is relatively constant in the +2 state. Additionally, the average direction of hydrogen bonds between waters in this region also shows tetrad-charge-dependent changes (**Figure 2.4b**). When the proton is out of the channel (CEC$_{Z'}$= -22.0 or 22.0 Å) the +0 state has significantly positive (pointing upward) hydrogen bonds in this region, averaging ~0.4, which need to reorient to achieve a negative direction for PT to the His37 tetrad. The +1 and +2 states, however, start negative and do not need to change as much.

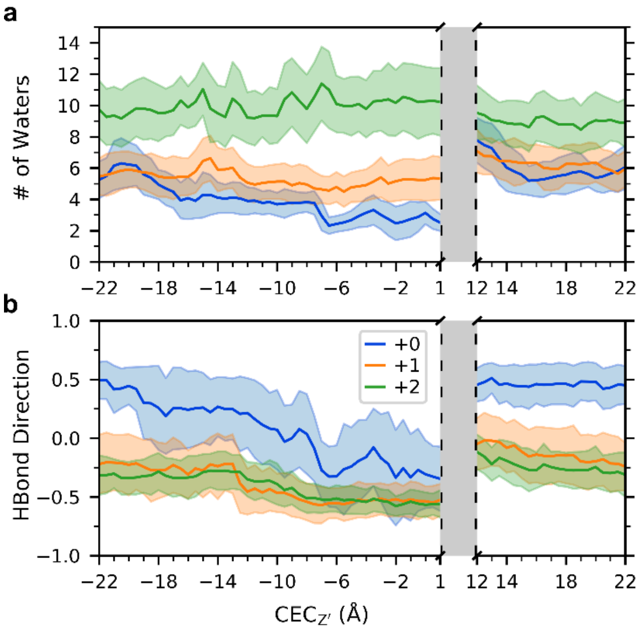

Figure 2.4. Analysis of water in the region below the His37 tetrad and above the Trp41 tetrad. See SI for specific definition. (a) The average and fluctuation (standard deviation) of the number of waters in this region as the proton moves through the channel. (b) The average and standard deviation of the direction of water-water hydrogen bonds in this region.

The same analysis was also done for the excess proton approaching the His37 tetrad from the viral exterior, focusing on the water in the region directly above the tetrad, below the Gly34 residues in **Figure 2.10**. The PMF deprotonation barriers as the proton moves from the His37 tetrad towards the viral exterior are relatively similar for the three charge states: ~12, 14, 11 kcal/mol in the +0, +1, and +2 states, respectively. We can consider how the water in this region responds to the proton approaching the tetrad by comparing the water when CEC$_{Z'}$ = 1.0 Å, which corresponds to the point at which the proton is completely past the His37 tetrad (by same logic as above). The number of waters in this region at $CEC_{Z'}$ = 1.0 is nearly the same as when the proton is out of the channel (ie, $CEC_{Z'}$= -22.0). In the +1 and +2 states, the number of waters above the His37 tetrad is relatively constant as the proton moves through the channel, but it varies in the +0 state by 3-4 waters. The average direction of hydrogen bonding in this region is also similar for all three charge-states. Before the excess proton enters the channel, all three have a slightly positive value corresponding to hydrogen bonds pointing away from the His37 tetrad, and all three show a similar response to the proton as it moves through the channel. The +0 state noticeably becomes more negative more quickly, due to the lack of positive charge on the His37 tetrad to strongly direct hydrogen bonding. The small changes in the number of waters and the similarity in the direction of hydrogen bonding across the three charge states suggest that, in contrast to below the His37 tetrad, the water structure does not significantly contribute to free energy differences of PT above the His37 tetrad in the different states.

Together, these results indicate how the +2 state is “primed” for PT. Physiologically, the +0, +1, and +2 states represent the dominate charge state in pH ranges of 8-9, 7-8, and 6-7 (according to recent constant pH simulations[58]); these results thus also imply the acid-activation mechanism of PT. For a proton to be released inward from the His37 tetrad, both the +0 and +1 states require more water to move into the region to solvate the excess charge. The water in the +0 state additionally must reorient to a greater degree than the water in the +1 and +2 states. The +2 state, however, does not need more water to enter this region, nor does the water need to completely reorient. This stability of the water as the proton passes through corroborates with the lower barrier in the +2 PMF, corresponding to proton conduction at low pH.

2.3.4 Water Structure Implications for Rectification Behavior

This water analysis can also be considered with regards to the fascinating rectification behavior (i.e., how the channel activates to allow for inward proton flux when the exterior pH is lowered, but does not activate to allow outward proton flux if the internal pH is lowered). Our previous work provides a qualitative explanation for rectification based on the asymmetry of the +0 PMF and the different relative free energy barriers between protonation and deprotonation in both the inward and outward directions. Essentially, the calculated rate for His37 protonation from the viral exterior is faster than that of His37 deprotonation towards the viral interior, such that the channel can move from the +0 to the +1 state, and similarly from the +1 to the activated +2 state. The relative rates of the reverse process, however, are flipped—the rate for His37 protonation from the viral *interior* is slower than that of His37 deprotonation towards the viral exterior. Thus, the rate of deprotonation (k_{off}) to the exterior outcompetes the rate of protonation from the interior (k_{on}) and the channel remains stuck in the +0 state instead of progressing to the +1 and the activated +2 states when the internal pH is lowered.

The water structure analysis provides further molecular insight into this rectification behavior. In the +0 state, proton movement from the interior to the His37 tetrad must be accompanied by several additional waters moving into the channel, requiring the channel to dilate around Trp41 (the radius also increases ~1Å, see below). Additionally, once the channel reaches the +1 state, the water hydrogen bond network surrounding the His37 tetrad is directed *away* from the tetrad, suggesting it would be easier for that proton to move off the tetrad (His37 to exterior) than an additional proton to move in. By comparing the average number of water molecules in the regions above and below the His37 tetrad when the tetrad is protonated (estimated from +1 simulations when the excess proton is out of the channel) versus when the excess proton is just above or below the tetrad ($CEC_{Z'}$= 1.0 or 12.0 Å respectively), we estimate that an additional ~1 waters must move into the region above the His37 tetrad for deprotonation towards the exterior, and ~2 below for interior deprotonation. From this analysis, one excess proton near the tetrad may more easily be solvated *above* the tetrad, as the change in the number of waters above the tetrad is smaller than that below the tetrad, requiring less rearrangement. This is consistent with our previous analysis of the rectification mechanism, showing that when the pH is lowered in the interior, the channel remains stuck in the +0 state because the k_{off} from the His37 tetrad to exterior bulk outcompetes the k_{on} from the interior to the His37 tetrad. This is not true in the opposite direction, so the channel jumps from the +0 to the +1 state only when the external pH is lowered. Here, we see the explanation of this asymmetry in the +0 PMF is due to the challenge of hydrating the region below the His tetrad and ease of hydrating the region above it.

2.3.5 The Proton Transiently Affects the Protein Structure

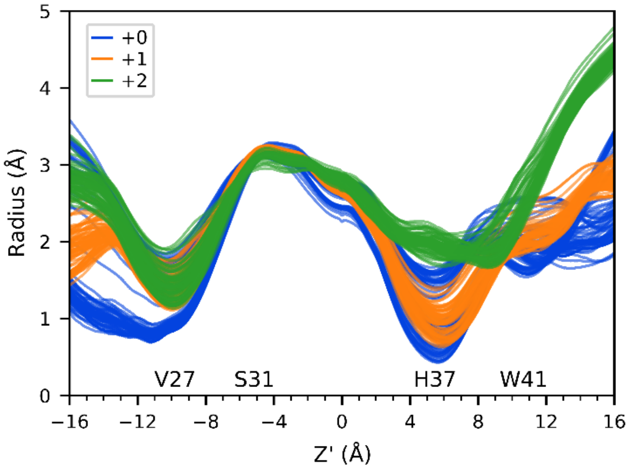

Figure 2.5. Radii profiles for 3 charge states. Each line is the average for one proton position binned every 0.5 Å, shown for $CEC_{Z'}$ = -22.0 to 1.0 and 10.5 to 22.0. The position of pore lining residues is indicated. See **Figure 2.12** for profiles plotted separately for each state.

The radii profiles of the channel were calculated using the program HOLE[65] for each excess proton position to determine correlations between the proton position and the protein structure. First, the radius of the channel was calculated for each frame and then averaged over each proton position. **Figure 2.5** shows these averaged radii profiles for all proton positions, for each charge state, where each line is the average for a proton position along the channel. Clearly, the +2 state is more open below the Trp41 tetrad and in the His37 tetrad region. The narrowest part of the channel in the +0 and +1 states occurs at the His37 tetrad (~Z' = 5 Å), and a secondary closure at the Val27 tetrad is seen near Z' = -10 Å. To more clearly distinguish how the channel radii profile changes are dependent on the proton position, the profiles are plotted separately for each state in **Figure 2.12**. Most striking is the variation seen in the +0 state that is not seen in the +1 and +2 states.

The +0 state can be divided into three main conformers based on clustering the radii profiles, shown in **Figure 2.6**. Following the excess proton's path from the top of the channel: when the proton is out of the channel and just starting to enter at the top, the channel is narrowest at the top near Val27, and slightly more open at the bottom of the channel (at this point the radius near the His37 tetrad is still less than that in the +2 state). As the excess proton approaches and passes through the Val27 gate, the top of the channel opens up and the bottom of the channel closes. As the proton moves through the middle of the channel, before reaching the His37 tetrad, both ends of the channel are closed. Finally, as the proton moves past the Trp41 gate, the channel returns to its initial conformation. The larger radius near the Trp41 residues when the proton passes by (seen in the variation near Z' = 12.0 Å in panel 3 of **Figure 2.6b** indicates how the channel must open to accommodate additional water molecules moving into that region in tandem with His37 tetrad deprotonation. **Figure 2.13** shows a difference plot of these profiles.

The +1 and +2 states also respond to the excess proton position, but to a much smaller degree. In the +1 state, the top of the channel opens slightly to accommodate the proton moving through the Val27 gate while the bottom of the pore becomes slightly narrower; the bottom of the channel opens slightly as the excess proton passes through that portion. The +2 state exhibits less dependence on the proton position. The different extent of response in each state is consistent with ssNMR results that indicate one conformational state with greater conformational fluctuations at low pH, but the presence of distinct states with small fluctuations at high pH in thick bilayers.[40] This suggests that the charged states (+1, +2) are more flexible and may have increased natural fluctuations that allow the proton to pass through, whereas the +0 state is more rigid and requires minor, but distinct, transitions to accommodate the proton.

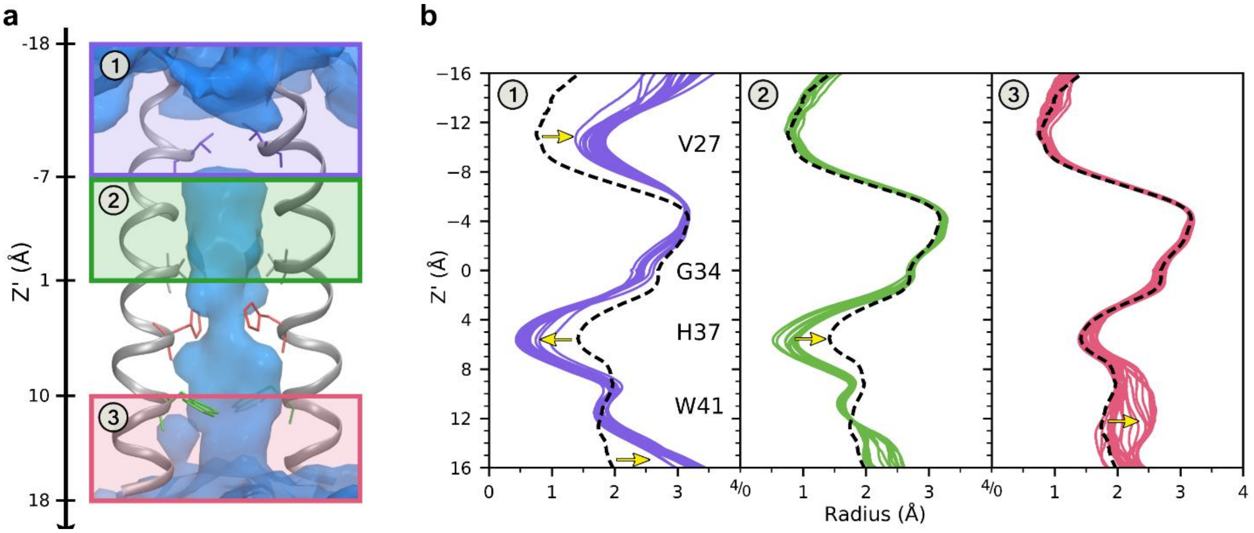

Figure 2.6. Distinct conformations observed in the +0 state. The channel can be divided into 3 main conformers when the excess proton is in different regions, labeled 1-3, based on radii profiles. (a) The three regions of the channel, indicated by purple, green, and pink boxes. The backbone of two helices are shown in gray, with pore lining residues as sticks. The average water oxygen density isosurface is shown in blue. (b) The radii profiles of the channel for all proton positions every 0.5 Å, separated based on which region the proton is located. Each colored line is an average radius profile for one proton position in that region. The black dashed line is the average radius profile when the proton is out of the channel. Yellow arrows highlight the largest deviations from the average.

2.3.6 Channel Asymmetry Distorts Proton Path

The three-dimensional (3D) path of the excess proton has not previously been studied, nor has the effect of the channel's asymmetry on the proton path. The M2 channel is not perfectly aligned with the z-axis, but instead is tilted as seen in SI **Figure 2.14**. In these simulations, the +2 state is most tilted and the +1 state is least titled, on average. The protein (initial structure from PDB: 4QKL) was initially aligned with the z-axis in the set-up for these simulations and became tilted during classical equilibration in fixed protonation states. It should be noted that on longer timescales the protein will sample tilts around the z-axis, such that the average principal axis would align with the z-axis. Additionally, we note that only the transmembrane portion of M2 is used in these simulations, which may affect the protein's degree of tilt in the membrane.

To investigate the 3D path of the proton through the channel, system configurations were first binned according to $\mathrm{CEC}_{Z'}$, and the two-dimensional (2D) histogram of CEC values (every 100 fs) was then calculated in the xy-plane for each bin and smoothed using Gaussian interpolation. Since CEC values were saved much more frequently than the coordinates of the entire system, the CEC coordinate was projected onto the Z' axis of the frame nearest in time. **Figure 2.7** depicts the 3D path of the excess proton based on most probable locations, built from 2D histograms such as those shown in **Figure 2.8**. It is clear that the proton's path is increasingly asymmetric in the +1 and +2 charge states. In combination with the average radius and center of the channel, as well as average positions of protein atoms involved in hydrogen bonding in each plane, we are able to identify the distribution of excess proton CEC values with respect to key pore-lining residues (see **Figure 2.8** and Movie 1 in the supplementary files). This again shows that the proton does not follow a symmetric path about the z-axis as it moves through the channel, but tends to cluster near important residues. This is not an artifact of insufficient sampling, but instead indicates a real preference for the proton to be near the wall of the channel in certain xy-planes. Interestingly, this preference also corresponds to the off-centered position of the ammonium group of amantadine and rimantadine,[66] which will be further explored in future work. The uneven distribution along the x- and y-axes is due to the asymmetric internal environment and density of water in the channel, which is evident in disparate hydrogen-bonding interactions between water and the channel (see **Table 2.1** in the Supporting Information). The slightly skewed orientation of each helix relative to the others as a result of the channel's tilt within the membrane may contribute to this asymmetry. The protonation of the His residues in the +1 and +2 also contributes to this asymmetry. A more detailed description of the excess proton's path is included in the SI.

Furthermore, due to the protein's tilt, this analysis shows that the cross-section of the pore is not well-defined as a circle all through the channel, but is more elliptical on the nanosecond timescale – the distributions at CECZ' = -10.0 Å, for example, appear to be outside the channel but in fact the Val27 residues transiently shift to open more space along the edge of the pore. The method used here to calculate the pore radius does not account for this asymmetry.

Understanding this asymmetric proton path could be helpful for structure-based drug design efforts, as it is can be used to identify particular "sticky" spots for the hydrated excess proton structure. This result also points to the dynamic shape of the pore, which could influence new drug shapes and sizes. It is critical as well to realize that the channel structure and helix tilts are dependent on lipid composition,[40, 67-69] and the precise asymmetries may vary depending on the composition.

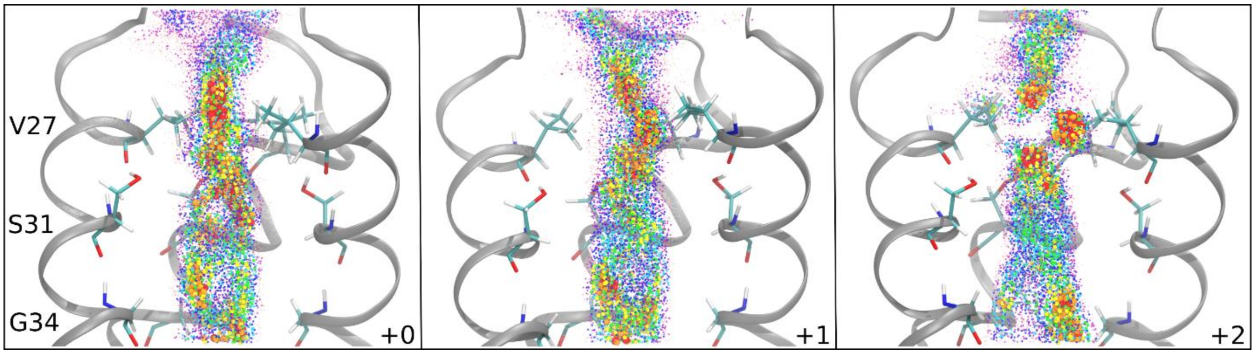

Figure 2.7. Histogram values of CEC density for each charge state. Each colored dot represents one 2D histogram bin value, colored and sized according to value. Values from high to low (indicating high to low probability of the CEC being found in that bin) are large to small, and red to purple. Bin densities less than 0.01 not shown. The structure is taken from a trajectory with the excess proton outside the channel and does not show changes associated with proton position. The front helix is removed for clarity. See movies 2-4 in SI for a better 3D perspective.

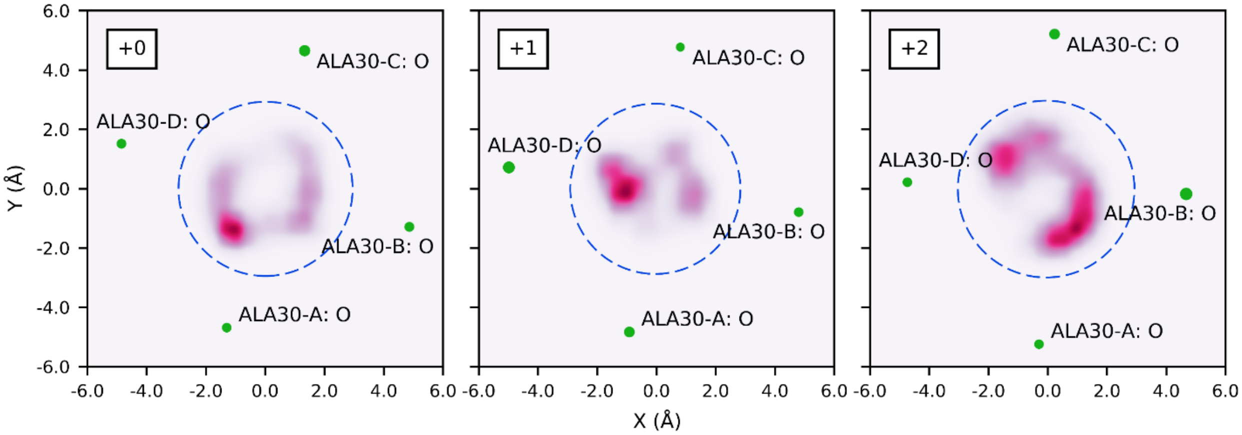

Figure 2.8. Density of CEC positions in the xy-plane for bin centered at $CEC_{Z'}$ = -1.6 Å. Blue dashed circles represent the average radius of the channel at Z' = -1.6 Å for this proton position, centered at the average center of the channel at Z' = -1.6 Å based on the center of mass of backbone atoms. Relevant nearby atoms shown in green, with size indicating proximity (larger atoms are closer to this plane). Atoms are labeled by residue and chain. See movie 1 in SI for equivalent slices moving through the channel.

2.4 Conclusions

In this work, extensive MS-RMD simulations were used to understand in molecular-level detail the specific interactions and structural perturbations that occur during PT in the M2 channel in the +0, +1, and +2 charge states of the His37 tetrad. Our results reveal how the proton causes subtle shifts in the protein structure to increase the pore volume and water density in its vicinity, as well as larger conformational changes in the +0 state. Asymmetry in the channel (persistent on the timescale of these simulations) is shown to influence the proton pathway, causing sharp deviations from a purely symmetric path. Additionally, water molecules move in and out of the channel and hydrogen-bond networks rearrange in response to the positive charge defect as the hydrated proton "paves its own path" through the channel. Specifically, the observed water rearrangements below the His37 tetrad provide further insight into the acid-activation mechanism, showing how, unlike the +2 state, the +0 and +1 states do not have enough optimally oriented water below the tetrad to facilitate proton conduction. In contrast, *above* the His37 tetrad, the hydration required for PT is reasonably facile in all charge states. This helps to explain the rectification mechanism: when the pH is lowered in the interior of the virus, the system remains stuck in the +0 state (it is not activated) because the barrier for proton uptake from the inside to the His37 tetrad is large and the barrier for release from the His37 tetrad is low. Thus, the rate of deprotonation of the His37 tetrad to the outside outcompetes protonation from the inside, all due to the asymmetry in the ease of hydration above and below the His37 tetrad.

Altogether, these results provide a more detailed look at the PT mechanism and give helpful insight for drug-design efforts to exploit channel asymmetry and hydration changes associated with acid activation. For example, designing molecules that destabilize hydration above the His37 tetrad could effectively block PT by switching the dominance of k_{off} versus $k_{on.}$ Also, designing molecules that do not just block the channel, but increase the barrier for PT past the Val27 region by binding above the channel and stabilizing a tightly closed structure, could be an effective strategy. Finally, we emphasize how this study again illustrates that PT is a highly dynamic process in which hydration and structural changes are induced during PT. Simulations that include an explicit, reactive excess proton are necessary to capture these correlated changes. Thus, looking only for “water wires” in experimental structures or empirical non-reactive MD simulations that omit an explicit excess proton may sometimes suggest misleading conclusions.

2.5 Supporting Information

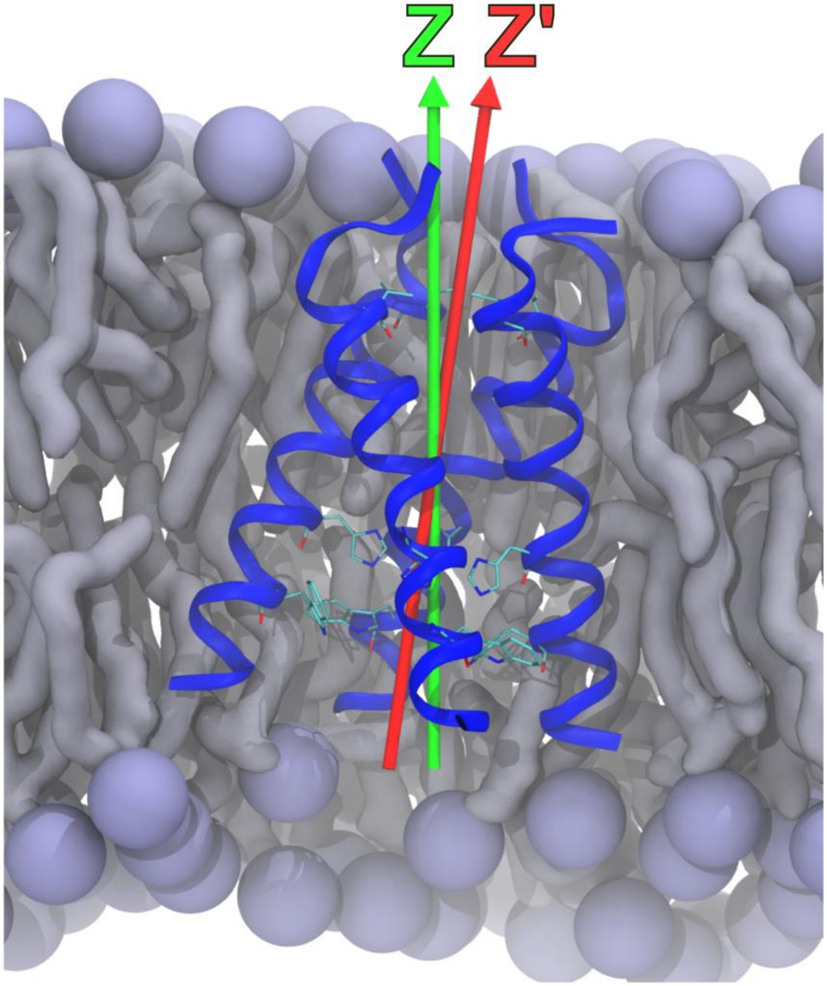

Figure 2.9. Illustration of the axis Z’ used throughout this study. The coordinate Z (green arrow) is the same as the Z-axis of the simulation box. The system is setup such that the membrane normal aligns with Z. Z’ is the principal axis of the protein (red arrow), calculated for each frame. This is used to redefine the reference frame in terms of the protein’s central axis for analysis. The protein helices for an example configuration are shown in blue, and the pore lining residues Val27, His37, and Trp41 are additionally shown as sticks. The lipids are light gray-blue.

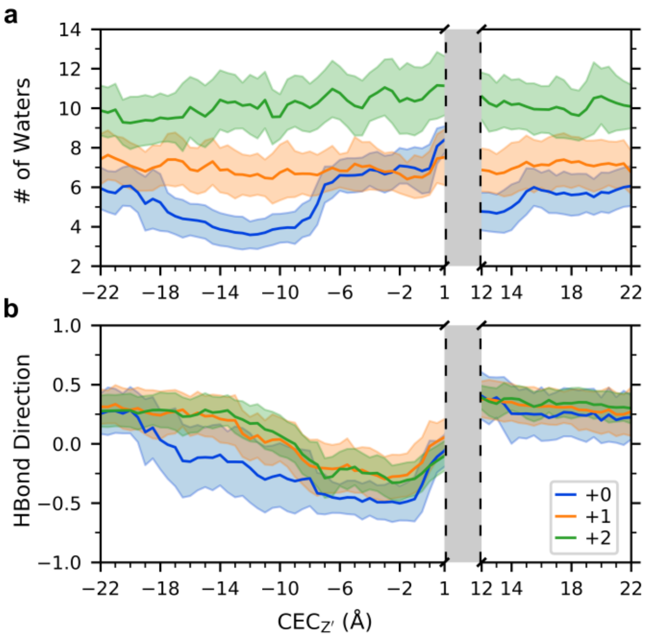

Figure 2.10. Analysis of water in the region above the His37 tetrad and below Gly34, see SI Methods for detailed definition. (a) Average and fluctuations (standard deviation) of the number of waters in this this region as a function of excess proton CEC position. (b) Average and standard deviation of the direction of water-water hydrogen bonds in this region as a function of CEC position.

2.5.1 Protein Sidechain Atoms Involved in Hydrogen Bonding

The sidechain atoms also play an important, but dynamic, role in stabilizing water structures through hydrogen bonding. As the hydrated excess proton moves through the channel, these interactions transiently shift both in strength and orientation. To understand the protein atoms' role in PT, we focus here on the sidechain hydroxyl group of Ser31, which is the only polar residue in the central portion of the channel.

The Ser31 hydroxyl oxygen atoms form hydrogen bonds as acceptors with water molecules (**Figure 2.11a**). The frequency of these bonds in the +1 and +2 His37 states is relatively constant as the proton moves through the channel, with +2 having a slightly higher frequency. The +0 state exhibits an increase in frequency when the proton is below the sidechain, near CEC$_{Z'}$= 0.0 Å, at which point all three states have the same frequency. Interestingly, the orientation of these hydrogen bonds is clearly altered as the proton passes through the Val27 gate. When the gate is closed, there is a water hydrogen bond with the Ser31 sidechain ~71, 76, and 78% of the time (averaged over the four helices) for the +0, +1, and +2 states, respectively. These water molecules form the top layer of water in the middle of the channel – one hydrogen bonded to each Ser31 sidechain in a ring-like structure, situated below Ser31 (**Figure 2.11a**, left panel). The opening of the gate, and subsequent formation of a continuous water wire to the top of the channel, disrupts and shifts the four water molecules such that they interact with Ser31 with an altered orientation as illustrated in the right panel of **Figure 2.11a** and calculated in **b**. In the +0 state, all four tend to interact from a slightly higher position, whereas in the +1 and +2 state the change is predominantly in one water-Ser31 interaction, pointing to the asymmetry in the channel's response to PT. The Ser31 hydroxyl group also donates a strong hydrogen bond to the Val27 backbone carbonyls on the same chain, present ~95% of the time when the excess proton is out of the channel. These interactions are also impacted by the Val27 gate opening, which distorts the geometry and decreases the frequency of backbone hydrogen bonding. These Ser31 interactions indicate how residues can act as a scaffold for PT in M2.

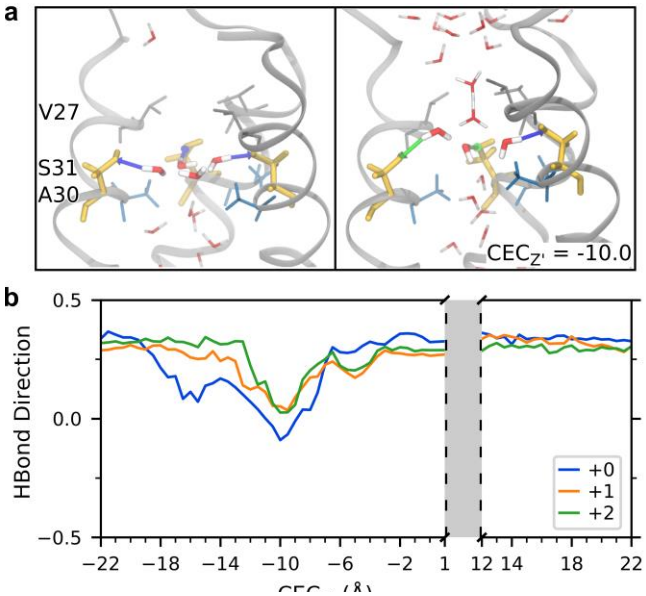

Figure 2.11. Protein-water hydrogen bonds affected by the CEC position. (a) Typical water structures in the +0 state when the excess proton is outside the channel (left) and above the Ser31 tetrad (right). Ser31 is shown in yellow, Ala30 in blue, hydrogen bonds are shown as arrows colored blue and green for bonding from below and above, respectively. Front helix does not have a Ser31-water hydrogen bond and is removed for clarity. (b) Ser31: Average direction of hydrogen bonds between water as donors and Ser31 sidechain oxygen atoms as acceptors.

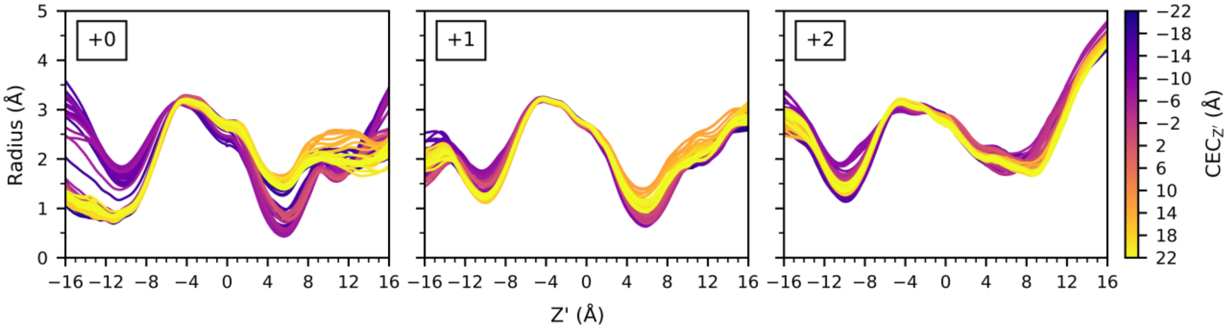

Figure 2.12. Radii profiles for each proton position as in **Figure 2.5**, but here separated by charge state and colored by excess proton position to highlight differences based on CEC$_{Z'}$ more clearly.

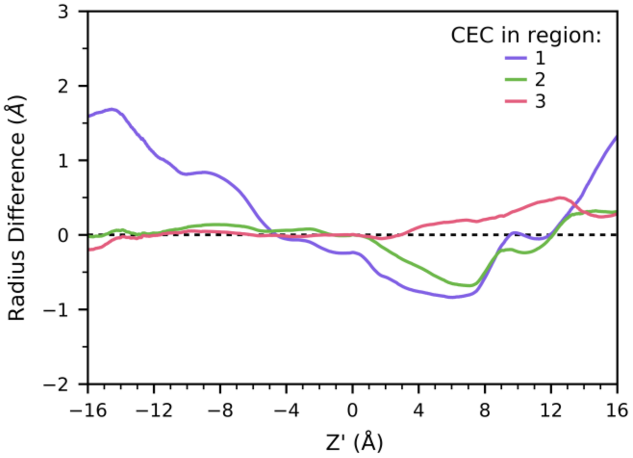

Figure 2.13. Difference plot of radii profiles for CEC in the three different regions of the +0 channel indicated in **Figure 2.6**. Region 1 (purple) corresponds to the range [-18.0, -7.0], from the top of the channel to just below Val27. Region 2 (green) corresponds to the range [-7.0,1.0], from below Val27 to just below Gly34. Region 3 (pink) corresponds to CEC$_{Z'}$ in the range [10.0, 18.0], from Trp41 to the bottom of the channel. The difference is calculated between the radius profile averaged over all CEC positions in a given region and the average radius profile when the CEC is out of the channel.

2.5.2 Channel Water Asymmetry

Table 2.1 shows the percent of frames, averaged over all proton positions, that have hydrogen bonds between water and the Gly34 backbone carboxyl oxygen of each helix. The disparity between the helices is highest in the +0 state – while helix C is hydrogen-bonded (74%) of the time, helix B only has such interactions (25%) of the time, indicating that the water structure is not four-fold symmetric in the dynamic system.

		Helix			
		A	B	C	D
His37 State	+0	55 %	25 %	74 %	41 %
	+1	69 %	56 %	63 %	86 %
	+2	68 %	66 %	72 %	78 %

Table 2.1. Gly34 backbone carbonyl and water hydrogen bond prevalence

2.5.3 Asymmetric Proton Path

As the excess proton approaches the channel, the water is more bulk-like until ~$CEC_{Z'}$ = -14.0 Å, and thus the CEC distribution above this point is more spread out and potentially out of the channel. After this point, the distribution narrows and is more focused in the center of the channel. The proton approaches and passes through the Val27 gate near $CEC_{Z'}$ = -12.0 through -8.0 Å. The CEC distribution is the narrowest in this region due to the small opening afforded by the Val27 gate – the sidechains fluctuate and shift up enough to let one water molecule bridge between the top-outer portion and the interior of the channel.

Around $CEC_{Z'}$ = -5.0 Å the distribution becomes ring-like, most prominent in the +0 and +1 states, as there are no waters directly in the center of the pore and the proton stays around the edge of the pore. The channel pore radius here is the largest of the internal portion of the channel (between Val27 and Trp41). In the +0 state, the proton has propensity towards one side of the channel near helix C. As the proton moves down, this propensity moves to the opposite side of the pore. This asymmetry can be attributed to the proximity of Ala30 backbone carbonyls and the water structure they influence. Past this, the channel narrows, the distribution again becomes more centralized. As the proton passes through the Gly34 region in the +2 state, approaching the His tetrad, there is slightly greater propensity for the proton to be on the sides of the channel above the uncharged His residues. This is likely due to reoriented His-water hydrogen bonds on protonated His residues that would not easily accept an excess proton.

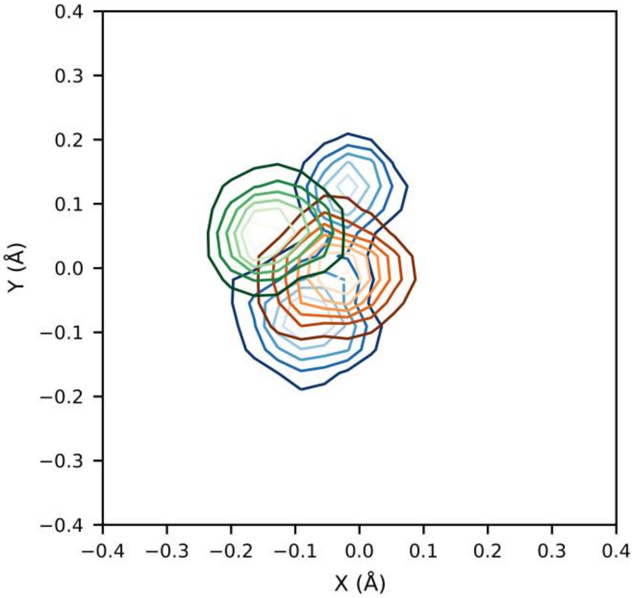

Figure 2.14. Contour plot of unit vectors describing the protein principal axis projected onto xy-plane. The principal axis and its unit vector were found for all frames and the histogram was calculated with bins of length 0.04 Å. The contour plots of the histograms for the +0, +1, and +2 charge states in blue, orange, and green respectively.

2.6 Additional Methods

Principal axis alignment. In order to calculate properties with respect to the protein's principal axis, the atom positions in all trajectories were rotated such that the protein principal axis was aligned with the z-axis. The principal axis was calculated from all non-hydrogen protein atoms, and the rotation matrix about the center of mass of Gly34 alpha-carbons was calculated from the resulting vector. The hydrated excess proton CEC positions were also rotated to get CEC$_{Z'}$.

Radii. Radii profiles were calculated using the program HOLE[65] for each frame of the aligned trajectories. Profiles were binned according to CEC$_{Z'}$ every 0.5 Å, and the average and standard deviation of the profile were calculated for each bin.

Hydrogen-bonding and water analysis. Frames from all aligned trajectories were binned according to CEC$_{Z'}$ every 0.5 Å.

Hydrogen bonds were identified by the following standard criteria, as has been used in previous M2 studies[59, 70, 71]: the distance between the donor and acceptor atoms was less than 3.5 Å, and the angle between the donor, hydrogen, and acceptor was greater than 150°.

Region assignment. The water was divided into four non-overlapping internal sections spanning the area between Val27 and Trp41, chosen in part based on experimental work showing layers of water in the crystal structure.[70] Hydrogen bonds were assigned to regions based on the donor's coordinates.

Abbreviations:

VALSC = geometric center of Val27 sidechain carbons atoms
SER = center of mass of 4 Ser31 alpha carbon atoms
GLY = center of mass of 4 Gly34 alpha carbon atoms
HISSC = geometric center of His37 sidechain atoms
TRPSC = geometric center of Trp41 sidechain atoms

The region definitions:

1. The region below VALSC and at or above SER.
2. The region below SER and at or above GLY.
3. The region below GLY and at or above HISSC, to account for water directly above the His37 tetrad.

4. The region below HISSC and at or above TRPSC, i.e., water between the His37 tetrad and Trp41 gate.

H-bond direction. To determine the direction of a hydrogen bond, the vector between the acceptor and donor atoms was found. The z-component of the unit vector (i.e., the cosine of the angle between the hydrogen bond vector and the z-axis) of this was taken as a measure of the hydrogen bond’s alignment with the protein principal axis (using aligned trajectories), yielding values between (-1.0,1.0).

Alpha-carbon distances. The backbone alpha-carbon distances from the center of the channel were calculated for each frame of each trajectory, for all residues on each helix. The center of the channel in a given simulation for a given residue was defined as the center of mass of the four alpha-carbons of that residue (one from each helix) averaged over that trajectory. Thus, these distances capture any outward-inward fluctuations at a given point along a helix and can distinguish the movements of individual helices. Frames were then binned according to $CEC_{Z'}$, and the averages and standard deviations of the distances for each helix were calculated. These averages and standard deviations were then averaged (this is the weighted average of all distances, where each frame is weighted such that each $CEC_{Z'}$ bin has equal contribution to the overall average).

Excess Proton CEC Histograms. The CEC xyz-coordinates were collected every 100 fs. The CEC position in the rotated frame of reference (with respect to Z’) was calculated based on the frame nearest in time. Coordinates were binned according to $CEC_{Z'}$ every 0.2 Å. The two-dimensional histogram was then calculated using the Python package NumPy for each bin, and images were generated with Gaussian interpolation to smooth over bin edges while preserving the main data features.

Figures of molecular structures were generated using Visual Molecular Dynamics (VMD) software.[72]

Chapter 3

Influenza A M2 Inhibitor Binding Understood through Mechanisms of Excess Proton Stabilization and Channel Dynamics

Reprinted with permission from:

Laura C. Watkins, William F. DeGrado, and Gregory A. Voth. Influenza A M2 Inhibitor Binding Understood through Mechanisms of Excess Proton Stabilization and Channel Dynamics. *Journal of the American Chemical Society* **2020** *142* (41), 17425-17433.

3.1 Introduction

Proton transport (PT) across cellular membranes is a critical component of many biomolecular systems, necessary, for example, to maintain pH gradients,[8, 9] to drive ATP synthesis,[10] and to facilitate the co- or anti-transport of other small molecules.[5, 6, 11] Because of their essential role in such systems, channels and transporters with PT functionality are often targets for drug design to inhibit or control PT—in the case of viruses and bacteria, to slow or prevent infection, but there are myriad other disease applications.[12-14] Drug design is notoriously challenging, as both thermodynamic and kinetic factors must be considered but are difficult to predict and control, and its success depends on high quality structures, an understanding of structural dynamics, and a knowledge of the protein’s function and its mechanism. Thus, beyond elucidating mechanisms of PT in order to understand how a specific channel or transporter works, studying the detailed interactions that facilitate PT can provide valuable insight to help guide drug design efforts.

The influenza virus kills up to 650,000 people each year,[73] and the impact of the recent global coronavirus pandemic[74] emphasizes how critical it is to maintain our focus on understanding and treating viral infections. The influenza A virus matrix 2 (M2) proton channel is a homo-tetrameric protein responsible for the acidification of the viral interior, a critical step in the influenza infection process.[34-36] It is the target of two of the three currently available oral antivirals, amantadine and rimantadine.[37, 75] While these are effective at blocking PT in wildtype M2, drug-resistant mutants have become the predominant strains, the majority of which contain a S31N mutation.[76, 77] This widespread resistance requires a continued drug design effort[78] informed by a deeper understanding of the PT and drug inhibition mechanisms. Additionally, M2 is considered an archetype for the viroporin family, a class of viral channels considered ideal drug targets.[79] The SARS-CoV-2 virus responsible for the COVID-19 pandemic contains two viroporins, protein E and 3.[80-82] Thus, viroporins are a critical class of proteins to study as potential therapeutic targets.

M2 is located in the viral capsid and is acid-activated: as the pH of the endosome encapsulating the virus is lowered, the M2 channel becomes activated and facilitates unidirectional proton flow to the viral interior, allowing the virus to escape the endosome and infect the cell. The key residue that controls activation is His37,[38, 53, 83] which can bind one additional proton and take on a +1 charge. One histidine from each helix forms the His37 tetrad, which can collectively hold a +0 to +4 excess charge, dependent on pH. The channel becomes activated and the C-terminal portion opens (adopting the $\text{Inward}_{\text{open}}$ conformation) upon reaching the +3 state, and PT occurs as the channel cycles through a transporter-like mechanism.[3, 23, 39, 40, 42, 43, 84, 85]

Amantadine and rimantadine belong to the adamantyl-amine class of inhibitors, binding in the upper-middle portion of the channel. These drugs were the predecessors of many related adamantane-based compounds featuring a relatively rigid, apolar group and an attached charged group.[86-93] Recently, Thomaston et al. published several high-resolution X-ray crystal structures of M2 with amantadine, rimantadine, and a novel spiro-adamantyl amine bound.[66] These structures provided the first opportunity to see the specific interactions that facilitate stable inhibitor binding and the disruption of the hydrogen-bonded water network otherwise present. Along with an earlier qualitative MD simulation study that guided the design of the spiro-adamantyl amine inhibitors,[94] the crystallographic analysis provided potential insights into the mechanism of inhibition, suggesting that the backbone carbonyls of pore-lining residues act as "physiochemical chameleons", able to engage in both hydrophobic and hydrophilic interactions, and that the drug is tilted off the channel's axis and interacts with waters in the Ala30 layer. Taken together, it is hypothesized that amantadine acts as a mechanism-based inhibitor, with the ammonium group functioning as a hydronium mimic. Computational studies to date have primarily focused on the means of entry into the channel and location of binding,[95-100] but have not deduced specific interactions between the drug, channel, and channel water involved in binding as they relate specifically to similar interactions seen in the PT mechanism.

Proton transport is an inherently quantum mechanical process, as the hydrated proton structure (hydronium-like) exists in a complex hydrogen-bonded network that rearranges dynamically as bonds break and form according to the Grotthuss shuttling mechanism.[15-17] Thus, classical molecular dynamics (MD) with fixed bonding topology cannot be used to study PT; moreover, *ab initio* methods are not efficient enough to reach the many nanosecond timescales necessary to obtain sufficient sampling in biomolecular systems that may have important degrees of freedom several orders of magnitude slower than proton shuttling. Multiscale Reactive Molecular Dynamics (MS-RMD)[18-21] (and Multistate Empirical Valence Bond, MS-EVB, before it) was developed to efficiently and accurately capture the solvation and delocalization of an excess proton in water, such that the quantum-chemical nature of the hydrated proton can be studied in the context of membrane proteins over the long timescales needed for accurate simulation of such systems. MS-RMD has been successfully applied in several protein systems to predict and explain mechanisms of PT.[3, 22-30]

In previous work,[3, 23] quantum mechanics/molecular mechanics (QM/MM) and MS-RMD were used to calculate potentials of mean force (PMFs, i.e., free energy profiles) of PT through the M2 channel in the +0-3 states, providing critical insight into the pH-dependent activation behavior and the role of the His37 tetrad in PT. Most recently, we further analyzed the MS-RMD simulations to explore the detailed interactions between the hydrated excess proton and the channel and found that the proton dynamically, as a function of its position, alters several properties of the protein and pore waters, including the hydrogen-bonding network and the protein structure.[30] This latter work illustrates how MS-RMD can be used successfully to investigate explicit, dynamic interactions between a hydrated proton and its immediate environment, as well as its indirect effects on other parts of the system. Here, we employ a similar approach as in this previous work to focus specifically on properties related to drug binding and how the position of the bound drug relates to the overall PT mechanism. Through this analysis, we examine the hypothesis that the adamantyl-amine drugs act as mechanism-based inhibitors—by identifying stabilizing interactions between the excess proton and the channel, we show how the drug may similarly be stabilized in support of this hypothesis. Additionally, by examining conformational fluctuations, we show that the drug binding pocket is an especially stable and symmetrical portion of the channel, conducive to binding a roughly spherical drug, and we reveal an additional minor hinge point towards the top of the channel which may be a relevant feature for future drug design efforts.

3.2 Methods

Simulations for calculating properties as the proton moves through the top of the channel were run as follows. Starting structures were taken from previous simulations, which were initiated from a crystal structure of the transmembrane portion of the M2 channel (this construct is referred to as M2TM) resolved at room-temperature and high pH (PDB: 4QKL[59]) embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer solvated with water. M2TM is the minimum construct necessary to retain proton conduction similar to full-length M2,[60] and it has been shown that the presence of amphipathic helices, included in the full-length M2 protein, do not significantly influence the PT mechanism.[3] The water and excess proton in the system were modelled using the MS-RMD method, which allows bonds to break and form by taking a linear combination of different bonding topology states at each timestep. The MS-EVB version 3.2 parameters[101] were used to describe the hydrated excess proton. We refer the reader to previous work for a full description of the method.[18-21] The excess proton center of excess charge (CEC) is defined as:[102]

$$\vec{r}_{CEC} = \sum_{i}^{N} c_i^2 \vec{r}_{COC}^{\,i}$$

where $\vec{r}_{COC}^{\,i}$ is the center-of-charge of the i^{th} diabatic MS-RMD state, and c_i^2 is the amplitude of that state. The sum is over all N states. The interactions in the remainder of the system were defined by the CHARMM36 forcefield. Simulations were run in the NVT ensemble at 308.0 K using LAMMPS[103] (http://lammps.sandia.gov) with the MS-RMD package. The collective variable (CV) defined for umbrella sampling (US) is the z-coordinate of the vector between the excess proton CEC and the center of mass of the four Gly34 alpha carbons, as in our previous work, such that the CV has negative values at the top (N-terminal end) of the protein and progresses to positive values at the bottom (C-terminal end). Simulations were run with the excess proton at every 0.5 Å along the CV coordinate between -18.0 and 1.0, generating 39 windows. To ensure that the proton remained in the channel, a cylindrical restraint was added at 8 Å with a force constant of 10 kcal/mol·$Å^2$ using the open-source, community-developed PLUMED library.[104, 105] After a 250 ps MS-RMD equilibration, the replica exchange umbrella sampling[61] technique was used to facilitate convergence. Production simulations were run for ~2-4 ns with frames saved every 1 ps.

For calculating hydrogen bond residence times, longer independent trajectories were run with the CEC restrained in 5 different positions using US as described above. Each trajectory was run for 1.75-2 ps with frames saved every 10fs.

Simulation frames were binned by excess proton CEC value for subsequent analyses, which were performed in Python[106] using the SciPy,[107] NumPy,[108] and pandas[109] libraries. For hydrogen bond analysis, values were averaged over the four helices. Hydrogen bonds were defined by the following criteria: the donor-acceptor distance must be less than 3.5Å, and the donor-hydrogen-acceptor angle must be greater than 150°. Several hydrogen bond definitions were tested and did not affect the conclusions (not shown). For calculating residence times, a hydrogen bond was considered in place as long as the particular water molecule remained the closest water to the protein atom and the hydrogen bond criteria were met.

Images of molecular structures were rendered in Visual Molecular Dynamics (VMD),[72] while other figures were generated using Matplotlib.[110]

3.3 Results and Discussion

If the adamantyl-amine drugs are acting as mechanism-based inhibitors as hypothesized, we would expect to see specific aspects of the PT mechanism taken advantage of or replicated by the drug upon drug binding. To test this, we performed MS-RMD simulations of M2 in the +0 His37 charge state with an explicit excess proton to evaluate the hydrogen-bonding networks, pore shape, and protein fluctuations throughout PT that relate to drug binding. By focusing on PT in the +0 state, we are studying the process of proton entry and diffusion to His37 in the first key step of channel activation, paralleling inhibitor entry into the channel. We additionally expect this to be a prevalent charge state in drug-bound structures due to the lowered His37 pKas.[111] Replica exchange umbrella sampling was used to obtain sufficient sampling of proton positions throughout the top portion of the channel, with windows from CEC_Z = -18.0 to 1.0 Å where the coordinate origin is defined as center of mass of the Gly34 alpha carbons. The channel is aligned along the z-axis for all subsequent analyses.

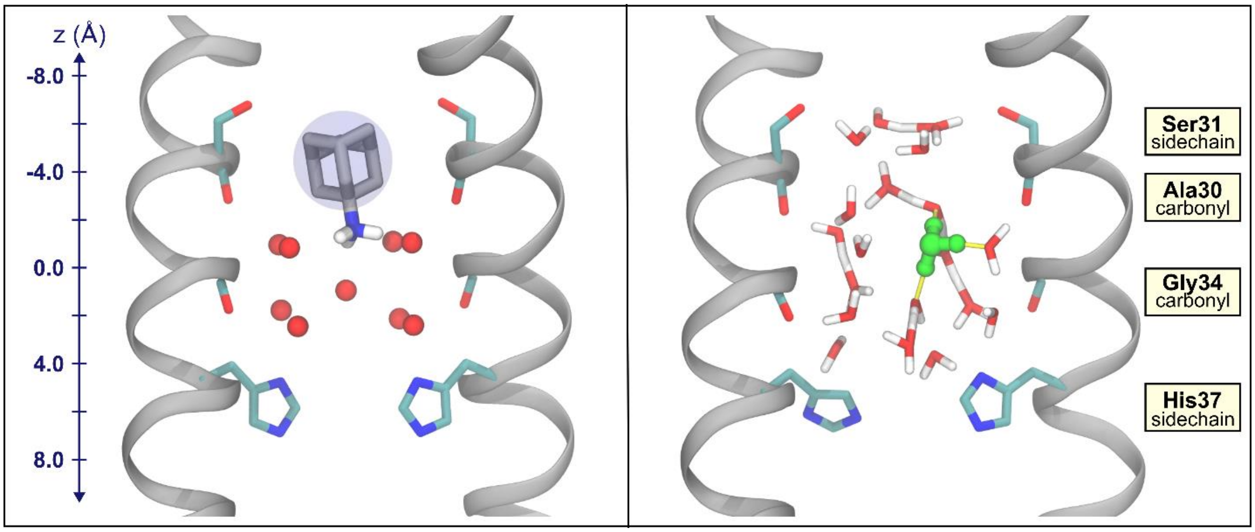

Figure 3.1. Comparison of M2 with amantadine and the hydrated excess proton. Left, x-ray crystal structure of amantadine-bound M2, PDB 6BKK. Water oxygens are shown in red. A blue circle around the adamantane group is depicted to reflect the spherical nature of this group. Hydrogens on the ammonium group of amantadine were added. Right, a snapshot from an MS-RMD trajectory with the most hydronium-like water indicated in green. In both, two opposing chains of M2 are shown in silver. The Ser31, His37 sidechains and the Gly34, Ala30 backbone carbonyls are shown. The z-coordinate for the system is included on the left, where z = 0 Å is defined as the center-of-mass of the Gly34 alpha-carbons.

To understand how properties of PT may provide insight into drug binding, we primarily examine variations dependent on proton position. We compared the values of each property when the proton is at the drugs' ammonium group positions versus other parts of the channel to determine if the drugs could be taking advantage of the channel's natural ability to stabilize a proton. This idea is highlighted in **Figure 3.1**, which shows both the drug-bound crystal structure and a snapshot of a hydrated excess proton in the channel from our simulations. We refer to the drugs' ammonium nitrogen position along the z-axis in the crystal structure as $\mathrm{AmmN_Z}$. This value is -1.7 and -1.5 Å for the $\mathrm{Inward_{closed}}$ amantadine and rimantadine bound structures, respectively (averaged over the two tetramers in each crystal structure).

3.3.1 Flexible Hydrogen Bonds Stabilize the Excess Proton near $\mathrm{AmmN_Z}$.

It has been shown in our previous work that hydrogen bonds within the channel, including those between water and protein atoms, help facilitate proton transport by altering their direction and frequency of interaction as the proton moves through the channel. Here, we focus specifically on water interactions that may help account for excess charge stabilization near $\mathrm{AmmN_Z}$. In **Figure 3.2**, we calculate the occupancy of three different hydrogen bonds between protein atoms and water as a function of the excess proton position in the channel. While the Ala30 hydrogen bond occupancy is constant as the excess proton enters and moves through the top of the channel, as it approaches the Ala30 carbonyls, the occupancy decreases ~20%. This dip indicates the Ala30 hydrogen-bonded waters can flexibly reduce their interaction with the protein as a result of an excess charge in their vicinity. Additionally, this dip is centered at -2.5 Å, near $\mathrm{AmmN_Z}$. At this point, the role of the waters near Ala30 carbonyls in hydrating the proton is maximized. In a previously published PMF of the +0 His37 charge state,[3] there is notably a local minimum near this point, further indicating that an excess charge is relatively stable here. This supports the hypothesis that amantadine and rimantadine are mechanism-based inhibitors and take advantage of the channel's natural ability to stabilize a hydrated excess proton in order to stabilize the drug's ammonium group.

The hydrogen bond occupancy of waters with the Gly34 carbonyls increases once the excess proton passes through the Val27 gate and remains fairly consistent across proton positions thereafter, exhibiting little dependence on the hydrated proton position once it is in the channel. The Ser31 sidechain water occupancies are shown for comparison, which do not show a noticeable trend based on proton position. Thus, this change in interactions is not a universal effect throughout the channel, but the Ala30 waters seem to be uniquely flexible in this manner. These differences are consistent with drug design studies —while compounds such as spiro-adamantyl amine have been able to displace the water in the Ala30 layer, no designed inhibitors have displaced the water around Gly34.

To further understand how the dynamics of the hydrogen-bond network may show how these drugs benefit from the channel's inherent excess-charge stabilization used in proton transport, we examined the average residence times of hydrogen bonds between water and several important protein atoms. To do this, independent trajectories were run for five different excess proton positions, including two trajectories with the proton completely outside the channel (CEC_Z = -24.0, 24.0 Å) and three when the proton is near $AmmN_Z$. These results are shown in **Figure 3.3**. The Ala30 water residence times slightly increase when the excess proton is near $AmmN_Z$, the Ser31 water residence times do not show any significant difference between the proton outside the channel and at $AmmN_Z$, and those of Gly34 waters decrease at $AmmN_Z$. The waters hydrogen bonded to His37 imidazole nitrogens show the greatest change in residence times and are shown to highlight the ability of this method to describe such differences.

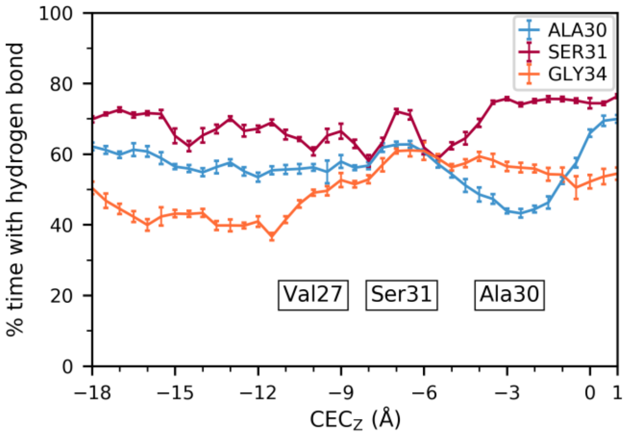

Figure 3.2. Hydrogen bond occupancy averages and error as a function of hydrated excess proton position between water and the backbone carbonyls of Ala30, Gly34, and the sidechain hydroxyl group of Ser31. Approximate average positions of Val27 sidechains, Ser31 sidechains, and Ala30 carbonyls are indicated.

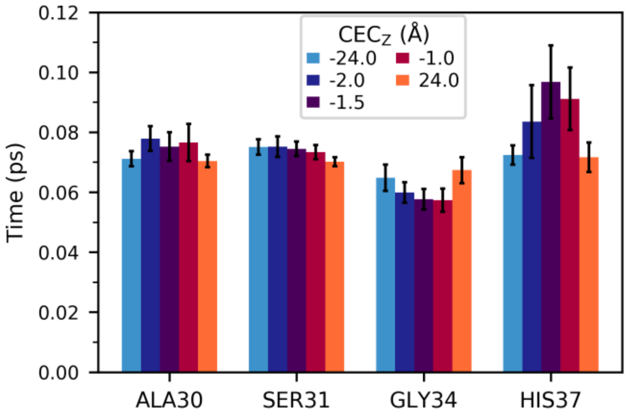

Figure 3.3. Average hydrogen bond residence times, calculated for the backbone carbonyls if Ala30, Ser31, and Gly34, and the unprotonated nitrogen of the His37 imidazole sidechain. The quantities are averaged over the interactions of all four helices, and error bars were calculated using block averaging. Each bar represents the value calculated from one trajectory with the proton restrained at the labeled Z coordinate.

With the above results for Ala30, this may indicate that several waters remain tightly hydrogen bonded to the Ala30 backbone carbonyls, while one or more are bonded less frequently. While the Gly34 hydrogen bonds do not form less frequently (as indicated in **Figure 3.2**) with an excess charge in this region, they do exhibit greater dynamics and flexibility. This change indicates an increase in water dynamics when the excess proton is near, which could help stabilize and solvate the excess charge in the Ala30 water layer.

Taken together, these results further support the hypothesis that amantadine and rimantadine act as mechanism-based inhibitors: the channel acts as a scaffold to facilitate PT by harboring flexible protein-water interactions that can adapt and respond to a positive excess charge, with specific ability to stabilize an excess proton near $\mathrm{AmmN_Z}$.

3.3.2 Drug Tilt Positions Ammonium Group in Highest CEC Density.

One prominent characteristic of the amantadine and rimantadine bound structures is the drug's tilt within the pore. This tilted conformation is also seen in solid-state NMR studies.[112] Given the drug's three-fold symmetry in a four-fold symmetric channel, the ammonium group cannot form hydrogen bonds with all 4 waters hydrogen-bonded to Ala30, leading in part to this tilt. Based on our previous work examining the proton's path through the channel, we used a similar analysis to examine the density of CEC positions when the excess proton is near the ammonium position in drug-bound structures. **Figure 3.4A** shows the difference in CEC density when the proton is near $\mathrm{AmmN_Z}$ compared with the average over all proton positions through the top portion of the channel. This 2-dimensional histogram of CEC positions in the XY-plane is calculated for $\mathrm{CEC_Z} = -1.7 \pm 0.2$ Å, minus the average over all normalized histograms for $\mathrm{CEC_Z}$ positions [-18.0, 1.0] Å binned every 0.2 Å. Interestingly, in this portion of the channel the excess proton prefers to be near the edge of the pore, unlike the predominant preference for the center of the pore throughout the rest of the channel, as indicated by the positive values around the edge and negative values in the center **Figure 3.4B** shows the radial density of the CEC in this same region of the channel. Possible positions of amantadine's ammonium group nitrogens were calculated based on the drug's position and tilt in the crystal structure, and their radii are included as dashed lines (these positions were also used to generate the image in **Figure 3.1**). These hydrogens can extend to a radius ~1.5 Å in this static crystal structure, which means that the slightly off-centered ammonium group directly positions 1-2 of its hydrogens in the region of the CEC's highest density. The CEC's propensity for the edge of the pore indicates that the drug's tilt in the channel may not only be a necessary component of its binding, but also a thermodynamic advantage. This tilt further allows the drug to act as a hydronium mimic, as the hydrogens of the ammonium group are in the favorable positions of the solvated excess proton.

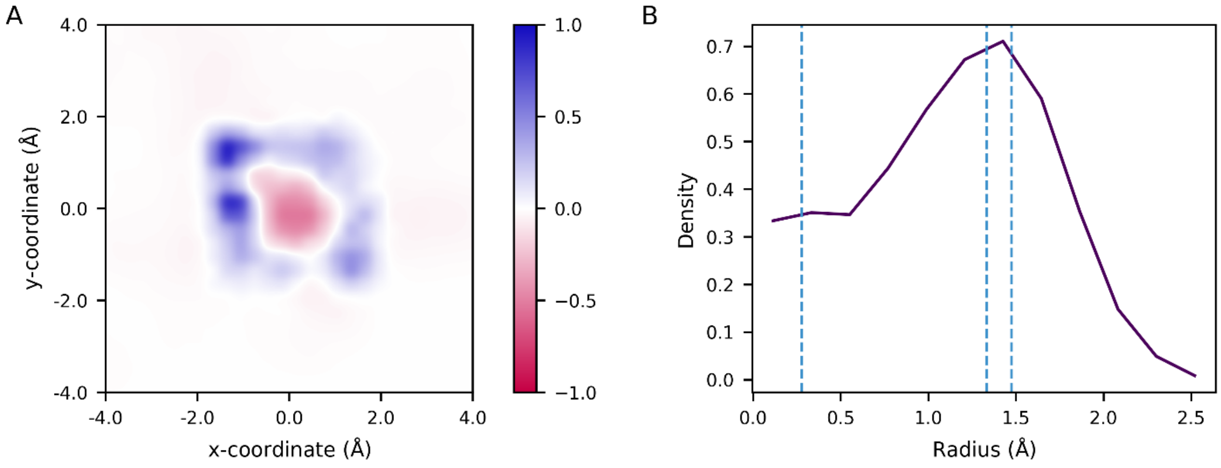

Figure 3.4. A. The difference in hydrated excess CEC density in the xy-plane when the CEC is located within CEC_Z=-1.7±0.2 Å compared with the average CEC density over all CEC positions in range z=[-18.0,1.0]. **B**. The average radial density of the proton when CEC_Z =-1.7±0.2 Å. One possible set of positions of amantadine's ammonium group hydrogens calculated from the drug-bound crystal structure are indicated by blue dashed lines.

The analysis of hydrogen bonding changes and proton densities indicates how the ammonium group is a functional addition to the adamantane scaffold, as the charged group is positioned in a region where the channel is especially adept to stabilize an excess charge. This stabilization relies on flexible water structures and hydrogen-bond interactions that can undergo minor changes to accommodate the proton, suggesting that the adamantyl-amine inhibitors are acting as hydronium mimics — they take advantage of these inherent features to help solvate the charged ammonium group. The identification of other regions of the channel with increased ability to stabilize an excess charge, such as areas of increased proton density or significantly flexible water interactions, could help provide new targets for drugs to act as hydronium-mimics.

3.3.3 Pore Shape and Stability Near Ser31 are Ideal for Adamantane Binding.

Another hypothesis about the adamantyl-amine class of inhibitors is that adamantane is effectively spherical and can freely rotate within the channel, but has no rotatable bonds, which minimizes the entropy lost upon binding. This rapid rotation can be seen on the NMR time scale[112] and is consistent with the recent Thomaston et al. crystallographic studies, in which the motion was indirectly inferred. Nevertheless, its significance depends on the dynamic nature of the channel — if the protein exhibits great structural fluctuations in the region where the drug binds, then drug binding may induce changes that greatly decrease the entropy and this hypothesis would not fully explain the drugs' efficacy. To better understand how the channel's natural dynamics may lend itself to favorable drug binding, we examined the pore shape throughout our trajectories. As an estimate of the asymmetry of the channel, we calculated the eccentricity, which is essentially a measure of how "circular" a given oval is. The eccentricity is defined as:

$$e = \sqrt{1 - \frac{b^2}{a^2}}$$

where *a* and *b* are the semi-major and semi-minor axes, respectively, which we approximate by the distance between alpha-carbons on opposing helices. A schematic of this is shown in SI **Figure 3.7**. Eccentricity can have values between 0 and 1, with 0 indicating a circle and 1 indicating a parabola.

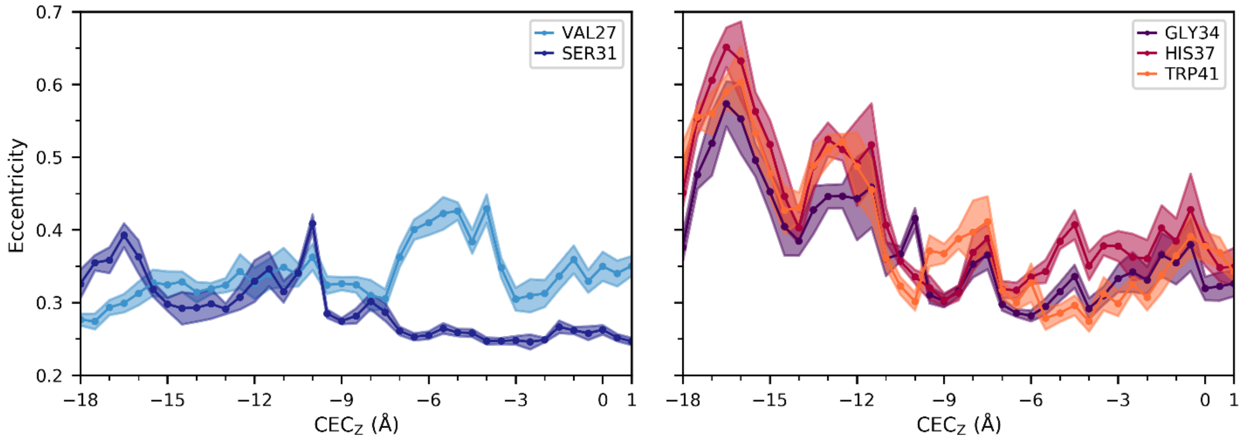

Figure 3.5. Average and standard deviation of the pore eccentricity estimated by alpha-carbon distances of the pore-lining residues as a function of the hydrated excess proton CEC position.

These results are shown in **Figure 3.5.** Average and standard deviation of the pore eccentricity estimated by alpha-carbon distances of the pore-lining residues as a function of the hydrated excess proton CEC position.. Eccentricity maximum, minimum, and RMSD values are calculated in SI **Table 3.1**. The pore-lining residues in the bottom half of the channel, Gly34, His37, and Trp41, all show a greater degree of asymmetry and a wider range of eccentricity values, dependent on the excess proton position, than the pore-lining residues in the top part of the channel. Interestingly, proton entry at CEC$_z$ = -17 Å has a pronounced effect on the channel near Trp41, greater than that when the proton nears the center of the channel. Ser31, however, has overall the smallest average eccentricity and the lowest minimum value than the other pore-lining residues during PT in this portion of the channel. Additionally, Ser31 and Val27 have smaller proton position dependent changes in eccentricity than the pore-lining residues in the bottom half of the channel. This result indicates that the Ser31 region is the most symmetrical and stable in the channel.

While analyzing the alpha-carbon distances and eccentricity, we also examined the correlation between these alpha-carbons distances on opposing helices, shown in **Figure 3.6**, to gain further insight into protein motion and conformational fluctuations on the nanosecond timescale. These motions captured here are equilibrium fluctuations in the +0, Inward$_{\mathrm{closed}}$ state, not necessarily motions driving the transition between Inward$_{\mathrm{open}}$ and Inward$_{\mathrm{closed}}$. The calculated correlations indicate that the channel's equilibrium structural fluctuations are dominated by alternating inward-outward motions of opposing helices. At each pore-lining residue, the distances are negatively correlated—that is, when helices A and C move farther apart, helices B and D move closer together, and vice-versa.

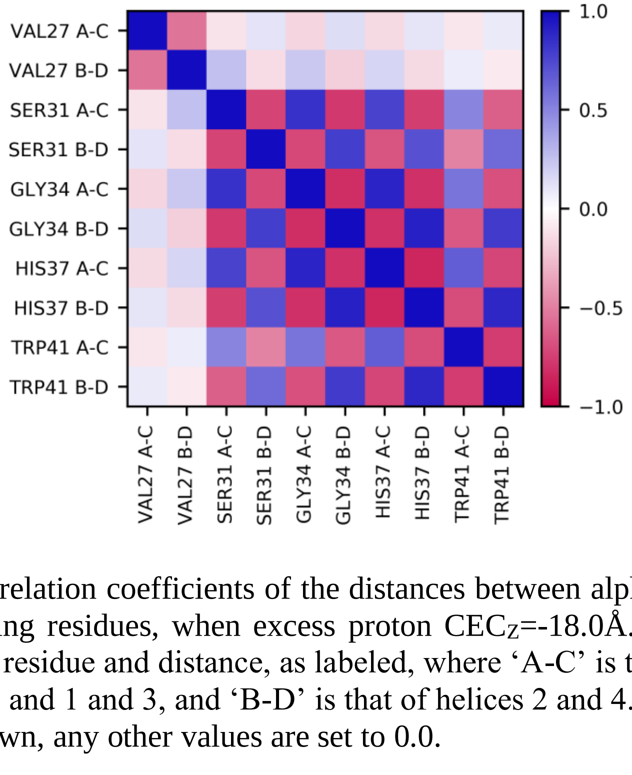

Figure 3.6. Pearson correlation coefficients of the distances between alpha-carbons on opposing helices, for all pore-lining residues, when excess proton CEC$_Z$=-18.0Å. Each row and column correspond to a specific residue and distance, as labeled, where 'A-C' is the distance between the alpha-carbons of helices and 1 and 3, and 'B-D' is that of helices 2 and 4. Only those values with a p-value < 0.05 are shown, any other values are set to 0.0.

Gly34 is known to be the hinge point whose kinking controls the large structural change between inward-open and inward-closed conformations, which may falsely lead to the conclusion that the conformational fluctuations at equilibrium above and below Gly34 are decorrelated, with a stable core centered at Gly34. Interestingly, however, the motions at Gly34 are strongly and similarly correlated with the motions at both Ser31 and His37. This correlation indicates that in this fixed charge state, the Gly34 kink is relatively rigid. Instead, there is a noticeable lack of correlation between Val27 and the other pore-lining residues, suggesting that there is a secondary, minor "hinge" between Val27 and Ser31 that decorrelate the inward-outward motions between the helices above and below this point. This natural hinge observed near Val27 furthers our understanding of Val27 acting as a secondary gate that opens to allow proton and water entry into the channel.[113, 114] In our simulations, this valve can readily hydrate, particularly in the presence of a nearby excess proton. Moreover, it is frequently closed, which may make passage of a hydrated sodium or chloride ion more difficult. This aspect of the Val27 gate and its relevance for PT and proton selectivity is likely an important feature of the M2 channel and could be further explored in future work.

Given that the adamantane group of the drug is centered in the Ser31 tetrad plane, we hypothesize that these facets of the channel's dynamics are critical for fully explaining the drugs' favorable binding. Because of the more circular shape of the pore at the Ser31 tetrad, the spherical adamantane group can fit snugly under the hydrophobic Val27 cleft and block PT. Additionally, the relative stability of the pore in the region of drug binding helps explain why drug binding is thermodynamically favorable. Because the channel exhibits smaller structural fluctuations here than in other regions of the channel, the adamantane based drugs are able to bind with minimal loss of entropy as the channel does not need to lose flexibility to create a stable drug-binding interface. This has important implications for designing drugs that use scaffolds different from the adamantane group,[90, 115-120] or that interact with drug-resistant mutants such as S31N. While drug binding to more flexible regions of the channel is possible, only modest changes in potency are often observed despite large changes in the size of the drugs. This is likely because of the need to counter the greater loss of entropy resultant from structural changes and reduced fluctuations.

One limitation of this study, however, is the homogeneous POPC bilayer, which is commonly used in experiments and is standard in computational studies, but does not capture the complexity of the viral membrane and may influence channel dynamics.[121] Thus, these features need to be examined in the more complex membrane to fully understand their physiological relevance.

3.4 Conclusions

Altogether, we have shown how the adamantyl-amine inhibitors of M2 are suited to exploit various inherent features of the M2 channel that naturally facilitate proton transport, further supporting the claim that they function as mechanism-based inhibitors. The flexible hydrogen bond interactions, measured in both hydrogen bond occupancy and residence times, indicate how the channel is suited to stabilizing an excess charge near $\mathrm{AmmN_Z}$. Thus, the ammonium group of these inhibitors can act as a hydronium-mimic by binding in this region. We also analyzed the pore shape throughout the channel by calculating the eccentricity of the pore based on alpha-carbon distances. The results from these calculations indicate that the drug binding pocket is an especially stable and symmetrical portion of the channel, conducive to binding a roughly spherical drug. Finally, by examining the correlations between these distances, we found an additional minor hinge point towards the top of the channel which may be a relevant feature for future drug design efforts.

Understanding these features as they relate to drug binding gives further insight into the specific interactions that stabilize the adamantyl-amine inhibitors in wildtype M2, and these results suggests that this approach could be used to aid future . drug-design efforts to methodically create new inhibitors for S31N mutants. With the recent publication of high resolution influenza B M2 (BM2) structures,[122] it is possible to conduct similar studies to elucidate the detailed PT mechanism in BM2 to guide drug design in this functionally similar protein.

This work also shows how similar analyses to understand the details of explicit proton transport mechanisms (not those inferred by water structures alone) could be used in other systems, and extended to ion transporters such as the SARS-CoV-2 viroporins, to help inform mechanism-based inhibitor design. Elucidating the inherent features of drug-targetable proton transporters, such as flexible water and hydrogen bonding interactions, preferred proton positions, dynamic pore shapes, and structural fluctuations, can help guide the design of drug scaffolds and added substituents. The MS-RMD simulation methodology utilized in this work has made these studies possible both for M2 and other important drug targets.

3.5 Supplementary Information

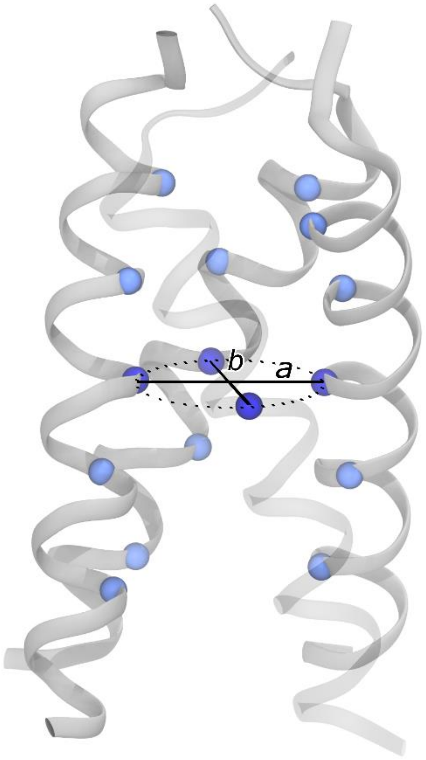

Figure 3.7. A schematic of eccentricity calculations, with pore-lining alpha-carbons shown as blue spheres. Dark blue spheres are Gly34 alpha-carbons. The major and minor axes of the oval are defined by the distances between alpha-carbons on opposing helices, indicated as lines *a* and *b*.

Table 3.1. Eccentricity values of the pore at each of the pore-lining residues, calculated using alpha-carbon positions as described in the text. Eccentricity was calculated for proton positions in the top half of the channel, CEC_z = [-18.0,-1.0] Å.

Residue	Average *e* over all CEC_z	Max value of *e* and position	Min value of *e* and position	Difference between max and min values	RMSD
Val27	0.34	0.43 CEC_z = -4.0 Å	0.27 CEC_z = -17.5 Å	0.16	0.13
Ser31	0.29	0.41 CEC_z = -10.0 Å	0.25 CEC_z = -2.5 Å	0.16	0.13
Gly34	0.38	0.57 CEC_z = -16.5 Å	0.28 CEC_z = -6.0 Å	0.29	0.17
His37	0.42	0.65 CEC_z = -16.5 Å	0.30 CEC_z = -9.0 Å	0.35	0.18
Trp41	0.40	0.60 CEC_z = -16.0 Å	0.27 CEC_z = -4.0 Å	0.33	0.17

Chapter 4

Multiscale Simulation of an Influenza A M2 Channel Mutant Reveals Key Features of Its Markedly Different Proton Transport Behavior

4.1 Introduction

Viroporins are small viral ion channels that are crucial for viral pathogenicity.[123] They are found in many clinically relevant viruses including coronaviruses, influenza A virus, HIV-1, and hepatitis C virus, and are ideal therapeutic targets. The influenza A matrix protein 2 (M2) proton channel is considered a prototype for viroporins and is the target of the antiviral drugs amantadine and rimantadine.[37] Following the encapsulation of the virus in a cellular endosome, M2 is responsible for the acidification of the viral interior via unidirectional proton transport (PT), which allows the virus to escape the endosome and infect the cell.[34, 124, 125] The prevalent Asp44 to Asn (D44N) mutation results in increased proton conduction, which helps protect more acid-labile hemagglutinin in certain influenza strains.[126-128] While much is known about PT in wildtype (WT) M2, the increased conduction in the D44N mutant is not as well understood. Studying this viroporin is important both for increasing our understanding of PT in M2, as well as for identifying the fundamental principles of PT through viroporins such as the SARS-CoV-2 ORF3a protein.[129]

M2 contains four central histidine residues (His37) that select for protons and control the channel's activation.[3, 23, 38] At high pH, all four residues are deprotonated, which we refer to as the +0 state, and the C-terminus of the channel is contracted in an inactive C_{closed} conformation. As the pH is lowered, each histidine residue can bind one additional proton, up to the +4 state. It is

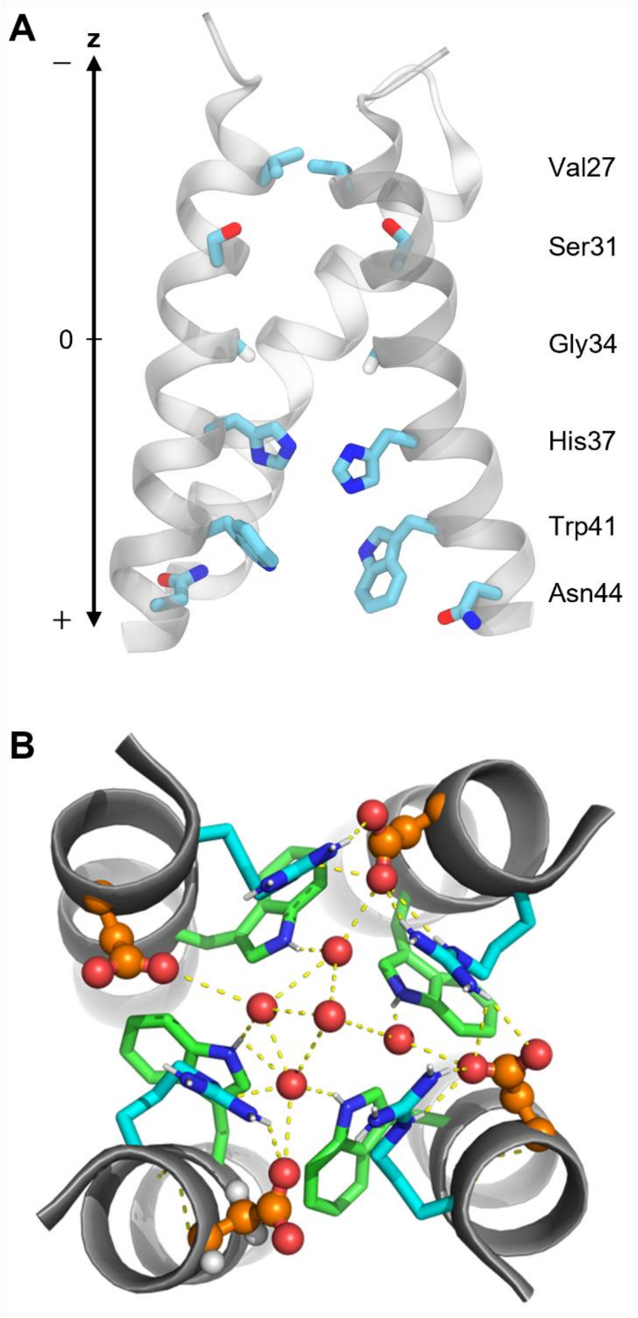

Figure 4.1. A. Equilibrated +0 D44N M2 structure with pore-lining residues. Front helix and sidechains on back helix removed for clarity. **B**. Structure of the Trp41 gate in WT M2 (PDB: 6US8[2]) looking from the C-terminus (interior, +z direction) towards the N-terminus (exterior, -z direction). The carbon atoms of Trp41, Asp44 and Arg 45 are, respectively in green, orange, and cyan. Water molecules are shown as red spheres.

believed that most proton conduction occurs by cycling through the +2 and +3 states based on experimentally determined pK_a's, pH-dependent conduction, and multiscale computer simulations. [3, 23, 49-53] As a third excess proton enters the channel in a +2 state and binds to His37, electrostatic repulsion between these residues causes the C-terminal ends of the helices to widen into a C_{open} conformation which opens the Trp41 gate,[39, 40, 42, 43, 84] increasing hydration and facilitating inward proton flux.[3, 23, 44-48] NMR experiments of M2 at low pH show a broadening of peaks, suggesting a dynamic and flexible ensemble of C_{open} conformers.[43, 130]

One critical characteristic of PT in M2 is its asymmetric conductance, or inward rectification — WT M2 only becomes activated and conducts protons when the exterior pH is low.[55, 131] When only the interior pH is lowered, M2 does not show any significant outward current, indicating that it cannot reach the necessary +2 charge state from internal protons. This asymmetry is moderated by the Trp41 gate, which forms a hydrophobic barrier in the C_{closed} state.[39] With little water access to the His37 tetrad, the channel does not easily take up an excess proton from the interior and does not activate. Mutation of Trp41 removes this gate, resulting in outward conduction under low pH_{in} conditions.[39] The D44N mutant also does not exhibit asymmetric conductance and can activate at low pH_{in}. Previous studies have shown that Asp44 plays a particularly important structural role in stabilizing the closed conformation of the Trp41 gate.[43, 46, 131] The four Asp carboxylates in the tetramer form an extensive network of hydrogen-bonds that include a salt bridge to Arg45 from neighboring helices, and a network of waters that hydrogen-bond to the Trp41 indole NH groups (**Figure 4.1**). The D44N mutation is expected to disrupt the salt bridge and the water-mediated hydrogen bonds to the Trp indole groups. Indeed, NMR studies by Ma, et al, show that this substitution destabilizes the Trp gate,[126] increasing its dynamics and hydration. This resulting destabilized gate was hypothesized to allow unimpeded protonation and deprotonation of His37 from either side of the bilayer. These conclusions were supported by relatively short MD simulations, but the dynamics of the D44N mutant impeded attempts to structurally characterize D44N.

Additional experiments by Ma, et al, confirmed that Asp44 is essential for asymmetric conductance.[126] This work also showed that the D44N mutant has increased conduction and that the inward conduction has an altered pH_{out} dependence relative to WT. In the physiologically relevant pH range 6-5 (the pH of the endosome is approximately 6.0-5.4), the D44N mutant M2 is 3- to 4-fold more conductive than WT M2. Additionally, the pK_a values of two distinct protonation states relevant to conduction (hypothesized to be the formation of the +3 and +4 states) are shifted up in the D44N mutant, indicating that channel activation occurs at a higher pH. NMR spectra in that study also showed that the D44N C_{open} and C_{closed} structures are very similar to those of WT M2, and that mutation induces conformational changes similar to those in WT at low pH.

It has not been studied and quantified, however, how the D44N mutation and corresponding structural changes affect PT on the molecular level, including the explicit proton transport. To this end, we have utilized extensive Multiscale Reactive Molecular Dynamics (MS-RMD)[18-21] and Quantum Mechanics/Molecular Mechanics (QM/MM) simulations with an explicitly transporting excess proton to study PT in the transmembrane portion of D44N M2. These simulation methods allow one to address dynamical questions that are not accessible using classical (non-reactive) MD, which does not allow the bond-breaking and bond-making steps involved in proton transfer. MS-RMD was developed to efficiently and accurately model the solvation and delocalization of a hydrated excess proton in water within a classical MD context for aqueous and biomolecular systems. With potentials fit from accurate QM calculations,[101] MS-RMD allows protons to form and break bonds with water molecules, capturing the underlying physics of the Grotthuss shuttling (proton hopping) mechanism,[15-17] while also reaching the many-nanosecond timescales necessary for studying PT in complex biomolecular systems. MS-RMD has successfully been applied to several protein systems to predict and explain PT mechanisms, including influenza A M2.[3, 22, 23, 26-30, 132-135]

In our previous work,[3, 23] MS-RMD and QM/MM simulations of the WT M2 channel were used to calculate potentials of mean force (PMFs) of explicit PT in the +0 through +3 protonation states, providing fundamental insight into the PT mechanism, the role of the His37 tetrad, and the mechanism of activation in the +2 state. Recently, we further used these simulations to explore in detail how the hydrated excess proton interacts with the channel, and found that the excess proton dynamically, as a function of position, alters the protein structure and water hydrogen-bonding network.[30] This approach was subsequently used to examine the precise interactions between the excess proton, water, and channel in the adamantine binding pocket.[135] There, we showed that the channel is especially stable in the binding region, and we demonstrated how amantadine inhibitors take advantage of the channel's natural ability to stabilize an excess proton charge.

Here, we employ MS-RMD and QM/MM simulations to deduce how the PT mechanism in D44N M2 differs from WT, how the mutation results in increased conductance, and why rectification is lost. The MS-RMD method is used to simulate PT in the +0, +1, and +2 protonation states through the channel. QM/MM simulations were used for PT through the His37 region, and together with MS-RMD were used to calculate the PMF of PT through the +2 state, the first protonation state relevant to conduction. Our results show that the +0 and +1 mutant states are structurally similar to the WT +2 state, and the +2 mutant state is significantly more open at the C-terminal. This opening causes the N-terminal end to be more closed, shifting the PMF barrier of entry to the Val27 gate. Additionally, we show how the altered water structures in D44N M2 facilitate PT, making the overall barrier for PT lower and thus increasing conductance, and enabling activation at low pH_{in}.

4.2 Methods

4.2.1 Classical MD Simulations

Because no experimental structures of D44N M2 are available, the initial M2 mutant structure was obtained by taking a previously equilibrated WT M2 structure from ref[3] and mutating the residue in Visual Molecular Dynamics (VMD).[136] Because NMR data has shown the WT and mutant structures to be very similar, this substitution is not expected to significantly alter the results. The WT starting structure was resolved at room temperature and high pH (PDB: 4QKL[59]), embedded in a 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) bilayer and solvated with water. This structure contains the transmembrane portion of M2, which is the minimum construct necessary to retain proton conduction similar to full-length M2.[60] The only titratable residues in the channel besides His37 are located at the ends and solvated in bulk water and were set in their neutral state. The mutant structure was initialized in the +0 state and equilibrated for at least 800 ns. The +1 and +2 states were generated by subsequently protonating one His37 residue and equilibrating for 800 ns twice.

4.2.2 MS-RMD Simulations

MS-RMD was used to model the water and excess proton in simulations used in the protein and water structure analyses. The MS-RMD method allows hydrogen-oxygen bonds to break and form by taking a linear combination of possible bonding topology states at each timestep; we refer the reader to previous work for a full description of the method.[18-21] The MS-EVB 3.2 parameters were used to describe the hydrated excess proton.[101] The excess proton center of excess charge (CEC) is defined as[102]

$$\vec{r}_{CEC} = \sum_{i}^{N} c_i^2 \, \vec{r}_{COC}^{\,i} \ , \tag{1}$$

where $\vec{r}_{COC}^{\,i}$ is the coordinate of the center of excess charge of the ith diabatic state and c_i^2 is amplitude of that state. The CEC allows one to track during the simulation the position of the most hydronium-like structure during the Grotthuss proton shuttling process through water molecules. The CHARMM36 force field was used to model the remaining interactions in the system. Simulations were run at 308 K in the NVT ensemble using LAMMPS[103] with the MS-RMD package.

The collective variable (CV) used to define the excess proton's progression through the channel was defined as the z-coordinate of the vector between the CEC and the center of mass of the four Gly34 α-carbons as in our previous work.[3, 23, 30, 135] Umbrella sampling (US) was performed using the replica exchange umbrella sampling (REUS)[137] technique to facilitate convergence. Simulations were run with the proton restrained every 0.5 Å along the CV coordinate throughout the channel, minus the His37 region, with a 10.0 kcal·mol^{-1}·Å^{-2} US force constant. Additionally, a cylindrical restraint was applied to ensure the proton remained in the channel at 8.0 Å and 12.0 Å for the top and bottom of the channel, respectively, with a force constant of 10.0 kcal·mol^{-1}·Å^{-2}using the open-source, community developed PLUMED library.[105, 138] Windows were equilibrated for 250 ps, followed by 1.5-5 ns production simulations.

4.2.3 QM/MM Simulations

While MS-RMD has been successfully used to model amino acid residue protonation and deprotonation in certain previous work, there is not currently a reliable method for fitting models of coupled residues in very complex environments, such as the His37 tetrad. Thus, to calculate the PMF for PT through the M2 channel in the +2 state, QM/MM simulations were run for PT through the His37-Trp41 tetrad region following the same protocol as in our previous work.[3, 23] The CEC coordinate used in the QM/MM simulations is fully defined in the Supporting Information. This formulation captures the delocalized nature of the excess proton; it has been shown to adequately describe the excess proton position in QM/MM simulations of biological PT channels[139-141] and has been used in our previous work with M2.[3, 23]

Simulation parameters are the same as in our previous work,[3] but briefly summarized here. The QM atoms included the His37 sidechains, the excess proton, and up to three solvation shells of water above and below His37 to ensure the proton was sufficiently surrounded by QM waters. The QM region was treated by Becke-Lee-Yang-Parr level density functional theory[142, 143] with empirical dispersion corrections,[144] under the Gaussian plane wave scheme.[145] Goedecker-Teter-Hutter pseudopotentials[146] were used and the Kohn-Sham orbitals were expanded in the Gaussian TZV2P basis set. The integration timestep was 0.5 fs, and the temperature was set at 308 K and controlled by a Nose-Hoover thermostat with a 0.1 ps relaxation time.

The QM/MM simulations were set up as follows. Three initial windows were set up from equilibrated MS-RMD simulations: two with the excess proton on either end of the region, and one with the excess proton bound to one of the His37 residues. Umbrella sampling windows were spaced every ~0.25 Å for a total of 43 windows, and the QM CEC was restrained with a 40 $\text{kcal}\cdot\text{mol}^{-1}\cdot\text{Å}^{-2}$ force constant. Simulation windows were pulled from adjacent equilibrated windows. Each window had an equilibration time of ~2ps and ~12-35ps of production sampling. The CEC position was collected every step. Because the +2 channel is more open, MM water would occasionally enter the QM region. To address this, waters were checked every ~3ps of simulation and appropriately reassigned to the QM or MM region if needed. Systems were equilibrated for 100fs following such an exchange before continuing the production run. All QM/MM simulations were performed using the CP2K package.[147]

4.2.4 PMF Calculations

These simulation results were combined with MS-RMD simulation results to calculate the hybrid PMF using the discrete transition-based reweighting analysis method (dTRAM)[148] as implemented in the PyEMMA python package.[149] Error bars were estimated using a block-averaging analysis.

4.2.5 Analysis

All analyses in this paper were performed using MS-RMD US trajectories to delineate changes as the excess proton moves through the channel. Analysis was done with respect to the principal axis of the channel, Z', to account for the channel's tilt.[30] The CEC position along this axis is labeled $\mathrm{CEC}_{Z'}$.

The direction of hydrogen bonds in the water analysis was calculated as the cosine of the angle between the donor-acceptor vector and the Z' axis. See our previous work for more detail.[30, 135]

4.3 Results and Discussion

4.3.1 D44N Mutation Lowers His37 Deprotonation Barrier and Shifts Proton Entry Barrier

Hybrid MS-RMD and QM/MM simulations were used to calculate the +2 mutant PMF shown in **Figure 4.2**, as described in the Methods. This PMF is the free energy associated with an additional excess proton moving through the channel in a +2 charge state. In the MS-RMD simulations, where the excess proton is out of the His37 tetrad region, the +2 charge is fixed and two histidine residues on opposing helices are protonated. In the QM/MM region, all four histidine residues and surrounding water are treated in the QM regime, allowing the protons to delocalize and the excess proton to form and break bonds with His37 residues. We note that due to the limited sampling possible in the QM/MM regime, the error calculated by block averaging likely underrepresents the true error; however, it is sufficiently well-defined to distinguish differences in the shape and position of the free energy minima and maxima compared with the WT PMF.

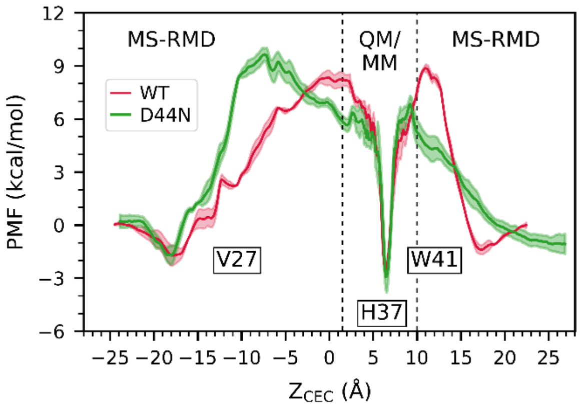

Figure 4.2. PMF for PT through the +2 WT and D44N mutant state, calculated from hybrid MS-RMD and QM/MM simulations. The regions calculated from each simulation method are indicated. Error bars are shown. The WT PMF is reproduced with permission from reference 3.

	Out -> His37	His37 -> In	In -> His37
WT	10.1	11.4	10.3
D44N	11.3	9.7	7.9

Table 4.1. Barriers for different steps in the forward and reverse PT process from calculated +2 PMFs. All numbers are in kcal/mol. The three columns from left to right are for 1) proton entry from the viral exterior to His37 protonation, 2) His37 deprotonation towards the viral interior, and 3) proton entry from the viral interior to His37 protonation.

There are two main features of the PMF that are distinct from the WT +2 PMF. First, the barrier for proton entry from the exterior is shifted from proton CEC position at Z = 0 Å in WT to Z = -10 Å in the mutant, at the location of Val27. This indicates that PT is more difficult through the Val27 tetrad, the primary gate at the top of the channel, relative to WT. Below, we show and discuss how the structural changes associated with the D44N mutation may make the top half more closed and rigid. Because we did not directly sample the water solvation through the Val27 gate or the gate structure, this rigidity may be artificially increasing the PMF shown here. Further studies are needed to verify this.

The second distinction from WT is a ~2.4 kcal/mol decrease in the barrier for proton entry from the interior and ~1.7 kcal/mol decrease for His37 deprotonation towards the viral interior, around Z = 9 Å, shown in **Table 4.1**. In WT M2, this deprotonation is the rate-limiting step in the PT process.[45, 150] The lower barrier here indicates that protons can pass through the His37 tetrad more readily, resulting in overall increased conduction, in agreement with experiments. This decrease also changes which PT step has the largest free energy barrier—now, proton entry from the viral exterior becomes the largest barrier, indicating that in D44N M2 the rate-limiting step may be proton entry through the Val27 gate rather than His37 deprotonation. In this case, the concentration of protons to the viral exterior will determine the rate of proton transport to His37—while this barrier is similar to the WT deprotonation barrier and the rate *constants* may be similar, the effective rates are not necessarily the same. The D44N rate is expected to be increased due to the higher concentration of protons to the exterior moving towards the His37 tetrad, compared with the low concentration of protons in the His37 tetrad region in WT M2 going over the barrier to the interior.

4.3.2 Structural Changes Due to D44N Mutation

The D44N mutation alters the protein's structure and its equilibrium dynamics, causing the channel to be wider at the bottom compared to WT. To first examine the difference in structure, we compared the pore radii profiles between the WT and D44N mutant structures for the +0-+2 His37 charge states with no excess proton in the channel (**Figure 4.3**, top). In WT M2, with no excess proton in the channel, there is a narrow constriction at the Trp41 gate in the +0 and +1 structures. The +2 structure is more open, with a pore radius ~ 3Å, which allows water to freely reach the His37 tetrad and thus PT is facilitated. In the D44N mutant +0 and +1 states, the Trp41 gate is open with a pore radius of ~3Å, similar to the WT +2 state. The +2 mutant state is further expanded: below the Trp41 gate, the +2 state is much more open, with a 3Å increase in radius near the D44N mutation. Additionally, the +2 state is noticeably wider in the region between Gly34 and His37. Between Val27 and just above Gly34, Z'=-10.0 to Z'=-2.0 Å, the pore is remarkably similar to WT. We previously showed the stability in this region in WT M2 is a critical determinant for successful amantadine inhibition.[135] Although the D44N mutation significantly alters the conformation of the channel, it surprisingly does not change this region. This finding is consistent with observation that the D44N M2 channel is sensitive to amantadine.[127] In this case, drug binding induces the closed conformation with packed Trp41 indoles, although there are greater structural fluctuation in D44N versus WT near the His37/Trp41 gate. Above Val27, the +1 and +2 mutant structures are more closed relative to their WT counterparts.

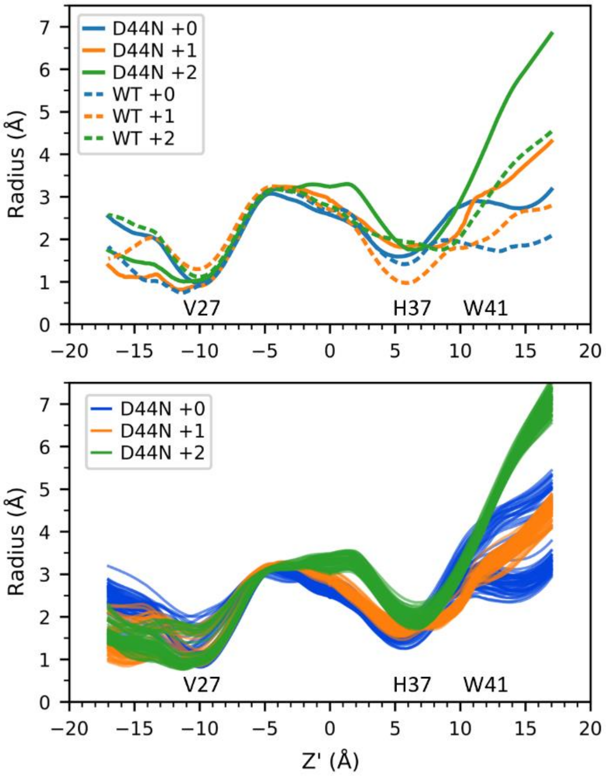

Figure 4.3. Top: radii profiles for WT and D44N M2 in the +0, +1, and +2 charge states. The x-axis is the Z' channel axis. WT data is taken from ref. 43. These profiles represent the average structures when the excess proton is out of the channel ($CEC_{Z'}$ = -24.0Å). Bottom: The average radii profiles calculated for excess proton positions every 0.25 Å in the mutant channel. Trajectory frames were binned by excess proton $CEC_{Z'}$ with bin widths of 0.25Å. Each line is the average radii profile for a given bin.

We also looked at how the pore radii change as the proton moves through the channel by calculating the average channel radii profile for different excess proton positions throughout the channel (**Figure 4.3,** bottom). Each thin line in this figure shows the calculated radii profile when the excess proton $CEC_{Z'}$ is fixed at a specific position. These profiles were calculated for $CEC_{Z'}$ values every 0.25 Å between -22 Å and 22 Å (minus the QM/MM region), for a total of 144 lines for each charge state. Here, we see changes in the pore radii at the top of the channel in all three states, dependent on proton position, as well as at the bottom of the channel in the +0 state. In our previous work, we showed increased variation at His37 in the +0 and +1 WT structures that is not observed in the +2 state. This variation seen most prominently in +0 WT M2 occurs as the proton moves through the Val27 gate — for a proton to pass through the Val27 gate, +0 WT M2 transiently adopts a conformation that widens the top portion of the channel while constricting the bottom portion. We attributed this correlated movement between the top and bottom halves of the channel to an increased rigidity at high pH. In contrast, all three charge states of the mutant show minimal structural variation due to proton position near His37, suggesting that the C-terminal end of D44N M2 is more flexible than WT at high pH, similar to the low pH WT structures. There is, however, noticeable variation below W41.

The mutant's shift towards a structure that is more open towards the viral interior and closed towards the viral exterior affects the pore hydration and free energy associated with PT — the free energy barrier for proton entry into the channel from the exterior increases, while the barrier for proton entry from the interior decreases, as reflected in the PMFs shown and discussed above.

4.3.3 D44N Mutation Increases Hydration and Alters the Water Network

To understand how the structural differences in the D44N mutant affect the water structures and PT throughout the channel, we analyzed the number of waters within different regions of the channel and the water hydrogen bond orientations. In our previous work, we showed how the increased water and optimal orientation of the water near His37 in the WT +2 state facilitated PT and "primed" the +2 state for PT. Conversely, fewer waters and less optimal hydrogen-bonding orientation in the region below His37 in the +0 and +1 states with no excess proton present indicate that the water structure must change significantly for PT to occur through this region, as seen when the proton nears the His37 tetrad. The greater rearrangement necessary for PT from the viral interior to His37 in high pH structures also helps explain the rectification behavior of WT M2 — this rearrangement requirement increases the energy barrier for a proton to enter from the viral interior, and the water hydrogen bond network orientated away from the His37 tetrad suggests it would be easier for a proton to move off the tetrad than an additional proton to move in.

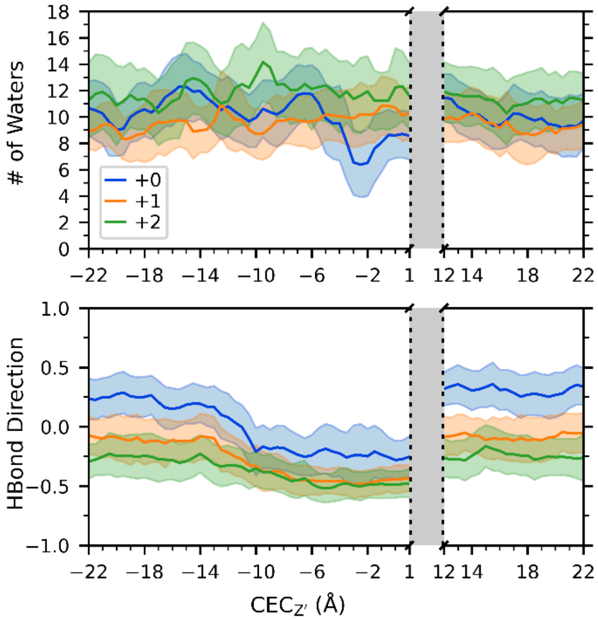

Figure 4.4. Water structure analysis for D44N in the region below His37 and above Trp41 from MS-RMD simulations. Top: The average and standard deviation of the number of waters in this region, as a function of the excess proton CEC position in the channel. Bottom: The average and standard deviation of the water-water hydrogen bonds direction. The grayed-out region corresponds to the QM/MM region.

One critical component of this priming in the WT +2 state is the water structure between His37 and Trp41, which facilitates His37 deprotonation towards the viral interior. Throughout the PT process the +2 state consistently has 9-11 water molecules in this region, whereas the +0 and +1 states have fewer than 6, except when the proton is in this region and it draws an additional 2-3 waters in.[30] In the mutant, all three charge states consistently have 8 or more waters in this region, with lower variability between states (**Figure 4.4**). The number of waters only dips below 8 in the +0 state when CEC$_{Z'}$ = -2.0 Å, at which point the His37-Trp41 region is less relevant. This difference from WT suggests that the likelihood of proton release from His37 to the virial interior at high pH is increased in D44N M2, as the high pH structures represented by +0 and +1 have similar water structures to the activated, low pH +2 state. This increased likelihood suggests more leaky conductance at high pH.

The hydrogen bond network below the His37 tetrad in D44N is similar to that in WT, with the +0 state taking on overall positive average values with no proton present, and the +1 and +2 states have negative values. The +0 state value in D44N, however, is lower than in WT by ~0.25. This lowering suggests that the increase in number of waters limits the His37's influence on water direction, which correlates with previous reports of a "parallel circuit" of hydrogen bonds in D44N M2 from classical MD simulations.[59] This effect is most clearly seen in the region above Gly34 and below Ser31, as shown in **Figure 4.5**. In this region in WT, each state has a distinct line with values ~0.12, 0.23, 0.43 when an excess proton is not in the channel. In the mutant, all three charge states have similar values throughout the PT process, indicating that the His37 tetrad charge has minimal effect on the water structure in this region. The greater similarity in water networks across charge states in D44N M2 indicates that there may be fewer differences in PT between states in this mutant.

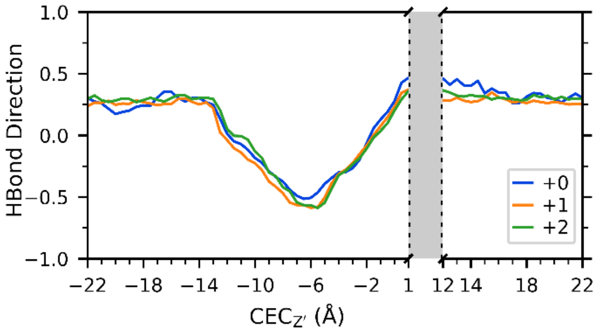

Figure 4.5. Average water-water hydrogen bond direction for D44N in the region above Gly34 and below Ser31 as a function of the excess proton CEC position in the channel. The grayed-out region corresponds to the QM/MM region.

The different water structure in D44N M2 also helps explain the loss of rectification behavior in this mutant, that is, that D44N M2 can become activated by protons entering from the viral interior at low pH_{in} and high pH_{out}. In WT M2, the +0 and +1 states require additional water molecules below His37 for an excess proton to move through that region, whereas there is no significant change in water number in the +0 and +1 mutant states. Because the region is sufficiently hydrated, protons can more easily enter from the viral interior and the rate of protonation (k_{on}) can compete with the rate of deprotonation towards the exterior (k_{off}). If k_{on} outcompetes k_{off}, it becomes significantly probable to reach an activated +2 state. This is also reflected in the PMFs, as discussed above, which show that the barrier for protonation from the interior is ~2 kcal/mol less in the mutant M2 than in WT M2.

4.4 Conclusions

In this work, we have used extensive MS-RMD and QM/MM simulations with an explicit excess proton to investigate how PT in the D44N M2 mutant differs from WT on the molecular level. First, we calculated the PMF for PT through the +2 state and discussed critical differences from WT—the barrier for His37 protonation from the viral interior is reduced, and the barrier for entry from the viral exterior is shifted to the Val27 gate, increasing conduction and potentially changing the rate-limiting step. We furthered our understanding of how this mutation alters the structure of M2 by examining channel radii profiles, showing that the high pH structures are similar to the low pH WT structure, and that opening the bottom of the channel results in narrowing the pore at the top near Val27. Finally, by analyzing the water hydrogen bond network, we showed that increased hydration facilitates PT and His37 binding in the low charge states and decreases the effect of the His37 charge on the water orientation. Together, these results also help explain the loss of rectification in the D44N mutant.

These results also have implications for other mutations in the lower half of the channel that affect the Trp41 gate and lose rectification behavior, such as W41F. Any mutation that disrupts the gate (such as Asp44 mutations) or makes the gate smaller (such as Trp41 mutations) will increase water access to the His37 residues and thereby increase the probability of activation from the viral interior. We expect the height of the protonation/deprotonation barrier from His37 towards the viral interior for a given charge state to be proportional to the hydration level of His37.

Studying the prevalent mutants of M2 is critical for continuing anti-flu drug-design efforts and is necessary to fully understand PT in M2 and the function of specific residues. Here, we show the significance of Asp44 and its naturally occurring variants in determining fundamental characteristics of PT, including the level of conduction, rectification, and pH dependence, and how this mutation alters these characteristics. As a prototype for viroporins more generally, these results indicate the importance of investigating mutations to deduce their effect on transport mechanisms and their physiological relevance.

Supporting Information

4.4.1 QM/MM CEC Definition

The CEC coordinate in the QM/MM simulations captures the delocalized nature of the excess proton, and is defined as[139]

$$\vec{\xi} = \sum_{i=1}^{N_H} \vec{r}^{H_i} - \sum_{j=1}^{N_X} w^{X_j} \vec{r}^{X_j} - \sum_{i=1}^{N_H}\sum_{j=1}^{N_X} f_{SW}\left(d_{X_jH_i}\right)(\vec{r}^{H_i} - \vec{r}^{X_j}) + \vec{\xi}_{correct}$$

where X_j's are the histidine nitrogen atoms and water oxygen atoms in the QM region, and H_i's are the hydrogen atoms bound to those heavy atoms in the QM region. The w^{X_j}'s are the hydrogen coordination numbers of the heavy atoms in its molecule's least protonated state during the PT process—for water oxygen atoms this value is 2 (two hydrogens per oxygen), and for His37 nitrogen atoms it is 0.50 in the +0 state (four hydrogens shared by eight His37 nitrogen atoms), 0.625 in the +1 state, and 0.75 in the +2 state. The $d_{X_jH_i}$ variable is the distance between atoms X_j and H_i. The function $f_{SW}\left(d_{X_jH_i}\right)$ measures the current coordination number of H_i to X_j: $f_{SW}\left(d_{X_jH_i}\right) = 1/1 + \exp\left[(d_{X_jH_i} - r_{SW})/d_{SW}\right]$. The parameters are set to $d_{SW} = 0.04$ Å and $r_{SW} = 1.25$ Å.[140]

The correction term $\vec{\xi}_{correct}$ was previously introduced to correct for the contribution due to the presence of multiple protons around multiple protonatable sites in the His37 tetrad[23]

$$\vec{\xi}_{correct} = \frac{1}{8}\sum_{i=1}^{8}\sum_{j=1}^{8} m_i(\vec{r}^{X_j} - \vec{r}^{X_i})$$

where X denotes one of the eight nitrogen atoms of the His37 tetrad, and m_i switches from 1 to 0 as nitrogen atom i is deprotonated:

$$m_i = \sum_{H_j \in \{H\}} f_{SW}\left(d_{X_i,H_j}\right)^{16} \Big/ \sum_{H_j \in \{H\}} f_{SW}\left(d_{X_i,H_j}\right)^{15} .$$

This formulation has been shown to accurately describe the excess proton position in QM/MM simulations of biological PT channels[139-141] and has been used successfully in our previous work with M2.[3, 23]

Chapter 5

Proton Transport in *De Novo* Designed Proteins

5.1 Introduction

The controlled transport of protons across cellular membranes through transmembrane (TM) proteins is a critical process in biological systems for maintaining electrochemical gradients and harnessing its energy. Proton transport occurs as a hydrated excess proton shuttles through water molecules, through the Grotthuss mechanism.[15-17] One important component of this process is a protein's ion selectivity. Proteins that are selective for protons typically have narrow pores, incapable of fitting both a larger ion and the water molecules necessary to solvate it. Additionally, these proteins may gate the water needed for proton transport by interspersing polar or acidic residues between hydrophobic stretches. The influenza A M2 proton channel, for example, uses a group of four histidine residues as a selectivity filter, preventing the transport of other ions.[38, 53]

Recently, Mravic, et al, in an effort to understand the role of apolar side-chain packing in membrane protein folding and assembly, designed and synthesized a stable pentameric membrane protein.[151] They first started by examining the protein phospholamban (PLN) and noting that the rigid apolar domain is critical for protein stability, in contrast to the polar domain which was more dynamic and became splayed throughout their simulations. Using this rigid apolar domain as guidance, they were then able to design and synthesize a PLN-like peptide (PL5) using an idealized model parameterized from PLN's rigid domain. They showed that the fully apolar PL5 formed stable pentamers, and from this, derived a steric code for stable interhelical packing, specific to five-helix bundles—based on the labeling of canonical alpha-helix seven residue repeats (abcdefg), where each letter indicates one residue, they show that the repeat LxxIXexxL forms stable pentamers, where Xe is either Ile, Val, Thr, or Cys.

This work was a monumental step forward in protein design, showing for the first time that apolar side-chain packing is sufficient to stabilize membrane protein assembly, without hydrogen-bonding or any other polar interactions. Furthermore, their design of a fully hydrophobic protein provides an excellent scaffold on which to experiment adding in and controlling specific functionality by mutating pore-lining residues.

Coworkers of Mravic in the DeGrado group have continued this study by attempting to design proteins capable of facilitating proton transport (PT), starting with this hydrophobic scaffold. Kratochvil, et al, targeted four pore lining residues (Leu3, Leu7, Leu10, and Leu14) for mutations, and found that several combinations of Gln residues mutated in these positions resulted in proton flux. As simple as one mutation in one of these four positions is sufficient to introduce proton conduction to the protein.

Here, we use computer simulations to study PT in one of these designed proteins, LQLL, with experimentally measured proton conductance, as well as in the non-conducting hydrophobic pentamer LLLL, to understand how the mutation of one residue can switch this PT functionality on and off. These structures are shown in **Figure 5.1** Extensive MS-RMD simulations with umbrella sampling is used to calculate two-dimensional (2D) potentials of mean force (PMFs) of PT for each system, revealing that the added Q residue acts as an "anchor" within the channel to stabilize water wires from Gln to either end of the channel, but not both simultaneously. This greatly reduces the free energy cost of hydrating the channel as the excess proton moves through the otherwise hydrophobic pore, as the water spans a smaller distance, and thereby facilitates PT.

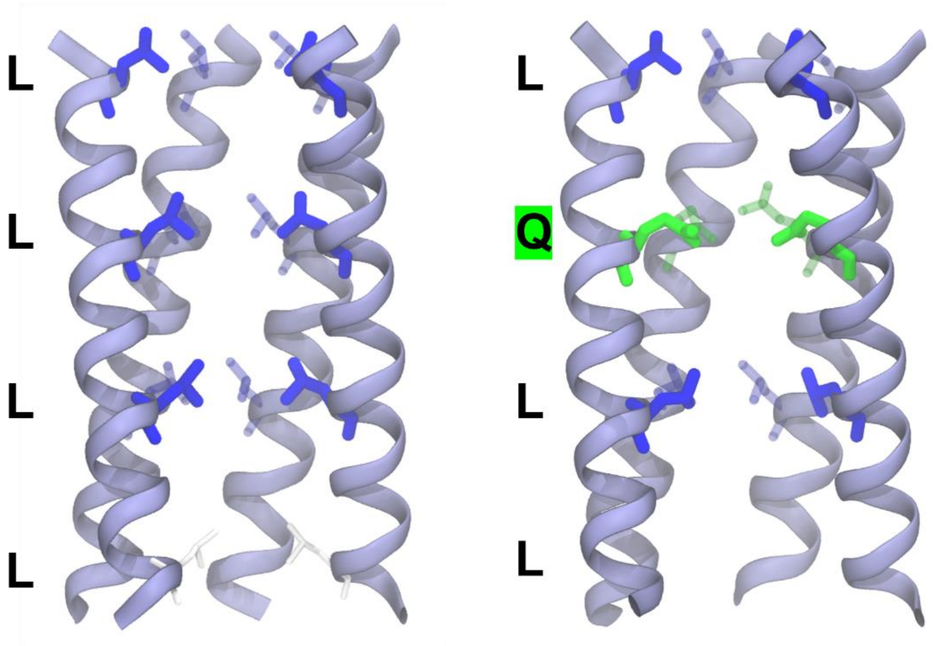

Figure 5.1. The two systems studied in this work: left, LLLL and right, LQLL. The three leucine residues mutated to glutamine in various combinations are shown in blue, glutamine is in green. The front helix is removed for clarity.

5.2 Methods

The x-ray crystal structures for LQLL and LLLL were used as the starting structures for simulation. Classical simulations were first performed to equilibrate the protein structures and the simulation systems. Each protein was embedded in a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) bilayer and solvated with water using the CHARMM GUI, and the membrane and water were equilibrated using a standard equilibration protocol. Classical equilibration with no restraints was performed for 500ns. The CHARMM36 forcefield was used to model all interactions, and simulations were run at 298K in the NPT ensemble using GROMACS.

MS-RMD[18-21] was subsequently used to model the water and excess proton in all simulations used in our analyses. The MS-RMD method captures proton delocalization in water by allowing hydrogen-oxygen bonds to break and form. This is done by taking a linear combination of possible bonding topology states at every timestep. See our previous work for a detailed description and theory. The MS-EVB 3.2[101] parameters were used to describe the hydrated excess proton. The excess proton center of excess charge (CEC) is defined as

$$\vec{r}_{CEC} = \sum_{i}^{N} c_i^2 \, \vec{r}_{COC}^{\,i} \ , \tag{1}$$

Where $\vec{r}_{COC}^{\,i}$ is the coordinate of the center of excess charge of the ith diabatic state and c_i^2 is the amplitude of that state. The CEC defines the position of the delocalized excess proton. The CHARMM36 forcefield was used to model the remaining interactions. Simulations were run at 298K in the NVT ensemble using LAMMPS[103] with the MS-RMD package.

Umbrella sampling simulations were performed in two dimensions to model the PT process. The first collective variable (CV) used is the CEC position along the channel axis, Z'_{CEC}. The channel axis for each system was defined as the average principal component of the protein from a 750ps MS-RMD equilibration simulation. The position along this axis is calculated in reference to the center of mass of the Ile13 alpha-carbons, such that $Z'_{cec} = 0$Å at that point. The second CV used, φ, is a recently developed CV that measures the water connectivity within a channel using graph theory.[152] This CV is a significant improvement over water density, which does not directly bias the formation of a continuous water wire and can result in unphysical water “clumps” as the bias increases. Instead, the new water connectivity CV measures the length of transient water wire formations on a scale from 0 to 1, where 0 is no water and 1 is water fully connected throughout the channel, agnostic to the number of water molecules. We refer the reader to reference 152 for further theoretical details.

Umbrella sampling windows were set up every 0.5Å in the range [-22, 22] Å for Z'_{CEC} and every 0.035 in the range [0.140,0.980] and [0.245,0.980] for φ for LLLL and LQLL, respectively, for a total of ~2000 windows for each system. Initial windows were generated by using steered MD to pull water into the channel, and the excess proton was placed at each point along the channel. Subsequent windows were pulled from nearby windows. Windows were equilibrated for 100ps, and then run for 0.5-3.5ns. A few additional windows were added to ensure sampling overlap.

The 2D PMFs were calculated using WHAM-2D.[153] Error bars were calculated using the block method with four blocks. The minimum free energy path (MFEP) was calculated using the string method.

5.3 Results and Discussion

The 2D PMFs for PT through LLLL and LQLL are shown in **Figure 5.2**. The MFEP for LQLL is indicated in white in **Figure 5.2B** and the free energy along the MFEP is shown in **Figure 5.3**. The LLLL PMF has a prohibitively high barrier of ~45 kcal/mol at Z'_{CEC} = 0Å, with nearly symmetrical free energies sloping down on either side. The addition of the glutamine residue has several noticeable effects on the LQLL PMF. First, it drastically reduces the overall free energy across the 2D PMF. With this, the LQLL PMF transition barrier is shifted down to Z'_{CEC} = ~-4Å and the overall barrier height is significantly decreased to ~20 kcal/mol. The Gln residue helps to flatten the PMF near Z'_{CEC} = ~5 Å, when the proton is directly near Gln, and it creates a nearly vertical contour at Z'_{CEC} = 2.5 Å—that is, when the excess proton CEC is at this position in the channel, the channel water can easily fluctuate between partially and fully connected with little free energy cost. This is seen in the 3D depiction of the PMF in **Figure 5.2D** (note that here we are looking *down* the φ CV). While the LLLL PMF is monotonically increasing on either side towards the barrier, the LQLL PMF has two distinct barriers. Gln introduces an additional free energy minimum, breaking the insurmountable LLLL free energy barrier down into two smaller free energy steps.

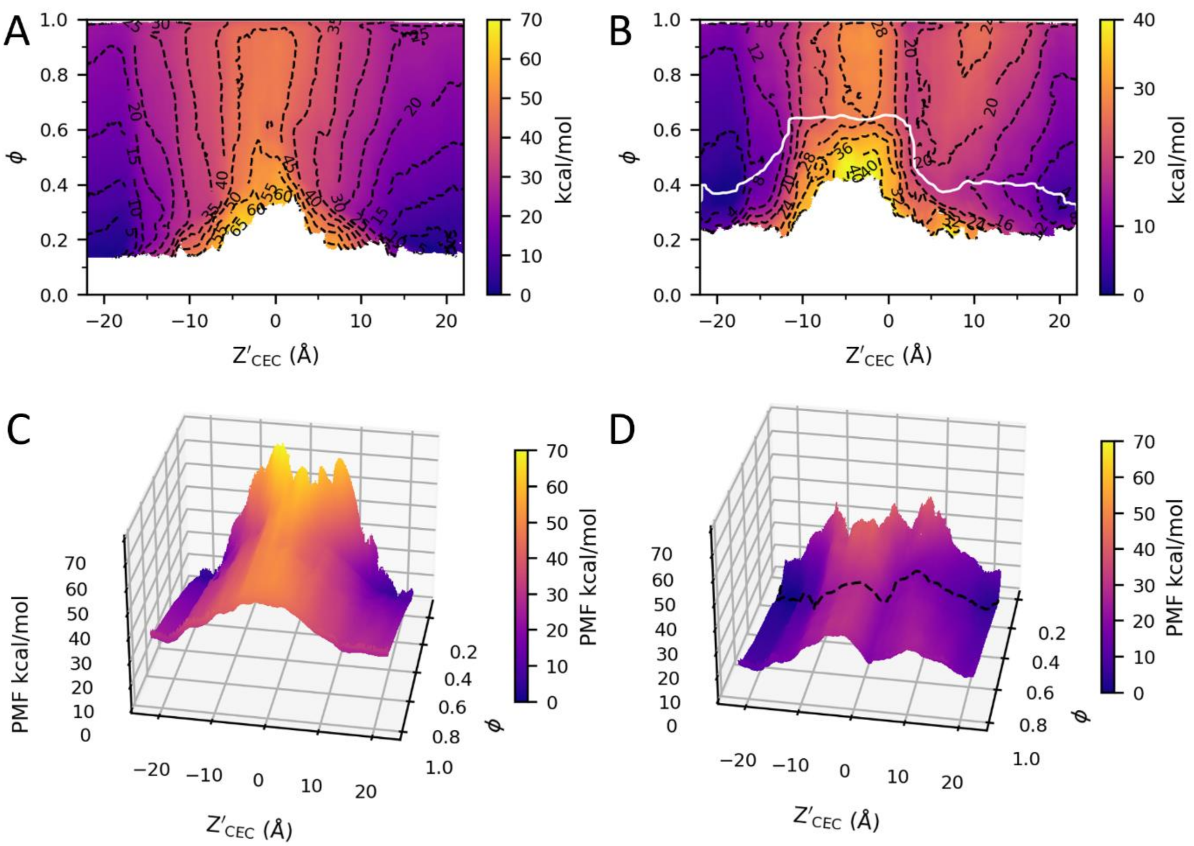

Figure 5.2. A and **B:** the 2D PMFs for LLLL and LQLL, respectively. Note that a different color scheme is used for each to better show the details of the free energy landscape. Contours are shown and labeled for each. In **B**, the white line is the MFEP. **C** and **D** are 3D depictions of the same PMFs for LLLL and LQLL, respectively, here using the same color scheme. In **D,** the dashed black line is the MFEP.

The MFEP in **Figure 5.3** further reveals how Gln influences the PT path. In the LQLL system, the proton moves through the top of the channel while maintaining a low water connectivity. Once the excess proton moves past the Gln residue, at Z'$_{CEC}$ = 2.5 Å, the water connectivity must increase before it can continue to move down. It does not, however, need to be fully connected through the channel.

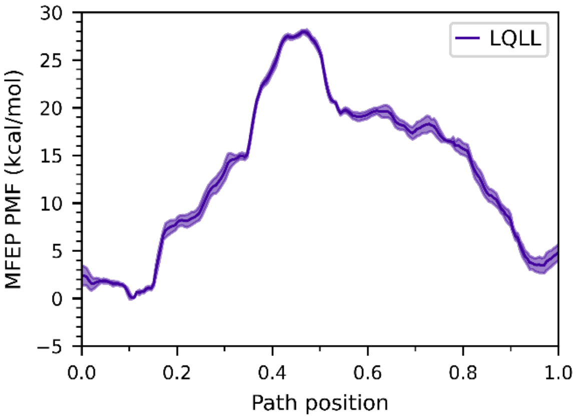

Figure 5.3. The free energy along the LQLL MFEP. Error bars shown calculated using block averaging analysis with four blocks.

The role of water in each PT path was further explored by looking at the average water volume throughout the channel along the MFEP, shown in **Figure 5.4**. As the proton moves through the top of the channel towards Gln, the channel water is connected from the outside (top) to Gln, and the bottom of the channel is dry. As the proton moves through the bottom two-thirds of the channel, however, the water shifts with it, connecting Gln with the bottom exterior bulk water and leaving the top portion is dry. The transition point for the water occurs at Z'$_{CEC'}$= 2.5 Å, when the excess proton is just below the Gln residue, corresponding to the vertical portion of the MFEP on the 2D PMF at that position—while the proton remains stabilized by the Gln residue, the waters transition

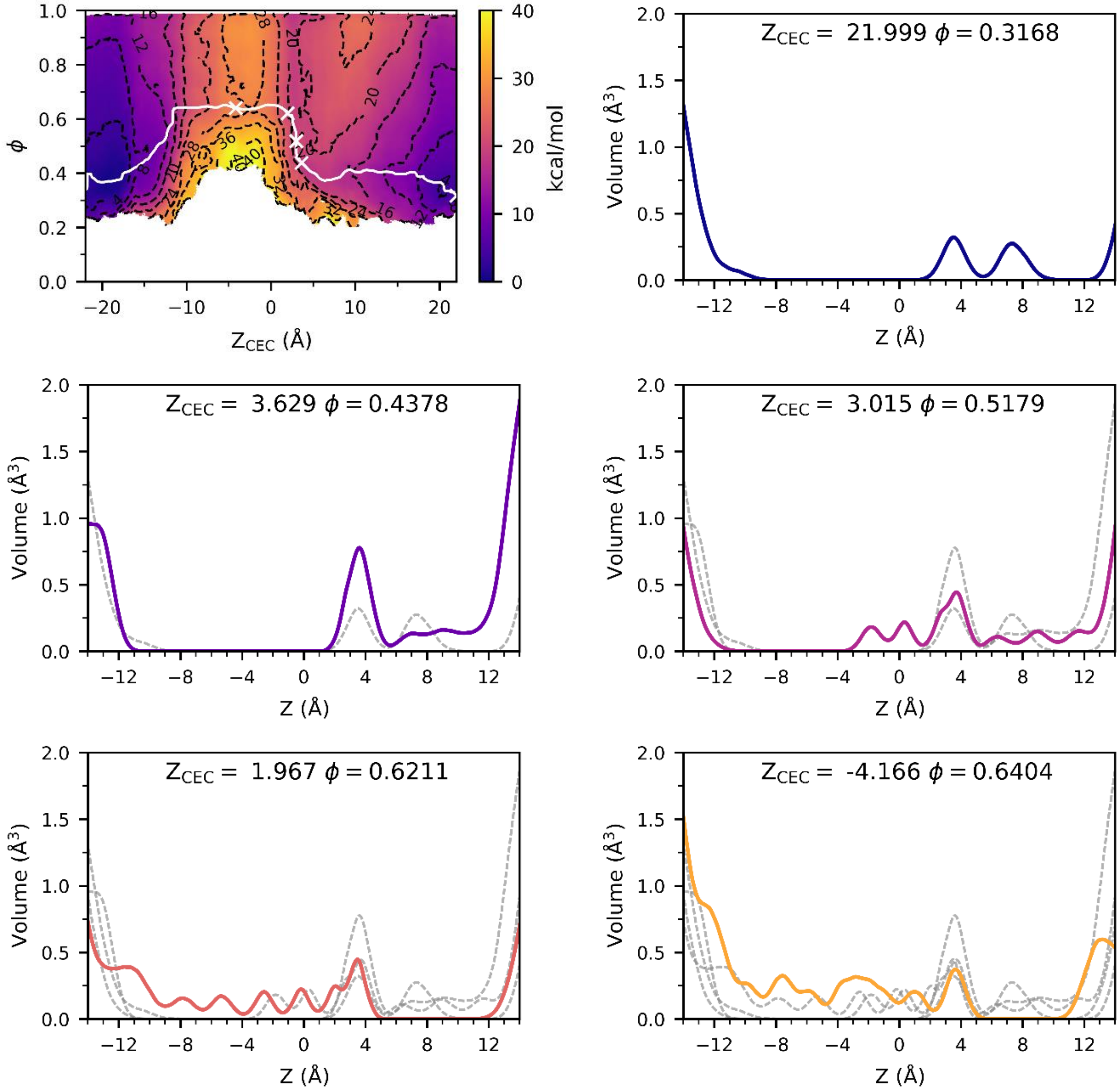

Figure 5.4. Average volume of water in the channel shown for 5 different points along the MFEP, as indicated by white X's in **A**. **B**, the excess proton is out of the channel and the few waters in the channel are not connected to bulk. **C**, the proton is just below Gln and water connects to the top. **D**, the water structure transitions from the top to the bottom while the proton remains below Gln. **E**, the water is full connected in the bottom. **F**, the proton moves through the bottom-connected water wire. The dashed lines in **C-F** are the curves shown in all previous panels, **B-E**.

from the top to the bottom of the channel and more water is drawn in to span the larger hydrophobic stretch. Thus, Gln acts as an “anchor” to stabilize water structures necessary for PT. By allowing and stabilizing the formation of water wires to one side only, the Gln residue greatly reduces the cost of forming connected water wires. This proposed mechanism is shown altogether in **Figure 5.5**.

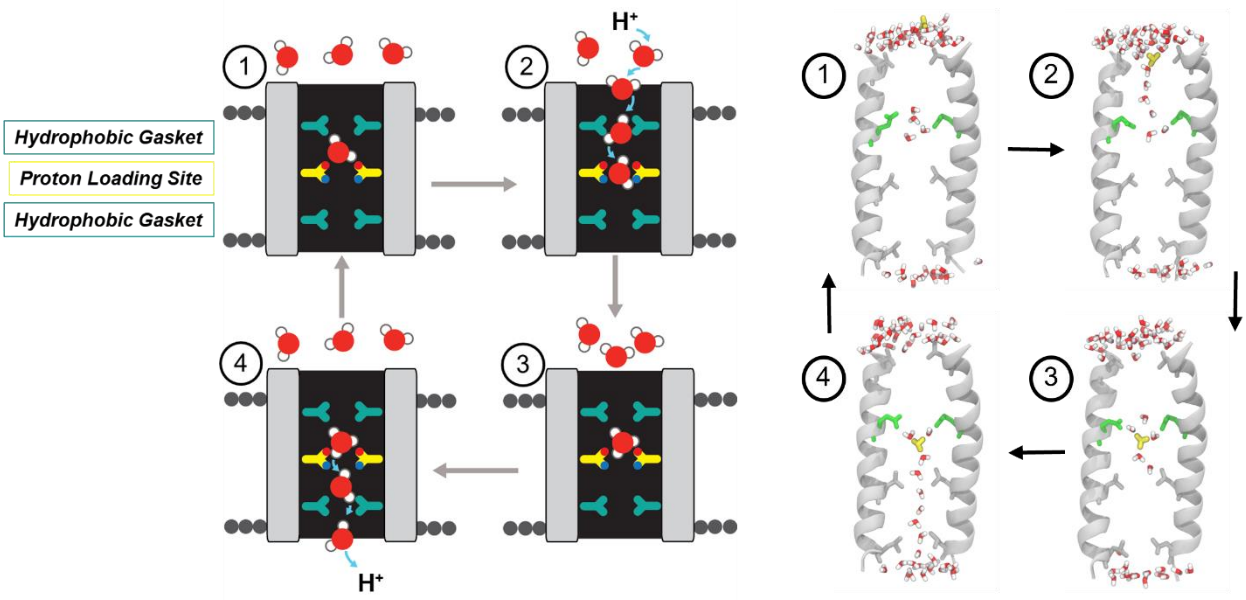

Figure 5.5 Formation of transient water wires through hydrophobic gaskets enables selective transport of protons. The presence of the polar Gln group within the hydrophobic lumen creates a PLS that enables the formation of transient water wires that transport the proton to interior of the channel lumen. 1) With no proton in the channel, there are a few waters near Gln. 2) The presence of the proton at the top of the channel induces the formation of a transient water wire that connects the Gln group with the bulk water to the top. 3) The water wire disappears when the hydrated excess proton is stabilized by the Gln group. 4) Before the excess proton moves through the bottom of the channel, a second network of transient water wires forms.

We also examined the conformation of the 5 Gln residues to determine how they are oriented in the channel and whether they exhibit symmetrical behavior. **Figure 5.6** shows the dihedral angles and distances to the channel center for all five Gln residues along the MFEP. We use the absolute value to ignore x/y tilts and focus on the degree of rotation relative to the z-axis, where 0° indicates the carbonyl oxygen is pointing downward toward the center of the channel, and 180° indicates it is pointing upward toward the top of the channel. This plot shows that the Gln residues have different conformations. When the excess proton is in the bottom of the channel, and there is no water in the top, there are consistently three residues primarily pointing down, one points up, and one is in between. When the proton is in the vicinity of the Gln residues, they shift to further stabilize the excess charge—in path positions near 0.4, when the proton is below the glutamines, they all shift to point downward. As the proton moves above the Gln residues, in path positions near 0.7, however, the residues do not all point directly up, but the average of the 5 in the right panel **Figure 5.6** indicates that there is an overall upward shift.

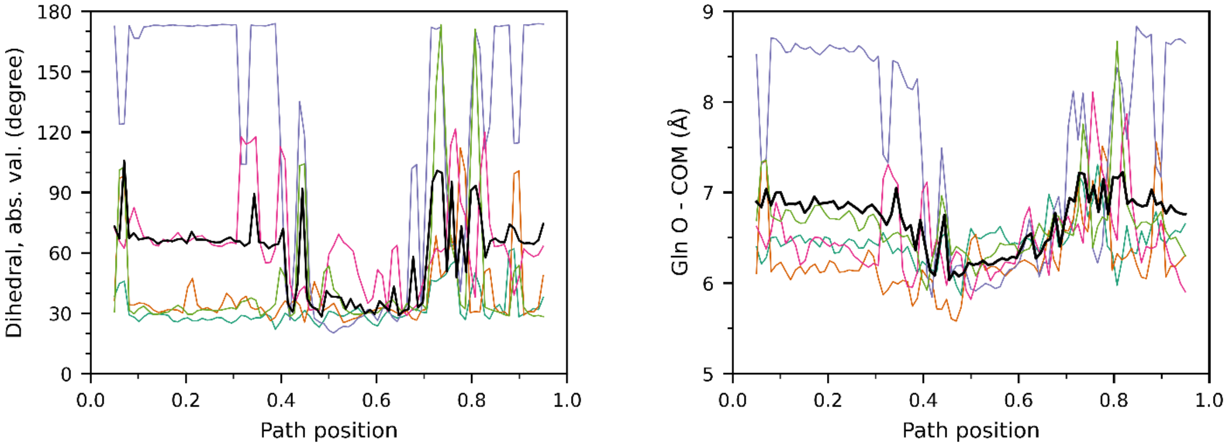

Figure 5.6. Left, the absolute value of the Gln sidechain dihedral for each Gln, as a function of the system position along the path. The average over all 5 residues is shown in black. Right, the distance between the Gln carbonyl oxygen and the Ile13 COM, shown for each Gln. The average over all 5 residues is shown in black.

5.4 Conclusion

Together, these results show how one polar residue in an otherwise hydrophobic TM protein is sufficient to create a proton channel with significant and selective proton conductance. Starting from the stable hydrophobic LLLL pentamer, experiments indicate that substituting in a glutamine residue for either of the interior leucine residues results in proton conduction. In this work, we focused on the LQLL construct to understand how the glutamine mutation enables PT. Using extensive MS-RMD simulations biasing both the CEC position and the water connectivity, we calculated 2D PMFs of PT through the LQLL and LLLL which revealed that Gln drastically reduces the free energy cost of hydration and PT through the channel.

Additional examination of the water structure during PT shows that the water dynamically hydrates the channel, with several stable states possible: with no proton present, Gln is hydrated but not connected with bulk water, when the proton is in the channel, water is connected from Gln to bulk on either the top or bottom side, dependent on where the proton is in the channel. Furthermore, the hydrophobic top or bottom part of the channel becomes fully hydrated and connected with bulk water before the proton passes through. Thus, the glutamine residue functions as an anchor for water in the channel, allowing water wires to form on either side connected with the bulk water with relatively low free energy cost when an excess proton is present.

These results are a critical step forward in *de novo* protein design with specific functionality. Particularly, these results indicate our ability to design proton-selective channels and to understand the mechanism of PT and role of side chain residues. This work could be extended to further tune the rate of proton conduction and to design channels capable of selective transport of other ions.

Chapter 6

Conclusions and Outlook

6.1 Proton Transport in Proteins is Dynamic

Altogether, the work in this thesis presents a dynamic picture of proton transport in proteins and reveals key characteristics that are critical for understanding PT in biological systems overall. Through studies on M2, I show PT is associated with dynamic water structures and hydrogen bond rearrangements that respond to an excess charge. These dynamic changes can be anchored and stabilized by protein residues, seen in M2 and in *de novo* designed channels. Additionally, the presence and movement of a proton can shift a protein away from its equilibrium structure—not only can it affect the residues near the proton, but also the backbone away from the proton, dependent on the rigidity of the structure.

The methods used here to study the dynamic processes associated with PT in M2 could be easily applied in other systems. While M2 has been studied extensively and much is known about the PT mechanism, there are many other proton channels and transporters where these details are unknown—studying them from this perspective may identify key features of its PT mechanism, while also help explaining other observables such pH dependence or activation. This type of approach could also be further extended to study other interactions in M2 and more complicated systems, such as in systems that co-transport other ions or molecules.

6.2 Studying Proton Transport Can Aid Drug Design

My detailed study of an excess proton in M2 as it relates to the efficacy of M2 inhibitors reveals underlying features of PT that adamantyl amine inhibitors are able to “take advantage of” to bind and block the M2 channel. The flexibility of protein-water hydrogen bonds, CEC “hotspots”, and channel shape and dynamics all contribute to the drugs’ stability in binding to the channel.

The results of this work could be further tested by running simulations of the inhibitors in the channel and altering various features to determine their importance for stable binding. For example, removing the charge of the amine group would reveal how necessary the positive charge is for binding. Furthermore, slight modifications to the inhibitors that place the charged group in a part of the channel that is not favorable for the CEC would also help indicate its importance.

Most importantly, this work could be extended to aid in drug design for M2 and other critical proton channels, such as ORF3a in SARS-CoV-2. Rather than aiming to explain the effectiveness of an inhibitor, this approach could be used in a forward manner to identify underlying features that designed drugs could match. This could be particularly helpful in guiding where to place charged groups on an inhibitor to most effectively use the channel’s natural ability to stabilize excess charges.

6.3 Proton Transport in Designed Proteins

My work on using computer simulation to model PT in *de novo* designed proteins revealed how a single polar residue in an otherwise hydrophobic channel can induce proton conduction. One key feature of PT clearly shown in this system is the ability of a proton to drive hydration in a channel ahead of it. With no proton present, it is unlikely for water to enter the channel and fully connect to bulk water on both sides; however, a proton induces the formation of water wires between bulk water and the key glutamine residue before moving through the hydrophobic regions. This highlights how critical it is to study PT with a reactive excess proton—static pictures from crystal structures and classical MD cannot depict the whole story of PT, nor can the mechanism of PT be assumed based on the presence or absence of water wires in these methods.

In this work, I studied just one designed protein with proton conductance. Future work could extend this to study systems with two glutamine residue substitutions, or other mutations with the aim to control conduction rates and selectivity more precisely. By taking a more bioinformatics approach and considering the wealth of available natural proteins and their functions, perhaps more complex features such as unidirectional conduction or even co-transport of other ions could also be designed and built into functional proteins.

Bibliography

1. Liang, R.; Swanson, J. M.; Madsen, J. J.; Hong, M.; DeGrado, W. F.; Voth, G. A., Acid activation mechanism of the influenza A M2 proton channel. *Proc Natl Acad Sci U S A* **2016**, 201615471.

2. Thomaston, J. L., DeGrado, W.F., Influenza A M2 proton channel wild type TM domain bound to S-rimantadine. *To be published* **2020**.

3. Liang, R.; Swanson, J. M. J.; Madsen, J. J.; Hong, M.; DeGrado, W. F.; Voth, G. A., Acid activation mechanism of the influenza A M2 proton channel. *Proc Natl Acad Sci U S A* **2016,** *113* (45), E6955-E6964.

4. Decoursey, T. E., Voltage-gated proton channels and other proton transfer pathways. *Physiol. Rev.* **2003,** *83* (2), 475-579.

5. Mitchell, P., Coupling of Phosphorylation to Electron and Hydrogen Transfer by a Chemi-Osmotic Type of Mechanism. *Nature* **1961,** *191*, 144-8.

6. Stryer, L., Chapters 17 and 22. In *Biochemistry*, W. H. Freeman & Co.: New York, 1988.

7. Wraight, C. A., Chance and design-Proton transfer in water, channels and bioenergetic proteins. *Bba-Bioenergetics* **2006,** *1757* (8), 886-912.

8. Decoursey, T. E., Voltage-gated proton channels and other proton transfer pathways. *Physiol Rev* **2003,** *83* (2), 475-579.

9. Deamer, D. W., Proton permeation of lipid bilayers. *J Bioenerg Biomembr* **1987,** *19* (5), 457-79.

10. Maloney, P. C.; Kashket, E. R.; Wilson, T. H., A protonmotive force drives ATP synthesis in bacteria. *Proc Natl Acad Sci U S A* **1974,** *71* (10), 3896-900.

11. Wraight, C. A., Chance and design--proton transfer in water, channels and bioenergetic proteins. *Biochim Biophys Acta* **2006,** *1757* (8), 886-912.

12. Aoi, W.; Marunaka, Y., Importance of pH homeostasis in metabolic health and diseases: crucial role of membrane proton transport. *Biomed Res Int* **2014,** *2014* (Figure 1), 598986.

13. Sachs, G., Proton pump inhibitors and acid-related diseases. *Pharmacotherapy* **1997,** *17* (1), 22-37.

14. Poet, M.; Kornak, U.; Schweizer, M.; Zdebik, A. A.; Scheel, O.; Hoelter, S.; Wurst, W.; Schmitt, A.; Fuhrmann, J. C.; Planells-Cases, R.; Mole, S. E.; Hubner, C. A.; Jentsch, T. J., Lysosomal storage disease upon disruption of the neuronal chloride transport protein ClC-6. *Proc Natl Acad Sci U S A* **2006,** *103* (37), 13854-9.

15. Knight, C.; Voth, G. A., The curious case of the hydrated proton. *Acc Chem Res* **2012,** *45* (1), 101-9.

16. de Grotthuss, C. J. T., Sur la décomposition de l'eau et des corps qu'elle tient en dissolution á l'aide de l'électricité galvanique. *Annales de Chimie* **1806,** *LVIII*, 54-74.

17. Agmon, N., The Grotthuss Mechanism. *Chem. Phys. Lett.* **1995,** *244*, 456-462.

18. Knight, C.; Lindberg, G. E.; Voth, G. A., Multiscale reactive molecular dynamics. *J Chem Phys* **2012,** *137* (22), 22A525.

19. Yamashita, T.; Peng, Y.; Knight, C.; Voth, G. A., Computationally Efficient Multiconfigurational Reactive Molecular Dynamics. *J Chem Theory Comput* **2012,** *8* (12), 4863-4875.

20. Nelson, J. G.; Peng, Y.; Silverstein, D. W.; Swanson, J. M., Multiscale Reactive Molecular Dynamics for Absolute pKa Predictions and Amino Acid Deprotonation. *J Chem Theory Comput* **2014,** *10* (7), 2729-2737.

21. Lee, S.; Liang, R.; Voth, G. A.; Swanson, J. M., Computationally Efficient Multiscale Reactive Molecular Dynamics to Describe Amino Acid Deprotonation in Proteins. *J Chem Theory Comput* **2016,** *12* (2), 879-91.

22. Lee, S.; Swanson, J. M.; Voth, G. A., Multiscale Simulations Reveal Key Aspects of the Proton Transport Mechanism in the ClC-ec1 Antiporter. *Biophys J* **2016,** *110* (6), 1334-45.

23. Liang, R.; Li, H.; Swanson, J. M.; Voth, G. A., Multiscale simulation reveals a multifaceted mechanism of proton permeation through the influenza A M2 proton channel. *Proc Natl Acad Sci U S A* **2014,** *111* (26), 9396-401.

24. Liang, R.; Swanson, J. M.; Peng, Y.; Wikstrom, M.; Voth, G. A., Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. *Proc Natl Acad Sci U S A* **2016,** *113* (27), 7420-5.

25. Lee, S.; Mayes, H. B.; Swanson, J. M.; Voth, G. A., The Origin of Coupled Chloride and Proton Transport in a Cl(-)/H(+) Antiporter. *J Am Chem Soc* **2016,** *138* (45), 14923-14930.

26. Liang, R.; Swanson, J. M. J.; Wikstrom, M.; Voth, G. A., Understanding the essential proton-pumping kinetic gates and decoupling mutations in cytochrome c oxidase. *Proc Natl Acad Sci U S A* **2017,** *114* (23), 5924-5929.

27. Parker, J. L.; Li, C.; Brinth, A.; Wang, Z.; Vogeley, L.; Solcan, N.; Ledderboge-Vucinic, G.; Swanson, J. M. J.; Caffrey, M.; Voth, G. A.; Newstead, S., Proton movement and coupling in the POT family of peptide transporters. *Proc Natl Acad Sci U S A* **2017,** *114* (50), 13182-13187.

28. Wang, Z.; Swanson, J. M. J.; Voth, G. A., Modulating the Chemical Transport Properties of a Transmembrane Antiporter via Alternative Anion Flux. *J Am Chem Soc* **2018,** *140* (48), 16535-16543.

29. Mayes, H. B.; Lee, S.; White, A. D.; Voth, G. A.; Swanson, J. M. J., Multiscale Kinetic Modeling Reveals an Ensemble of Cl(-)/H(+) Exchange Pathways in ClC-ec1 Antiporter. *J Am Chem Soc* **2018,** *140* (5), 1793-1804.

30. Watkins, L. C.; Liang, R.; Swanson, J. M. J.; DeGrado, W. F.; Voth, G. A., Proton-Induced Conformational and Hydration Dynamics in the Influenza A M2 Channel. *J Am Chem Soc* **2019,** *141* (29), 11667-11676.

31. Knight, C.; Voth, G. A., The Curious Case of the Hydrated Proton. *Acc. Chem. Res.* **2012,** *45* (1), 101-109.

32. Swanson, J. M.; Maupin, C. M.; Chen, H.; Petersen, M. K.; Xu, J.; Wu, Y.; Voth, G. A., Proton solvation and transport in aqueous and biomolecular systems: insights from computer simulations. *J Phys Chem B* **2007,** *111* (17), 4300-14.

33. Taraphder, S.; Maupin, C. M.; Swanson, J. M.; Voth, G. A., Coupling Protein Dynamics with Proton Transport in Human Carbonic Anhydrase II. *J Phys Chem B* **2016,** *120* (33), 8389-404.

34. Pinto, L. H.; Lamb, R. A., The M2 proton channels of influenza A and B viruses. *J Biol Chem* **2006,** *281* (14), 8997-9000.

35. Pinto, L. H.; Holsinger, L. J.; Lamb, R. A., Influenza virus M2 protein has ion channel activity. *Cell* **1992,** *69* (3), 517-28.

36. Chizhmakov, I. V.; Geraghty, F. M.; Ogden, D. C.; Hayhurst, A.; Antoniou, M.; Hay, A. J., Selective proton permeability and pH regulation of the influenza virus M2 channel expressed in mouse erythroleukaemia cells. *J Physiol* **1996,** *494 (Pt 2)*, 329-36.

37. Wang, C.; Takeuchi, K.; Pinto, L. H.; Lamb, R. A., Ion channel activity of influenza A virus M2 protein: characterization of the amantadine block. *J Virol* **1993,** *67* (9), 5585-94.

38. Wang, C.; Lamb, R. A.; Pinto, L. H., Activation of the M2 ion channel of influenza virus: a role for the transmembrane domain histidine residue. *Biophys J* **1995,** *69* (4), 1363-71.

39. Tang, Y.; Zaitseva, F.; Lamb, R. A.; Pinto, L. H., The gate of the influenza virus M2 proton channel is formed by a single tryptophan residue. *J Biol Chem* **2002,** *277* (42), 39880-6.

40. Hu, F.; Luo, W.; Cady, S. D.; Hong, M., Conformational plasticity of the influenza A M2 transmembrane helix in lipid bilayers under varying pH, drug binding, and membrane thickness. *Biochim Biophys Acta* **2011,** *1808* (1), 415-23.

41. Hu, F.; Luo, W.; Hong, M., Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. *Science* **2010,** *330* (6003), 505-508.

42. Li, C.; Qin, H.; Gao, F. P.; Cross, T. A., Solid-state NMR characterization of conformational plasticity within the transmembrane domain of the influenza A M2 proton channel. *Biochim Biophys Acta* **2007,** *1768* (12), 3162-70.

43. Schnell, J. R.; Chou, J. J., Structure and mechanism of the M2 proton channel of influenza A virus. *Nature* **2008,** *451* (7178), 591-5.

44. Sharma, M.; Yi, M.; Dong, H.; Qin, H.; Peterson, E.; Busath, D. D.; Zhou, H. X.; Cross, T. A., Insight into the mechanism of the influenza A proton channel from a structure in a lipid bilayer. *Science* **2010,** *330* (6003), 509-12.

45. Polishchuk, A. L.; Lear, J. D.; Ma, C.; Lamb, R. A.; Pinto, L. H.; Degrado, W. F., A pH-dependent conformational ensemble mediates proton transport through the influenza A/M2 protein. *Biochemistry* **2010,** *49* (47), 10061-71.

46. Acharya, R.; Carnevale, V.; Fiorin, G.; Levine, B. G.; Polishchuk, A. L.; Balannik, V.; Samish, I.; Lamb, R. A.; Pinto, L. H.; DeGrado, W. F.; Klein, M. L., Structure and mechanism of proton transport through the transmembrane tetrameric M2 protein bundle of the influenza A virus. *Proc Natl Acad Sci U S A* **2010,** *107* (34), 15075-80.

47. Wei, C.; Pohorille, A., Activation and proton transport mechanism in influenza A M2 channel. *Biophys J* **2013,** *105* (9), 2036-45.

48. Williams, J. K.; Zhang, Y.; Schmidt-Rohr, K.; Hong, M., pH-dependent conformation, dynamics, and aromatic interaction of the gating tryptophan residue of the influenza M2 proton channel from solid-state NMR. *Biophys J* **2013,** *104* (8), 1698-708.

49. Hong, M.; DeGrado, W. F., Structural basis for proton conduction and inhibition by the influenza M2 protein. *Protein Sci* **2012,** *21* (11), 1620-33.

50. Miao, Y.; Fu, R.; Zhou, H. X.; Cross, T. A., Dynamic Short Hydrogen Bonds in Histidine Tetrad of Full-Length M2 Proton Channel Reveal Tetrameric Structural Heterogeneity and Functional Mechanism. *Structure* **2015,** *23* (12), 2300-2308.

51. Liao, S. Y.; Yang, Y.; Tietze, D.; Hong, M., The influenza m2 cytoplasmic tail changes the proton-exchange equilibria and the backbone conformation of the transmembrane histidine residue to facilitate proton conduction. *J Am Chem Soc* **2015,** *137* (18), 6067-77.

52. Colvin, M. T.; Andreas, L. B.; Chou, J. J.; Griffin, R. G., Proton association constants of His 37 in the Influenza-A M218-60 dimer-of-dimers. *Biochemistry* **2014,** *53* (38), 5987-94.

53. Hu, J.; Fu, R.; Nishimura, K.; Zhang, L.; Zhou, H. X.; Busath, D. D.; Vijayvergiya, V.; Cross, T. A., Histidines, heart of the hydrogen ion channel from influenza A virus: toward an understanding of conductance and proton selectivity. *Proc Natl Acad Sci U S A* **2006,** *103* (18), 6865-70.

54. Chizhmakov, I. V.; Ogden, D. C.; Geraghty, F. M.; Hayhurst, A.; Skinner, A.; Betakova, T.; Hay, A. J., Differences in conductance of M2 proton channels of two influenza viruses at low and high pH. *Journal of Physiology* **2003,** *546*, 427-438.

55. Mould, J. A.; Li, H. C.; Dudlak, C. S.; Lear, J. D.; Pekosz, A.; Lamb, R. A.; Pinto, L. H., Mechanism for proton conduction of the M(2) ion channel of influenza A virus. *J Biol Chem* **2000,** *275* (12), 8592-9.

56. Hu, F.; Schmidt-Rohr, K.; Hong, M., NMR detection of pH-dependent histidine-water proton exchange reveals the conduction mechanism of a transmembrane proton channel. *J Am Chem Soc* **2012,** *134* (8), 3703-13.

57. Leiding, T.; Wang, J.; Martinsson, J.; DeGrado, W. F.; Arskold, S. P., Proton and cation transport activity of the M2 proton channel from influenza A virus. *Proc Natl Acad Sci U S A* **2010,** *107* (35), 15409-14.

58. Chen, W.; Huang, Y.; Shen, J., Conformational Activation of a Transmembrane Proton Channel from Constant pH Molecular Dynamics. *J Phys Chem Lett* **2016,** *7* (19), 3961-3966.

59. Thomaston, J. L.; Alfonso-Prieto, M.; Woldeyes, R. A.; Fraser, J. S.; Klein, M. L.; Fiorin, G.; DeGrado, W. F., High-resolution structures of the M2 channel from influenza A virus reveal dynamic pathways for proton stabilization and transduction. *Proc Natl Acad Sci U S A* **2015,** *112* (46), 14260-5.

60. Ma, C.; Polishchuk, A. L.; Ohigashi, Y.; Stouffer, A. L.; Schon, A.; Magavern, E.; Jing, X.; Lear, J. D.; Freire, E.; Lamb, R. A.; DeGrado, W. F.; Pinto, L. H., Identification of the functional core of the influenza A virus A/M2 proton-selective ion channel. *Proc Natl Acad Sci U S A* **2009,** *106* (30), 12283-8.

61. Sugita, Y.; Kitao, A.; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations. *Journal of Chemical Physics* **2000,** *113* (15), 6042-6051.

62. Day, T. J. F.; Soudackov, A. V.; Čuma, M.; Schmitt, U. W.; Voth, G. A., A second generation multistate empirical valence bond model for proton transport in aqueous systems. *The Journal of Chemical Physics* **2002,** *117* (12), 5839-5849.

63. Peng, Y.; Swanson, J. M.; Kang, S. G.; Zhou, R.; Voth, G. A., Hydrated Excess Protons Can Create Their Own Water Wires. *J Phys Chem B* **2015,** *119* (29), 9212-8.

64. Li, H.; Chen, H.; Zeuthen, T.; Conrad, C.; Wu, B.; Beitz, E.; Voth, G. A., Enhancement of Proton Conductance by Mutations of the Selectivity Filter of Aquaporin-1. *J. Mol. Biol.* **2011,** *407*, 607-620.

65. Smart, O. S.; Neduvelil, J. G.; Wang, X.; Wallace, B. A.; Sansom, M. S. P., HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics* **1996,** *14* (6), 354-360.

66. Thomaston, J. L.; Polizzi, N. F.; Konstantinidi, A.; Wang, J.; Kolocouris, A.; DeGrado, W. F., Inhibitors of the M2 Proton Channel Engage and Disrupt Transmembrane Networks of Hydrogen-Bonded Waters. *J Am Chem Soc* **2018,** *140* (45), 15219-15226.

67. Zhou, H.-X.; Cross, T. A., Modeling the membrane environment has implications for membrane protein structure and function: Influenza A M2 protein. *Protein Science* **2013,** *22* (4), 381-394.

68. Luo, W.; Cady, S. D.; Hong, M., Immobilization of the influenza A M2 transmembrane peptide in virus envelope-mimetic lipid membranes: a solid-state NMR investigation. *Biochemistry* **2009,** *48* (27), 6361-8.

69. Lin, C. W.; Mensa, B.; Barniol-Xicota, M.; DeGrado, W. F.; Gai, F., Activation pH and Gating Dynamics of Influenza A M2 Proton Channel Revealed by Single-Molecule Spectroscopy. *Angew Chem Int Ed Engl* **2017,** *56* (19), 5283-5287.

70. Thomaston, J. L.; Woldeyes, R. A.; Nakane, T.; Yamashita, A.; Tanaka, T.; Koiwai, K.; Brewster, A. S.; Barad, B. A.; Chen, Y.; Lemmin, T.; Uervirojnangkoorn, M.; Arima, T.;

Kobayashi, J.; Masuda, T.; Suzuki, M.; Sugahara, M.; Sauter, N. K.; Tanaka, R.; Nureki, O.; Tono, K.; Joti, Y.; Nango, E.; Iwata, S.; Yumoto, F.; Fraser, J. S.; DeGrado, W. F., XFEL structures of the influenza M2 proton channel: Room temperature water networks and insights into proton conduction. *Proc Natl Acad Sci U S A* **2017,** *114* (51), 13357-13362.

71. Gianti, E.; Carnevale, V.; Degrado, W. F.; Klein, M. L.; Fiorin, G., Hydrogen-bonded water molecules in the m2 channel of the influenza a virus guide the binding preferences of ammonium-based inhibitors. *Journal of Physical Chemistry B* **2015,** *119* (3), 1173--1183.

72. Humphrey, W.; Dalke, A.; Schulten, K., {VMD} -- {V}isual {M}olecular {D}ynamics. *Journal of Molecular Graphics* **1996,** *14*, 33-38.

73. World Health Organization. Influenza (Seasonal). https://www.who.int/en/news-room/fact-sheets/detail/influenza-(seasonal) (accessed March 16, 2020).

74. World Health Organization. Coronavirus disease (COVID-2019) Situation Report - 143. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200611-covid-19-sitrep-143.pdf (accessed June 11, 2020).

75. Shaw, M. L., The Next Wave of Influenza Drugs. *ACS Infect Dis* **2017,** *3* (10), 691-694.

76. Dong, G.; Peng, C.; Luo, J.; Wang, C.; Han, L.; Wu, B.; Ji, G.; He, H., Adamantane-resistant influenza a viruses in the world (1902-2013): frequency and distribution of M2 gene mutations. *PLoS One* **2015,** *10* (3), e0119115.

77. Musharrafieh, R.; Ma, C.; Wang, J., Discovery of M2 channel blockers targeting the drug-resistant double mutants M2-S31N/L26I and M2-S31N/V27A from the influenza A viruses. *Eur J Pharm Sci* **2020,** *141*, 105124.

78. Artem, O.; Timur, G.; Tatiana, G.; Tatiana, P., Prospects for Specific Influenza Treatment. *SRP* **2020,** *11* (2), 242.

79. Scott, C.; Griffin, S., Viroporins: structure, function and potential as antiviral targets. *J Gen Virol* **2015,** *96* (8), 2000-2027.

80. Wilson, L.; McKinlay, C.; Gage, P.; Ewart, G., SARS coronavirus E protein forms cation-selective ion channels. *Virology* **2004,** *330* (1), 322-31.

81. Lu, W.; Zheng, B. J.; Xu, K.; Schwarz, W.; Du, L.; Wong, C. K.; Chen, J.; Duan, S.; Deubel, V.; Sun, B., Severe acute respiratory syndrome-associated coronavirus 3a protein forms an ion channel and modulates virus release. *Proc Natl Acad Sci U S A* **2006,** *103* (33), 12540-5.

82. Issa, E.; Merhi, G.; Panossian, B.; Salloum, T.; Tokajian, S., SARS-CoV-2 and ORF3a: Nonsynonymous Mutations, Functional Domains, and Viral Pathogenesis. *mSystems* **2020,** *5* (3), e00266-20.

83. Venkataraman, P.; Lamb, R. A.; Pinto, L. H., Chemical rescue of histidine selectivity filter mutants of the M2 ion channel of influenza A virus. *J Biol Chem* **2005,** *280* (22), 21463-72.

84. Hu, F.; Luo, W.; Hong, M., Mechanisms of proton conduction and gating in influenza M2 proton channels from solid-state NMR. *Science* **2010,** *330* (6003), 505-8.

85. DiFrancesco, M. L.; Hansen, U. P.; Thiel, G.; Moroni, A.; Schroeder, I., Effect of cytosolic pH on inward currents reveals structural characteristics of the proton transport cycle in the influenza A protein M2 in cell-free membrane patches of Xenopus oocytes. *PLoS One* **2014,** *9* (9), e107406.

86. Jalily, P. H.; Duncan, M. C.; Fedida, D.; Wang, J.; Tietjen, I., Put a cork in it: Plugging the M2 viral ion channel to sink influenza. *Antiviral Res* **2020,** *178*, 104780.

87. Wang, Y.; Hu, Y.; Xu, S.; Zhang, Y.; Musharrafieh, R.; Hau, R. K.; Ma, C.; Wang, J., In Vitro Pharmacokinetic Optimizations of AM2-S31N Channel Blockers Led to the Discovery of

Slow-Binding Inhibitors with Potent Antiviral Activity against Drug-Resistant Influenza A Viruses. *J Med Chem* **2018,** *61* (3), 1074-1085.

88. Barniol-Xicota, M.; Gazzarrini, S.; Torres, E.; Hu, Y.; Wang, J.; Naesens, L.; Moroni, A.; Vazquez, S., Slow but Steady Wins the Race: Dissimilarities among New Dual Inhibitors of the Wild-Type and the V27A Mutant M2 Channels of Influenza A Virus. *J Med Chem* **2017,** *60* (9), 3727-3738.

89. Gordon, N. A.; McGuire, K. L.; Wallentine, S. K.; Mohl, G. A.; Lynch, J. D.; Harrison, R. G.; Busath, D. D., Divalent copper complexes as influenza A M2 inhibitors. *Antiviral Res* **2017,** *147*, 100-106.

90. Duque, M. D.; Torres, E.; Valverde, E.; Barniol, M.; Guardiola, S.; Rey, M.; Vzquez, S., *Editor: Diego Mu \~ n oz-Torrero 2. Inhibitors of the M2 channel of influenza A virus*. 2011; Vol. 37, p 978--81.

91. Wang, J.; Ma, C.; Balannik, V.; Pinto, L. H.; Lamb, R. A.; Degrado, W. F., Exploring the Requirements for the Hydrophobic Scaffold and Polar Amine in inhibitors of M2 from Influenza A Virus. *ACS Med Chem Lett* **2011,** *2* (4), 307-312.

92. Hu, W.; Zeng, S.; Li, C.; Jie, Y.; Li, Z.; Chen, L., Identification of hits as matrix-2 protein inhibitors through the focused screening of a small primary amine library. *J Med Chem* **2010,** *53* (9), 3831-4.

93. Drakopoulos, A.; Tzitzoglaki, C.; McGuire, K.; Hoffmann, A.; Konstantinidi, A.; Kolokouris, D.; Ma, C.; Freudenberger, K.; Hutterer, J.; Gauglitz, G.; Wang, J.; Schmidtke, M.; Busath, D. D.; Kolocouris, A., Unraveling the Binding, Proton Blockage, and Inhibition of Influenza M2 WT and S31N by Rimantadine Variants. *ACS Med Chem Lett* **2018,** *9* (3), 198-203.

94. Wang, J.; Ma, C.; Fiorin, G.; Carnevale, V.; Wang, T.; Hu, F.; Lamb, R. A.; Pinto, L. H.; Hong, M.; Klein, M. L.; DeGrado, W. F., Molecular dynamics simulation directed rational design of inhibitors targeting drug-resistant mutants of influenza A virus M2. *J Am Chem Soc* **2011,** *133* (32), 12834-41.

95. Van Nguyen, H.; Nguyen, H. T.; Le, L. T., Investigation of the free energy profiles of amantadine and rimantadine in the AM2 binding pocket. *Eur Biophys J* **2016,** *45* (1), 63-70.

96. Llabres, S.; Juarez-Jimenez, J.; Masetti, M.; Leiva, R.; Vazquez, S.; Gazzarrini, S.; Moroni, A.; Cavalli, A.; Luque, F. J., Mechanism of the Pseudoirreversible Binding of Amantadine to the M2 Proton Channel. *J Am Chem Soc* **2016,** *138* (47), 15345-15358.

97. Homeyer, N.; Ioannidis, H.; Kolarov, F.; Gauglitz, G.; Zikos, C.; Kolocouris, A.; Gohlke, H., Interpreting Thermodynamic Profiles of Aminoadamantane Compounds Inhibiting the M2 Proton Channel of Influenza A by Free Energy Calculations. *J Chem Inf Model* **2016,** *56* (1), 110-26.

98. Gianti, E.; Carnevale, V.; DeGrado, W. F.; Klein, M. L.; Fiorin, G., Hydrogen-bonded water molecules in the M2 channel of the influenza A virus guide the binding preferences of ammonium-based inhibitors. *J Phys Chem B* **2015,** *119* (3), 1173-83.

99. Leonov, H.; Astrahan, P.; Krugliak, M.; Arkin, I. T., How do aminoadamantanes block the influenza M2 channel, and how does resistance develop? *J Am Chem Soc* **2011,** *133* (25), 9903-11.

100. Ioannidis, H.; Drakopoulos, A.; Tzitzoglaki, C.; Homeyer, N.; Kolarov, F.; Gkeka, P.; Freudenberger, K.; Liolios, C.; Gauglitz, G.; Cournia, Z.; Gohlke, H.; Kolocouris, A., Alchemical Free Energy Calculations and Isothermal Titration Calorimetry Measurements of Aminoadamantanes Bound to the Closed State of Influenza A/M2TM. *J Chem Inf Model* **2016,** *56* (5), 862-76.

101. Biswas, R.; Tse, Y. L.; Tokmakoff, A.; Voth, G. A., Role of Presolvation and Anharmonicity in Aqueous Phase Hydrated Proton Solvation and Transport. *J Phys Chem B* **2016,** *120* (8), 1793-804.

102. Day, T. J. F.; Soudackov, A. V.; Cuma, M.; Schmitt, U. W.; Voth, G. A., A second generation multistate empirical valence bond model for proton transport in aqueous systems. *Journal of Chemical Physics* **2002,** *117* (12), 5839-5849.

103. Plimpton, S., Fast Parallel Algorithms for Short-Range Molecular Dynamics. *Journal of Computational Physics* **1995,** *117* (1), 1-19.

104. consortium, P., Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **2019,** *16* (8), 670-673.

105. Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G., PLUMED 2: New feathers for an old bird. *Comput Phys Commun* **2014,** *185* (2), 604-613.

106. Rossum, G. *Python reference manual*; CWI (Centre for Mathematics and Computer Science): 1995.

107. Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, I.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy, C., SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* **2020,** *17* (3), 261-272.

108. Travis E, O., A guide to NumPy. Trelgol Publishing: USA, 2006.

109. McKinney, W., Data Structures for Statistical Computing in Python. **2010**, 51--56.

110. Hunter, J. D., Matplotlib: A 2D graphics environment. *Comput Sci Eng* **2007,** *9* (3), 90-95.

111. Hu, J.; Fu, R.; Cross, T. A., The chemical and dynamical influence of the anti-viral drug amantadine on the M2 proton channel transmembrane domain. *Biophys J* **2007,** *93* (1), 276-83.

112. Cady, S. D.; Schmidt-Rohr, K.; Wang, J.; Soto, C. S.; Degrado, W. F.; Hong, M., Structure of the amantadine binding site of influenza M2 proton channels in lipid bilayers. *Nature* **2010,** *463* (7281), 689-92.

113. Khurana, E.; Devane, R. H.; Dal Peraro, M.; Klein, M. L., Computational study of drug binding to the membrane-bound tetrameric M2 peptide bundle from influenza A virus. *Biochim Biophys Acta* **2011,** *1808* (2), 530-7.

114. Yi, M.; Cross, T. A.; Zhou, H. X., A secondary gate as a mechanism for inhibition of the M2 proton channel by amantadine. *J Phys Chem B* **2008,** *112* (27), 7977-9.

115. Moorthy, N. S.; Poongavanam, V.; Pratheepa, V., Viral M2 ion channel protein: a promising target for anti-influenza drug discovery. *Mini Rev Med Chem* **2014,** *14* (10), 819-30.

116. Duque, M. D.; Ma, C.; Torres, E.; Wang, J.; Naesens, L.; Juarez-Jimenez, J.; Camps, P.; Luque, F. J.; DeGrado, W. F.; Lamb, R. A.; Pinto, L. H.; Vazquez, S., Exploring the size limit of templates for inhibitors of the M2 ion channel of influenza A virus. *J Med Chem* **2011,** *54* (8), 2646-57.

117. Zhao, X.; Zhang, Z. W.; Cui, W.; Chen, S. W.; Zhou, Y.; Dong, J. H.; Jie, Y. L.; Wan, J. T.; Xu, Y.; Hu, W. H., Identification of camphor derivatives as novel M2 ion channel inhibitors of influenza A virus. *Medchemcomm* **2015,** *6* (4), 727-731.

118. Zhao, X.; Jie, Y.; Rosenberg, M. R.; Wan, J.; Zeng, S.; Cui, W.; Xiao, Y.; Li, Z.; Tu, Z.; Casarotto, M. G.; Hu, W., Design and synthesis of pinanamine derivatives as anti-influenza A M2 ion channel inhibitors. *Antiviral Res* **2012,** *96* (2), 91-9.

119. Rey-Carrizo, M.; Torres, E.; Ma, C.; Barniol-Xicota, M.; Wang, J.; Wu, Y.; Naesens, L.; DeGrado, W. F.; Lamb, R. A.; Pinto, L. H.; Vazquez, S., 3-Azatetracyclo[5.2.1.1(5,8).0(1,5)]undecane derivatives: from wild-type inhibitors of the M2 ion channel of influenza A virus to derivatives with potent activity against the V27A mutant. *J Med Chem* **2013,** *56* (22), 9265-74.

120. Li, F.; Hu, Y.; Wang, Y.; Ma, C.; Wang, J., Expeditious Lead Optimization of Isoxazole-Containing Influenza A Virus M2-S31N Inhibitors Using the Suzuki-Miyaura Cross-Coupling Reaction. *J Med Chem* **2017,** *60* (4), 1580-1590.

121. Duong-Ly, K. C.; Nanda, V.; Degrado, W. F.; Howard, K. P., The conformation of the pore region of the M2 proton channel depends on lipid bilayer environment. *Protein Sci* **2005,** *14* (4), 856-61.

122. Mandala, V. S.; Loftis, A. R.; Shcherbakov, A. A.; Pentelute, B. L.; Hong, M., Atomic structures of closed and open influenza B M2 proton channel reveal the conduction mechanism. *Nat Struct Mol Biol* **2020,** *27* (2), 160-167.

123. Nieva, J. L.; Madan, V.; Carrasco, L., Viroporins: structure and biological functions. *Nat Rev Microbiol* **2012,** *10* (8), 563-74.

124. Chizhmakov, I. V.; Geraghty, F. M.; Ogden, D. C.; Hayhurst, A.; Antoniou, M.; Hay, A. J., Selective proton permeability and pH regulation of the influenza virus M2 channel expressed in mouse erythroleukaemia cells. *J Physiol* **1996,** *494 (Pt 2)* (2), 329-36.

125. Pinto, L. H.; Holsinger, L. J.; Lamb, R. A., Influenza virus M2 protein has ion channel activity. *Cell* **1992,** *69* (3), 517-528.

126. Ma, C.; Fiorin, G.; Carnevale, V.; Wang, J.; Lamb, R. A.; Klein, M. L.; Wu, Y.; Pinto, L. H.; DeGrado, W. F., Asp44 stabilizes the Trp41 gate of the M2 proton channel of influenza A virus. *Structure* **2013,** *21* (11), 2033-41.

127. Betakova, T.; Ciampor, F.; Hay, A. J., Influence of residue 44 on the activity of the M2 proton channel of influenza A virus. *J Gen Virol* **2005,** *86* (Pt 1), 181-184.

128. Grambas, S.; Bennett, M. S.; Hay, A. J., Influence of amantadine resistance mutations on the pH regulatory function of the M2 protein of influenza A viruses. *Virology* **1992,** *191* (2), 541-9.

129. Kern, D. M.; Sorum, B.; Mali, S. S.; Hoel, C. M.; Sridharan, S.; Remis, J. P.; Toso, D. B.; Kotecha, A.; Bautista, D. M.; Brohawn, S. G., Cryo-EM structure of SARS-CoV-2 ORF3a in lipid nanodiscs. *Nat Struct Mol Biol* **2021,** *28* (7), 573-582.

130. Hu, J.; Asbury, T.; Achuthan, S.; Li, C.; Bertram, R.; Quine, J. R.; Fu, R.; Cross, T. A., Backbone structure of the amantadine-blocked trans-membrane domain M2 proton channel from Influenza A virus. *Biophys J* **2007,** *92* (12), 4335-43.

131. Chizhmakov, I. V.; Ogden, D. C.; Geraghty, F. M.; Hayhurst, A.; Skinner, A.; Betakova, T.; Hay, A. J., Differences in conductance of M2 proton channels of two influenza viruses at low and high pH. *J Physiol* **2003,** *546* (Pt 2), 427-38.

132. Liang, R.; Swanson, J. M. J.; Peng, Y.; Wikström, M.; Voth, G. A., Multiscale simulations reveal key features of the proton-pumping mechanism in cytochrome c oxidase. *Proceedings of the National Academy of Sciences* **2016,** *113,* 7420-7425.

133. Li, C.; Yue, Z.; Espinoza-Fonseca, L. M.; Voth, G. A., Multiscale Simulation Reveals Passive Proton Transport Through SERCA on the Microsecond Timescale. *Biophys J* **2020,** *119* (5), 1033-1040.

134. Wang, Z.; Swanson, J. M. J.; Voth, G. A., Local conformational dynamics regulating transport properties of a Cl(-) /H(+) antiporter. *J Comput Chem* **2020,** *41* (6), 513-519.

135. Watkins, L. C.; DeGrado, W. F.; Voth, G. A., Influenza A M2 Inhibitor Binding Understood through Mechanisms of Excess Proton Stabilization and Channel Dynamics. *J Am Chem Soc* **2020,** *142* (41), 17425-17433.

136. Humphrey, W.; Dalke, A.; Schulten, K., VMD -- V isual M olecular D ynamics. *Journal of Molecular Graphics* **1996,** *14*, 33--38.

137. Sugita, Y.; Kitao, A.; Okamoto, Y., Multidimensional replica-exchange method for free-energy calculations. *The Journal of Chemical Physics* **2000,** *113* (15), 6042-6051.

138. Bonomi, M. B., G.; Camilloni, C.; et al., Promoting transparency and reproducibility in enhanced molecular simulations. *Nat Methods* **2019,** *16* (8), 670-673.

139. Konig, P. H.; Ghosh, N.; Hoffmann, M.; Elstner, M.; Tajkhorshid, E.; Frauenheim, T.; Cui, Q., Toward theoretical analysis of long-range proton transfer kinetics in biomolecular pumps. *J Phys Chem A* **2006,** *110* (2), 548-63.

140. Riccardi, D.; Konig, P.; Prat-Resina, X.; Yu, H.; Elstner, M.; Frauenheim, T.; Cui, Q., "Proton holes" in long-range proton transfer reactions in solution and enzymes: A theoretical analysis. *J Am Chem Soc* **2006,** *128* (50), 16302-11.

141. Liang, R.; Swanson, J. M.; Voth, G. A., Benchmark Study of the SCC-DFTB Approach for a Biomolecular Proton Channel. *J Chem Theory Comput* **2014,** *10* (1), 451-462.

142. Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys Rev B Condens Matter* **1988,** *37* (2), 785-789.

143. Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys Rev A Gen Phys* **1988,** *38* (6), 3098-3100.

144. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H., A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J Chem Phys* **2010,** *132* (15), 154104.

145. Lippert, B. G.; Parrinello, J. H.; Michele, A hybrid Gaussian and plane wave density functional scheme. *Molecular Physics* **2010,** *92* (3), 477-488.

146. Hartwigsen, C.; Goedecker, S.; Hutter, J., Relativistic separable dual-space Gaussian pseudopotentials from H to Rn. *Physical Review B* **1998,** *58* (7), 3641-3662.

147. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J., Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comput Phys Commun* **2005,** *167* (2), 103-128.

148. Wu, H.; Mey, A. S.; Rosta, E.; Noe, F., Statistically optimal analysis of state-discretized trajectory data from multiple thermodynamic states. *J Chem Phys* **2014,** *141* (21), 214106.

149. Scherer, M. K.; Trendelkamp-Schroer, B.; Paul, F.; Perez-Hernandez, G.; Hoffmann, M.; Plattner, N.; Wehmeyer, C.; Prinz, J. H.; Noe, F., PyEMMA 2: A Software Package for Estimation, Validation, and Analysis of Markov Models. *J Chem Theory Comput* **2015,** *11* (11), 5525-42.

150. Zhou, H. X., A theory for the proton transport of the influenza virus M2 protein: extensive test against conductance data. *Biophys J* **2011,** *100* (4), 912-21.

151. Mravic, M.; Thomaston, J. L.; Tucker, M.; Solomon, P. E.; Liu, L.; DeGrado, W. F., Packing of apolar side chains enables accurate design of highly stable membrane proteins. *Science* **2019,** *363* (6434), 1418-1423.

152. Li, C.; Voth, G. A., A Quantitative Paradigm for Water Assisted Proton Transport Through Proteins and Other Confined Spaces. *bioRxiv* **2021**.

153. Grossfield, A., “WHAM: the weighted histogram analysis method”, version 2.0.9, http://membrane.urmc.rochester.edu/wordpress/?page_id=126.