UC Davis
UC Davis Electronic Theses and Dissertations

## Title

Computational Approaches for Advancing Ion Channel Drug Discovery and Precision Medicine

## Permalink

https://escholarship.org/uc/item/1w64f2jf

## Author

Ngo, Khoa Tran Anh

## Publication Date

2024

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Computational Approaches for Advancing
Ion Channel Drug Discovery and Precision Medicine

By

KHOA NGO
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Biophysics

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

Colleen E. Clancy, Chair

Igor Vorobyov

Jon Sack

Committee in Charge

2024

## Abstract

Ion channels are integral membrane proteins facilitating permeation of ions through aqueous pores across cell membranes. They are essential to numerous physiological processes such as nerve impulses and heartbeats and are key therapeutic targets due to their involvement in various diseases. This dissertation employs advanced computational techniques to investigate the structure and function of ion channels, specifically focusing on the human (denoted by the prefix "h") voltage-gated sodium channel hNa$_V$1.7, voltage-gated potassium channel hERG, and small-conductance calcium-activated potassium channel hSK2, which are critical therapeutic targets for pain management and the treatment of cardiac arrhythmias. A brief overview of ion channels and lipid membranes as well as their experimental and computational studies is provided in Chapters 1 and 2, followed by three specific ion channel related stories in Chapters 3 through 5 and finishing with closing remarks and future outlook in Chapter 6.

The first story focuses on the human voltage-gated sodium channel hNa$_V$1.7, a prime target for pain therapy due to its predominant expression in peripheral nociceptive neurons. We explored the interaction between the hNa$_V$1.7 channel's voltage sensing domains (VSDs) II and IV in different conformational states and the tarantula venom peptide protoxin-II, a potent and selective inhibitor of the channel. Our computational models revealed that protoxin-II exhibits state-specific binding modes to different states of the channel's VSDs, illuminating the molecular basis of its inhibitory effects through enhanced hNa$_V$1.7 channel deactivation and inactivation. Additionally, we identified key channel's residues involved in the toxin binding and their energetic contributions. These findings are crucial for the rational design of novel pain therapeutics with improved selectivity and potency.

The second story shifts focus to the human voltage-gated potassium channel hK$_V$11.1, also known as hERG (human *Ether-à-go-go*-Related Gene), a vital component of cardiac repolarization and a notorious drug anti-target. This makes the hERG channel a primary target for assessing the safety of new drugs due to association of its inhibition with drug-induced arrhythmias. Utilizing AlphaFold2, a state-of-the-art deep learning tool trained to predict protein structures from amino acid sequences, we modeled the elusive inactivated

and closed conformations of the hERG channel not available experimentally. Molecular dynamics and drug docking simulations validated these models, aligning closely with existing experimental data. Drug docking simulations revealed state-specific drug interactions, underscoring the critical importance of the inactivated state for drug binding and the potential for drug entrapment in the closed state. Analysis of interaction networks across different conformational states of the hERG channel provided insights into potential molecular mechanisms underlying state transitions. These findings are pivotal not only for understanding the functioning of this crucial channel but also for predicting drug-induced arrhythmias from drug chemical structures and thus developing cardiac-safe medications.

The third story explores the regulation of the human small-conductance calcium-activated potassium channel hSK2 by phosphatidylinositol 4,5-bisphosphate (PIP2) lipid in cardiac cells, aiming to enhance therapeutic strategies for cardiac arrhythmias like atrial fibrillation. Using optogenetic constructs to deplete PIP2, our collaborative study demonstrated that PIP2 depletion reduces hSK2 current in both cultured Chinese hamster ovary (CHO) cells and rabbit ventricular myocytes, underscoring PIP2's crucial role in hSK2 channel activation. Homology modeling and molecular dynamics simulations revealed the dynamic binding sites and mechanisms of PIP2 interaction with hSK2 channels, identifying key residues like R395 and E398 involved in channel activation by PIP2. This detailed analysis provides insights into the molecular mechanisms regulating cardiac excitability, offering potential targets for new anti-arrhythmic therapies.

In summary, this doctoral dissertation delves into the intricate world of membrane proteins, in particular, ion channels, exploring their structural and functional intricacies including their interactions with the lipid bilayer environment, pharmacological agents, and toxin peptides. By employing cutting-edge computational approaches, we provided novel insights and strategies that hold promise for accelerating drug discovery and advancing precision medicine, ultimately aiming to develop more targeted and effective therapies for cardiac and neurological disorders.

## Acknowledgements

First and foremost, I would like to extend my deepest gratitude to my advisor, Colleen Clancy, whose guidance and steadfast support have been pivotal throughout my PhD journey. Her dedication to my professional success has been a constant source of motivation and has profoundly shaped my development as a researcher. Colleen's insights and encouragement have been invaluable, and I am immensely grateful for her unwavering commitment to my mentorship.

I am equally thankful to my co-advisor, Igor Vorobyov, for his mentorship and invaluable advice. Igor has always been remarkably responsive, generously offering his expertise and support at any hour, even in the early mornings of 3 AM. His willingness to engage deeply with my questions and challenges has greatly contributed to my growth and confidence as a scientist.

My sincere appreciation also goes to Pei-Chi Yang and Gonzalo Hernandez-Hernandez for their essential support and for enriching my understanding of the computational and theoretical frameworks critical to my graduate studies. Their assistance has been crucial in navigating the complexities of my research.

I would also like to extend my gratitude to all the members of the Clancy and Vorobyov Labs. Each member's unique contribution has not only enhanced my academic experience but also provided a supportive and collaborative environment that enriched my doctoral studies.

I could not have embarked on this journey without the unconditional love and support of my family. My deepest thanks to Tao Ngo, Ly Tran, Vu Ngo, and Ngam Ngo, whose unwavering encouragement and belief in my abilities have been my foundation throughout this challenging and rewarding process. Your love has been my constant reassurance and source of strength, and I am eternally grateful for each of you.

Finally, special acknowledgments are reserved for my cats, Momo and Orange, as well as my cherished pets who preceded them. Thank you for your companionship and the countless moments of joy you have brought into my life. While your presence may mean a less tidy home and a somewhat lighter wallet, my heart would indeed be empty without you.

## Table of Contents

# CHAPTER 1:

# Molecular Calligraphy on the Lipid Bilayer Canvas: The Art of Membrane Transport and Electrical Signaling

*"Every cell is a triumph of natural selection, and we are made of trillions of cells. Within us, is a little universe."* – Carl Sagan (1934 – 1996)

---

## 1.1) The Cellular Membrane: A Gateway to Life's Mysteries

The cell membrane, a semi-permeable barrier that delineates the boundaries of all living cells, is a remarkably complex and dynamic structure. Far from being a mere inert envelope, the membrane orchestrates many vital cellular processes and serves as a platform for intricate molecular interactions[1–4]. This beautifully organized assembly of lipids, proteins, and carbohydrates maintains the delicate equilibrium between the intracellular and extracellular environments[1–4]. In addition, it also facilitates various functions essential for cellular survival, communication, and homeostasis[1–4].

At the core of the cell membrane lies the lipid bilayer, a fluid mosaic of amphipathic molecules that form the structural foundation of this biological barrier. This dynamic matrix comprises various lipid species, including phospholipids, glycolipids, and cholesterol, which self-assemble into a two-dimensional sheet through hydrophobic and hydrophilic interactions[1,2,5,6]. The unique biophysical properties of the lipid bilayer, such as its fluidity, curvature, and lateral organization, play a crucial role in modulating the activities of membrane-associated proteins and regulating critical cellular processes[2,5,6].

Embedded within this lipid matrix is a diverse collection of proteins, each meticulously sculpted by evolution to perform specialized functions. These membrane-bound proteins can be broadly classified into two categories: integral proteins, which span the entirety of the lipid bilayer, and peripheral proteins, which are loosely associated with the membrane surface[7,8]. The intricate interplay between these proteins and the lipid environment is a fundamental determinant of membrane structure and function.

Among the integral membrane proteins, ion channels and transporters serve as molecular gatekeepers that regulate the precise flow of ions, nutrients, and other essential molecules across the membrane[9,10]. These protein complexes possess exquisitely designed pores and gates that selectively allow the passage of specific molecular species while maintaining precise control over their transport kinetics. By governing the movement of charged particles and metabolites, these membrane proteins orchestrate a wide range of physiological processes, from electrical signaling in neurons and muscle contraction to nutrient uptake and waste elimination[9,10].

Another significant class of membrane proteins comprises the receptors, which act as molecular sentinels that detect and respond to a diverse array of extracellular signals[11]. These proteins span the lipid bilayer, presenting specific binding sites on their extracellular domains that recognize and interact with ligands, such as hormones, neurotransmitters, or growth factors[11]. Upon ligand binding, receptors undergo conformational changes that trigger intracellular signaling cascades, ultimately leading to various cellular responses, including gene expression, metabolic regulation, and cell proliferation or differentiation[11].

In addition to proteins, the cell membrane is adorned with a diverse array of carbohydrate moieties, collectively referred to as glycocalyx[5,12]. These sugar residues can be found attached to lipids (glycolipids) or proteins (glycoproteins), forming intricate glycoconjugates that play crucial roles in cell-cell recognition, signaling, and adhesion[5,12]. The glycocalyx serves as a molecular interface between the cell and its environment, mediating interactions with other cells, pathogens, and the extracellular matrix[5,12].

### 1.2) Membrane Proteins: Architects of Cellular Functionality

Membrane proteins are integral components of the cell membrane, playing crucial roles in various biological processes and remarkably represent over 50% of all targets for drugs[13,14]. They constitute approximately 20–30% of all cellular proteins and can be broadly classified into two categories: integral (or intrinsic) and peripheral (or extrinsic) membrane proteins[7,13–15].

Integral membrane proteins are embedded within the lipid bilayer, where their hydrophobic regions interact with the fatty acid chains of membrane lipids, and their hydrophilic regions extend into the aqueous environment[1,16]. These proteins include transmembrane proteins, which span the membrane one or more times, and lipid-anchored proteins, which are covalently attached to membrane lipids[17]. Transmembrane proteins are particularly crucial for the structural integrity of the membrane and for mediating transport across it, while lipid-anchored proteins often play roles in signaling and membrane dynamics[17].

Peripheral membrane proteins, on the other hand, are not embedded in the lipid bilayer. Instead, they are loosely attached to the membrane's surface, often through interactions with integral membrane proteins or membrane lipids[17,18]. These proteins can be easily dissociated from the membrane without disrupting the lipid bilayer[17,18].

Membrane proteins are versatile molecules that perform a variety of essential functions within the cell. They are involved in transport, facilitating the movement of ions, nutrients, and waste products across the membrane to maintain cellular homeostasis[17,19,20]. As receptors, they detect extracellular signals such as hormones, neurotransmitters, and growth factors, initiating intracellular signaling pathways[21]. Some membrane proteins also exhibit enzymatic activity, catalyzing crucial biochemical reactions at the membrane surface[22]. Additionally, adhesion molecules among these proteins play a vital role in mediating interactions between cells and the extracellular matrix, which is fundamental for tissue formation and immune responses[23]. Lastly, certain membrane proteins provide structural support to the cell by anchoring the cytoskeleton to the membrane, ensuring cellular stability and shape[23].

Ion channels are a subclass of membrane proteins that form pore-like structures, allowing the selective passage of ions across the cell membrane down their electrochemical gradients[17,19,20]. They play a pivotal role in generating and propagating electrical signals in excitable cells, such as neurons and muscle cells, i.e. skeletal, smooth, and cardiac myocytes[17,19,20].

Ion channels can be classified based on their gating mechanisms (how they open and close), ion selectivity (which ions they allow to pass), and their physiological roles. Voltage-gated ion channels, such as sodium ($Na_V$), potassium ($K_V$), and calcium ($Ca_V$) channels, open and close in response to changes in membrane potential, playing a crucial role in the generation and propagation of electrical signals in excitable cells[24]. Ligand-gated ion channels, also known as ionotropic receptors, are activated by the binding of specific ligands, such as neurotransmitters, facilitating rapid communication across synapses[25]. Mechanosensitive ion channels respond to mechanical forces or changes in membrane tension, enabling cells to sense and respond to physical stimuli[26]. Lastly, thermosensitive ion channels are triggered by temperature variations, contributing to the detection of thermal stimuli and the regulation of temperature-dependent physiological processes[27]. Each type of ion channel contributes to the intricate regulation of cellular activities, highlighting the diversity and complexity of these membrane proteins.

Ion channels play a pivotal role in various physiological processes. They are crucial for neuronal communication, where they enable the initiation and propagation of action potentials in neurons, facilitating rapid signal transmission within the nervous system[24]. In muscle cells, ion channels regulate the flow of ions necessary for muscle contraction[19,24]. They are also essential for maintaining cardiac rhythm, as the coordinated opening and closing of ion channels in cardiac cells ensure a regular heartbeat[28]. Furthermore, ion channels are involved in sensory perception, converting stimuli such as light, sound, touch, and temperature into electrical signals[26,27]. Additionally, they contribute to the regulation of cell volume by controlling the movement of ions and osmotically active substances across the cell membrane, maintaining cellular homeostasis[29].

Given their central role in cellular function and their involvement in various diseases, ion channels are a major focus of biomedical research and drug development[30]. Understanding the structure, function, and regulation of these membrane proteins is key to developing targeted therapies for a range of disorders, including neurological conditions, cardiovascular diseases, and pain management[30,31].

## 1.3) Voltage-Gated Ion Channels: Structures and Functions

Voltage-gated ion channels (VGICs) are transmembrane proteins that play a critical role in the generation and propagation of electrical signals in excitable cells, such as neurons, skeletal muscle cells, and cardiac cells. These channels are sensitive to changes in membrane potential and operate by opening or closing in response to depolarization or hyperpolarization, allowing the selective flow of specific ions across the cell membrane down their electrochemical gradients. This activity is fundamental to rapid electrical signaling between and within cells, and sometimes across long distances, as it initiates major cellular events like synaptic vesicle exocytosis and muscle contraction[24,32,32].

VGICs are tightly regulated by various mechanisms, including phosphorylation, interaction with accessory proteins, and modulation by intracellular signaling molecules[33]. These mechanisms finely tune the opening and closing of these channels, ensuring precise control over cellular excitability and signal transmission. Post-translational modifications such as ubiquitination play a key role in the regulation of the number of channels on the cell surface, thus modulating neuronal excitability[33,34].

VGICs play a crucial role in various physiological processes. They are vital for neuronal communication, as they are responsible for the initiation and propagation of action potentials in neurons, enabling rapid signal transmission within the nervous system[24,32,32]. In cardiac cells, VGICs are key regulators of the heartbeat, controlling the rhythmic depolarization and repolarization of the heart muscle[28]. Additionally, in skeletal, cardiac, and smooth muscle cells, VGICs are also involved in the excitation-contraction coupling process, leading to muscle contraction[24,35].

Dysfunction or mutations in VGICs can lead to a range of pathological conditions, including epilepsy, cardiac arrhythmias, and muscle disorders[30]. Consequently, understanding the structure, function, and regulation of VGICs is of paramount importance for developing targeted therapies to treat these and other related diseases.

### a. Voltage-Gated Sodium Channels

Voltage-gated sodium channels ($Na_V$ channels) are crucial transmembrane proteins that enable electric signaling in cells. They are responsible for the rapid depolarization phase of the action potential in excitable cells, such as neurons, cardiac, smooth, and skeletal muscle cells, by selectively allowing the flow of sodium ions ($Na^+$) across the cell membrane. This function is essential for initiating and propagating action potentials, which are vital for numerous physiological processes[36].

$Na_V$ channels are composed of a large α-subunit, which forms the pore through which $Na^+$ ions pass, and one or two smaller β-subunits that modulate channel function[36]. The α-subunit consists of four homologous domains (I-IV), each containing six transmembrane segments (S1-S6). The S4 segment acts as the voltage sensor, detecting changes in membrane potential and triggering channel opening. The segments S5 and S6, along with the intervening loop, form the pore region that determines ion selectivity and conductance. The inactivation gate, located in the intracellular loop between domains III and IV, temporarily blocks the channel after activation to prevent excessive $Na^+$ influx. This fast inactivation is necessary for the cell to repolarize and prepare the channel for reactivation[36].

Upon membrane depolarization, $Na_V$ channels open, allowing $Na^+$ ions to enter the cell, leading to further depolarization. This creates a positive feedback loop that rapidly raises the membrane potential. Shortly after opening, the channels inactivate, halting $Na^+$ influx and contributing to the repolarization of the membrane. The precise regulation of $Na_V$ channel opening, closing, and inactivation is essential for the proper functioning of the nervous and cardiovascular systems. This process underlies the generation of rapid, high-frequency action potentials necessary for fast signal transmission across excitable tissues[32,36].

There are nine subtypes of $Na_V$ channels, labeled $Na_v1.1$ to $Na_v1.9$, expressed in neuronal, cardiac, smooth and skeletal muscle cells[37]. In humans, the major VGSC isoform in the heart is the $hNa_v1.5$ channel that is responsible for initiating the cardiac action potential[28,38]. $Na_V1.5$ channels are predominantly expressed in cardiac myocytes, the muscle cells

responsible for the contraction of the heart. During the cardiac action potential, $Na_V1.5$ channels open quickly, causing a rapid influx of $Na^+$ ions. This leads to depolarization and the initiation of Phase 0 of the cardiac action potential, crucial for spreading electrical signals in the heart and ensuring synchronized chamber contractions for effective blood pumping.

Mutations or dysfunctions in $Na_V1.5$ channels can lead to various cardiac arrhythmias, such as long QT syndrome, Brugada syndrome, and atrial fibrillation[39]. These conditions often arise from alterations in the channel's ability to conduct sodium ions properly, affecting the electrical activity of the heart and leading to arrhythmic disorders. The study of these mutations has significantly advanced the understanding of the molecular mechanisms underlying these cardiac conditions[39].

Recent research has provided significant insights into the role of the $Na_V1.5$ channel in cardiac electrophysiology and its potential as a target for treating heart rhythm disorders. For example, using computational modeling and simulations, Nguyen *et al.* identified several key binding sites within the pore lumen of $Na_V1.5$ that can accommodate drug molecules[40]. Their simulations revealed a hydrophilic access pathway through the intracellular gate and a hydrophobic pathway through a fenestration between domains III and IV, facilitating drug access to the channel[40]. The findings significantly enhance understanding of the molecular mechanisms by which antiarrhythmic and local anesthetic drugs modulate $Na_V1.5$ activity[40].

Certain subtypes of $Na_V$ channels, such as $Na_V1.7$, $Na_V1.8$, and $Na_V1.9$, are particularly important in pain sensation[41]. They are involved in the generation and propagation of action potentials in nociceptive neurons, which are specialized sensory neurons responsible for detecting painful stimuli. When these neurons are activated by noxious stimuli, $Na_V$ channels open, allowing an influx of $Na^+$ ions that leads to the rapid depolarization of the membrane and the initiation of an action potential. This electrical signal is then transmitted along the sensory nerves to the central nervous system, where it is interpreted as pain.

Mutations or dysregulation of these channels can lead to altered pain sensitivity. For example, gain-of-function mutations in $Na_V1.7$ are associated with conditions like erythromelalgia and paroxysmal extreme pain disorder, which are characterized by

episodes of severe pain[42,43]. Conversely, loss-of-function mutations in $Na_V1.7$ can result in congenital insensitivity to pain, where individuals do not feel pain at all[44]. Given their crucial role in pain signaling, $Na_V$ channels are potential targets for the development of analgesic drugs. Blocking these channels can prevent the initiation and propagation of action potentials in nociceptive neurons, thereby reducing pain sensation. Natural peptide toxins, such as those derived from spider and scorpion, have been found to bind to specific sites on the $Na_V1.7$ channel, inhibiting its function and effectively reducing pain sensation[45–47]. Research into the mechanism of action of these toxins has provided valuable insights for the development of new pain therapies[45–47]. Understanding how these natural toxins interact with $Na_V1.7$ channels can facilitate the development of more effective and selective $Na_V1.7$ blockers. These blockers have the potential to provide pain relief without the side effects associated with traditional pain medications, offering a promising avenue for the treatment of chronic pain conditions[48,49].

Certain $Na_V$ subtypes, particularly $Na_V1.1$, play a significant role in proprioception, the sensory ability to perceive the position and movement of the body[50]. $Na_V1.1$ channels are essential for the proper function of mammalian proprioceptors, which are sensory neurons that relay information about muscle stretch and tension. These channels significantly contribute to the intrinsic excitability of proprioceptors. Genetic deletion of $Na_V1.1$ in sensory neurons causes profound motor coordination deficits, akin to those seen in conditions with impaired proprioceptive signaling. This is because $Na_V1.1$ is crucial for transmitting the receptor potentials generated during sustained muscle stretch, highlighting its indispensable role in proprioceptive function and related motor behaviors[50].

**b. Voltage-Gated Potassium Channels**

Voltage-gated potassium channels ($K_V$ channels) are critical transmembrane proteins that facilitate the selective passage of potassium ions ($K^+$) across cell membranes. These channels are essential for maintaining the resting membrane potential, shaping action potentials, and regulating cellular excitability. They play a fundamental role in the generation and propagation of the action potential, helping to regulate the electrical activity of neurons and

other excitable cells[51]. Currently, there are 40 known human voltage-gated potassium channel alpha subunits[31].

$K_V$ channels are homo-tetramers, where each subunit includes six transmembrane segments (S1-S6)[52]. The S5 and S6 segments, along with the intervening loop (P-loop), form the pore-forming region of the channel, which contains the selectivity filter responsible for the high specificity for $K^+$ ions. The S1-S4 segments make up the voltage-sensing domain, with the S4 segment acting as the voltage sensor. This arrangement enables the channels to open or close in response to changes in membrane potential, providing a mechanism for their voltage-dependent gating[52]. $K_V$ channels exhibit a variety of gating mechanisms[53,54]. These include voltage-dependent opening and closing, which are crucial for repolarizing the membrane following an action potential, thereby terminating the electrical signal and preparing the cell for the next stimulus. Additionally, $K_V$ channels also demonstrate inactivation mechanisms[55–60]. Inactivation can be rapid or slow and typically involves a conformational change in the channel that temporarily reduces its conductivity even in the continued presence of a voltage stimulus[55–60]. This inactivation process is critical for shaping the action potential and controlling the firing frequency of neurons and cardiac cells, thereby fine-tuning cellular excitability and signal transmission.

$K_V$ channels play a crucial role in a myriad of physiological processes beyond their well-known function in regulating action potentials. For example, they are instrumental in hormone secretion, where they help in the release of various hormones from endocrine cells[61]. This secretion process often involves facilitating the repolarization phase of action potentials which reset the membrane potential after a depolarization, enabling cells to respond to further stimuli and regulate hormone release efficiently[61].

In the context of cell volume regulation, $K_V$ channels contribute significantly to the maintenance of cell volume under varying osmotic conditions. These channels participate in regulatory volume decrease, a process essential for cells to recover from swelling by expelling $K^+$ ions and restoring their original volume[62]. This mechanism is particularly vital in immune cells like T lymphocytes, where proper volume regulation is crucial for cell function and survival[62].

They are also deeply involved in modulating immune responses. They regulate the activity of immune cells by controlling the membrane potential and calcium signaling, which are critical for the activation and function of these cells. This regulation is crucial for the proper functioning of the immune system, influencing processes from lymphocyte activation to cytokine secretion and cytotoxic responses[63].

The hERG channel, or $K_V11.1$, is a critical component of cardiac cell physiology, contributing significantly to the cardiac action potential's repolarization phase[64]. Dysfunctions or mutations in hERG can lead to serious cardiac arrhythmias such as Long QT Syndrome[65]. Beyond the heart, hERG channels are also expressed in various brain regions, smooth muscle cells, and endocrine cells, indicating their broader physiological relevance[66]. Importantly, hERG channels are implicated in cancer, where they may influence cell proliferation and apoptosis, suggesting a potential target for anticancer therapy[67].

hERG is a notorious anti-target in drug development due to its association with serious cardiac side effects[66,68]. Blockade of the hERG channel can lead to QT prolongation, increasing the risk of potentially fatal arrhythmias such as *torsades de pointes*[69]. Many drugs have been withdrawn or faced restricted use due to their unintended interactions with hERG, which underscores the importance of screening for hERG activity early in the drug development process[69–71].

$K_V1.3$ channels, part of the Shaker family of potassium channels, play diverse roles across the immune system and are crucial in T lymphocyte activation and proliferation[72]. Inhibition of $K_V1.3$ has been explored as a therapeutic approach in treating autoimmune diseases due to its role in immune modulation[73]. Moreover, $K_V1.3$ channels are involved in various non-immune functions such as metabolism, insulin resistance, and sensory perception, underscoring their multifunctional nature[72].

**c. Voltage-Gated Calcium Channels**

Voltage-gated calcium channels ($Ca_V$ channels) are pivotal membrane proteins that mediate the influx of calcium ions ($Ca^{2+}$) into cells, serving as key regulators of various cellular

processes. These include muscle contraction, neurotransmitter release, gene expression, and cell proliferation[74]. The influx of $Ca^{2+}$ through these channels triggers numerous cellular responses, vital for physiological functions across different tissue types[74].

A typical $Ca_V$ channel comprises a central pore-forming α1-subunit and auxiliary β, α2δ, and sometimes γ subunits, enhancing the functional diversity and regulatory complexity of these channels[75]. The α1-subunit contains four homologous domains (I-IV), each consisting of six transmembrane segments (S1-S6), similar to the structure of $Na_V$ channels. The S4 segments function as voltage sensors, responding to changes in membrane potential to trigger channel opening. The pore region, formed by the S5 and S6 segments and the intervening P-loop, includes the selectivity filter which is critical for the high specificity for $Ca^{2+}$ ions, ensuring that these channels selectively conduct calcium ions over other types[75].

$Ca_V$ channels are activated by depolarization of the membrane potential, leading to the opening of the channel and an influx of $Ca^{2+}$ ions[74]. This calcium influx is crucial for various physiological processes such as muscle contraction in cardiac, smooth, and skeletal muscle cells, the release of neurotransmitters in neurons, and the activation of intracellular signaling pathways that control gene expression and cell growth[74]. The precise regulation of $Ca^{2+}$ influx through these channels is critical for maintaining cellular homeostasis and proper physiological function, with implications for a wide array of cellular and systemic physiological processes[74].

The $Ca_V1.2$ channel, also known as an L-type calcium channel, is a well-studied example that highlights the critical roles of $Ca_V$ channels in cellular processes[76]. $Ca_V1.2$ channels play a vital role in the cardiovascular system, particularly in the regulation of cardiac muscle contraction[28]. They facilitate the influx of $Ca^{2+}$ necessary for the excitation-contraction coupling in cardiac myocytes. This process is essential for the proper functioning of the heart's pumping action, as the entry of $Ca^{2+}$ through $Ca_V1.2$ channels triggers the contraction of heart muscles. Dysfunctions or mutations in $Ca_V1.2$ can lead to cardiac arrhythmias, underscoring their importance in maintaining cardiac health and rhythm[76].

$Ca_V1.2$ channels also regulate contraction in smooth muscle cells. They are involved in various smooth muscle functions such as vasoconstriction and gastrointestinal motility, which are vital for blood pressure regulation and digestive processes[76]. The proper functioning of $Ca_V1.2$ channels ensures that smooth muscles can respond appropriately to physiological stimuli[76].

### 1.4) Experimental Approaches in Protein Structure Determination

Protein structure determination is crucial for deciphering the function and biological roles of proteins. Understanding the three-dimensional structures of proteins is essential as it directly relates to their functions in biological processes. Throughout the decades, a variety of experimental techniques have been developed to reveal the complex three-dimensional structures of proteins, essential for getting insights into their function and molecular mechanisms of its regulation.

1. ***X-ray Crystallography:*** The historical journey of protein structure determination began predominantly with X-ray crystallography[77]. X-ray crystallography involves directing X-rays at a crystallized protein and analyzing the diffraction patterns that result[78]. These patterns provide detailed information about the atomic structure of the protein, allowing scientists to construct models of the protein at atomic resolution. The precision and detail offered by this technique have made it a fundamental tool in the field of structural biology, contributing to numerous scientific advances and medical discoveries. This method remains one of the most widespread and accurate techniques for elucidating protein structures. The foundations of protein crystallography were laid in the early 20$^{th}$ century, but it was not until the 1950s that the method achieved major milestones with the determination of the structures of hemoglobin and myoglobin[79–81]. These groundbreaking studies marked the beginning of structural biology and provided the first clear insights into how protein structures could explain their functions and interactions.
2. ***Nuclear Magnetic Resonance (NMR) Spectroscopy:*** NMR spectroscopy is a sophisticated technique used to determine protein structures in solution and in membrane environment as well, revealing not only their static shapes but also dynamic

changes. By exploiting the magnetic properties of atomic nuclei, NMR spectroscopy provides detailed insights into protein structures based on the specific frequencies at which these nuclei resonate in a magnetic field[82]. Since its adaptation for biological molecules in the 1980s[83], NMR spectroscopy has become essential for studying proteins, especially for understanding the structure and dynamics of transient, complex, or disordered protein states that other methods might miss. Techniques like inferential structure determination, which combines NMR data with statistical models, further enhance our ability to explore protein dynamics in depth[84].

3. ***Cryo-Electron Microscopy (Cryo-EM):*** Cryo-EM has transformed the landscape of structural biology, especially for studying biomolecules that are challenging to crystallize[85]. This technique involves rapidly freezing samples and then examining them under an electron microscope to prevent the formation of ice crystals that can damage the structure[85]. Unlike traditional methods that require crystallization, cryo-EM allows scientists to observe proteins and other biological complexes in a state close to their natural environment. The origins of cryo-EM trace back to the 1970s and 1980s when the foundational processes were developed[86,87]. However, it was the introduction of direct electron detectors and automated data collection techniques in the 2000s that significantly improved the resolution and usability of cryo-EM[88], leading to what is often called the resolution revolution[89]. These advancements enabled the detailed visualization of biomolecules at near-atomic resolutions, which was previously only possible with X-ray crystallography[88,89]. Cryo-EM is particularly advantageous for studying large protein complexes, membrane proteins, and viruses that are difficult or impossible to crystallize[90]. This method has been crucial for understanding the mechanisms of complex molecular machines and has had profound implications in drug design and other biotechnological applications.

The acquisition of detailed structural data through experimental techniques is invaluable across various scientific domains, significantly enhancing our understanding and manipulation of biological systems. The Protein Data Bank (PDB) serves as a crucial repository for these data, housing the three-dimensional structural information of proteins, nucleic acids, and complex assemblies obtained from experiments[91]. This freely accessible

database enables researchers worldwide to study, model, and simulate the molecular machinery of life.

The benefits of acquiring detailed structural data are manifold. In drug development, knowing the precise structure of target molecules allows for the rational design of drugs that can bind with high specificity and efficacy, potentially reducing side effects and improving therapeutic outcomes[92]. Similarly, in bioengineering and enzyme engineering, structural insights enable the redesign of enzymes to catalyze reactions more efficiently or under novel conditions, which is crucial for industrial applications such as biofuel production[93]. Additionally, understanding the structures of biological molecules aids in elucidating the mechanisms of diseases at the molecular level, facilitating the development of more effective treatments by targeting specific molecular pathways involved in disease processes[94].

### 1.5) Lipid Diversity and Functionality in the Cell Membrane

The cell membrane is a complex mosaic of lipids, each contributing to its structural integrity, dynamic behavior, and functional diversity. Indeed, in most animal cell membranes, lipid molecules make up 50% of the mass, while proteins account for most of the remaining half[1]. Each lipid molecule found in cell membranes exhibits an amphipathic (or amphiphilic) characteristic, having a hydrophilic (polar) region and a hydrophobic (nonpolar) region[1,6]. This dual characteristic enables the lipid molecules to self-assemble into membranes. This process is driven by entropy, as the hydrophobic parts of the molecules avoid water and cluster together, while the hydrophilic parts seek to interact with the surrounding water. This organization reduces the system's overall energy, effectively creating a barrier that defines the cell's boundaries and provides structure to the membrane[6]. The lipid composition of the membrane is heterogeneous, varying between different cell types and even within different regions of the same cell[6,95]. Understanding the variety and roles of these lipids is crucial for deciphering membrane physiology and its implications in health and disease. The major classes of membrane lipids are:

1. ***Phospholipids:*** These are the most abundant and thus considered the primary building blocks of the cell membrane[1,6]. Phospholipids are amphipathic molecules with a

hydrophilic phosphate head group and two hydrophobic fatty acid tails. The most common phospholipids include phosphatidylcholine (PC), phosphatidylethanolamine (PE), phosphatidylserine (PS), and phosphatidylinositol (PI). Each phospholipid contributes to membrane properties; for example, PC is the most abundant eukaryotic phospholipid and provides structural stability in addition to cell-signaling functions[6,96], while PS is negatively charged and involved in interactions with proteins and mediation of intracellular and extracellular signaling pathways[97]. 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) is a specific phospholipid often studied for its prevalence in biological membranes and its role in maintaining membrane fluidity[98]. Phosphatidylinositol 4,5-bisphosphate (PIP2) is an integral component of many cellular signaling pathways[99], serving as a substrate for phospholipase C (PLC), which catalyzes its hydrolysis to produce second messengers, inositol 1,4,5-trisphosphate (IP3), and diacylglycerol (DAG). These second messengers play essential roles in intracellular calcium release and the activation of protein kinase C, respectively[100,101].

2. ***Sterols:*** These lipids are critical components of the cell membrane, modulating its fluidity, thickness, and permeability[102]. As an example, cholesterol is crucial in regulating various biological processes[103], such as passive permeation[104], and protein and enzyme activity[105]. Cholesterol intercalates between phospholipids, reducing membrane fluidity by restricting the movement of fatty acid tails and increasing membrane rigidity[102]. This balance is essential for maintaining membrane integrity and optimizing the function of membrane proteins[102].
3. ***Glycolipids:*** These lipids, composed of a lipid moiety and one or more carbohydrate residues, are primarily found in the outer leaflet of the plasma membrane[107]. Glycolipids play crucial roles in cell recognition, adhesion, and communication, as well as in the immune response[107]. Gangliosides, a type of glycolipid, are involved in cell signaling and are markers for cellular differentiation[107]. These sialic acid-containing glycosphingolipids are primarily found in the nerve tissues but are also present in various other tissues, highlighting their importance across multiple biological functions and diseases[107]. Gangliosides facilitate essential cell-cell interactions and play regulatory roles in natural killer cell activity, myelin-axon interactions, and inflammation, often through their interactions with proteins like Siglecs and selectins[107].

4. ***Sphingolipids:*** This class of lipids includes sphingomyelin, ceramide, and glycosphingolipids[108–110]. Sphingolipids are essential for membrane structure and are involved in various signaling pathways[108–110]. Sphingomyelin, for example, is a major component of lipid rafts and is involved in signal transduction and apoptosis[109]. Ceramide, another significant sphingolipid, functions as a bioactive molecule that regulates cell death, differentiation, and proliferation[110]. Glycosphingolipids, including gangliosides, are essential for cellular recognition processes and are involved in the modulation of cell signaling, contributing to the structural diversity and functionality of cellular membranes[108].

The intricate interplay between different lipids in the cell membrane is essential for its diverse functions. The specific arrangement and interactions of these lipids influence membrane fluidity, curvature, and the formation of microdomains, which are critical for membrane dynamics and the regulation of cellular activities. Understanding the diversity and roles of membrane lipids is fundamental for exploring membrane biology and its implications in health and disease.

### 1.6) The Electrical Symphony: Cellular Currents and the Action Potential

The action potential is a fundamental mechanism in cellular physiology, representing a brief and rapid change in the membrane potential of excitable cells such as neurons and muscle cells. This electrical event underlies the transmission of nerve impulses and muscle contraction, orchestrating a complex interplay of ionic currents across the cell membrane, essential for communication and function in many bodily systems.

At the core of the action potential lies the delicate balance between two primary types of ionic currents: inward sodium currents ($I_{Na}$) and outward potassium currents ($I_K$). These currents are mediated by voltage-gated sodium ($Na_V$) and potassium ($K_V$) channels, which respond dynamically to changes in membrane potential.

The action potential is a masterful orchestration of ionic currents, a testament to the precision and complexity of cellular processes. Understanding this electrical symphony not

only sheds light on the fundamental mechanisms of cellular communication but also provides insights into the pathophysiology of various neurological and cardiac disorders, paving the way for targeted therapeutic interventions.

**a. Neuronal Action Potential**

The neuronal action potential embodies the quintessence of neural communication, facilitating the rapid and precise transmission of information within the nervous system[39]. This electrical impulse traverses the neuron's axon, carrying signals across vast networks of neural pathways[39].

The neuronal action potential is characterized by distinct phases[111–114]:

1. ***Resting Membrane Potential:*** The stage is set with the neuron at rest, harboring a membrane potential of approximately -70 to -60 mV. This negative potential is maintained chiefly by the selective permeability of the membrane to potassium ions ($K^+$) and the diligent efforts of the $Na^+/K^+$ ATPase pump, which extrudes three sodium ions ($Na^+$) for every two potassium ions brought into the cell, thereby sustaining the electrochemical gradient.
2. ***Depolarization:*** Triggered by a stimulus, this phase sees the dramatic opening of voltage-gated sodium channels ($Na_V$), which allows a swift influx of $Na^+$ into the neuron. This influx shifts the membrane potential in a more positive direction, marching towards the threshold required to fire an action potential.
3. ***Rising Phase:*** Once the threshold is crossed, $Na_V$ channels fully open, permitting a more substantial entry of $Na^+$ through a positive feedback loop mechanism. This rapid inflow propels the membrane potential upward, creating the rising phase of the action potential, a critical step in the neural signal's journey.
4. ***Overshoot:*** The culmination of the rising phase is the overshoot, where the membrane potential momentarily reverses its sign, achieving a positive peak between +30 to +40 mV. This reversal underscores the neuron's full commitment to transmitting the signal.
5. ***Repolarization:*** The narrative takes a turn as voltage-gated potassium channels ($K_V$) open, ushering $K^+$ out of the neuron. Concurrently, $Na_V$ channels start to inactivate,

stemming the $Na^+$ influx. These actions collectively pull the membrane potential back down towards its resting value.

6. ***Afterhyperpolarization:*** An overcorrection occurs, with the membrane potential dipping below the resting level due to the continued outward flow of $K^+$ and the delayed closure of $K_V$ channels. This hyperpolarization ensures a unidirectional flow of the action potential and contributes to the refractory period, during which the neuron is less susceptible to firing another action potential.
7. ***Return to Resting State:*** Gradually, the membrane potential is restored to its resting level, prepared for the initiation of a new action potential. This restoration is facilitated by the $Na^+/K^+$ ATPase pump as well as leak channels, which rectify the ion concentrations disrupted during the action potential.

**b. Cardiac Action Potential**

In stark contrast to neurons, cardiac cells orchestrate a longer and more complex action potential, essential for the rhythmic contraction of the heart. The cardiac action potential encompasses five phases, each contributing to the heart's ability to pump blood effectively[28,115,116]:

***Phase 0 (Rapid Depolarization):*** Initiated by the opening of voltage-gated sodium channels, this phase mirrors the depolarization in neurons but is pivotal for the synchronous contraction of cardiac muscle.

***Phase 1 (Initial Repolarization):*** A brief outward flux of $K^+$ follows, slightly lowering the membrane potential in a nod to the impending plateau phase.

***Phase 2 (Plateau):*** Characteristically unique to cardiac cells, this phase is mediated by the slow, prolonged opening of voltage-gated calcium channels ($Ca_V$), allowing $Ca^{2+}$ to enter the cell. This influx is counterbalanced by a slow efflux of $K^+$, maintaining a prolonged depolarized state crucial for cardiac muscle contraction.

***Phase 3 (Rapid Repolarization):*** The closure of $Ca_V$ channels and an enhanced outward flow of $K^+$ through $K_V$ channels bring about a swift repolarization, resetting the cardiac cell's membrane potential.

***Phase 4 (Resting Membrane Potential):*** The cardiac cell returns to its resting state, readying itself for the next action potential. This state is maintained by the $Na^+/K^+$ ATPase pump and leak channels, ensuring the heart's readiness for another cycle of contraction.

**c. Comparing and Contrasting Neuronal and Cardiac Action Potentials**

Neuronal and cardiac action potentials, while both essential for their respective systems, exhibit several key distinctions.

Neuronal action potentials are brief and sharp, designed to facilitate rapid and precise signal transmission across the nervous system[111–114]. These action potentials typically last only a few milliseconds and involve a rapid influx of sodium ions followed by a swift efflux of potassium ions, allowing neurons to quickly reset and fire at high frequencies[111]. This rapid signaling is crucial for processing complex information and responding to environmental stimuli rapidly[111].

In contrast, cardiac action potentials feature a longer duration (200-400 ms compared to 1-5 ms of action potentials in the nerves or muscles[117]) with a unique plateau phase, critical for the heart's pumping function[28,115,116]. This plateau phase, maintained mainly by calcium influx, ensures a prolonged contraction of the heart muscle, necessary for effective blood circulation[28]. The longer duration and extended refractory period of cardiac action potentials prevent premature re-excitation of the heart cells, ensuring coordinated and sequential contractions essential for maintaining a steady heartbeat[28].

When it comes to the ionic contributions, both types of action potentials rely on the flow of sodium ($Na^+$) and potassium ($K^+$) ions. However, calcium ($Ca^{2+}$) ions play a more prominent role in cardiac action potentials, contributing significantly to the plateau phase and the excitation-contraction coupling in heart muscle cells[118]. The refractory period also differs between the two; neurons have a shorter refractory period, allowing for rapid, successive firing of action potentials[111], whereas cardiac cells have an extended refractory period that matches with the longer duration of the action potential[28]. This extended refractory period

is crucial for preventing the possibility of tetanic contractions and ensuring the orderly progression of the heartbeat[28,115,116].

Another distinction lies in the regenerative mechanism; neuronal action potentials are typically initiated by synaptic inputs or sensory stimuli[28], while cardiac action potentials are often initiated by pacemaker cells, such as those in the sinoatrial node[47], which possess the intrinsic ability to generate rhythmic depolarizations. Furthermore, the propagation of these action potentials differs; neuronal action potentials propagate along axons to transmit information between different regions of the nervous system, whereas cardiac action potentials propagate through the interconnected network of cardiac muscle cells, ensuring synchronized contraction of the heart chambers[28,111–116].

In summary, despite sharing some fundamental mechanisms, neuronal and cardiac action potentials exhibit distinct characteristics and roles that reflect the specialized functions of the nervous and cardiovascular systems. Understanding these differences is crucial for appreciating the intricate workings of these vital physiological processes.

**d. Patch-Clamp Technique for Studying Ionic Currents**

The patch-clamp technique, developed by Erwin Neher and Bert Sakmann in the late 1970s, revolutionized the field of electrophysiology by enabling high-resolution recordings of membrane currents from single ion channels and whole cells. This method is essential for understanding the electrophysiological properties of neurons and cardiac cells, particularly in measuring action potentials and the underlying ionic currents. The significant impact of this technique was recognized with the awarding of the Nobel Prize in Physiology or Medicine to Neher and Sakmann in 1991 for their groundbreaking work in this area[119].

Patch-clamping involves using a glass micropipette with a polished tip to create a tight seal (giga ohm seal or gigaseal) creating a resistance in the 10-100 giga-ohms range with the cell membrane. This technique enables the isolation and precise recording of electrical currents across the cell membrane, providing high sensitivity and minimal noise interference. The micropipette is filled with an electrolyte solution, and a silver-silver chloride (Ag-AgCl)

electrode is inserted to transmit the electrical signals to an amplifier, facilitating detailed electrophysiological studies[119–121].

There are several different configurations of the patch-clamp setup, each designed to achieve specific experimental goals[122,123]. The primary configurations are:

1. ***Cell-Attached Mode:*** The pipette forms a seal with a small patch of the cell membrane, allowing for the recording of currents from individual ion channels within that patch.
2. ***Whole-Cell Mode:*** After forming the cell-attached configuration, suction is applied to rupture the patch of membrane under the pipette, establishing electrical access to the entire cell. This allows for the recording of whole-cell currents and action potentials.
3. ***Inside-Out and Outside-Out Modes:*** These configurations involve the excision of a patch of membrane with the pipette. In the inside-out mode, the intracellular side of the membrane is exposed to the bath solution, while in the outside-out mode, the extracellular side faces the bath. These configurations are useful for studying the effects of intracellular or extracellular factors such as different voltage protocols, ionic bath compositions and drug molecules on ion channel activity.

The patch-clamp technique has been instrumental in advancing our understanding of the electrical properties of cells and the function of ion channels under both normal and pathological conditions[121]. Its precision and versatility make it a cornerstone method in electrophysiological research[119].

**e. Measuring Action Potentials with Patch-Clamping**

Patch-clamping, especially in whole-cell mode, is excellently suited to measure the membrane potential in excitable cells. The procedure includes[124,125]:

1. ***Preparation:*** Cells are typically cultured or freshly isolated and placed in a recording chamber filled with an appropriate physiological solution. The glass pipette is filled with an electrolyte solution that mimics the intracellular environment.

2. ***Gigaseal Formation:*** The pipette is carefully positioned against the cell membrane using a micromanipulator. Gentle suction is applied to form a tight seal between the pipette and the membrane.
3. ***Membrane Rupture:*** A brief suction or an electrical pulse ruptures the membrane patch, allowing the pipette to access the cell's interior. This configuration now allows the measurement of the cell's overall ionic currents.
4. ***Recording:*** Using a patch clamp amplifier, the membrane potential is held at a set value (voltage clamp) or allowed to vary naturally (current clamp). In current clamp mode, the amplifier records changes in membrane potential as the cell generates action potentials in response to stimuli.
5. ***Data Analysis:*** The recorded action potentials are analyzed to determine key parameters such as amplitude, duration, and frequency. This information provides valuable insights into the function of ion channels, the effects of pharmacological agents, and the mechanisms underlying various physiological and pathological conditions.

## 1.7) Benchtop to Desktop: The Next Frontier in Human Electrophysiology

In the opening chapter of this dissertation, we have explored the complex architecture and dynamic roles of the cell membrane, highlighting its critical function as the orchestrator of membrane transport and electrical signaling. We examined the detailed structure of the lipid bilayer, the diverse functionalities of embedded proteins, and the intricate dance of ions and molecules across this biological barrier. This foundation underscores the membrane's pivotal role in regulating cellular life processes, from ion exchange to receptor signaling, all of which are fundamental to understanding human physiology.

As we move forward, we will pivot to a discussion of computational approaches that complement the experimental insights gained in the first chapter. These computational models are invaluable in bridging the gap between theoretical knowledge and practical application, especially in the complex field of human electrophysiology. This is particularly crucial in the study of ion channels and transport mechanisms, where computational

simulations can predict how changes in environment, structure, or intermolecular interactions affect cellular functionality.

The interplay between computational and experimental methodologies exemplifies the cutting-edge trends of modern biomedical science, where innovation in one area sparks advancements in another. Through this interdisciplinary approach, we strive to unlock deeper insights that could redefine the future of medicine and human well-being.

## 1.8) References


1. Alberts, B. *et al. Molecular biology of the cell*. vol. 3 (Garland New York, 1994).
2. Coskun, U. & Simons, K. Cell membranes: the lipid perspective. *Struct. Lond. Engl. 1993* **19**, 1543–1548 (2011).
3. Gould, S. B. Membranes and evolution. *Curr. Biol. CB* **28**, R381–R385 (2018).
4. Simons, K. Cell membranes: A subjective perspective. *Biochim. Biophys. Acta* **1858**, 2569–2572 (2016).
5. Harrison, R. & Lunt, G. G. Glycoproteins, Glycolipids and Cellular Recognition. in *Biological Membranes: Their Structure and Function* (eds. Harrison, R. & Lunt, G. G.) 130–165 (Springer Netherlands, 1980). doi:10.1007/978-94-011-6857-1_7.
6. van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**, 112–124 (2008).
7. Boes, D. M., Godoy-Hernandez, A. & McMillan, D. G. G. Peripheral Membrane Proteins: Promising Therapeutic Targets across Domains of Life. *Membranes* **11**, 346 (2021).
8. Sargiacomo, M., Lisanti, M., Graeve, L., Le Bivic, A. & Rodriguez-Boulan, E. Integral and peripheral protein composition of the apical and basolateral membrane domains in MDCK cells. *J. Membr. Biol.* **107**, 277–286 (1989).
9. Almeida, J. G., Preto, A. J., Koukos, P. I., Bonvin, A. M. J. J. & Moreira, I. S. Membrane proteins structures: A review on computational modeling tools. *Biochim. Biophys. Acta Biomembr.* **1859**, 2021–2039 (2017).
10. Hedin, L. E., Illergård, K. & Elofsson, A. An introduction to membrane proteins. *J. Proteome Res.* **10**, 3324–3331 (2011).
11. Westerfield, J. M. & Barrera, F. N. Membrane receptor activation mechanisms and transmembrane peptide tools to elucidate them. *J. Biol. Chem.* **295**, 1792–1814 (2020).

12. Cooper, G. M. Structure of the Plasma Membrane. in *The Cell: A Molecular Approach. 2nd edition* (Sinauer Associates, 2000).

13. Levental, I. & Lyman, E. Regulation of membrane protein structure and function by their paralipidomes. *Nat. Rev. Mol. Cell Biol.* **24**, 107–122 (2023).

14. Liszewski, K. Dissecting the Structure of Membrane Proteins. *Genet. Eng. Biotechnol. News* **35**, 1, 14, 16–17 (2015).

15. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E. L. L. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes1. *J. Mol. Biol.* **305**, 567–580 (2001).

16. Whitelegge, J. P. Integral Membrane Proteins and Bilayer Proteomics. *Anal. Chem.* **85**, 2558–2568 (2013).

17. Voet, D. & Voet, J. G. *Biochemistry*. (John Wiley & Sons, 2010).

18. Scott, J. L., Musselman, C. A., Adu-Gyamfi, E., Kutateladze, T. G. & Stahelin, R. V. Emerging methodologies to investigate lipid–protein interactions. *Integr. Biol.* **4**, 247–258 (2012).

19. Di Resta, C. & Becchetti, A. Introduction to Ion Channels. in *Integrins and Ion Channels: Molecular Complexes and Signaling* (eds. Becchetti, A. & Arcangeli, A.) 9–21 (Springer, 2010). doi:10.1007/978-1-4419-6066-5_2.

20. Roux, B. Ion channels and ion selectivity. *Essays Biochem.* **61**, 201–209 (2017).

21. Nishida, N. *et al.* Functional dynamics of cell surface membrane proteins. *J. Magn. Reson.* **241**, 86–96 (2014).

22. Hochstadt, J. & Quinlan, D. The function and activity of certain membrane enzymes when localized on- and off- the membrane. *J. Cell. Physiol.* **89**, 839–852 (1976).

23. Haier, J. & Nicolson, G. L. Cell biology and clinical implications of adhesion molecules in colorectal diseases: Colorectal cancers, infections and inflammatory bowel diseases. *Clin. Exp. Metastasis* **18**, 623–638 (2000).

24. Hille, B. *Ion Channels of Excitable Membranes*. (Oxford University Press, 2001).

25. Colquhoun, D. & Lape, R. Allosteric coupling in ligand-gated ion channels. *J. Gen. Physiol.* **140**, 599–612 (2012).

26. Haswell, E. S., Phillips, R. & Rees, D. C. Mechanosensitive channels: what can they do and how do they do it? *Struct. Lond. Engl. 1993* **19**, 1356–1369 (2011).

27. Chen, Y. & Song, K. The Role of Thermosensitive Ion Channels in Mammalian Thermoregulation. *Adv. Exp. Med. Biol.* **1349**, 355–370 (2021).

28. Jespersen, T. The Cardiac Ion Channels. in *Cardiac Arrhythmias - New Considerations* (IntechOpen, 2012). doi:10.5772/26530.

29. Bortner, C. D. & Cidlowski, J. A. Ion channels and apoptosis in cancer. *Philos. Trans. R. Soc. B Biol. Sci.* **369**, 20130104 (2014).

30. Kaczorowski, G. J., McManus, O. B., Priest, B. T. & Garcia, M. L. Ion Channels as Drug Targets: The Next GPCRs. *J. Gen. Physiol.* **131**, 399–405 (2008).

31. Wulff, H., Castle, N. A. & Pardo, L. A. Voltage-gated potassium channels as therapeutic targets. *Nat. Rev. Drug Discov.* **8**, 982–1001 (2009).

32. Ahern, C. A., Payandeh, J., Bosmans, F. & Chanda, B. The hitchhiker's guide to the voltage-gated sodium channel galaxy. *J. Gen. Physiol.* **147**, 1–24 (2016).

33. Laedermann, C. J., Abriel, H. & Decosterd, I. Post-translational modifications of voltage-gated sodium channels in chronic pain syndromes. *Front. Pharmacol.* **6**, (2015).

34. Laedermann, C. J., Decosterd, I. & Abriel, H. Ubiquitylation of voltage-gated sodium channels. *Handb. Exp. Pharmacol.* **221**, 231–250 (2014).

35. Chanda, B. & Bezanilla, F. Tracking voltage-dependent conformational changes in skeletal muscle sodium channel during activation. *J. Gen. Physiol.* **120**, 629–645 (2002).

36. de Lera Ruiz, M. & Kraus, R. L. Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. *J. Med. Chem.* **58**, 7093–7118 (2015).

37. International Union of Pharmacology. XLVII. Nomenclature and Structure-Function Relationships of Voltage-Gated Sodium Channels | Pharmacological Reviews. https://pharmrev.aspetjournals.org/content/57/4/397.

38. Li, Z. *et al.* Structural Basis for Pore Blockade of the Human Cardiac Sodium Channel Nav1.5 by the Antiarrhythmic Drug Quinidine**. *Angew. Chem. Int. Ed.* **60**, 11474–11480 (2021).

39. Song, W. & Shou, W. Cardiac Sodium Channel Nav1.5 Mutations and Cardiac Arrhythmia. *Pediatr. Cardiol.* **33**, 943–949 (2012).

40. Nguyen, P. T., DeMarco, K. R., Vorobyov, I., Clancy, C. E. & Yarov-Yarovoy, V. Structural basis for antiarrhythmic drug interactions with the human cardiac sodium channel. *Proc. Natl. Acad. Sci.* **116**, 2945–2954 (2019).

41. Bennett, D. L., Clark, A. J., Huang, J., Waxman, S. G. & Dib-Hajj, S. D. The Role of Voltage-Gated Sodium Channels in Pain Signaling. *Physiol. Rev.* **99**, 1079–1151 (2019).

42. Estacion, M. *et al.* NaV1.7 Gain-of-Function Mutations as a Continuum: A1632E Displays Physiological Changes Associated with Erythromelalgia and Paroxysmal Extreme Pain Disorder Mutations and Produces Symptoms of Both Disorders. *J. Neurosci.* **28**, 11079–11088 (2008).

43. Theile, J. W., Jarecki, B. W., Piekarz, A. D. & Cummins, T. R. Nav1.7 mutations associated with paroxysmal extreme pain disorder, but not erythromelalgia, enhance Navβ4 peptide-mediated resurgent sodium currents. *J. Physiol.* **589**, 597–608 (2011).

44. Goldberg, Y. P. *et al.* Loss-of-function mutations in the Nav1.7 gene underlie congenital indifference to pain in multiple human populations. *Clin. Genet.* **71**, 311–319 (2007).

45. Ngo, K. *et al.* Elucidating molecular mechanisms of protoxin-II state-specific binding to the human NaV1.7 channel. *J. Gen. Physiol.* **156**, e202313368 (2024).

46. Xu, H. *et al.* Structural Basis of Nav1.7 Inhibition by a Gating-Modifier Spider Toxin. *Cell* **176**, 702-715.e14 (2019).

47. Clairfeuille, T. *et al.* Structural basis of α-scorpion toxin action on Nav channels. *Science* **363**, eaav8573 (2019).

48. Nguyen, P. T. *et al.* Computational design of peptides to target NaV1.7 channel with high potency and selectivity for the treatment of pain. *eLife* **11**, e81727 (2022).

49. Wulff, H., Christophersen, P., Colussi, P., Chandy, K. G. & Yarov-Yarovoy, V. Antibodies and venom peptides: new modalities for ion channels. *Nat. Rev. Drug Discov.* **18**, 339–357 (2019).

50. Espino, C. M. *et al.* NaV1.1 is essential for proprioceptive signaling and motor behaviors. *eLife* **11**, e79917 (2022).

51. Kim, D. M. & Nimigean, C. M. Voltage-Gated Potassium Channels: A Structural Examination of Selectivity and Gating. *Cold Spring Harb. Perspect. Biol.* **8**, a029231 (2016).

52. Jensen, M. Ø. *et al.* Mechanism of Voltage Gating in Potassium Channels. *Science* **336**, 229–233 (2012).

53. Treptow, W., Marrink, S.-J. & Tarek, M. Gating Motions in Voltage-Gated Potassium Channels Revealed by Coarse-Grained Molecular Dynamics Simulations. *J. Phys. Chem. B* **112**, 3277–3282 (2008).

54. Costa, F., Guardiani, C. & Giacomello, A. Molecular dynamics simulations suggest possible activation and deactivation pathways in the hERG channel. *Commun. Biol.* **5**, 1–11 (2022).

55. Cuello, L. G., Jogini, V., Cortes, D. M. & Perozo, E. Structural mechanism of C-type inactivation in K+ channels. *Nature* **466**, 203–208 (2010).

56. Li, J., Shen, R., Reddy, B., Perozo, E. & Roux, B. Mechanism of C-type inactivation in the hERG potassium channel. *Sci. Adv.* **7**, eabd6203 (2021).

57. Miranda, W. E. *et al.* Selectivity filter modalities and rapid inactivation of the hERG1 channel. *Proc. Natl. Acad. Sci.* **117**, 2795–2804 (2020).

58. Reddi, R., Matulef, K., Riederer, E. A., Whorton, M. R. & Valiyaveetil, F. I. Structural basis for C-type inactivation in a Shaker family voltage-gated K+ channel. *Sci. Adv.* **8**, eabm8804 (2022).

59. Tan, X.-F. *et al.* Structure of the Shaker Kv channel and mechanism of slow C-type inactivation. *Sci. Adv.* **8**, eabm7814.

60. Chandy, K. G., Sanches, K. & Norton, R. S. Structure of the voltage-gated potassium channel KV1.3: Insights into the inactivated conformation and binding to therapeutic leads. *Channels* **17**, 2253104.

61. Abbott, G. W. Biology of the KCNQ1 Potassium Channel. *New J. Sci.* **2014**, e237431 (2014).

62. Bobak, N. *et al.* Volume regulation of murine T lymphocytes relies on voltage-dependent and two-pore domain potassium channels. *Biochim. Biophys. Acta BBA - Biomembr.* **1808**, 2036–2044 (2011).

63. Bachmann, M. *et al.* Voltage-Gated Potassium Channels as Regulators of Cell Death. *Front. Cell Dev. Biol.* **8**, (2020).

64. Butler, A., Helliwell, M. V., Zhang, Y., Hancox, J. C. & Dempsey, C. E. An Update on the Structure of hERG. *Front. Pharmacol.* **10**, (2020).

65. Havakuk, O. & Viskin, S. A Tale of 2 Diseases: The History of Long-QT Syndrome and Brugada Syndrome. *J. Am. Coll. Cardiol.* **67**, 100–108 (2016).

66. Vandenberg, J. I. *et al.* hERG K(+) channels: structure, function, and clinical significance. *Physiol. Rev.* **92**, 1393–1478 (2012).

67. Jehle, J., Schweizer, P. A., Katus, H. A. & Thomas, D. Novel roles for hERG K+ channels in cell proliferation and apoptosis. *Cell Death Dis.* **2**, e193–e193 (2011).

68. Raschi, E., Vasina, V., Poluzzi, E. & De Ponti, F. The hERG K+ channel: target and antitarget strategies in drug development. *Pharmacol. Res.* **57**, 181–195 (2008).

69. Hancox, J. C., McPate, M. J., El Harchi, A. & Zhang, Y. H. The hERG potassium channel and hERG screening for drug-induced torsades de pointes. *Pharmacol. Ther.* **119**, 118–132 (2008).

70. Witchel, H. J. Drug-induced hERG Block and Long QT Syndrome. *Cardiovasc. Ther.* **29**, 251–259 (2011).

71. Kamiya, K., Niwa, R., Morishima, M., Honjo, H. & Sanguinetti, M. C. Molecular Determinants of hERG Channel Block by Terfenadine and Cisapride. *J. Pharmacol. Sci.* **108**, 301–307 (2008).

72. Colley, B. S., Biju, K. C., Visegrady, A., Campbell, S. & Fadool, D. A. Neurotrophin B receptor kinase increases Kv subfamily member 1.3 (Kv1.3) ion channel half-life and surface expression. *Neuroscience* **144**, 531–546 (2007).

73. Chandy, G., Strong, M., Aiyar, J. & Gutman, G. A. Structural and Biochemical Features of the Kv1.3 Potassium Channel: An Aid to Guided Drug Design. *Cell. Physiol. Biochem.* **7**, 135–147 (2008).

74. Catterall, W. A. Voltage-gated calcium channels. *Cold Spring Harb. Perspect. Biol.* **3**, a003947 (2011).

75. Wu, J. *et al.* Structure of the voltage-gated calcium channel $Ca_v1.1$ complex. *Science* **350**, aad2395 (2015).

76. Hofmann, F., Flockerzi, V., Kahl, S. & Wegener, J. W. L-Type CaV1.2 Calcium Channels: From In Vitro Findings to In Vivo Function. *Physiol. Rev.* **94**, 303–326 (2014).

77. Holmes, K. C. & Blow, D. M. The use of x-ray diffraction in the study of protein and nucleic acid structure. *Methods Biochem. Anal.* **13**, 113–239 (1965).

78. Ringe, D. & Petsko, G. A. Study of protein dynamics by X-ray diffraction. *Methods Enzymol.* **131**, 389–433 (1986).

79. Perutz, M. F. *et al.* Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* **185**, 416–422 (1960).

80. Kendrew, J. C. *et al.* A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **181**, 662–666 (1958).

81. Kendrew, J. C. *et al.* Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature* **185**, 422–427 (1960).

82. Marion, D. An Introduction to Biological NMR Spectroscopy. *Mol. Cell. Proteomics MCP* **12**, 3006–3025 (2013).

83. Williamson, M. P., Havel, T. F. & Wüthrich, K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by 1H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **182**, 295–315 (1985).

84. Brookes, D. H. & Head-Gordon, T. Experimental Inferential Structure Determination of Ensembles for Intrinsically Disordered Proteins. *J. Am. Chem. Soc.* **138**, 4530–4538 (2016).

85. Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).

86. Dubochet, J. & McDowall, A. w. Vitrification of Pure Water for Electron Microscopy. *J. Microsc.* **124**, 3–4 (1981).

87. Newmark, P. Cryo-transmission microscopy: Fading hopes. *Nature* **299**, 386–387 (1982).

88. Kuster, D. J., Liu, C., Fang, Z., Ponder, J. W. & Marshall, G. R. High-Resolution Crystal Structures of Protein Helices Reconciled with Three-Centered Hydrogen Bonds and Multipole Electrostatics. *PLoS ONE* **10**, e0123146 (2015).

89. Kühlbrandt, W. The Resolution Revolution. *Science* **343**, 1443–1444 (2014).

90. Piper, S. J., Johnson, R. M., Wootten, D. & Sexton, P. M. Membranes under the Magnetic Lens: A Dive into the Diverse World of Membrane Protein Structures Using Cryo-EM. *Chem. Rev.* **122**, 13989–14017 (2022).

91. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

92. Lin, X., Li, X. & Lin, X. A Review on Applications of Computational Methods in Drug Screening and Design. *Molecules* **25**, 1375 (2020).

93. Ngo, K., Bruno da Silva, F., Leite, V. B. P., Contessoto, V. G. & Onuchic, J. N. Improving the Thermostability of Xylanase A from Bacillus subtilis by Combining Bioinformatics and Electrostatic Interactions Optimization. *J. Phys. Chem. B* **125**, 4359–4367 (2021).

94. Zimmermann, M. T. Molecular Modeling is an Enabling Approach to Complement and Enhance Channelopathy Research. *Compr. Physiol.* **12**, 3141–3166 (2022).

95. Casares, D., Escribá, P. V. & Rosselló, C. A. Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. *Int. J. Mol. Sci.* **20**, 2167 (2019).

96. Tang, W., Keesler, G. A. & Tabas, I. The Structure of the Gene for Murine CTP:Phosphocholine Cytidylyltransferase, Ctpct. *J. Biol. Chem.* **272**, 13146–13151 (1997).

97. Kay, J. & Grinstein, S. Phosphatidylserine-mediated cellular signaling. *Adv. Exp. Med. Biol.* **991**, 177–93 (2013).

98. Yepuri, N. *et al.* Synthesis of Perdeuterated 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine ([D82 ]POPC) and Characterisation of Its Lipid Bilayer Membrane Structure by Neutron Reflectometry. *ChemPlusChem* **81 3**, 315–321 (2016).

99. Bulley, S. J., Clarke, J. H., Droubi, A., Giudici, M.-L. & Irvine, R. F. Exploring phosphatidylinositol 5-phosphate 4-kinase function. *Adv. Biol. Regul.* **57**, 193–202 (2015).

100. Lee, S. B. & Rhee, S. G. Significance of PIP2 hydrolysis and regulation of phospholipase C isozymes. *Curr. Opin. Cell Biol.* **7**, 183–189 (1995).

101. Cockcroft, S. & Raghu, P. Topological organisation of the phosphatidylinositol 4,5-bisphosphate–phospholipase C resynthesis cycle: PITPs bridge the ER–PM gap. *Biochem. J.* **473**, 4289–4310 (2016).

102. Chakraborty, S. *et al.* How cholesterol stiffens unsaturated lipid membranes. *Proc. Natl. Acad. Sci.* **117**, 21896–21905 (2020).

103. Maxfield, F. R. & van Meer, G. Cholesterol, the central lipid of mammalian cells. *Curr. Opin. Cell Biol.* **22**, 422–429 (2010).

104. Corvera, E., Mouritsen, O. G., Singer, M. A. & Zuckermann, M. J. The permeability and the effect of acyl-chain length for phospholipid bilayers containing cholesterol: theory and experiment. *Biochim. Biophys. Acta* **1107**, 261–270 (1992).

105. de Meyer, F. J.-M., Rodgers, J. M., Willems, T. F. & Smit, B. Molecular simulation of the effect of cholesterol on lipid-mediated protein-protein interactions. *Biophys. J.* **99**, 3629–3638 (2010).

106. Deng, Y., Almsherqi, Z. A., Ng, M. M. L. & Kohlwein, S. D. Do viruses subvert cholesterol homeostasis to induce host cubic membranes? *Trends Cell Biol.* **20**, 371–379 (2010).

107. Ikami, T., Ishida, H. & Kiso, M. [46] Synthesis and biological activity of glycolipids, with a focus on gangliosides and sulfatide analogs. in *Methods in Enzymology* vol. 311 547–568 (Academic Press, 2000).

108. Taniguchi, M. & Okazaki, T. The role of sphingomyelin and sphingomyelin synthases in cell death, proliferation and migration—from cell and animal models to human disorders. *Biochim. Biophys. Acta BBA - Mol. Cell Biol. Lipids* **1841**, 692–703 (2014).

109. Chakraborty, M. & Jiang, X.-C. Sphingomyelin and Its Role in Cellular Signaling. in *Lipid-mediated Protein Signaling* (ed. Capelluto, D. G. S.) 1–14 (Springer Netherlands, 2013). doi:10.1007/978-94-007-6331-9_1.

110. Iessi, E. *et al.* Chapter Five - On the role of sphingolipids in cell survival and death. in *International Review of Cell and Molecular Biology* (eds. Spetz, J. K. E. & Galluzzi, L.) vol. 351 149–195 (Academic Press, 2020).

111. Wareham, A. C. Action potential: generation and propagation. *Anaesth. Intensive Care Med.* **6**, 200–203 (2005).

112. Chen, I. & Lui, F. Neuroanatomy, Neuron Action Potential. in *StatPearls* (StatPearls Publishing, 2024).

113. Fink, C. Using Phase Response Curves to Understand Neuronal Synchronization and Sleep. (2012).

114. Yalçınkaya-Selimbeyoglu, B. H., Yilmaz, B. & Ozilgen, M. *Thermodynamic Analyses of Information Dissipation along the Passive Dendritic Trees and Active Action Potential*. (2017).

115. Jha, K. & Jha, K. Biomedical Signal Analysis and Its Physio-Clinical Perspectives. *https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-5225-2829-6.ch001* https://www.igi-global.com/gateway/chapter/www.igi-global.com/gateway/chapter/188147 (1AD).

116. Bastiaenen, R. & Behr, E. R. Sudden death and ion channel disease: pathophysiology and implications for management. *Heart* **97**, 1365–1372 (2011).

117. Ackerman, M. J. & Clapham, D. E. chapter 18 - Excitability and Conduction. in *Molecular Basis of Cardiovascular Disease (Second Edition)* (ed. CHIEN, K. R.) 311–335 (W.B. Saunders, 2004). doi:10.1016/B978-0-7216-9428-3.50023-8.

118. Morad, M. & Tung, L. Ionic events responsible for the cardiac resting and action potential. *Am. J. Cardiol.* **49**, 584–594 (1982).

119. Carmeliet, E. From Bernstein's rheotome to Neher-Sakmann's patch electrode. The action potential. *Physiol. Rep.* **7**, e13861 (2019).

120. Elzenga, J. T. M. Patch Clamp Techniques for Plant Cells. in *Plant Electrophysiology: Methods and Cell Electrophysiology* (ed. Volkov, A. G.) 225–243 (Springer, 2012). doi:10.1007/978-3-642-29119-7_10.

121. A Practical Guide to Patch Clamping | SpringerLink. https://link.springer.com/chapter/10.1007/978-1-4419-1229-9_1.

122. Franciolini, F. Patch clamp technique and biophysical study of membrane channels. *Experientia* **42**, 589–594 (1986).

123. Kodirov, S. A. Whole-cell patch-clamp recording and parameters. *Biophys. Rev.* **15**, 257–288 (2023).

124. Perkins, K. L. Cell-attached voltage-clamp and current-clamp recording and stimulation techniques in brain slices. *J. Neurosci. Methods* **154**, 1 (2006).

125. Souza, A. C. *et al.* Chapter 2 - Recording Day and Night: Advice for New Investigators in the Sleep and Memory Field. in *Handbook of Behavioral Neuroscience* (ed. Manahan-Vaughan, D.) vol. 28 43–62 (Elsevier, 2018).

# CHAPTER 2:

# From Micro to Macro: Computational Techniques for Multiscale Physiological Studies

*"Nothing in life is to be feared, it is only to be understood.*
*Now is the time to understand more, so that we may fear less."* – Marie Curie (1867 – 1934)

---

## 2.1) Overview of Computational Modeling and Simulation in Physiology

### a. The Need for Computational Models

Computational models are invaluable in physiology, providing platforms for formulating hypotheses about how biological systems operate under different conditions and interventions. These models offer virtual environments for the predictive testing and refinement of hypotheses, enhancing our comprehension of intricate biological processes in a reproducible and controlled setting. Developed based on well-understood biological mechanisms, these models undergo rigorous validation against empirical data, confirming their accuracy and reliability. Such validations are essential, reinforcing the models' ability to mimic and predict real-world biological behaviors accurately[1].

Moreover, scientists can use computational models to explore scenarios that might be difficult or unethical to test *in vivo*, such as the impact of various drug dosages on human tissues[2] or the progression of diseases under different conditions[3,4]. By simulating these conditions, they can gather preliminary data that can guide the design of future experimental studies, ultimately leading to more efficient research efforts.

As technology advances, human knowledge also expands. The iterative nature of computational modeling allows for continuous improvement of the models as new data becomes available[5]. This adaptability ensures that the models remain relevant and accurate over time. The integration of computational models with experimental data not only

enhances the robustness of the research findings but also accelerates the pace of scientific advancements by providing deeper insights into the complexities of biological systems.

**b. Types of Models**

1. ***Whole-body:*** These models simulate the interactions of all physiological systems within an entire organism, providing a comprehensive view of overall health and disease states. Such models are essential for understanding how various systems of the body interact and influence each other, which is crucial in diagnosing systemic diseases and in the development of treatments that might affect multiple body systems. Whole-body models often integrate data from various organs to simulate how changes in one affect the others, and they are especially useful in the study of drug's pharmacokinetics and pharmacodynamics across the human body[6].
2. ***Organs and Organ Systems:*** These models focus on a specific organ, detailing its functionality and pathophysiology under various conditions. By concentrating on one organ, these models allow researchers and clinicians to delve deeply into the specific diseases affecting that organ, such as heart disease in cardiac models[7]. Organ-specific models are critical in developing targeted therapies and are widely used in the development of medical devices and surgical interventions that require precise understanding of an organ's response to treatments.
3. ***Tissue:*** Tissue models simulate the behavior of specific types of tissues, such as epithelial, connective, or muscle tissues. These models are particularly valuable in studying diseases that primarily affect specific tissue types, such as fibrosis or cancer. Tissue models help in understanding the progression of these diseases and the impact of potential treatments at the tissue level. They provide insights into cellular interactions within the tissue and how these interactions contribute to the overall function or dysfunction of the tissue[8].
4. ***Cell:*** Cell models focus on the activities within individual cells, such as action potential generation[9], metabolic pathways[10], genetic regulation[11], or signaling pathways[12]. These models are fundamental in molecular biology and genetic engineering as they provide insights into how changes at the cellular level affect cell behavior and disease

progression. Cell models are essential for drug development, particularly for understanding how drugs interact with cellular components and influence cell function[9].

5. ***Subcellular:*** These models delve into the subcellular level, focusing on proteins and other macromolecules as well as specific subcellular signaling pathways. They are used to study the structural properties of proteins, their interactions, and their role in cellular functions. Subcellular models are crucial for understanding diseases at the molecular level and for developing targeted therapies that interact with specific proteins or molecular pathways. They often employ techniques from molecular dynamics[13,14] and quantum chemistry[15] to simulate the behavior of molecules in high detail, which is critical for drug design and understanding enzyme catalysis.

**c. Classifications of Simulations**

1. ***Deterministic:*** These simulations employ predefined equations and parameters to predict specific outcomes under controlled conditions, allowing for precise predictions and analyses. Deterministic models are particularly useful in situations where variables are well understood and can be accurately measured. They excel in scenarios where the biological systems are large enough that random fluctuations average out and do not significantly affect the system's behavior. A notable example is molecular dynamics simulation, which is extensively used for studying the dynamic structural behavior of biomolecules like proteins and nucleic acids. Molecular dynamics simulations are particularly effective in exploring the motion and interactions at the atomic level over time, such as examining the folding process of proteins[14] or the interactions between ligands and their target protein sites, providing precise insights into molecular mechanisms[13,16]. Additionally, deterministic models are crucial in cardiac cell functional modeling, particularly for simulating the precise interactions within cardiac cells during heart rhythm propagation and potential arrhythmia mechanisms[9].
2. ***Stochastic:*** Incorporating elements of randomness and probability, stochastic simulations offer a more accurate reflection of the inherent variability found in biological processes. Stochastic models are particularly suited for examining biological systems where the discrete nature and the small number of molecules can lead to

significant variations in the behavior of the system. For instance, Brownian dynamics simulations, a subset of stochastic methods, are utilized to calculate reaction rates in biochemical pathways, thereby offering insights into cellular processes[17,18].

3. ***Agent-based**:* These models simulate the actions and interactions of individual autonomous agents within a system, such as cells or molecules, to study the collective effects of individual behaviors on the system as a whole. This approach is particularly valuable for understanding emergent behaviors and the dynamics of complex systems, where the global system behavior results from the interactions at a micro-level. As an example, agent-based modeling is instrumental in exploring the dynamics of tumor microenvironments and their response to therapeutic interventions, providing a deeper understanding of cancer progression and treatment efficacy[19].

## 2.2) Biomolecular Structural Modeling

The evolution of protein structural modeling reflects a significant shift from early empirical methods[20,21], which relied heavily on direct insights and physical experimentation, to advanced computational approaches powered by artificial intelligence[22–24]. This transformation has been pivotal in enabling researchers to unravel the complexities of biological systems, accurately predict protein behaviors, and accelerate the pace of drug discovery. Particularly, the integration of machine learning technologies marks a profound advancement, enhancing the precision, accuracy, and efficiency of modeling techniques and broadening the scope of their application in modern biosciences.

The journey began in the mid-20th century with the first protein structures determined by X-ray crystallography. The pioneering work by Max Perutz and John Kendrew on hemoglobin and myoglobin, respectively, laid the groundwork for structural biology[25–27]. By the 1970s, as more protein structures were solved, researchers could predict protein structures based on amino acid sequence homology. This method, known as homology modeling, relies on the assumption that protein structures are more conserved through evolution than their amino acid sequences. Tools such as MODELLER, developed in the 1990s, facilitated these predictions by constructing models based on known template structures[28].

By the early 2000s, the field saw further advancements with the development of the Rosetta platform – a comprehensive software suite for protein structure prediction and design[20,21,29]. It uses both experimental data and theoretical models to predict the three-dimensional structures of proteins[30], design new proteins[21], model protein-protein interactions[13,31], and simulate protein-ligand binding[32]. Rosetta has been pivotal in various biomedical applications, such as vaccine design where it was used to predict the structure of surface proteins of viruses to create more effective vaccine candidates[33]. It has also been instrumental in enzyme design, where researchers have used it to engineer enzymes with new functionalities, such as the creation of enzymes capable of degrading plastic, thereby addressing environmental issues[34]. Additionally, Rosetta has contributed to understanding the molecular basis of genetic disorders by modeling the effects of mutations on protein structural stability[35]. I-TASSER (Iterative Threading ASSEmbly Refinement) is another powerful tool for protein structure and function predictions[36]. It works by constructing a protein model through iterative template-based fragment assembly, followed by atomic-level structure refinement.

The most recent revolution in protein structural modeling has been driven by deep learning, epitomized by DeepMind's AlphaFold[23]. AlphaFold uses a deep learning architecture based on the attention mechanism that allows it to interpret the spatial relationships between amino acid residues in a protein[23]. AlphaFold's approach involves predicting the distances and angles between pairs of residues, which it uses to construct a highly accurate 3D model of the protein structure[23]. AlphaFold's capabilities were highlighted in the Critical Assessment of protein Structure Prediction (CASP) competition, where it outperformed other methods by a significant margin[37,38]. The accuracy of AlphaFold's predictions is often comparable to that achieved with experimental methods, such as X-ray crystallography and cryo-electron microscopy, making it a powerful tool for researchers[37,38]. The system's ability to predict structures within days, a process that might otherwise take years in the lab, marks a monumental shift in structural biology. The tool has profound implications for biomedical research, such as expediting the process of drug discovery by providing detailed structural data that can help in the design of drugs that precisely target disease-related proteins[39].

ESMFold, developed by researchers at Meta AI, extends the capabilities of machine learning in protein structure prediction by incorporating evolutionary scale modeling[24]. This method utilizes large language models that are traditionally used in natural language processing to understand the evolutionary relationships between different protein sequences. By training on extensive databases of known protein structures, ESMFold can predict the structure of related proteins even if they are not closely related to any sequence with a known structure[24]. ESMFold emphasizes leveraging evolutionary data, which helps in modeling proteins that have few known analogues, potentially filling gaps left by other methods like AlphaFold[24].

RoseTTAFold, developed by the Baker lab at the University of Washington, shares similarities with AlphaFold in its use of deep learning but introduces a novel three-track neural network architecture that processes information about the sequence, structure, and the relations between sequence and structure[22]. This architecture allows RoseTTAFold to efficiently predict protein structures as well as interactions between proteins and RNA or between multiple protein subunits. The technique strikes a balance between computational efficiency and accuracy, making it particularly useful for researchers with limited computational resources. RoseTTAFold has been applied to various problems beyond mere structure prediction, including complex modeling scenarios involving multiple interacting chains[40].

The primary differences among these approaches employing artificial intelligence lie in their architectural distinctions and specific applications. AlphaFold uses a deep learning model based on transformers and excels in accuracy, making it suitable for creating detailed models of protein structures[23]. ESMFold applies evolutionary data through large language models, which can enhance predictions for proteins with fewer known relatives[24]. In contrast, RoseTTAFold's flexible architecture allows it to handle a broader range of biological problems, including multi-chain interactions, not just single protein folding[22].

Today, the integration of artificial intelligence (AI) and deep learning into structural modeling and molecular docking continues to evolve, promising not only faster and more accurate predictions but also the capability to handle larger and more complex molecular

systems. This evolution from empirical and knowledge-based methods to computational and AI-driven techniques highlights the dynamic and rapidly advancing nature of the field, continually expanding the boundaries of what can be achieved in computational biology and molecular medicine. As these tools evolve, they will undoubtedly unveil new dimensions of biological complexity, opening up further opportunities for scientific discovery and therapeutic invention.

## 2.3) Molecular Docking of Drugs and Peptides

Molecular docking is a computational technique that models the interaction between two or more molecules, typically between a small molecule (like a drug or peptide) with a larger molecule (such as a protein), to predict the orientation and affinity of the molecules when they are bound to each other. This method is crucial for understanding molecular mechanisms of action and for facilitating the design of new therapeutic agents with higher efficacy and specificity.

Molecular docking involves two primary components: the docking algorithm and the scoring function. The docking algorithm predicts the possible binding modes of the drug molecule to its target protein, while the scoring function evaluates the stability and strength of the interaction. More realistic docking considers the flexibility of both the ligand and the target molecule[41], although accounting for protein flexibility remains computationally challenging and is often simplified.

The history of molecular docking, a vital computational technique in drug discovery, dates back to the early 1980s when the first algorithms were developed to predict how small molecules, such as drugs, interact with proteins[42]. This method emerged from the need to understand and predict the interactions at the molecular level without relying solely on labor-intensive experimental methods. This pioneering software, named DOCK, was developed by Tack Kuntz and colleagues at the University of California, San Francisco in 1982[42]. It was aimed to simulate the docking process of small molecules into the binding sites of target proteins, akin to fitting a key into a lock[42].

Over the years, molecular docking has evolved with advancements in computational power and algorithms, including the development of more sophisticated scoring functions and the incorporation of protein and ligand flexibility to improve accuracy[32,43–46]. As an example, RosettaDock is a commonly used and powerful tool for modeling protein-protein interactions today[43–45]. It refines the interface between proteins by optimizing the side-chain conformations and slight adjustments to the backbone, allowing a high degree of accuracy in predicting protein complexes[43]. RosettaDock is particularly useful in the design of protein interfaces and engineering novel protein interactions[43–45]. For example, it was used in our recent study to investigate the interaction between the tarantula peptide Protoxin-II and the $Na_V1.7$ sodium channel, important in pain perception[13] (see Chapter 3 for more details). The tool provided detailed insights into the binding interactions between the toxin and the channel, highlighting crucial interaction points for drug design. This research supports the development of Protoxin-II analogs and offers a foundation for creating targeted pain management therapies[13]. Moreover, RosettaDock has been instrumental in designing antibodies that can selectively bind to and neutralize the influenza virus, providing a basis for potential flu vaccines[47].

As another example, GALigandDock is a molecular docking software that uses a genetic algorithm to optimize the docking simulations[32]. Genetic algorithms are search algorithms based on the mechanics of natural selection and genetics. They are particularly useful in molecular docking for exploring large search spaces and finding the best fit or highest affinity of a ligand to a protein's active site. GALigandDock iteratively improves the fit of the drug molecule to the binding site by mutations and crossover in genetic algorithm reflecting natural evolution[32]. This method is effective in cases where the ligand flexibility needs to be considered, such as in docking large and complex molecules.

Commonly used in the pharmaceutical industry, GLIDE and GOLD are widely recognized for their efficacy in virtual screening and docking accuracy[48–50]. GLIDE, for example, is praised for its ability to consistently yield high enrichment factors (i.e., the concentration of the active ligands) among the top-scoring docking hits, making it an excellent choice for virtual screening applications[48,51]. It successfully identifies active compounds, showing a notable

advantage in handling various protein targets and ligand sets[52]. Similarly, GOLD has demonstrated strong performance across multiple assessments, especially noted for its superior sampling power and the ability to accurately reproduce known inhibitor conformations in the docking process[49,50]. These tools have become industry standards due to their robust algorithms and consistent results in complex docking scenarios.

In conclusion, the advancements in molecular docking have undeniably established it as an indispensable method within the pharmaceutical industry. Its critical role in the development and optimization of new therapeutic agents continues to influence groundbreaking research and innovation in drug discovery, particularly in the context of both small molecule drugs and peptide-based therapies. As this technique evolves, it promises to further enhance the accuracy and efficacy of therapeutic interventions, marking its significance not only as a tool for today but also as a foundation for future pharmaceutical advancements.

### 2.4) Molecular Dynamics Simulations of Biomolecules

Molecular dynamics (MD) simulations are advanced computational techniques designed to study the physical movements of atoms and molecules within various systems. These simulations are instrumental for scientists as they enable detailed observation and analysis of how particles interact and evolve over time. The process involves numerically solving Newton's equations of motion for each particle, providing a dynamic view of molecular behavior and interactions.

The development of MD simulations can be traced back to the mid-20$^{th}$ century when scientists first utilized computers to study the dynamics of hard spheres[53]. Initially, these simulations were rudimentary, limited to modeling only a handful of particles due to the computational constraints of the time. However, the field of MD simulations has undergone significant transformations due to technical advancements in computing power[54]. By the 1970s, researchers were able to perform simulations of proteins and other complex biological molecules, marking a significant milestone in the field[55]. Today's simulations leverage enhanced computational power and sophisticated algorithms, allowing for the

intricate and detailed modeling of complex biomolecular systems. This evolution has expanded the applicability of MD simulations, making them a crucial tool in areas such as drug discovery and materials science, where they help predict molecular behaviors and interactions with high accuracy[56,57].

A variety of software packages are available for conducting MD simulations, each with its own unique features and capabilities. Prominent among these are GROMACS[58], AMBER[59], and NAMD[60], which are widely used in the scientific community for a range of applications from biomolecular studies to materials science. These programs are designed to facilitate detailed and scalable simulations, providing researchers with the tools needed to model complex systems at the molecular level.

The predictive power of MD simulations largely depends on the use of accurate forcefields. Forcefields are mathematical models that describe the forces and potential energies between atoms in a molecule. They include sets of equations and empirical parameters that describe various types of interactions —such as bonded interactions (e.g., covalent bonds) and non-bonded interactions (e.g., electrostatic and van der Waals).

The selection of an appropriate forcefield is critical because it directly affects the simulation's ability to replicate real-world molecular dynamics. Different forcefields may be optimized for specific types of molecules or conditions, thus affecting the results of the simulations[61]. For instance, the AMBER ff14SB is often preferred for exploring protein dynamics[62], while CHARMM27 can be used for the investigations of nucleic acid interactions[63]. The development and refinement of these forcefields are ongoing processes, as they are continually updated and validated against experimental data to improve their accuracy and applicability[61,64].

MD simulations are extensively employed in the study of biomolecules, including proteins and nucleic acids[57]. By simulating the atomic-level interactions within these complex molecular systems, MD simulations provide invaluable insights into their structures, dynamics, and functions. This detailed understanding is crucial for unraveling the mechanisms underlying various biological processes and phenomena.

In pharmaceutical research, MD simulations play a pivotal role in drug discovery and development. They are used to model and predict the interactions between drugs/peptides and their molecular targets, aiding in the design of molecules with optimal binding properties[13]. This computational approach not only speeds up the drug development process but also enhances its accuracy by allowing scientists to foresee how modifications in drug structure might impact efficacy and safety.

## 2.5) Advanced Techniques in Molecular Dynamics Simulations

MD simulations employ several advanced techniques to enhance sampling efficiency and navigate the complex energy landscapes of molecular systems. These techniques are crucial for studying dynamic processes that occur over extended timescales or involve significant free energy barriers.

1. ***Coarse Graining:*** Coarse-grained models in MD simulations reduce the number of degrees of freedom by simplifying the molecular systems, thus extending simulation timescales and enabling the study of larger systems efficiently[65]. Coarse-grained models simplify molecules into fewer interaction groups compared to all-atom representations, such as combining all atoms in an amino acid residue as a single particle, which allows the simulation to cover longer timescales due to fewer interactions being calculated[65]. Coarse graining serves as a bridge between detailed molecular behavior and larger-scale biological functions, making it possible to explore complex systems that were previously intractable with traditional simulation methods. For example, Maisuradze *et al.* used coarse-grained MD simulations to explore the folding dynamics of proteins[66]. Their work demonstrates that coarse-grained models can effectively capture essential aspects of protein folding kinetics and thermodynamics, providing insights into folding pathways and intermediate states[66].
2. ***Umbrella Sampling:*** Umbrella sampling is a powerful technique used in MD simulations to calculate free energy profiles along a reaction coordinate by performing multiple simulations at various values of the coordinate using harmonic restraints[67]. This method is particularly useful for exploring the free energy

landscapes of biomolecular systems, where it helps understand various biochemical processes by providing insights into the energy barriers and stable states along a reaction pathway. For example, umbrella sampling has been employed in studies such as those by Kästner to calculate the free-energy differences across biomolecular reactions[67]. This method involves applying a series of biased potentials, typically harmonic, to drive the system through different states along a predefined reaction coordinate. Each state is explored within a window set by reference coordinates and restraints, and the data collected from these windows are then integrated using techniques like weighted histogram analysis method to construct the overall free energy profile across the reaction coordinate[68].

3. ***Metadynamics:*** Metadynamics is an advanced simulation technique that employs a time-dependent, history-dependent bias potential to effectively navigate free energy landscapes, making it particularly suitable for complex molecular systems where local minima can trap the system[69]. This method functions by adding a Gaussian potential at each step of the simulation, effectively filling these minima and allowing the system to explore other states that might otherwise be inaccessible due to high energy barriers. The adaptive nature of the bias potential in metadynamics aids in exploring these landscapes efficiently by continuously modifying the potential in response to the state of the system.
4. ***Replica Exchange MD:*** Replica exchange MD is a sophisticated simulation method that improves sampling efficiency by running multiple replicas of a system concurrently at various temperatures or modifications of the energy function (Hamiltonian)[70]. This technique enables the replicas to exchange configurations among themselves, which helps in overcoming energy barriers and facilitates a more thorough exploration of the energy landscape. Replica exchange MD is especially beneficial because it allows systems trapped in local minima to escape and explore other areas of the energy landscape that might be energetically favorable[70]. The configuration exchanges between replicas ensure that each replica can traverse a wider range of energy states than would be possible in a single, isolated simulation, leading to more robust and reliable sampling.

## 2.6) Computational Models of Cellular Electrophysiology

Moving from the molecular intricacies of proteins and ions to the cellular level, the field of electrophysiology utilizes mathematical modeling to bridge our understanding of how individual cellular components contribute to larger, system-wide functions. In electrophysiology, these models transform our insights from molecular dynamics into a broader context, focusing on the electrical activities within cells. By applying principles from biophysics and electrical circuit theory, we treat the cell membrane as an electrical circuit, where ion channels act as essential components. This methodological shift allows us to apply mathematical laws to simulate and predict the dynamics of action potentials, providing a comprehensive view that links molecular interactions to cellular behavior. Although this approach simplifies complex biological processes, it effectively exposes the fundamental patterns and governing principles of action potentials, enhancing our understanding of cellular function in health and disease.

### a. Modeling the Cell Membrane as an Electrical Circuit

The cell membrane can be modeled as a parallel electrical circuit composed of resistors and capacitors. The resistors represent the ion channels, each with a specific conductance for different ions (e.g., $Na^+$, $K^+$, and $Ca^{2+}$), while the capacitor represents the lipid bilayer's capacity to store electrical charge.

Ohm's Law states that the current (I) through a conductor is directly proportional to the voltage difference (V) across it and inversely proportional to its resistance (R):

$$I = \frac{V}{R}$$

In the context of ion flow, the current for a specific ion ($I_{\text{ion}}$) can be expressed as:

$$I_{\text{ion}} = g_{\text{ion}} \cdot (V_m - E_{\text{ion}})$$

where $g_{\text{ion}}$ is the conductance of the ion channel, $V_m$ is the membrane potential, and $E_{\text{ion}}$ is the equilibrium potential for that ion (also known as the reversal or Nernst potential).

The Nernst potential describes the voltage across a cell's membrane necessary to balance the net diffusion of a particular ion X into and out of the cell. It can be calculated using the Nernst equation:

$$E_X = \frac{RT}{zF} \ln\left(\frac{[X]_{out}}{[X]_{in}}\right)$$

Where $E_X$ is the Nernst potential, $R$ is the universal gas constant, $T$ is the temperature in Kelvin, $z$ is the valance (net charge) of the ion, $F$ is the Faraday's constant, and $[X]_{out}$ and $[X]_{in}$ are the concentrations of the ion outside and inside the cell, respectively.

According to Kirchhoff's Current Law, the total current entering a junction equals the total current leaving it. For the total ionic current ($I_{\text{ionic}}$) across the membrane, this implies:

$$I_{\text{ionic}} = I_{\text{Na}} + I_{\text{K}} + I_{\text{Ca}} + I_{\text{leak}}$$

where $I_{\text{Na}}$, $I_{\text{K}}$, and $I_{\text{Ca}}$ are the currents for sodium, potassium, and calcium ions, respectively, and $I_{\text{leak}}$ represents leak currents through other ion channels.

The capacitive current ($I_C$) is associated with the charging and discharging of the membrane capacitor. It is related to the rate of change of the membrane potential:

$$I_C = C_m \frac{dV_m}{dt}$$

where $C_m$ is the membrane capacitance and $\frac{dV_m}{dt}$ is the rate of change of the membrane potential. The total membrane current is the sum of the ionic and capacitive currents:

$$I_{\text{total}} = I_{\text{ionic}} + I_C$$

**b. The Hodgkin-Huxley Model: First Mathematical Model of Neuronal Action Potential**

The Hodgkin-Huxley model, developed by Alan Hodgkin and Andrew Huxley in 1952, is a seminal mathematical framework used to describe the electrical characteristics of excitable

cells, such as neurons and muscle cells[71]. This model provides a quantitative description of the action potential, the rapid rise and fall in membrane potential that occurs in response to a stimulus. To do so, it represents the cell membrane as an electrical circuit with ion channels modeled as voltage-dependent conductances.

The model is constructed from a set of interrelated equations that collectively describe the dynamic behavior of the action potential in excitable cells. These equations can be categorized into different components, each representing a specific aspect of the cellular electrophysiology:

### *i. Membrane Potential*

The membrane potential equation in the Hodgkin-Huxley model describes how the membrane potential ($V$) changes over time ($t$). It is based on the principle of conservation of electrical charge, taking into account the capacitive current ($I_C$) and the ionic currents ($I_{\mathrm{Na}}$, $I_{\mathrm{K}}$, and $I_{\mathrm{L}}$):

$$\frac{dV}{dt} = \frac{1}{C_m}(I_{\mathrm{ext}} - I_{\mathrm{Na}} - I_{\mathrm{K}} - I_{\mathrm{L}})$$

where:

- $V$ is the membrane potential.
- $C_m$ is the membrane capacitance, which represents the ability of the membrane to store charge.
- $I_{\mathrm{ext}}$ is the externally applied stimulus current, which can be used to stimulate the cell.
- $I_{\mathrm{Na}}$, $I_{\mathrm{K}}$, and $I_{\mathrm{L}}$ are the sodium, potassium, and leak currents, respectively, which represent the flow of ions through their respective channels.

### *ii. Gating Variables*

The gating variable equations describe the time-dependent changes in the probabilities of ion channels being in certain states. These probabilities are represented by the gating

variables $n$, $m$, and $h$. The variable $n$ corresponds to the activation of potassium channels, while $m$ and $h$ correspond to the activation and inactivation of sodium channels, respectively:

$$\frac{dn}{dt} = \alpha_n(V) \cdot (1 - n) - \beta_n(V) \cdot n$$

$$\frac{dm}{dt} = \alpha_m(V) \cdot (1 - m) - \beta_m(V) \cdot m$$

$$\frac{dh}{dt} = \alpha_h(V) \cdot (1 - h) - \beta_h(V) \cdot h$$

where:

- $n, m,$ and $h$ represent the proportion of channels in the open state or available to open.
- $\alpha$ and $\beta$ are voltage-dependent rate constants that determine the transitions between open and closed or inactivated states of the channels.

Each of the rate constant for the gating variables has an associated voltage-dependent function that describes how quickly the ion channels respond to changes in membrane potential. The functions for these constants are derived from experimental data[71], typically involving voltage-clamp experiments where the membrane potential is clamped at different voltages and the resulting currents are measured (see Chapter 1), and are determined as follows:

$$\alpha_n(V) = \frac{0.01(10 - V)}{\exp\left(\frac{10 - V}{10}\right) - 1}$$

$$\beta_n(V) = 0.125 \exp\left(-\frac{V}{80}\right)$$

$$\alpha_m(V) = \frac{0.1(25 - V)}{\exp\left(\frac{25 - V}{10}\right) - 1}$$

$$\beta_m(V) = 4\exp\left(-\frac{V}{18}\right)$$

$$\alpha_h(V) = 0.07\exp\left(-\frac{V}{20}\right)$$

$$\beta_h(V) = \frac{1}{\exp\left(\frac{30-V}{10}\right)+1}$$

### *iii. Ionic Currents*

The ionic currents in the Hodgkin-Huxley model are defined as:

- Sodium Current ($I_{\text{Na}}$): $I_{\text{Na}} = \bar{g}_{\text{Na}} \cdot m^3 \cdot h \cdot (V - E_{\text{Na}})$
- Potassium Current ($I_{\text{K}}$): $I_{\text{K}} = \bar{g}_{\text{K}} \cdot n^4 \cdot (V - E_{\text{K}})$
- Leak Current ($I_{\text{L}}$): $I_{\text{L}} = \bar{g}_{\text{L}} \cdot (V - E_{\text{L}})$

where:

- $\bar{g}_{\text{Na}}$, $\bar{g}_{\text{K}}$, and $\bar{g}_{\text{L}}$ are the maximum conductances for sodium (1.2 mS/cm$^2$), potassium (0.36 mS/cm$^2$), and leak currents (0.003 mS/cm$^2$), respectively[72], where S is siemens or 1/ohm.
- $E_{\text{Na}}$, $E_{\text{K}}$, and $E_{\text{L}}$ are the reversal (Nernst) potentials for sodium (55.17 mV), potassium (-72.14 mV), and leak currents (-49.42 mV), respectively[72].

The Hodgkin-Huxley model stands as a cornerstone in the field of neuroscience and electrophysiology, providing a comprehensive mathematical framework for understanding the biophysical underpinnings of the action potential. Since its inception in the early 1950s, the model has not only deepened our grasp of neuronal and muscular excitability but has also paved the way for a myriad of advancements in both theoretical and applied science.

## c. Solving the Hodgkin-Huxley Model with Numerical Integration: The Forward Euler Method

Numerical integration is a crucial technique used to solve the differential equations in the Hodgkin-Huxley model, as analytical solutions might not be effective due to the model's complexity. Methods such as the forward Euler method, Runge-Kutta methods, and the predictor-corrector method are commonly employed[72]. These methods provide approximate solutions to the equations by discretizing time and iteratively calculating the membrane potential and gating variables. The choice of integration method and step size can impact the accuracy and computational efficiency of the simulations.

The forward Euler method is one of the simplest and most commonly used numerical methods for integrating ordinary differential equations (ODEs)[73]. It is a first-order numerical procedure for solving ODEs with a given initial value and is based on the idea of using the slope of the tangent line at a known point to estimate the value of the function at a nearby point. For a general ODE:

$$\frac{dy}{dt} = f(t, y)$$

the Euler method approximates the solution at a subsequent time step $(t + \Delta t)$ as:

$$y(t + \Delta t) = y(t) + \Delta t \cdot f(t, y(t))$$

where:

- $y(t)$ is the value of the function at time $t$.
- $\Delta t$ is the time step.
- $f(t, y(t))$ is the rate of change of the function at time $t$.

In the context of the Hodgkin-Huxley model, the Euler method can be used to integrate the membrane potential equation and the gating variable equations over time. For example, the membrane potential equation:

$$\frac{dV}{dt} = \frac{1}{C_m}(I_{\text{ext}} - I_{\text{Na}} - I_{\text{K}} - I_{\text{L}})$$

can be discretized using the Euler method as:

$$V(t + \Delta t) = V(t) + \Delta t \cdot \frac{1}{C_m}(I_{\text{ext}} - I_{\text{Na}} - I_{\text{K}} - I_{\text{L}})$$

Similarly, the gating variable equations can be discretized as:

$$n(t + \Delta t) = n(t) + \Delta t \cdot (\alpha_n(V) \cdot (1 - n) - \beta_n(V) \cdot n)$$

$$m(t + \Delta t) = m(t) + \Delta t \cdot (\alpha_m(V) \cdot (1 - m) - \beta_m(V) \cdot m)$$

$$h(t + \Delta t) = h(t) + \Delta t \cdot (\alpha_h(V) \cdot (1 - h) - \beta_h(V) \cdot h)$$

While the Euler method is straightforward and easy to implement, it has some limitations. It is a first-order method, meaning its accuracy is proportional to the step size ($\Delta t$). Smaller time steps increase accuracy but also computational cost. For systems with rapid changes, such as action potentials, a small time step is necessary to capture the dynamics accurately (e.g., 0.001 ms). Higher-order methods, such as the Runge-Kutta methods, can provide better accuracy for a given time step but are more complex to implement[72].

**d. Integration and Uses of Markov Models in Computational Electrophysiology**

Markov models have emerged as a powerful tool in computational electrophysiology, enabling the detailed simulation of state-specific drug interactions and the modulation of ion channel gating. These models are particularly useful in safety pharmacology, where understanding the effects of drugs on ion channels is crucial for assessing their potential cardiotoxicity[2,74].

A Markov model represents an ion channel as a system of discrete states, including different conformational states such as open, closed, and inactivated. Transitions between these states are governed by rate constants, which can be modulated by voltage, ligands, or drugs. This allows for the simulation of complex gating kinetics and the effects of state-dependent drug interactions on ion channel behavior.

A simple example of a Markov model can be illustrated using a two-state system for an ion channel, which can exist in either an open ($O$) or closed ($C$) state. The transitions between these states are governed by rate constants: the opening rate ($\alpha$) and the closing rate ($\beta$).

The states and transitions can be represented as:

$$C \underset{\beta}{\overset{\alpha}{\leftrightarrow}} O$$

The behavior of this system can be described by a set of differential equations that represent the time-dependent probabilities of the ion channel being in each state. Let $P_C(t)$ be the probability of the ion channel being in the closed state at time $t$, and $P_O(t)$ be the probability of the ion channel being in the open state at time $t$. The sum of these probabilities must be equal to 1 at all times, as the ion channel must be in one of these two states:

$$P_C(t) + P_O(t) = 1$$

The rate of change of the probabilities can be described by the following differential equations:

$$\frac{dP_C}{dt} = \beta P_O - \alpha P_C$$

$$\frac{dP_O}{dt} = \alpha P_C - \beta P_O$$

These equations state that the rate of change of the probability of being in the closed state ($dP_C/dt$) is equal to the rate of channels closing ($\beta P_O$) minus the rate of channels opening ($\alpha P_C$), and vice versa for the open state ($dP_O/dt$).

The steady-state probabilities of the open and closed states can be found by setting the derivatives to zero and solving for $P_O$ and $P_C$:

$$0 = \beta P_O - \alpha P_C$$

$$0 = \alpha P_C - \beta P_O$$

Adding these equations gives $\beta P_O + \alpha P_C = \alpha P_C + \beta P_O$, which simplifies to $P_C + P_O = 1$, as expected. Solving for $P_O$ and $P_C$ in terms of the transition rate constants $\alpha$ and $\beta$ gives:

$$P_O = \frac{\alpha}{\alpha + \beta}$$

$$P_C = \frac{\beta}{\alpha + \beta}$$

These equations provide the steady-state probabilities of the channel being in the open and closed states, respectively, based on the opening and closing rates. This simple Markov model can be extended to include more states and transitions to capture the complexity of real ion channel behavior[2,74].

Drugs often interact with ion channels in a state-dependent manner, binding preferentially to certain conformational states[2,74,75]. Markov models can simulate these interactions by incorporating drug-bound states and modifying the transition rates between states based on the presence of the drug. This enables the prediction of how drugs affect the probability of channels being in the open, closed, or inactivated states, and consequently, their impact on the action potential.

Markov models are integral to a multi-scale approach to safety pharmacology[2], which spans from the atomic level to the tissue level and beyond:

1. ***Atomic Level:*** At the atomic scale, molecular dynamics simulations and structural studies inform the construction of Markov models by providing insights into the conformational changes of ion channels and their interaction sites for drugs and thus may provide identity of distinct states in the Markov model and transitions between them. Also, molecular dynamics simulations can estimate rates of conformational transitions between ion channel states as well as drug binding and unbinding (so called “on” and “off” rates), which can be used as Markov model parameters.
2. ***Channel Level:*** Markov models simulate the gating kinetics of ion channels at the channel function level, taking into account the state-specific binding of drugs and

their effects on channel transitions using experimental and molecular dynamics simulation data as model parameters.

3. ***Cellular Level:*** The ion channel Markov models including those describing effects of drugs on ion channels are integrated into models of the action potential, such as the Hodgkin-Huxley neuron[71] or Luo-Rudy cardiac myocyte[76] models, to predict how these ion channels and their drug modulation alter cellular excitability and electrical signaling.
4. ***Tissue Level:*** At the tissue level, Markov models incorporating the effects of specific ion channels and their interacting drugs on cellular action potentials are incorporated into models of cardiac or neural tissues to simulate their impact on the propagation of electrical waves and the emergence of arrhythmias or other dysfunctions. These models can be further integrated into the whole-organ (e.g., heart) or organ-system (e.g., cardiovascular system) models[77,78].

This multi-scale approach allows for a comprehensive assessment of drug cardiac safety, from the molecular mechanisms of drug-channel interactions to their potential proarrhythmic effects in cardiac tissue[2]. By bridging the gap between different scales of biological organization, Markov models and computational safety pharmacology provide a powerful framework for predicting the safety and efficacy of drugs in a preclinical setting.

## 2.7) Applications of Computational Electrophysiology: Past, Present, and Future

Computational electrophysiology has emerged as a pivotal tool in the study of action potentials, providing deep insights into the intricate electrical behavior of excitable cells. These models have a wide range of applications, from unraveling the fundamental mechanisms of physiology to contributing to the development of medical treatments.

A key application of computational modeling lies in unraveling the biophysical processes that drive the generation and propagation of action potentials. The Hodgkin-Huxley model, for example, has shed light on the role of voltage-gated ion channels in modulating the swift fluctuations in the membrane potential that define action potentials. By adjusting model

parameters, scientists can explore how factors like ion channel mutations or shifts in extracellular ion concentrations influence the characteristics of action potentials.

In the pharmaceutical realm, computational models play a crucial role in forecasting the impact of drugs on the excitability of cardiac and neuronal cells. Models that incorporate the Markov model of ion channels are particularly adept at simulating how drugs selectively bind to different states of ion channels, influencing the duration and shape of action potentials. This capability is crucial for evaluating the cardiotoxic potential of drugs, as prolonged cardiac action potentials can trigger arrhythmias. The O'Hara-Rudy ventricular action potential model[9], for instance, is employed to assess the risk of drug-induced arrhythmias in a virtual environment[2], aiding in the creation of safer pharmaceuticals.

Furthermore, computational modeling is instrumental in investigating the electrophysiological changes associated with various diseases. By tailoring models to replicate the specific conditions seen in disorders such as epilepsy, cardiac arrhythmias, or ion channelopathies, researchers can gain a better understanding of their pathophysiology and develop targeted treatments. Additionally, models tailored to individual patients' data can predict the effectiveness of various therapeutic approaches, heralding a new era of personalized medicine.

In summary, computational electrophysiology stands at the forefront of scientific discovery, offering a powerful lens through which to view the complexities of cellular electrical activity. As we continue to refine these models and integrate them with emerging technologies, we can expect even greater advancements in our understanding of biological systems and the development of innovative medical solutions.

The journey of computational electrophysiology is a testament to human curiosity and ingenuity. From the early days of manual calculations and rudimentary experiments, where pioneering scientists like Hodgkin and Huxley meticulously charted the course of action potentials, to the modern era of powerful computer processors and simulations, where intricate models come to life on formidable computers, we have witnessed a remarkable

evolution. This journey reflects our relentless pursuit of understanding the intricate electrical language of life.

As we continue to unravel the mysteries of the electric universe within us, we stand on the shoulders of giants, armed with advanced computational tools and a deepening appreciation for the delicate interplay of ions, membranes, and molecules that orchestrate the symphony of life. The future of computational electrophysiology promises to further bridge the gap between abstract mathematical equations and the tangible biological phenomena they represent, opening new horizons for exploration and discovery of the universe hidden within each of us.

## 2.8) References


1. Southern, J. *et al.* Multi-scale computational modelling in biology and physiology. *Prog. Biophys. Mol. Biol.* **96**, 60–89 (2008).
2. Yang, P.-C. *et al.* A Computational Pipeline to Predict Cardiotoxicity: From the Atom to the Rhythm. *Circ. Res.* **126**, 947–964 (2020).
3. Flandoli, F., Leocata, M. & Ricci, C. The Mathematical modeling of Cancer growth and angiogenesis by an individual based interacting system. *J. Theor. Biol.* **562**, 111432 (2023).
4. Jenner, A. L., Aogo, R. A., Davis, C. L., Smith, A. M. & Craig, M. Leveraging Computational Modeling to Understand Infectious Diseases. *Curr. Pathobiol. Rep.* **8**, 149–161 (2020).
5. Moore, M. N. & Noble, D. Editorial: Computational modelling of cell & tissue processes & function. *J. Mol. Histol.* **35**, 655–658 (2004).
6. Thiele, I. *et al.* Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. *Mol. Syst. Biol.* **16**, e8982 (2020).
7. Trayanova, N. A. & Prakosa, A. Up digital and personal: How heart digital twins can transform heart patient care. *Heart Rhythm* **21**, 89–99 (2024).

8. Mayourian, J., Sobie, E. A. & Costa, K. D. An Introduction to Computational Modeling of Cardiac Electrophysiology and Arrhythmogenicity. *Methods Mol. Biol. Clifton NJ* **1816**, 17–35 (2018).
9. O'Hara, T., Virág, L., Varró, A. & Rudy, Y. Simulation of the Undiseased Human Cardiac Ventricular Action Potential: Model Formulation and Experimental Validation. *PLOS Comput. Biol.* **7**, e1002061 (2011).
10. Berndt, N. & Holzhütter, H.-G. Mathematical Modeling of Cellular Metabolism. *Recent Results Cancer Res. Fortschritte Krebsforsch. Progres Dans Rech. Sur Cancer* **207**, 221–232 (2016).
11. Ay, A. & Arnosti, D. N. Mathematical modeling of gene expression: a guide for the perplexed biologist. *Crit. Rev. Biochem. Mol. Biol.* **46**, 137–151 (2011).
12. Rangamani, P. & Iyengar, R. Modelling cellular signalling systems. *Essays Biochem.* **45**, 83–94 (2008).
13. Ngo, K. *et al.* Elucidating molecular mechanisms of protoxin-II state-specific binding to the human NaV1.7 channel. *J. Gen. Physiol.* **156**, e202313368 (2024).
14. Ngo, K., Bruno da Silva, F., Leite, V. B. P., Contessoto, V. G. & Onuchic, J. N. Improving the Thermostability of Xylanase A from Bacillus subtilis by Combining Bioinformatics and Electrostatic Interactions Optimization. *J. Phys. Chem. B* **125**, 4359–4367 (2021).
15. Sheng, X. & Himo, F. The Quantum Chemical Cluster Approach in Biocatalysis. *Acc. Chem. Res.* **56**, 938–947 (2023).
16. Han, Y. *et al.* Elucidation of a dynamic interplay between a beta-2 adrenergic receptor, its agonist, and stimulatory G protein. *Proc. Natl. Acad. Sci.* **120**, e2215916120 (2023).
17. Huber, G. A. & McCammon, J. A. Browndye: A Software Package for Brownian Dynamics. *Comput. Phys. Commun.* **181**, 1896–1905 (2010).
18. Długosz, M. & Trylska, J. Diffusion in crowded biological environments: applications of Brownian dynamics. *BMC Biophys.* **4**, 1–9 (2011).

19. West, J., Robertson-Tessi, M. & Anderson, A. R. A. Agent-based methods facilitate integrative science in cancer. *Trends Cell Biol.* **33**, 300–311 (2023).

20. Leman, J. K. *et al.* Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods* **17**, 665–680 (2020).

21. Fleishman, S. J. *et al.* RosettaScripts: A Scripting Language Interface to the Rosetta Macromolecular Modeling Suite. *PLOS ONE* **6**, e20161 (2011).

22. Krishna, R. *et al.* Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science* **384**, eadl2528 (2024).

23. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

24. Lin, Z. *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* **379**, 1123–1130 (2023).

25. Perutz, M. F. *et al.* Structure of Hæmoglobin: A Three-Dimensional Fourier Synthesis at 5.5-Å. Resolution, Obtained by X-Ray Analysis. *Nature* **185**, 416–422 (1960).

26. Kendrew, J. C. *et al.* A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **181**, 662–666 (1958).

27. Kendrew, J. C. *et al.* Structure of Myoglobin: A Three-Dimensional Fourier Synthesis at 2 Å. Resolution. *Nature* **185**, 422–427 (1960).

28. Šali, A., Potterton, L., Yuan, F., van Vlijmen, H. & Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins Struct. Funct. Bioinforma.* **23**, 318–326 (1995).

29. Rohl, C. A., Strauss, C. E. M., Misura, K. M. S. & Baker, D. Protein Structure Prediction Using Rosetta. in *Methods in Enzymology* vol. 383 66–93 (Academic Press, 2004).

30. Song, Y. *et al.* High resolution comparative modeling with RosettaCM. *Struct. Lond. Engl. 1993* **21**, 10.1016/j.str.2013.08.005 (2013).

31. Mateos, D. L. & Yarov-Yarovoy, V. Structural Modeling of Peptide Toxin - Ion Channel Interactions using RosettaDock. 2022.06.29.498146 Preprint at https://doi.org/10.1101/2022.06.29.498146 (2022).

32. Park, H., Zhou, G., Baek, M., Baker, D. & DiMaio, F. Force Field Optimization Guided by Small Molecule Crystal Lattice Data Enables Consistent Sub-Angstrom Protein-Ligand Docking. *J. Chem. Theory Comput.* **17**, 2000–2010 (2021).

33. Kalita, P., Tripathi, T. & Padhi, A. K. Computational Protein Design for COVID-19 Research and Emerging Therapeutics. *ACS Cent. Sci.* **9**, 602–613 (2023).

34. Li, Q. *et al.* Computational design of a cutinase for plastic biodegradation by mining molecular dynamics simulations trajectories. *Comput. Struct. Biotechnol. J.* **20**, 459–470 (2022).

35. Thieker, D. F. *et al.* Stabilizing proteins, simplified: A Rosetta-based webtool for predicting favorable mutations. *Protein Sci. Publ. Protein Soc.* **31**, e4428 (2022).

36. Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).

37. AlQuraishi, M. AlphaFold at CASP13. *Bioinforma. Oxf. Engl.* **35**, 4862–4865 (2019).

38. Jumper, J. *et al.* Applying and improving AlphaFold at CASP14. *Proteins* **89**, 1711–1721 (2021).

39. Nussinov, R., Zhang, M., Liu, Y. & Jang, H. AlphaFold, allosteric, and orthosteric drug discovery: Ways forward. *Drug Discov. Today* **28**, 103551 (2023).

40. Baek, M. *et al.* Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nat. Methods* **21**, 117–121 (2024).

41. Teague, S. J. Implications of protein flexibility for drug discovery. *Nat. Rev. Drug Discov.* **2**, 527–541 (2003).

42. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R. & Ferrin, T. E. A geometric approach to macromolecule-ligand interactions. *J. Mol. Biol.* **161**, 269–288 (1982).

43. Marze, N. A., Roy Burman, S. S., Sheffler, W. & Gray, J. J. Efficient flexible backbone protein–protein docking for challenging targets. *Bioinformatics* **34**, 3461–3469 (2018).

44. Gray, J. J. *et al.* Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **331**, 281–299 (2003).

45. Wang, C., Bradley, P. & Baker, D. Protein-protein docking with backbone flexibility. *J. Mol. Biol.* **373**, 503–519 (2007).

46. Knegtel, R. M. A., Kuntz, I. D. & Oshiro, C. M. Molecular docking to ensembles of protein structures1. *J. Mol. Biol.* **266**, 424–440 (1997).

47. Pedotti, M., Simonelli, L., Livoti, E. & Varani, L. Computational Docking of Antibody-Antigen Complexes, Opportunities and Pitfalls Illustrated by Influenza Hemagglutinin. *Int. J. Mol. Sci.* **12**, 226–251 (2011).

48. Friesner, R. A. *et al.* Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J. Med. Chem.* **47**, 1739–1749 (2004).

49. Verdonk, M. L., Cole, J. C., Hartshorn, M. J., Murray, C. W. & Taylor, R. D. Improved protein-ligand docking using GOLD. *Proteins* **52**, 609–623 (2003).

50. Wang, Z. *et al.* Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys.* **18**, 12964–12975 (2016).

51. Repasky, M. P., Shelley, M. & Friesner, R. A. Flexible ligand docking with Glide. *Curr. Protoc. Bioinforma.* **Chapter 8**, Unit 8.12 (2007).

52. Zhou, Z., Felts, A. K., Friesner, R. A. & Levy, R. M. Comparative Performance of Several Flexible Docking Programs and Scoring Functions: Enrichment Studies for a Diverse Set of Pharmaceutically Relevant Targets. *J. Chem. Inf. Model.* **47**, 1599–1608 (2007).

53. Alder, B. J. & Wainwright, T. E. Studies in Molecular Dynamics. I. General Method. *J. Chem. Phys.* **31**, 459–466 (1959).

54. Ciccotti, G., Dellago, C., Ferrario, M., Hernández, E. R. & Tuckerman, M. E. Molecular simulations: past, present, and future (a Topical Issue in EPJB). *Eur. Phys. J. B* **95**, 3 (2022).

55. McCammon, J. A., Gelin, B. R. & Karplus, M. Dynamics of folded proteins. *Nature* **267**, 585–590 (1977).

56. Tidor, B. Molecular dynamics simulations. *Curr. Biol.* **7**, R525–R527 (1997).

57. Karplus, M. Molecular Dynamics Simulations of Biomolecules. *Acc. Chem. Res.* **35**, 321–323 (2002).

58. Pronk, S. *et al.* GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).

59. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

60. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).

61. Guvench, O. & MacKerell, A. D. Comparison of protein force fields for molecular dynamics simulations. *Methods Mol. Biol. Clifton NJ* **443**, 63–88 (2008).

62. Maier, J. A. *et al.* ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).

63. MacKerell, A. D., Banavali, N. & Foloppe, N. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**, 257–265 (2000).

64. Todorova, N., Legge, F. S., Treutlein, H. & Yarovsky, I. Systematic Comparison of Empirical Forcefields for Molecular Dynamic Simulation of Insulin. *J. Phys. Chem. B* **112**, 11137–11146 (2008).

65. Liwo, A. *et al.* Theory and Practice of Coarse-Grained Molecular Dynamics of Biologically Important Systems. *Biomolecules* **11**, 1347 (2021).

66. Maisuradze, G. G., Senet, P., Czaplewski, C., Liwo, A. & Scheraga, H. A. Investigation of Protein Folding by Coarse-Grained Molecular Dynamics with the UNRES Force Field. *J. Phys. Chem. A* **114**, 4471–4485 (2010).

67. Kästner, J. Umbrella sampling. *WIREs Comput. Mol. Sci.* **1**, 932–942 (2011).

68. Shirts, M. R. & Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **129**, 124105 (2008).

69. Pfaendtner, J. Metadynamics to Enhance Sampling in Biomolecular Simulations. *Methods Mol. Biol. Clifton NJ* **2022**, 179–200 (2019).

70. Qi, R., Wei, G., Ma, B. & Nussinov, R. Replica Exchange Molecular Dynamics: A Practical Application Protocol with Solutions to Common Problems and a Peptide Aggregation and Self-Assembly Example. *Methods Mol. Biol. Clifton NJ* **1777**, 101–119 (2018).

71. Hodgkin, A. L. & Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **117**, 500–544 (1952).

72. https://www.math.mcgill.ca/gantumur/docs/reps/RyanSicilianoHH.pdf. https://www.math.mcgill.ca/gantumur/docs/reps/RyanSicilianoHH.pdf.

73. 1.2: Forward Euler method. *Mathematics LibreTexts* https://math.libretexts.org/Bookshelves/Differential_Equations/Numerically_Solving_Ordinary_Differential_Equations_(Brorson)/01%3A_Chapters/1.02%3A_Forward_Euler_method (2022).

74. Moreno, J. D., Lewis, T. J. & Clancy, C. E. Parameterization for In-Silico Modeling of Ion Channel Interactions with Drugs. *PLoS ONE* **11**, e0150761 (2016).

75. Perrin, M. J., Kuchel, P. W., Campbell, T. J. & Vandenberg, J. I. Drug Binding to the Inactivated State Is Necessary but Not Sufficient for High-Affinity Binding to Human Ether-à-go-go-Related Gene Channels. *Mol. Pharmacol.* **74**, 1443–1452 (2008).

76. Luo, C. H. & Rudy, Y. A dynamic model of the cardiac ventricular action potential. I. Simulations of ionic currents and concentration changes. *Circ. Res.* **74**, 1071–1096 (1994).

77. Yang, P.-C. *et al.* A multiscale predictive digital twin for neurocardiac modulation. *J. Physiol.* **601**, 3789–3812 (2023).

78. Moreno, J. D. *et al.* A Computational Model to Predict the Effects of Class I Anti-Arrhythmic Drugs on Ventricular Rhythms. *Sci. Transl. Med.* **3**, 98ra83-98ra83 (2011).

# CHAPTER 3:

# From Tarantula Venom to Therapeutics: Unveiling Protoxin-2's Binding to Pain-Sensing Human $Na_V1.7$ Channels

Khoa Ngo*[1,2], Diego Lopez Mateos*[1,2], Yanxiao Han[2], Kyle C. Rouen[1,2], Surl-Hee Ahn[3], Heike Wulff[4], Colleen E. Clancy[2,4,5], Vladimir Yarov-Yarovoy**[2,6], and Igor Vorobyov**[2,4]

[1]*Biophysics Graduate Group, UC Davis, CA*

[2]*Department of Physiology and Membrane Biology, UC Davis, CA*

[3]*Department of Chemical Engineering, UC Davis, CA*

[4]*Department of Pharmacology, UC Davis, CA*

[5]*Center for Precision Medicine and Data Science, UC Davis, CA*

[6]*Department of Anesthesiology and Pain Medicine, UC Davis, CA*

** These authors contributed equally*

*** These authors are co-corresponding authors*

**Author Contributions:**

My role in this study involved conducting molecular dynamics simulations of the human $Na_V1.7$ channel in complex with protoxin-2. In addition, I analyzed and visualized the interactions between the peptide and the channel and calculated the binding energetics. I was responsible for creating figures 2, 6, 7, 8, 9, 10, and 11 in the main manuscript, as well as all figures in the supplementary material. The structural models used in the research were developed by the co-first author, Diego Lopez Mateos.

Adapted from "Elucidating molecular mechanisms of protoxin-II state-specific binding to the human NaV1.7 channel". J Gen Physiol 5 February 2024; 156 (2): e202313368. doi: https://doi.org/10.1085/jgp.202313368

*"It is very easy to answer many of these fundamental biological questions; you just look at the thing!"* – Richard P. Feynman (1918-1988)

## 3.1) Introduction

Voltage-gated sodium ($Na_V$) channels are responsible for electric signaling and play a key role in various physiological processes[1,2]. $Na_V$ channels are involved in the functioning of the central and peripheral nervous systems and the contraction of skeletal and cardiac muscles[2]. There are nine known members in the $Na_V$ family, named $Na_V1.1$ to $Na_V1.9$[2]. A pore-forming α subunit of a eukaryotic $Na_V$ channel is composed of four homologous domains (DI-DIV), each containing six transmembrane segments (S1-S6), with S1 through S4 serving as the voltage-sensing domain (VSD). The S4 segment contains four or five positively charged amino acid residues known as gating charge residues and acting together as a voltage sensor. The S5 and S6 segments, intervening loops, and pore helix form the channel pore containing the ion selectivity filter (SF) region. Membrane depolarization induces the charged residues in the S4 segment to move outwards, a movement that is transmitted to the pore domain through the S4-S5 linker leading to channel opening and a rapid influx of sodium ions into the cell. This sodium influx further depolarizes the cell membrane and leads to the generation of an action potential. $Na_V$ channel opening depends on the sequential activation of VSD I through III[3,4]. In contrast, the activation of VSD IV is coupled with the rapid inactivation of the channel through the release of the isoleucine – phenylalanine – methionine (IFM) or similar motif in the DIII-DIV intracellular loop (inactivation gate)[3–6]. This inactivation process is crucial to stop the influx of sodium ions and enable the repolarization of the membrane to terminate the action potential.

$Na_V$ channel dysfunction arising from mutations can lead to various pathophysiological conditions, including arrhythmias, epilepsy, chronic pain, and insensitivity to pain[7]. $Na_V1.3$, $Na_V1.7$, $Na_V1.8$, and $Na_V1.9$ channels have been implicated in pain signaling[2]. Transgenic mice lacking $Na_V1.7$- and $Na_V1.8$-positive nociceptors have been shown in experiments to exhibit increased mechanical and thermal pain thresholds[8]. In particular, the mice showed reduced or no response to inflammatory pain responses evoked by various stimuli[8]. Moreover, humans born with loss-of-function mutations in the SCN9A gene encoding for the $Na_V1.7$ α subunit were discovered to have congenital insensitivity to pain[9]. In contrast, gain-of-function mutations have been associated with aberrant chronic pain sensation and related

disorders such as inherited erythromelalgia and paroxysmal extreme pain disorder[10,11]. These observations further emphasize the importance of the $Na_V1.7$ channels to pain perception.

Local anesthetics, such as lidocaine, reduce the transmission of pain signals to the central nervous system by non-selectively blocking various $Na_V$ subtypes[12]. Due to their lack of $Na_V$ selectivity, side effects are often reported concerning the loss of other sensations, such as the sense of touch. The use of opioid analgesics is associated with side effects such as respiratory depression, constipation, and development of dependence[13]. Therefore, developing drugs that selectively inhibit $Na_V1.7$ function may result in strong analgesia without undesirable side effects[14,15]. $Na_V$ channels implicated in pain signaling have been discovered to be targets of natural toxins found in species such as spiders and scorpions[7]. The toxins modulate channel activity by blocking the ion permeation pathway (i.e., pore blockers) or binding to the VSD to alter channel gating kinetics (i.e., gating modifier toxins)[16].

Protoxin-2 (PTx2), a 30-residue peptide derived from the venom of the Peruvian Green Velvet tarantula (*Thrixopelma pruriens*), is a gating modifier toxin and has a moderate selectivity for $Na_V1.7$ versus other $Na_V$ subtypes, making it a suitable scaffold for a peptide-based pain therapeutic[17]. This toxin interferes with the activation of $Na_V$ channels by binding to the S3-S4 loop in the deactivated VSD II and shifting the voltage dependence of activation to a more positive potential[17]. PTx2 is 100-fold more potent for $Na_V1.7$ ($IC_{50}$ = 0.3-1.0 nM) versus other $Na_V$ subtypes[14,18–20]. Additionally, PTx2 binds to the deactivated $hNa_V1.7$ VSD IV inhibiting fast inactivation and producing sustained sodium currents, similar to the effect of α-scorpion toxins[4], but with lower potency ($IC_{50}$ = 0.24 μM)[19]. However, this effect of PTx2 is typically masked by the preferential inhibition of $Na_V$ channel activation due to its higher affinity binding to VSD II[19].

By using peptide toxins as a scaffold for designing novel ion channel modulators, researchers can potentially develop more effective and targeted therapies for chronic pain[15,21]. Potential strategies for pain therapy via $hNa_V1.7$ gating inhibition include the trapping of VSD II in the deactivated state[22] (to prevent the channel from activating) or the trapping of VSD IV in the activated state[23] (to keep the channel in the inactivated, non-conducting state). However,

current structural studies[4,17] have yet to result in a complete atomic-resolution understanding of how PTx2 interacts with $Na_V1.7$ VSDs in a state-specific manner and with the surrounding lipids. Recently, experimental structures of PTx2 bound to $Na_V1.7$ VSD II in various states have been resolved using chimeric constructs, grafting parts of the human $Na_V1.7$ channel sequence into a bacterial $Na_V$ channel[22]. These studies resulted in X-ray and cryogenic electron-microscopy (cryo-EM) structures of the PTx2-h$Na_V1.7$-VSD II-$Na_V$Ab complexes in activated and deactivated states[22]. Yet, these structures do not fully represent a human $Na_V$ channel. Notably, Shen *et al.*[17] obtained the first cryo-EM structure of the full h$Na_V1.7$, resolving structures of PTx2 bound to activated VSD II and VSD IV. However, due to the relatively low resolution of PTx2 densities, these structures did not allow complete atomic reconstruction of the peptide toxin. In addition, the studies did not provide energetic insights into toxin – channel and lipids interactions, which are crucial to designing improved peptides as ion channel modulators.

Computational methods can provide structural insights into the atomic-level mechanisms involved in ion channel function, modulation[24] and toxin binding[25–27]. Moreover, they are useful tools to guide the rational design process towards enhanced peptide variants[14,28]. Our study demonstrated that computational structural modeling and molecular dynamics simulations could accurately capture the atomic-resolution molecular mechanisms by which PTx2 binds to h$Na_V1.7$ VSD II and IV in the activated and deactivated states. Our molecular modeling and simulation insights will be useful to inform future peptide design studies and help researchers to identify potential peptide mutations to improve PTx2 selectivity and potency for h$Na_V1.7$. Moreover, our methodology can be expanded to study other toxin – ion channel interactions to design potential peptide-based therapeutics for diseases such as cancer[29], cardiac arrhythmia[30,31], epilepsy[32,33], and neuropathic pain[14,21,30].

## 3.2) Materials and Methods

### a. Structural Modeling and Peptide Docking

In this study, we used computational modeling methods guided by experimental structural data to investigate molecular mechanisms of PTx2 interaction with the h$Na_V1.7$ VSD II and

VSD IV in multiple conformational states **(Figure 1)**. Our primary objective was to generate models of PTx2 bound to VSD II and VSD IV in deactivated and activated conformations as input for molecular dynamics (MD) simulations.

To achieve this goal, we first used cryo-EM structures of PTx2 bound to hNa$_V$1.7 to both the activated VSD II and VSD IV (PDBs: 6J8J)[17]. However, the resolution of the toxin EM density, at ~5 Å, did not permit complete atomic structure reconstruction in the experimental structure[17]. To overcome this limitation, we utilized the Rosetta structural modeling software suite[34] and the PTx2 - hNa$_V$1.7 complex electron-microscopy (EM) density data to model the interactions of PTx2 with activated hNa$_V$1.7 VSD II and VSD IV. For VSD II, we used Rosetta Loop Modeling[35,36] for modeling missing residues in the S3-S4 loop which form receptor site for binding of PTx2 based on experimental structural data for both VSD II[17] and PTx2[22].

The PTx2 docking protocol comprised a first step using RosettaDock[37,38] with the Rosetta membrane energy function[39]. RosettaDock is a Monte Carlo-based multiscale algorithm that optimizes both rigid-body orientation and side chain conformations of the docked protein partners. In this step, the ion channel VSD's structural model was transformed into membrane coordinates by superimposing it into a reference hNa$_V$1.7 experimental structure downloaded from the Protein Data Bank of Transmembrane Proteins (PDBTM) database[40], which stores experimental structures of membrane proteins in membrane coordinates. The toxin was placed manually in three different initial locations around the binding site in the corresponding VSD, using the toxin's EM density as a reference. The structure of PTx2 was obtained from the PTx2-hNa$_V$1.7-VSD II-Na$_V$Ab X-ray structure (PDB: 6N4I)[22]. Subsequently, 20,000 docked models per input were generated with RosettaDock making a total of 60,000 models. The top 10% of models based on Rosetta *total_score* (total computed Rosetta energy of the protein) were extracted, and the top 100 *interface_score* (Rosetta interaction energy) models were used as an input for the next step. The second step involved EM density fitting of top-docked models to the experimental EM map using Rosetta EM density refinement.[41] During EM density refinement, toxin backbone and sidechain conformations are optimized using a modified energy function that accounts for the agreement between the model and the provided experimental EM map. In this step, 500 models were generated per input,

making a total of 50,000 models. The top 10% of models based on *total_score* was initially extracted. The final models were selected based on the recalculated *interface_score* after EM fitting, and on the *elec_dens_fast* score term (agreement between the Rosetta model and the experimental EM maps).

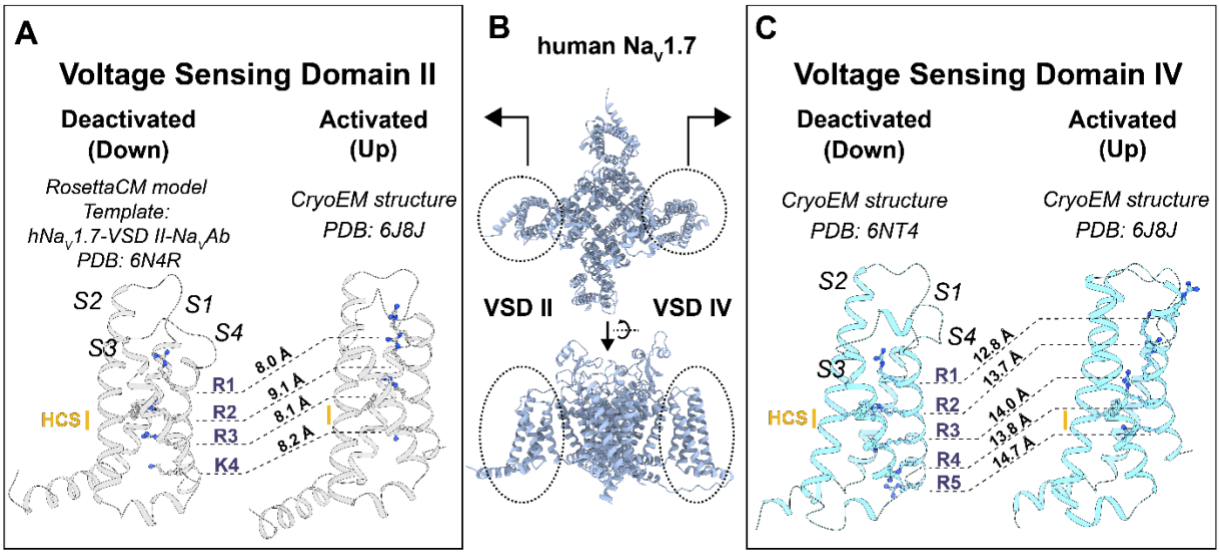



**Figure 1.** ***Structural comparison of different conformational states of hNaV1.7 VSD II and IV.*** *A) Snapshots of deactivated (left) and activated (right) VSD II structures. For each state, S1 to S4 VSD segments are shown, highlighting the gating charge residues ("R1", "R2", "R3", "K4", "R5") and the hydrophobic constriction site (HCS) in yellow. The distance between the $C_\alpha$ atoms of each gating charge in each state after superimposition of the VSDs is shown. B) Global view of cryo-EM NaV1.7 structure from PDB 6J8J[17] highlights the VSD II and IV positions. C) Snapshots of deactivated (left) and activated (right) VSD IV structures (same information as in panel A).*

Furthermore, we docked PTx2 onto the structure of hNaV1.7 VSD IV trapped in a deactivated state by an α-scorpion toxin (PDB: 6NT4)[4], thereby constructing a model of PTx2 bound to deactivated hNaV1.7 VSD IV. Note that this experimental structure comprises a human – cockroach chimeric NaV channel, in which the VSD IV has a human sequence while the remaining part is derived from the cockroach NaVPaS channel. To model the interactions between PTx2 and VSD IV, the NaVPaS portion was removed, leaving only the VSD IV. The docking protocol was similar to the two cases above, but without any EM density-related refinement or scoring, since no experimental map is available for this interaction. Finally, we

employed the cryo-EM structure of the chimera channel hNaV1.7-VSD II-NaVAb (PDB: 6N4R), [22] which has the VSD II in a deactivated state, as a template for RosettaCM [42] homology modeling. This allowed us to generate a model of the deactivated VSD II with the entire hNaV1.7 VSD II sequence (sequence identity between template and target: 63.7%). We generated 5,000 RosettaCM models, extracted the top 10% based on *total_score*, and selected the final model based on both low root-mean-square deviation (RMSD) for the template and low *total_score*. Then, we used RosettaDock[37,38] with membrane energy function[39] to dock PTx2 to the deactivated VSD II of the hNaV1.7 model described above.

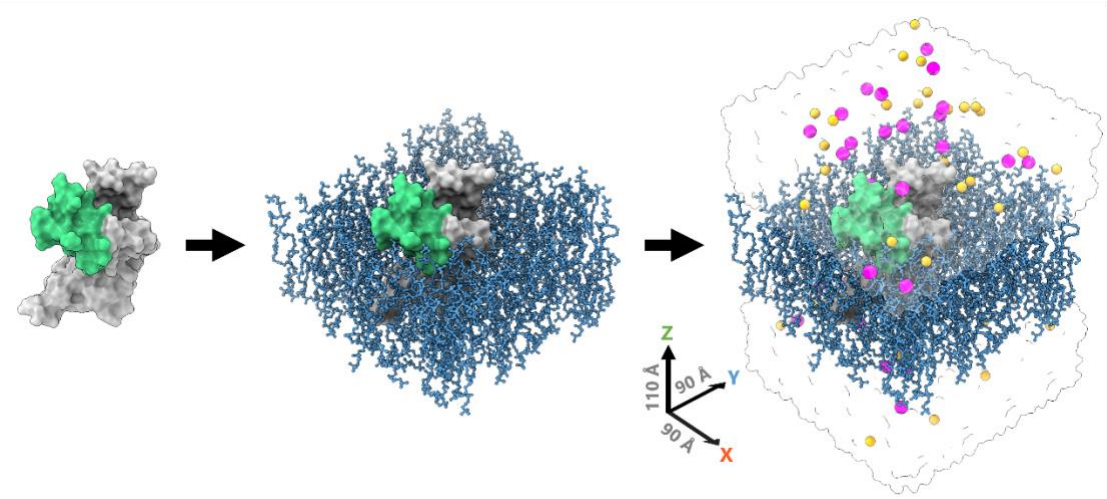



**Figure 2.** ***Creation process of a molecular simulation system consisting of the PTx2 – hNaV1.7 activated VSD II complex embedded in a POPC bilayer and solvated with aqueous 0.15 M NaCl solution.*** *PTx2 and VSD II are shown as green and dark gray surfaces, respectively. The lipid membrane composed of POPC lipids is shown in blue sticks. Water is displayed as a transparent surface. $Na^+$ and $Cl^-$ ions are displayed as purple and gold spheres, respectively.*

**b. Atomistic MD simulations**

Using the CHARMM-GUI web server[43], the VSD-bound-toxin complexes were inserted into tetragonal phospholipid bilayer patches composed of 160 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) molecules in total. The systems were then solvated by a 0.15 M aqueous NaCl solution to mimic the physiological extracellular environment, resulting in molecular systems of ~63,000 atoms as illustrated in **Figure 2**. The protonation

state of each residue was assigned under neutral pH conditions. Standard N- and C-termini were set for PTx2 and VSD II/IV. For PTx2, disulfide bonds were added between cystine residues C2 and C16, C9 and C21, and lastly, C15 and C25.

All-atom molecular dynamics (MD) simulations were performed using the Amber20 software package[44] and the standard all-atom Chemistry at Harvard Macromolecular Mechanics (CHARMM) force fields for proteins (CHARMM36m), lipids (C36), and ions[45–47] as well as the TIP3P water model[48]. During system equilibration, a 1 kcal/mol/Å$^2$ harmonic restraint was applied to all atoms in the system and was gradually reduced to 0.1 kcal/mol/Å$^2$ over a period of 90 ns. Subsequently, a production run of 910 ns was conducted, resulting in a total simulation time of 1 μs for each system. All MD simulations were run in the isobaric-isothermal or NPT ensemble at 310.15 K and 1 atm pressure, which were maintained using the Langevin temperature equilibration scheme and Berendsen barostat. Non-bonded interactions were computed up to a 9 Å cutoff. Long-range electrostatic interactions were computed using Particle mesh Ewald (PME) method[49], while no long-range correction was applied to van der Waals interactions as suggested for C36 lipid force field. All covalent bonds involving hydrogen atoms were constrained using the SHAKE algorithm[50] allowing to use of a 2-fs time step. The simulation protocol was replicated three times, each with different initial lipid positions and starting velocities of atoms. The replication accounts for potential variability in the lipid arrangement and system dynamics.

Afterward, protein-ligand interactions were characterized in each set of simulations using in-house Python scripts incorporating the protein–ligand interaction profiler (PLIP) Python module[51]. Binding free energy calculations were performed using the Molecular Mechanics Poisson-Boltzmann Surface Area (MM/PBSA) approach[52–54] with all-atom MD simulation trajectories by MMPBSA.py program in Amber Tools[55]. MM/PBSA was selected due to its computational efficiency, which facilitates the generation of qualitative insights into the energetic components of binding. Moreover, it enables the incorporation of an implicit membrane in its calculations. This feature is particularly advantageous as it fosters a more accurate approximation of the energetics underlying the interaction between PTx2 and hNa$_V$1.7 embedded in the membrane.

The Chamber module of ParmEd program was used to convert CHARMM-style forcefield files to Amber-style forcefield files without changes in the force field parameters[56]. The aqueous solution (with ionic strength of 0.15 M) and lipid membrane were treated implicitly using dielectric constants, denoted ε (water $\varepsilon_w$ = 80, lipid bilayer $\varepsilon_l$ = 2, and protein $\varepsilon_p$ = 4). The solvent probe radius was set to 1.4 Å, and the atomic radii were set according to the converted force field parameters.

To obtain the solvation free energy and the gas-phase electrostatic energy contributions without entropy, the particle-particle particle-mesh (P3M) procedure was used[57]. These calculations were performed with an implicit membrane, where the electrostatic energy includes both reaction field and Coulombic electrostatic energies. Entropy was calculated separately by the interaction entropy method,[58] where interaction energy between PTx2 and hNa$_V$1.7 VSD II/IV, including only Coulombic electrostatic and van der Waals energies was needed. To obtain the Coulombic energy separated from the reaction field energy, each system energy was recalculated using the dielectric boundary surface charges method in the implicit ionic solution.

To calculate the energy contribution of each residue to the binding process for each system, we computed the electrostatic and van der Waals interaction energies of PTx2 with POPC lipids and VSD II/IV using AMBER linear interaction energy analysis[44,59].

### c. Molecular Graphics

Molecular graphics visualization was performed using UCSF ChimeraX[60] to depict all resulting models and illustrate toxin – VSD protein residue interactions. Structural comparison was carried out through the superposition of the structures using the MatchMaker tool within ChimeraX.

## 3.3) Results

### a. General Overview of Different PTx2 - Na$_V$1.7 VSD Structures

In this study, we employed computational modeling, informed by experimental structural data, to investigate the molecular interactions between PTx2 and hNa$_V$1.7's Voltage Sensor

Domains (VSDs) II and IV in different conformational states **(Figure 1)**. It should be noted that our focus was solely on individual VSDs and not the whole channel alpha subunit, an approach selected to optimize the use of computational resources and consistently use available structural information. While experimental structures of activated/deactivated VSD IV and activated VSD II of human $Na_V1.7$ are available [17,22], the experimental structure of deactivated VSD II of $hNa_V1.7$ has yet to be determined. To acquire the necessary structural information, we sought to obtain structures of both $hNa_V1.7$ VSD II and VSD IV in both activated and deactivated conformations. To obtain a structure of deactivated VSD II, we used the experimental structure of the $hNa_V1.7$-VSD II-$Na_VAb$[22] as a template for homology modeling (sequence identity: 63.7%) using the RosettaCM[42]. We then compared the positions of gating charges in each VSD by calculating the distances between equivalent $C_\alpha$ atoms after VSD superimposition **(Figure 1)**. Our analysis revealed that gating charges of VSD II move ~8.5 Å, on average, between the deactivated and activated states. In the deactivated state, only the first gating charge ("R1") is above the hydrophobic constriction site (HCS), while in the activated state "R1", "R2" and "R3" are above the HCS. VSD IV showed a more dramatic displacement of the gating charges (~13.8 Å) between the activated and deactivated states. Just as in VSD II, in the deactivated conformation, only "R1" is located above the HCS; however, in the activated conformation of VSD IV "R4" also reaches above the HCS, which contrasts with the activated VSD II structure where only "R1" to "R3" lie above the HCS. These results suggest that the VSD II in the $Na_V1.7$ cryo-EM structure (PDB: 6J8J)[17] is not fully activated, which might be explained by the presence of PTx2 or the lack of membrane voltage during the structure determination process.

Subsequently, structural models of PTx2 in complex with VSD II and VSD IV were generated in both deactivated and activated states as starting points in MD simulations. The methodology utilized for generating these models is outlined in **Figure 3** and briefly summarized as follows: RosettaDock with a membrane energy function was employed for docking PTx2 to the VSDs. The activated conformations of the VSDs (PDB: 6J8J)[17] with bound PTx2 were further refined using the available experimental EM density maps.

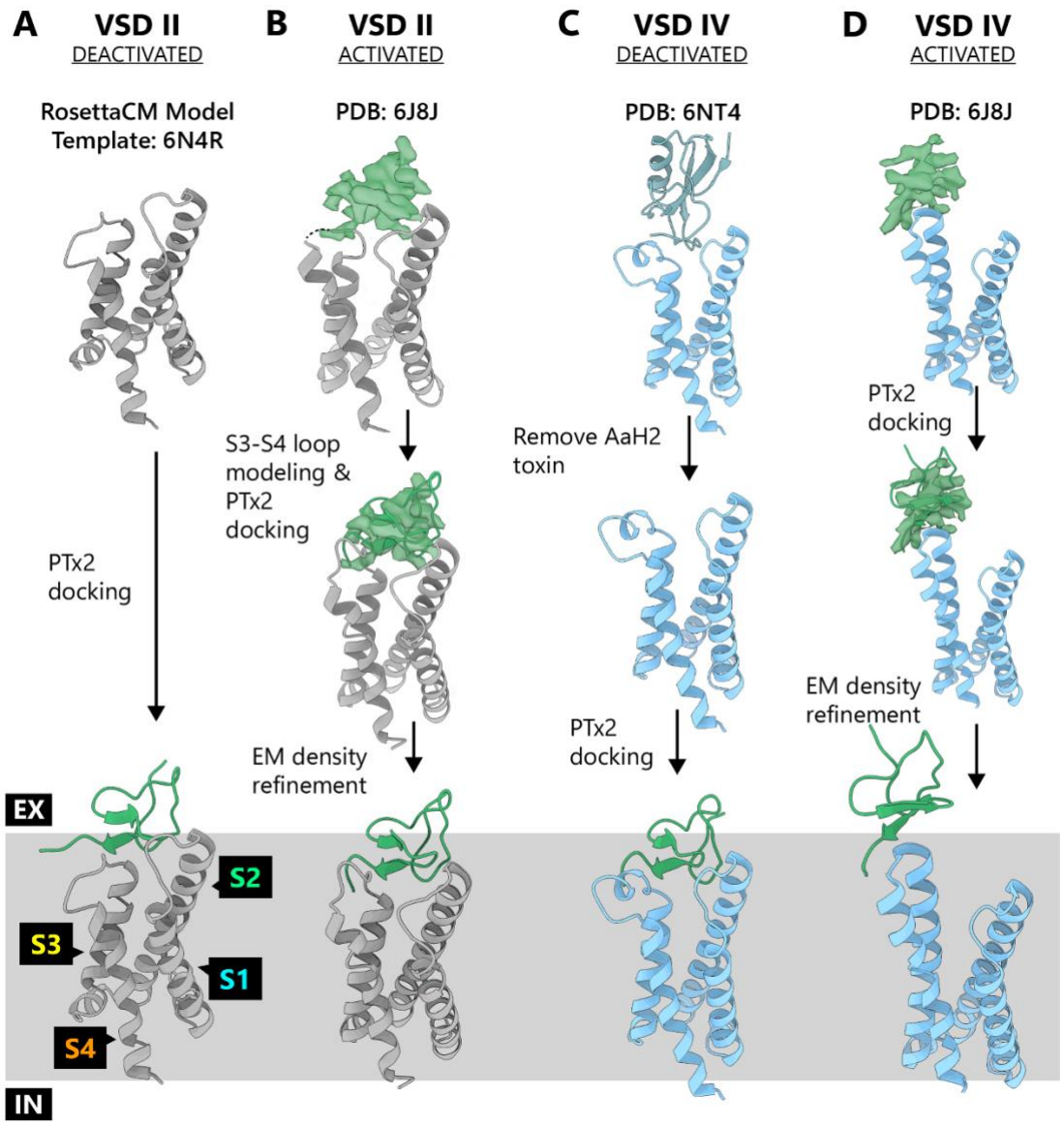



**Figure 3.** ***Workflow of the structural refinement, modeling, and docking process for different PTx2 – hNaV1.7 VSD II and IV structures****. The membrane is represented as a gray box, with the extracellular- and intracellular-facing sides labeled as "EX" and "IN" respectively.*

Previously, it was determined that PTx2 interferes with NaV1.7 activation by binding potently ($IC_{50}$ = 0.3-1.0 nM) to the S3-S4 loop in VSD II and shifting the voltage dependence of activation to more depolarized potentials[14,17–20]. Analysis of the structural models generated here provided important insights into the conformational-dependent binding of PTx2 to NaV1.7 VSD II & VSD IV. PTx2 binds to the VSD II at the “LFLAD” motif, which comprises residues L812 to D816 and is located within the S3 segment and S3-S4 loop[22]. The

toxin binding interface can be divided into two distinct regions: polar and hydrophobic. In the PTx2 - deactivated VSD II model (**Figure 4A, panel i**), the first gating charge ("R1") is observed to be located below the channel-toxin polar binding interface. The electronegative pocket created by VSD II residues E811 and D816 forms hydrogen bonds and salt bridges with the positively charged toxin residues K26 and R22 (**Figure 4A, panel iii**). The hydrophobic interface (**Figure 4A, panel ii**) is composed of M6 and tryptophan residues W5, W24, and W30 (not shown) from the toxin and L812 and F813 from the channel's VSD II S3 segment. Additionally, W24 was observed to be positioned into the VSD's top wedge, making hydrophobic contacts with channel's residues A776 and L770 from VSD II segment S2 (**Figure 4A, panel iv**).

In the PTx2 – activated VSD II model (**Figure 4B**), the upward movement of the S4 segment and subsequent relocation of the "R1" gating charge residue disturbs the electronegative pocket, and the toxin experiences a global shift that pulls it away from the VSD. The main interaction in our PTx2 – activated VSD II model at the polar interface is between toxin's R22 and K26 residues and channel's D816 residue (**Figure 4B, panel iii**). Also, channel's residue E818 interacts with K28 from the toxin (**Figure 4B, panel i**), although this interaction is not present in any of the experimental structures. The hydrophobic interface (**Figure 4B, panel ii**) comprises π-stacking interactions between toxin's W5 and W30 residues and channel's F813 residue (**Figure 4B, panel iv**). The overall shift of the toxin resulted in toxin's residue W24 no longer interacting with channel's residues of VSD II S2 segment.

Experimental results show that PTx2 can also bind to $Na_V1.7$ VSD IV and modulate channel inactivation,[19] although with much lower potency ($IC_{50}$ = 0.24 μM). Our models of PTx2 bound to both deactivated and activated VSD IV presented binding interfaces, which can also be separated into a polar interface and a hydrophobic interface. Our models revealed that the position and interface of PTx2 changed significantly more between the deactivated and activated states of VSD IV compared to VSD II.

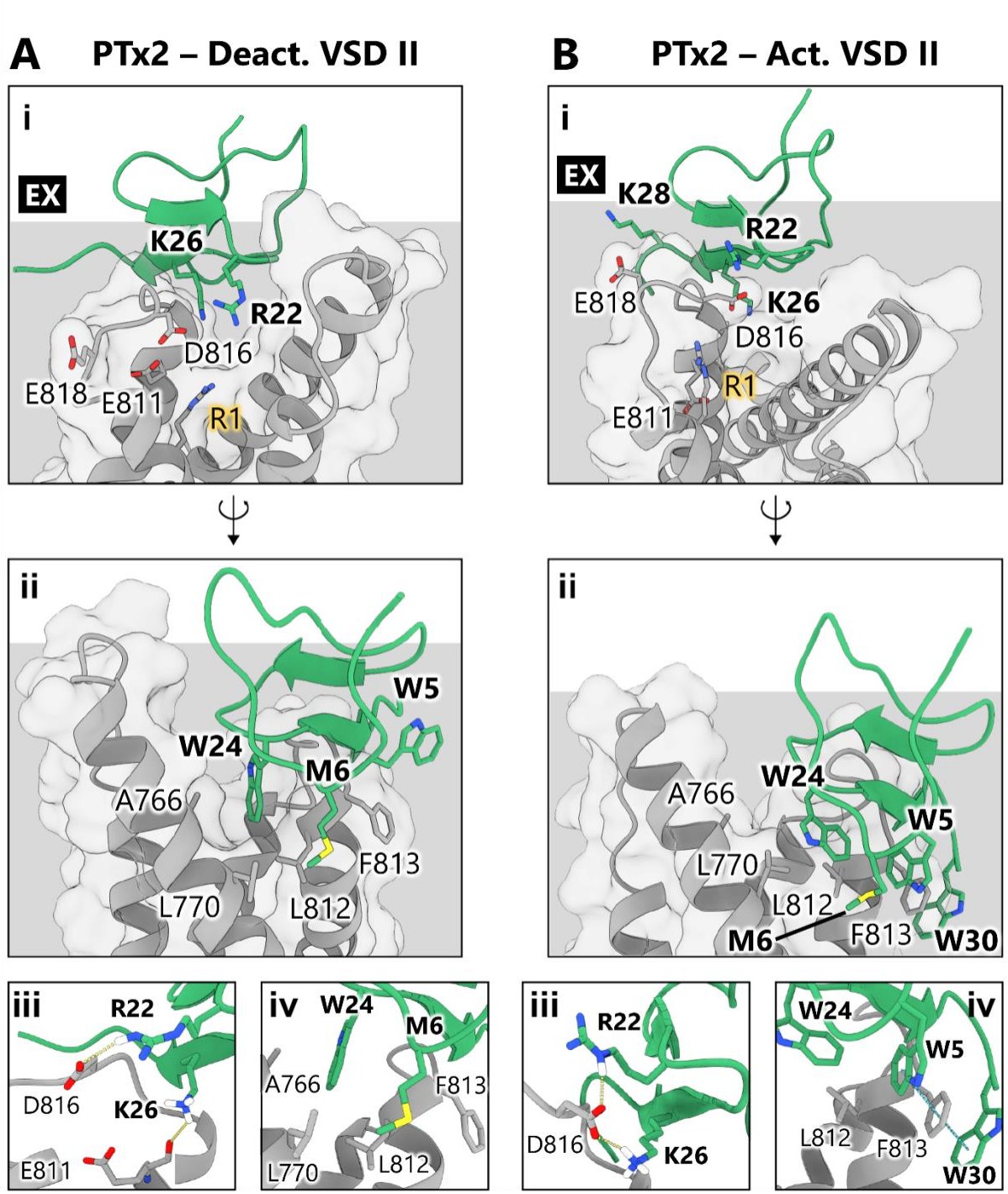



**Figure 4.** ***Binding interfaces of PTx2 – deactivated VSD II (A) and PTx2 – activated VSD II (B) complexes of top scoring Rosetta structural models.*** *i) Front view of the complex highlighting interactions with the VSD that comprise the polar binding interface; "R1" highlights the position of the first gating charge. ii) Back view of the complex highlighting*

*interactions with the VSD that comprise the hydrophobic or membrane-interacting interface. iii) Detailed interactions at the polar interface; yellow dashed lines indicate hydrogen bonds. iv) Detailed interactions at the hydrophobic interface; blue dashed lines highlight π - stacking interactions. The membrane is represented as a gray box, with the extracellular-facing side labeled as "EX".*

Our model of PTx2 bound to the deactivated VSD IV was built using the experimental structure of the hNa$_V$1.7 VSD IV in a deactivated state induced by binding of an α-scorpion toxin (PDB: 6NT4)[4], and it is important to note that the actual VSD IV deactivated state to which PTx2 binds may differ. In our final top Rosetta model, which was selected for subsequent MD simulations, PTx2 is embedded in the membrane between the VSD IV S3 and S2 segments. hNa$_V$1.7 VSD IV residues D1586 and E1589 create an electronegative pocket that accommodates R22 and K26 from the toxin (**Figure 5A, panel i**). Toxin's R22 extends into the VSD and establishes hydrogen bonds with channel's residue D1586, while K26 interacts with G1581 at the top of the S3 segment (**Figure 5A, panel iii**). The hydrophobic interface is more complex in this case (**Figure 5A, panel ii**). PTx2 makes contacts with residues in the VSD IV S2 segment, which are mainly hydrophobic interactions mediated by toxin's residue W24 and channel's residues Y1539 and W1538. Finally, toxin's residues W5, M6, and W30, which also serve as membrane anchors, interact with several hydrophobic residues from the channel located in the VSD IV S3 segment: M1582, F1583 and L1584 (**Figure 5A, panels ii and iv**). In our models, a large conformational rearrangement of the VSD IV occurs between the deactivated and the activated state. As a result, the location of PTx2 also undergoes a dramatic change.

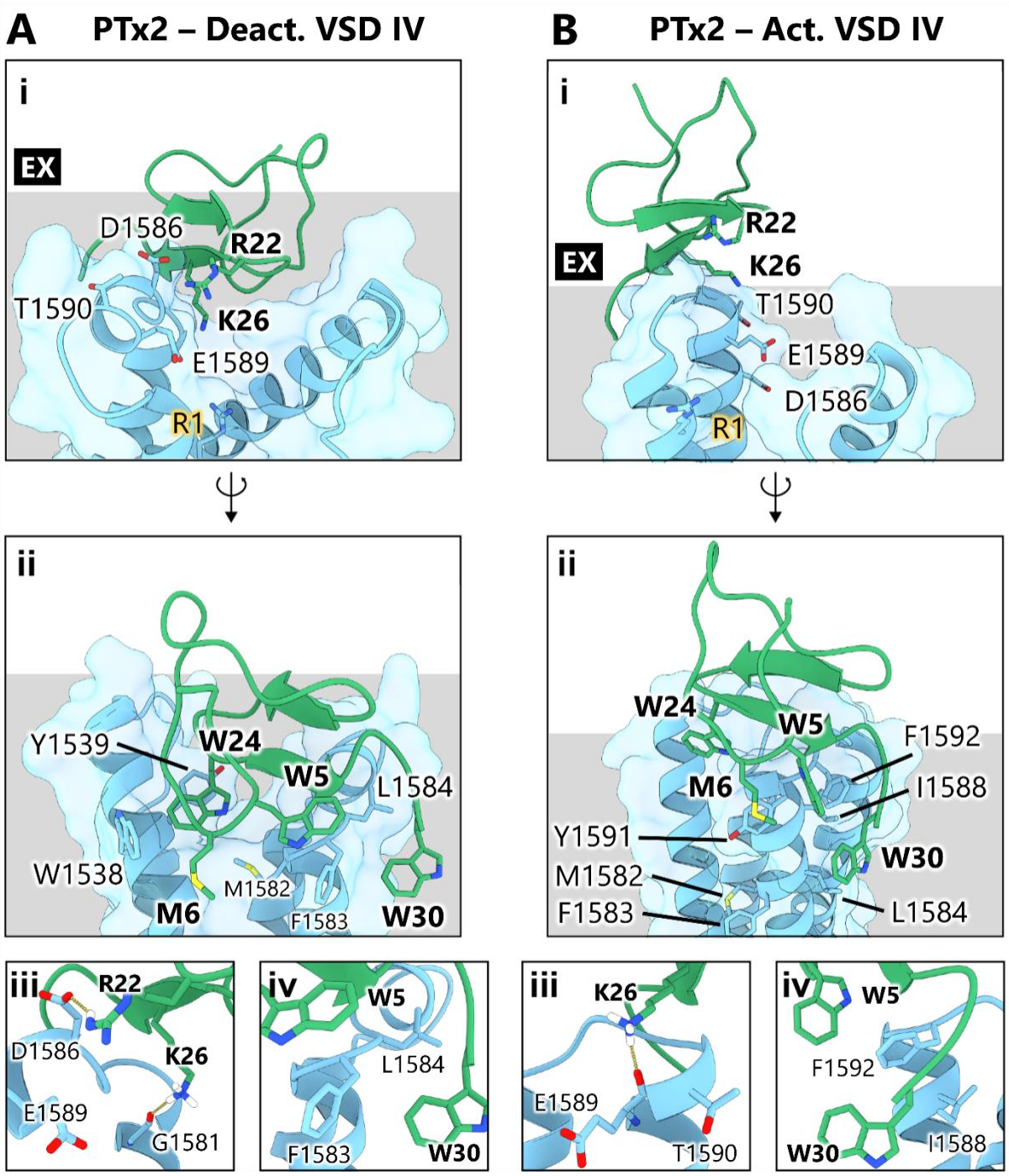



**Figure 5.** ***Binding interfaces of PTx2 – deactivated VSD IV (A) and PTx2 – activated VSD IV (B) complexes of top-scoring Rosetta structural models.*** *i) Front view of the complex highlighting interactions with the VSD that comprise the polar binding interface; "R1" highlights the position of the first gating charge. ii) Back view of the complex highlighting*

*interactions with the VSD that comprise the hydrophobic or membrane-interacting interface. iii) Detailed interactions at the polar interface; yellow dashed lines indicate hydrogen bonds. iv) Detailed interactions at the hydrophobic interface. The membrane is represented as a gray box, with the extracellular-facing side labeled as "EX".*

In the activated VSD IV state, the toxin is situated at the top of the S3 segment and does not establish interactions with the S2 segment. The polar interface observed in the activated VSD IV state (**Figure 5B, panel i**) is primarily mediated by the interaction between the sidechain of toxin's residue K26 and the backbone carbonyl oxygen of channel's S3 segment residue E1589 (**Figure 5B, panel iii**), with PTx2's residue R22 not appearing to play a significant role in this interaction in contrast to the deactivated state. In the hydrophobic interface (**Figure 5B, panel ii**), the toxin uses its residues W5, M6, and W30 to anchor to the membrane and establish hydrophobic contacts with residues I1588, Y1591, F1592 from the channel's S3 segment (**Figure 5B, panel iv**). Notably, these hydrophobic channel's VSD IV residues are positioned far from the binding interface in the deactivated state. In contrast, the channel's residues previously involved in the hydrophobic interface in the deactivated state (M1582, F1583 and L1584) now lie below the interface and do not interact with the toxin (**Figure 5B, panel ii**).

**b. Molecular Dynamics Simulations on PTx2 Interactions with hNa$_V$1.7 VSD II and IV**

We conducted molecular dynamics (MD) simulations on the docked PTx2 – hNa$_V$1.7 VSD structures to assess their stability in an environment more closely resembling physiological conditions. These simulations, carried out in three independent 1-μs long replicates, were instrumental in quantifying the energetics of the binding process and investigating the impact of lipids on binding. The simulations showed converging behavior during these timeframes, the details of which are discussed further below. To analyze the simulations, we utilized the atomic coordinates from the entire MD simulation trajectories, excluding 90-ns-long equilibration phases. From this data, contact maps were generated, revealing the residues involved in PTx2 – hNa$_V$1.7 VSD interactions, along with the type (e.g., hydrophobic, hydrogen bond, salt bridges, π-stacking, or cation-π) and duration of these interactions. The

criteria used to quantify these interactions are indicated in **Supplementary Table S1**. The contact maps are displayed in **Supplementary Figures S1-S4**.

**PTx2 – hNav1.7 deactivated and activated VSD II:** Throughout the simulation, PTx2 consistently remained bound to VSD II by lodging itself in a cleft located between the S1-S2 and S3-S4 loops, as depicted in **Figure 6A, D**. Hydrophobic contacts constitute the majority of interactions between the toxin and deactivated or activated VSD II (57 ± 4% or 55 ± 5%), followed by hydrogen bond formation (30 ± 2% or 29 ± 6%) and salt bridges (13 ± 2% or 16 ± 8%), as shown in **Figure 6B, E**.

The backbone RMSD profiles, as depicted in **Figure 6C** and **6F**, suggest that the toxin underwent only minor structural changes during the 1-μs simulations, approximately between 1 to 2 Å RMSD. In contrast, the VSDs underwent more considerable structural alterations during the simulations, about 3 to 4 Å RMSD. The overall RMSD for the complex varied between 3 to 5 Å, with deactivated VSD II systems being on the lower end and the activated VSD II systems being on the higher end. These instabilities could be due to their disconnection from the rest of the ion channel in addition to the flexibility of the loops as well as the pre-S1 helix. However, the RMSD of the toxin-binding region, defined as the area within a 6 Å radius (max distance for interaction detection) of where the toxin interacts with the VSD, experienced only minor deviations from the initial structures, with RMSDs between approximately 1 to 2.5 Å.

**Supplementary Figures S1 and S2** present the contact maps, averaged over three replicas, showcasing the interactions between PTx2 and VSD II in the deactivated and activated states, respectively, along with any interactions with lipids. In both cases, PTx2 formed hydrogen bonds, hydrophobic interactions, and salt bridges primarily with residues on the S3-S4 loop of VSD II (including E811, L812, F813, L814, A815, D816), as depicted in **Figure 7**.

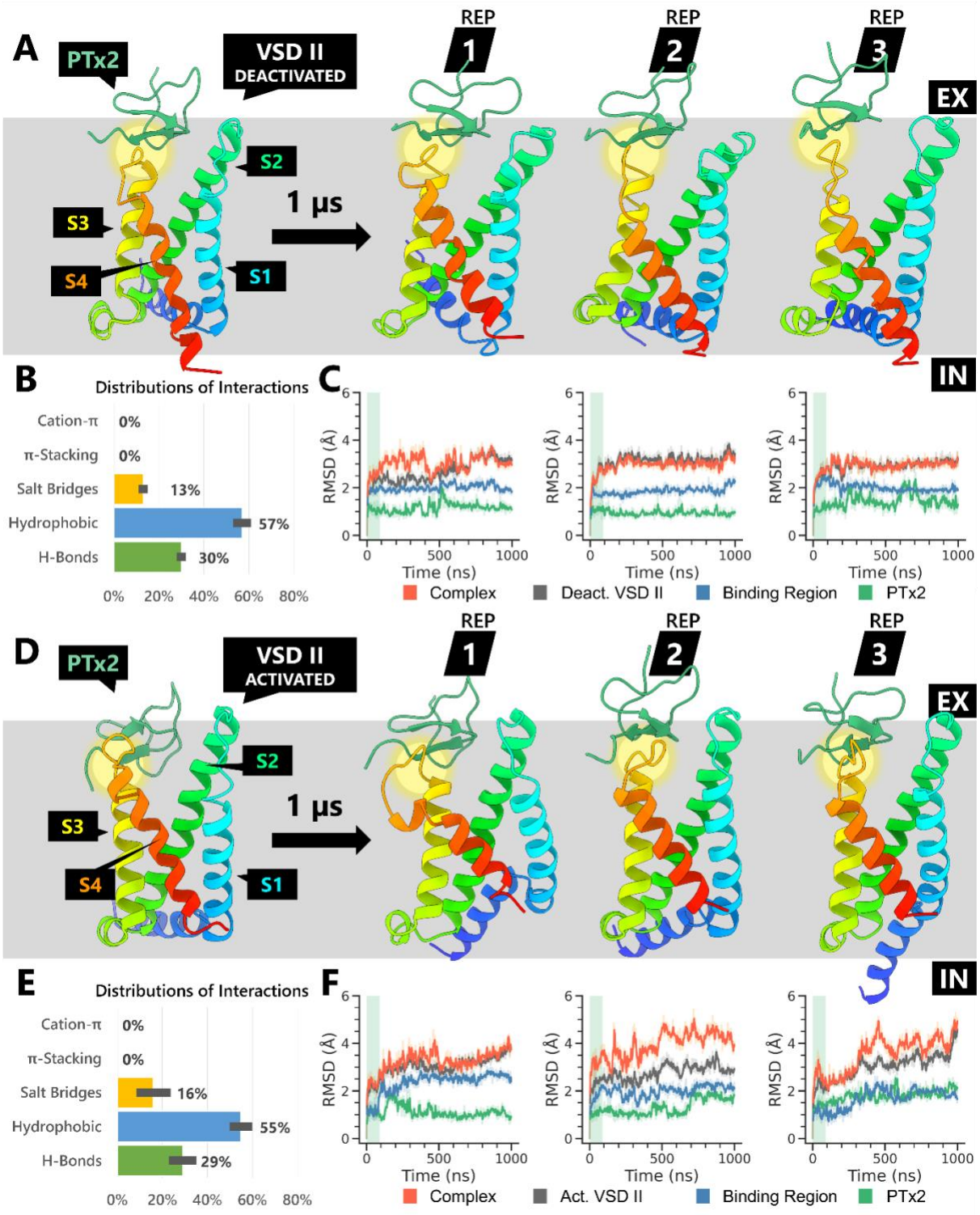



**Figure 6.** ***General overview of the three replicas (REP) of 1-µs molecular dynamics (MD) simulations of PTx2 – hNav1.7 VSD II in the deactivated (A, B, C) and activated states (D, E, F).*** *The lipid membrane is represented as a gray box, with the extracellular- and intracellular-*

*facing sides labeled as "EX" and "IN" respectively. PTx2 – VSD binding sites are highlighted in yellow circles. A, D) Conformational changes of the toxin – VSD complex following the 1-μs long MD simulation for each replica. B, E) Distribution of all toxin – VSD non-bonded interaction types encountered during the simulations, computed by dividing the sum of a specific interaction type by the sum of all interactions across all MD simulation frames. C, F) Time series of the root-mean-square deviation (RMSD) of PTx2 – VSD complex, VSD, toxin's binding region, and PTx2 backbone atoms compared to the initial structures. The plotted values represent the moving averages calculated over intervals of 5 ns each. The MD simulation equilibration period is shown as a light green box.*

In most cases, the root-mean-square deviation (RMSD) of the PTx2 binding site showed stabilization during the latter half of the simulation (≥ 500 ns). Based on this observation, we classified a contact between two residues as dominant if the average contact duration derived from three replicas, plus standard errors of the mean, exceeds 75% (or 50% for protein-lipid interactions) of the time in this latter half of the simulation. Our subsequent analyses primarily centered on characterizing and visualizing these dominant interactions that drove the binding process. That said, we also paid attention to any noteworthy non-dominant contacts that emerge in the study.

During PTx2 binding to the deactivated VSD II, several key interactions were observed. PTx2 K26 formed a hydrogen bond with VSD II E811 carbonyl located on the S3 helix, while K27 backbone engaged in hydrogen bonding with F813 carbonyl and A815 amide backbone groups on the VSD II S3-S4 loop (**Figure 7A**). PTx2's V20 participated in a hydrophobic interaction with VSD II residue A815, and toxin's residue R22 formed a salt bridge with VSD II residue D816, as depicted in **Figure 7C**. Notably, as shown in **Figure 7B**, PTx2 was observed to anchor itself to the membrane through hydrophobic interactions involving residue W24 and the surrounding lipids. Additionally, its residue K27 formed salt bridges with nearby lipid head groups. In addition to these dominant interactions, **Supplementary Figure S1** shows toxin's residues L23 and W24 also sporadically engaged in hydrophobic interactions with VSD II S1-S2 loop residues N763, A766, I767, and L770, effectively creating two anchors that positioned the toxin between the VSD loops.

Intriguingly, when PTx2 binds to the activated VSD II, it exhibited a stronger reliance on interactions with the surrounding lipids rather than VSD II residues. Similar to the previous case, toxin's residue K27 backbone engaged in hydrogen bonding with F813 carbonyl and A815 amide backbone groups on the S3-S4 loop (**Figure 7E**). As shown in **Figure 7D**, the PTx2 T8 side chain hydroxyl group engaged in hydrogen bonding and to a lesser extent, the K4 side chain amino group formed a salt bridge with nearby POPC phosphate ($PO_4$) groups. Moreover, as shown in **Figure 7F**, toxin's residue W30 engaged in hydrophobic interactions with lipids and to a lesser extent, K27 created a salt bridge with POPC $PO_4$ groups near the VSD II S3 location. At the same time, PTx2 residue W7 participated in cation-π interactions with neighboring lipids close to VSD II S2 segment (**Figure 7F**). Such interactions contributed to the formation of hydrophobic anchors that helped orient the toxin binding to the VSD II. Although to a lesser extent than the previous case, some hydrophobic interactions were observed involving PTx2 residues W5 and V20 with VSD II residues F813 and A815 (**Supplementary Figure 2**). Additionally, toxin's R22 sporadically formed salt bridges with VSD II residues E753 (S1-S2 loop) and D816 (S3-S4 loop).

**PTx2–hNa$_V$1.7 deactivated and activated VSD IV:** PTx2 exhibits a dynamic range of binding poses when interacting with VSD IV in the deactivated and activated states. Upon binding the deactivated VSD IV during the simulations, PTx2 progressively adjusted its binding pose, positioning itself deeper into the binding pocket nestled between the S1-S2 and S3-S4 loops as displayed in **Figure 8A**. Hydrophobic contacts formed most of the interactions between PTx2 and the deactivated VSD IV, accounting for 45 ± 5% as demonstrated in **Figure 8B**. These are followed by hydrogen bonds and salt bridges, constituting 32 ± 1% and 22 ± 5% of the interactions respectively. **Figure 8C** depicts considerable fluctuations in the backbone RMSD profiles of the whole complex, the VSD and the toxin's binding region throughout the simulations, varying between 2 to 3.5 Å. This significant variability primarily resulted from the flexible S3-S4 loop's movement as PTx2 embedded itself into the crevice formed between the two VSD loops.

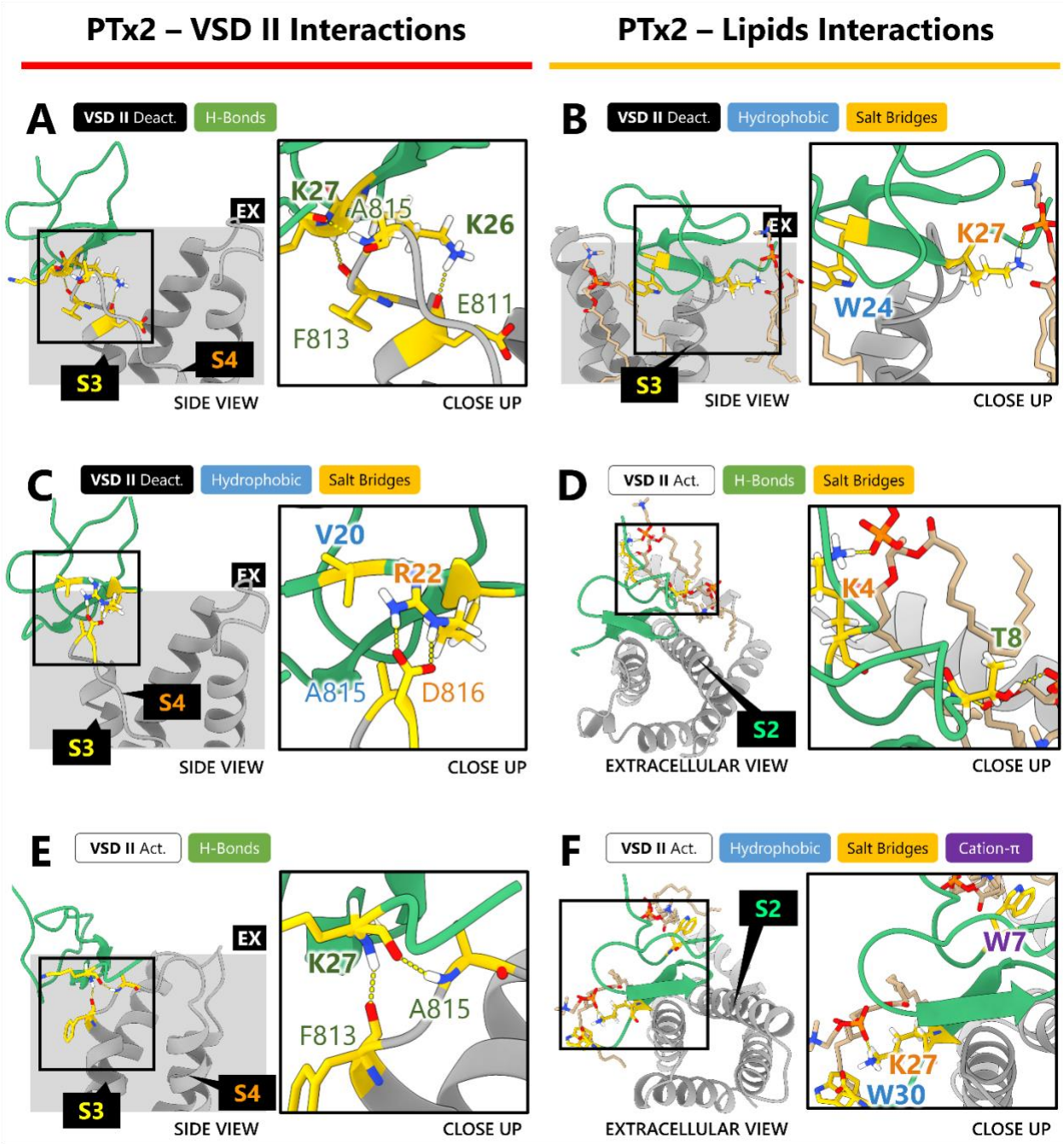



**Figure 7.** ***Dominant non-bonded interactions recorded from molecular dynamics (MD) simulations between PTx2 and hNav1.7 VSD II (A, C, E) or PTx2 and POPC membrane (B, D, F).*** *Residues engaged in binding are highlighted in gold with oxygen atoms participated in those interactions colored in red, nitrogen atoms in blue and hydrogen atoms in white. PTx2 is represented in green, while the VSD is depicted in gray. Each panel features interaction types*

*presented in color-coded boxes whose color corresponds to the involved residues' labels. The structures were visualized at the end of the 1-μs long MD simulations.*

**Supplementary Figures S3 and S4** present the contact maps, averaged from three replicas, showcasing the interactions between PTx2 and VSD IV in the deactivated and activated states, respectively, along with interactions with lipids. Throughout the MD simulations, PTx2 was bound to the deactivated VSD IV via multiple interactions with the residues on the S3-S4 loop of the VSD. As depicted in **Figure 9A**, PTx2's K27 formed several hydrogen bonds with F1583 carbonyl and A1585 amide backbone groups, both situated at the VSD IV S3 helix's terminus where the S3-S4 loop commences. Toxin's residues K26 and K28 established salt bridges with VSD IV residues E1589 and D1586, respectively, located higher on the S3-S4 loop, as indicated in **Figure 9B**. Additionally, PTx2's R22 contributed to salt bridge formation with D1586, albeit to a lesser extent than K28 (**Supplementary Figure S3**).

**Figure 9E** illustrates how the toxin bound to the deactivated VSD IV embedded itself within the membrane. PTx2 residues K4 and K27 generated salt bridges with lipid headgroups positioned proximal to the toxin's binding region on the S3-S4 loop. At the same time, PTx2's W30 initiated hydrophobic interactions with lipids, effectively functioning as a membrane anchor.

In addition to these dominant interactions, PTx2's K26 was observed to form a hydrogen bond with VSD IV residue G1581, located lower in the S3 helix (**Supplementary Figure S3**). Moreover, a limited number of interactions of PTx2 with the S1-S2 VSD IV residues were recorded. This includes hydrogen bonding between toxin's and VSD IV residues T8 and E1534, hydrophobic interactions of L23 and W24 with Y1537 and V1541, respectively, and the formation of a salt bridge between R13 and E1534 as shown in **Supplementary Figure S3**.

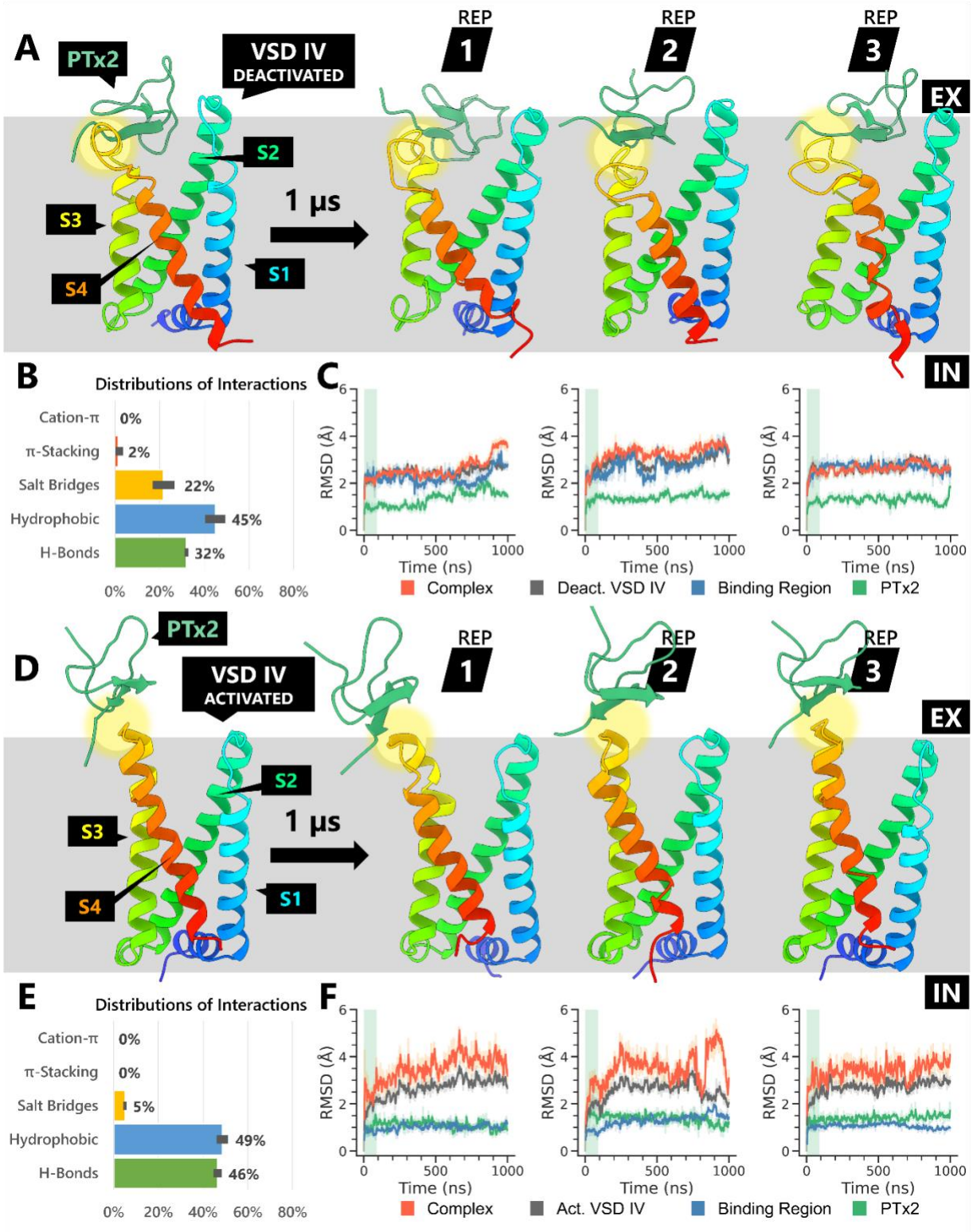



**Figure 8.** ***A general overview of the three replicas (REP) of 1-µs molecular dynamics (MD) simulations of PTx2 – hNav1.7 VSD IV in the deactivated (A, B, C) and activated***

***states (D, E, F).*** *The lipid membrane is represented as a gray box, with the extracellular- and intracellular-facing sides labeled as "EX" and "IN", respectively. PTx2 – VSD binding sites are highlighted in yellow circles. A, D) Conformational changes of the toxin – VSD complex following the 1-μs-long MD simulations for each replica. B, E) Distribution of all non-bonded toxin – VSD interaction types encountered during the simulations, computed by dividing the sum of a specific interaction type by the sum of all interactions across all simulation frames. C, F) Time series of the root-mean-square deviation (RMSD) of PTx2 – VSD complex, VSD, toxin's binding region, and PTx2 backbone atoms compared to the initial structures. The plotted values represent the moving averages calculated over intervals of 5 ns each. The MD simulation equilibration period is shown as a light green box.*

Upon binding with the activated VSD IV, as illustrated in **Figure 8D**, the toxin ascended above the membrane, thereby making only a limited number of interactions with lipids. As represented in **Figure 8E**, the interactions established by the toxin with the activated VSD IV were predominantly hydrophobic (49 ± 3%) and hydrogen bonds (46 ± 2%), with salt bridges making up a minor fraction (5 ± 1%). Throughout the MD simulations, the activated VSD IV underwent subtle conformational adjustments, as evidenced by a ~2-3 Å RMSD relative to the initial structure, shown in **Figure 8F**. However, the toxin's binding region remained largely unchanged, exhibiting a minor ~1.5 Å RMSD shift – a fluctuation nearly identical to that of the toxin itself. Interestingly, the RMSD of the complex displayed a greater shift, ranging from 3 to 4.5 Å. Although the structures of the PTx2 and VSD themselves did not exhibit any significant changes, a rotation in the toxin binding position (**Figure 8D**) resulted in the higher complex's RMSDs observed (**Figure 8F**).

With PTx2 situated atop the activated VSD IV S3-S4 loop, as depicted in **Figure 9C**, the toxin's K26 side chain amino group formed a hydrogen bond with E1589 backbone carbonyl group of the residue E1589 located on the S3 helix of VSD IV. Concurrently, toxin's K27 backbone amide and carbonyl groups established hydrogen bonds with Y1591 carbonyl and V1593 amide backbone groups (**Figure 9C**), respectively, both positioned at the start of the VSD IV S3-S4 loop. Furthermore, PTx2's V20 engaged in a hydrophobic interaction with VSD IV residue V1593, as visualized in **Figure 9D.**

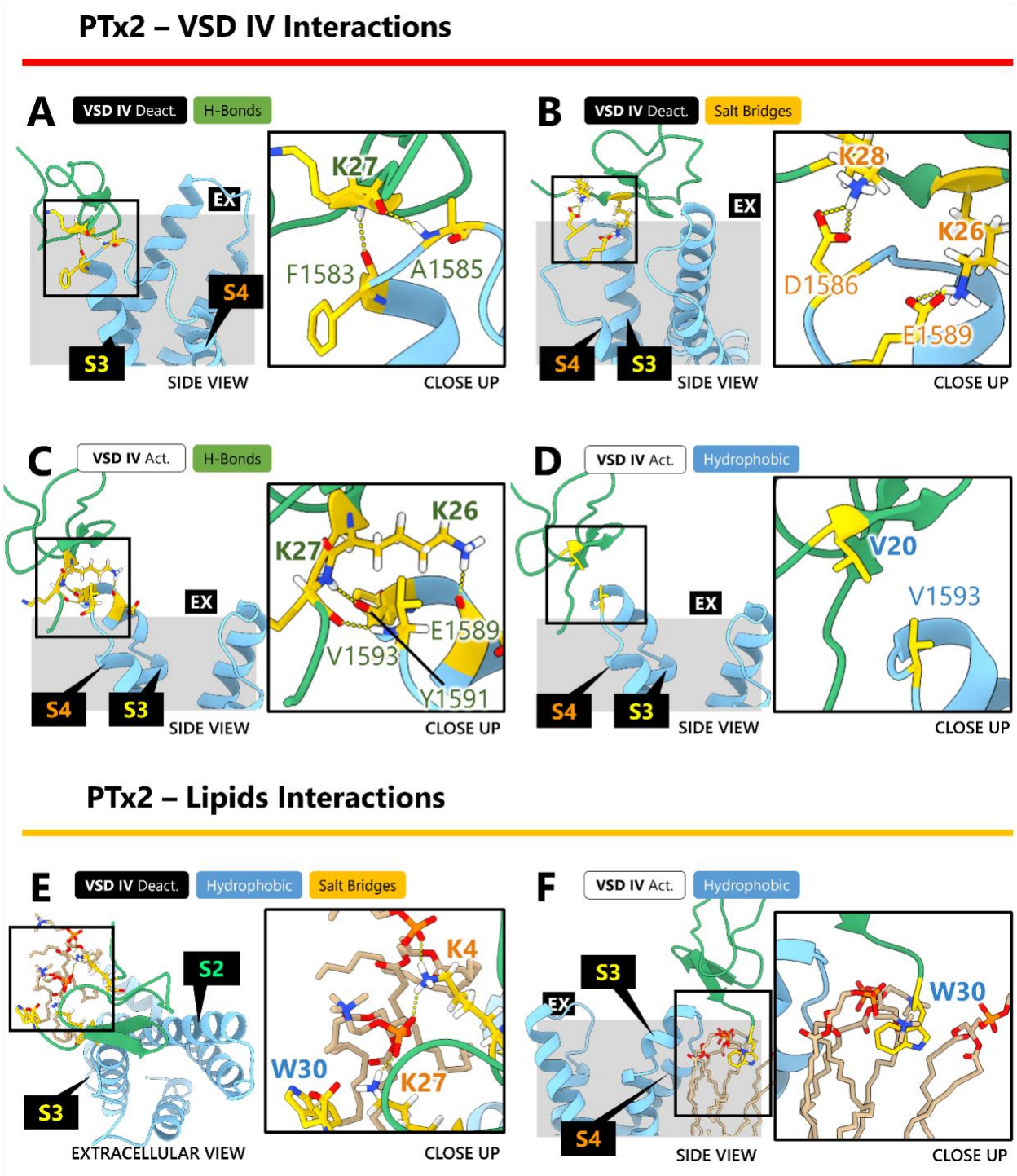



**Figure 9.** ***Dominant non-bonded interactions recorded from molecular dynamics (MD) simulations between PTx2 and hNav1.7 VSD IV (A, B, C, D) or PTx2 and POPC membrane (E, F).*** *Residues engaged in binding are highlighted in gold with oxygen atoms participated in those interactions colored in red, nitrogen atoms in blue and hydrogen atoms in white. PTx2 is*

*represented in green, while the VSD is depicted in light blue. Each panel features non-bonded interaction types presented in color-coded boxes whose colors correspond to the labels of involved residues. The structures were visualized at the end of the 1-μs-long MD simulations.*

Despite PTx2's significant elevation above the membrane, which resulted in substantial exposure to the solvent, its C-terminal tail reached down into the lipid bilayer. Here, toxin's residue W30 engaged in hydrophobic interactions with nearby lipids in proximity to the VSD IV S3-S4 loop (**Figure 9F**), maintaining sufficient toxin's binding stability. Additionally, marginal hydrophobic contacts involving PTx2's residue W5 and residue Y1591 on the VSD IV S3-S4 loop were detected as shown in **Supplementary Figure S4**.

**c. Analysis of Binding Energetics**

We calculated the free energy of binding between PTx2 and different states of hNa$_V$1.7 VSD II and IV using the implicit-solvent MM/PBSA methodology[55], based on our all-atom MD simulation trajectories. We analyzed each component of the free energy by block averages, taken during the latter half of each simulation (500 – 1000 ns) when binding stability was typically achieved as assessed by looking at the RMSD profiles (see **Figure 6** and **Figure 8**), in addition to the enthalpy and the interaction energy time series across all three replicas (**Supplementary Figure S12**). These block averages as well as standard errors of the mean were calculated from all three MD simulation replicas of each system, as illustrated in **Figure 10**.

The binding of PTx2 to the activated VSD II and deactivated VSD IV induced the most favorable binding enthalpy ($\Delta H_{MM/PBSA}$). In contrast, PTx2's binding to the activated VSD IV, marked by fewer intermolecular interactions in MD simulations, was characterized by the least negative enthalpy. When accounting for the entropic contributions ($-T\Delta S$), the resulting binding free energy of PTx2, $\Delta G_{bind}$, was the most favorable for the deactivated VSD II, with a value of -23.1 ± 0.9 kcal/mol. This was followed by the activated VSD IV (-21.1 ± 1.0 kcal/mol), deactivated VSD IV (-20.1 ± 1.6 kcal/mol), and activated VSD II (-18.5 ± 2.5 kcal/mol) in descending order of favorability. Through paired t-tests, we established that a statistically significant difference in $\Delta G_{bind}$ exists ($p < 0.05$) between PTx2 binding with the deactivated VSD II compared to the activated VSD II and the deactivated VSD IV, respectively.

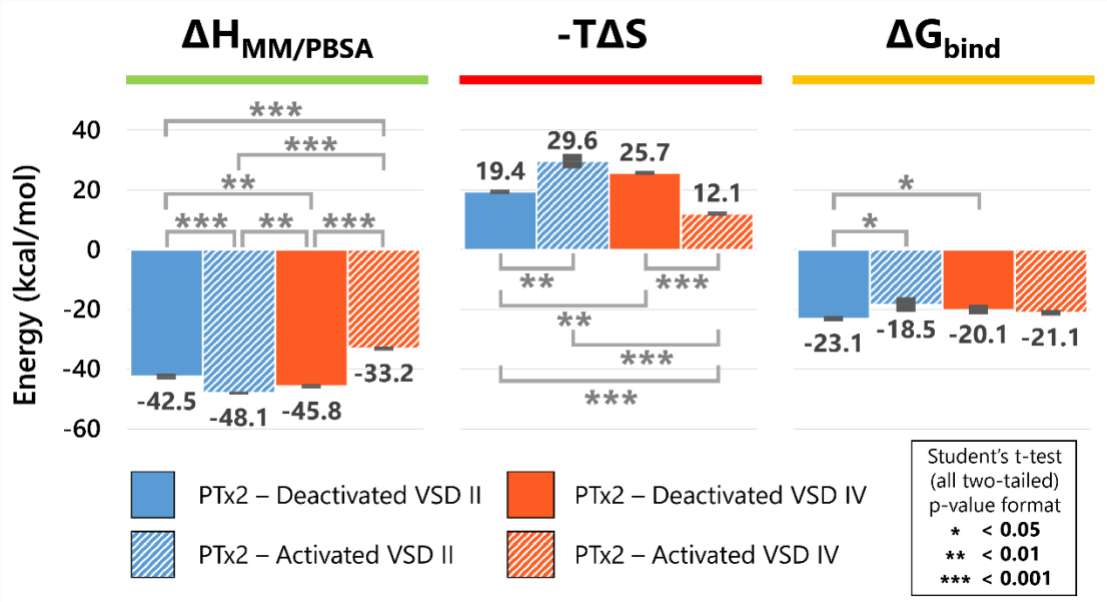



**Figure 10.** ***MM/PBSA binding energies in PTx2 – hNav1.7 VSD systems averaged over three replicas from the second half (500 ns onward) of the all-atom molecular dynamics simulations.*** *The optimal block sizes for the systems were selected to minimize the standard errors of the mean shown as error bars. Statistical significance was determined using unpaired two-tailed Student's t-test.*

Our MM-PBSA calculations, showing that PTx2 binds more favorably to the deactivated state of VSD II, align with previous electrophysiology experiments[19,22]. Those studies demonstrated that PTx2's antagonism of $Na_V1.7$ occurs with around 60-fold reduced potency when the membrane potential is held at a voltage that favors VSD activation[22]. Moreover, another research study suggested that the estimated $IC_{50}$ for PTx2's inhibition of $Na_V1.7$ activation (namely, binding to the deactivated VSD II) is roughly 400-fold lower than the observable $IC_{50}$ for the inhibition of inactivation (that is, binding to the deactivated VSD IV)[19], corresponding to a ~3.7 kcal/mol free energy difference in good agreement with our MM-PBSA estimate of 3.0 ± 1.8 kcal/mol.

The time series of MM/PBSA computed enthalpy and interaction energies are shown in **Supplementary Figure S12**. Among the four studied systems, the PTx2 – activated VSD IV system displayed the most stability while the remaining systems showed some degrees of variability across multiple replicas. These variances could stem from the binding mode of

PTx2 to the VSD in these cases where the toxin wedged into the space between the two VSD loops. This insertion process caused bending of the S3-S4 loop for the deactivated and activated VSD II as well as deactivated VSD IV systems (as shown in **Figures 6A, 6D, 8A**), where the attachment occurs, leading to a pattern of contact disruptions and formations. In contrast, the PTx2 – activated VSD IV system displayed the most stability, likely because PTx2 rose above the membrane and engaged in fewer interactions with the activated state of VSD IV (as shown in **Figure 8D)**. The limited interactions with the toxin helped preserve the VSD's structural integrity and kept the binding area largely unaltered.

Then we conducted a linear interaction energy analysis on our all-atom MD simulations using Amber20[44,59] and the same CHARMM force field parameters as used in the simulations. This allowed us to calculate the electrostatic and van der Waals interaction energies between PTx2 and VSD II/IV residues as well as explicit POPC lipids (not possible with the implicit-membrane MM/PBSA approach), and as depicted in **Figure 11** (for a more exhaustive depiction of specific interactions involving additional residues, refer to **Supplementary Figures S5-S7**).

**Figure 11A** shows the top 15 PTx2 residues that exhibit the most favorable total interaction energies with activated and deactivated VSD II/IV residues and lipids combined (refer to **Supplementary Figure S5** to see the full list). The interaction energy for each toxin residue with different VSD states displays a variety of patterns. However, the general trend suggests that most residues have stronger interactions (i.e., more negative interaction energy values) with both the VSD II and VSD IV in their deactivated states compared to their activated states. This could imply that the electrostatic environment of the VSDs in their deactivated state is more favorable for binding those residues. From an energetic perspective, PTx2's residues R22, K26, K27, K28, and W30 are the main contributors to PTx2 – VSD II and IV binding through the formation of hydrogen bonds, hydrophobic interactions, and salt bridges. The most apparent difference is observed in the case of PTx2 binding to the activated state of the VSD IV as the binding pose excludes many toxin residues from forming interactions with the VSD.

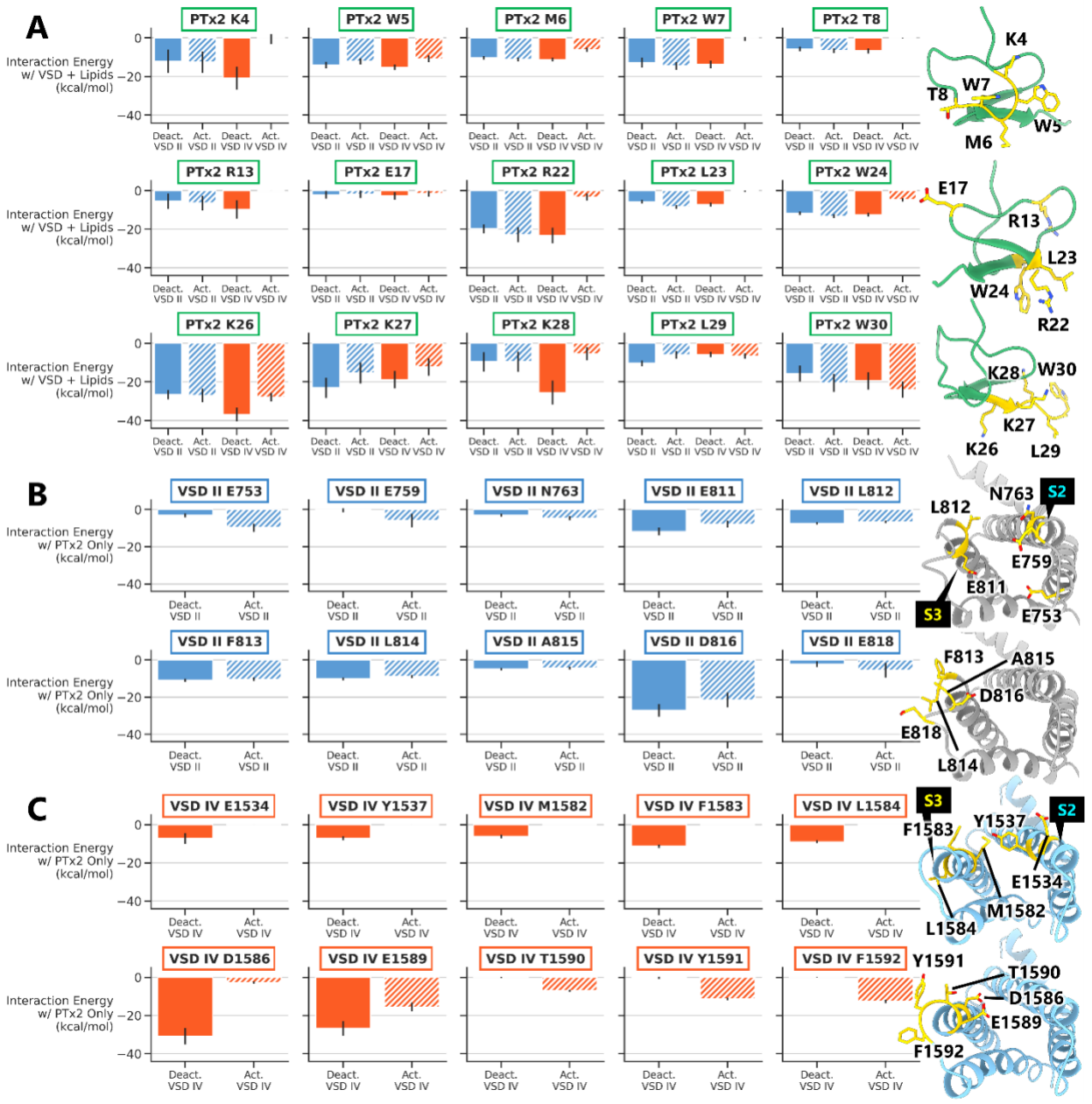



**Figure 11.** ***Interaction energies contributed by each residue toward the binding process for different PTx2 – hNav1.7 VSD MD systems.*** *Residues involved in protein-protein interactions are visualized in yellow on the corresponding structures to the right of each row. The VSD structures are viewed from the extracellular side. The top 15 residues for PTx2 (A) and top 10 residues for VSD II (B) or IV (C) with the most favorable interaction energies (van der Waals + electrostatic) averaged from all three MD replicas are displayed.*

The PTx2's residue K4 also exhibits a significantly more favorable interaction energy when binding to the deactivated state of VSD IV compared to other cases. This can be attributed to the specific binding pose of the toxin in this state, which positions K4 favorably for forming salt bridges with surrounding lipids. The interaction energies with lipids depicted in **Supplementary Figure S1-S4** revealed the main PTx2 – lipid membrane interaction surface, highlighting the important roles played by toxin residues K4, W5, M6, W7, K27, and W30 in anchoring PTx2 to the cell membrane. The tight binding of PTx2 to the outer leaflet of the cell membrane seems to be critical for the inhibition of PTx2 on hNa$_V$1.7 channels, which was supported by based a mutagenesis study[26] that produced W5Y, W7Y, W24Y, and W30Y PTx2 mutants. In that experimental study, the mutations to tyrosine were performed since tyrosine is more polar than tryptophan and is less suited to bind at the water – lipid bilayer interface[61]. These W-to-Y mutations led to 6- to 291-fold reduction in potency in the inhibition of hNa$_V$1.7[26], indicating the importance of these residues for PTx2 binding to hNa$_V$1.7 in line with our MD simulation predictions.

Despite their significant roles in binding to both states of VSD II and the deactivated state of VSD IV, PTx2 residues K4, W7, T8, R13, R22, and L23 did not form substantial interactions when binding to the activated state of VSD IV as shown in **Figure 11A**. This lack of significant interactions could be attributed to the orientation of the bound toxin, which positioned these residues at a considerable distance from potential binding partners. As a result, their contribution to the binding process is diminished or negligible in this state.

**Figure 11B** provides an overview of the interaction energies between notable VSD II residues with PTx2. When comparing the activated and deactivated states, minor variations are observed in terms of the interaction energies of VSD II residues involved in the binding of PTx2. Notably, residues E811 and D816, located on the S3-S4 loop, exhibit slightly more favorable interaction energies with PTx2 when VSD II is in the deactivated state compared to the activated state. Conversely, residues E753 and E759 on the S1-S2 loop, as well as E818 on the S3-S4 loop, show stronger interactions with PTx2 when VSD II is activated compared to when it is in the deactivated state.

Interestingly, **Figure 11C**, which illustrates the interaction energies of notable VSD IV residues with PTx2, demonstrates a distinct difference in their pattern when PTx2 binds to the activated and deactivated states of VSD IV. This disparity in interaction energies can be attributed to the positioning of PTx2 within the membrane. In the deactivated state (**Figure 8A**), PTx2 binds deeper within VSD IV cleft, leading to a stronger interaction with the VSD residues. Conversely, in the activated state, PTx2 binds more superficially to VSD IV, resulting in fewer interactions and primarily involving VSD residues located at the top of the S3-S4 loop (**Figure 8D**). Additionally, the analysis of PTx2's interactions with the lipid membrane reveals that when binding to VSD IV in the activated state, PTx2 exhibits significantly fewer interactions with the lipid membrane compared to the deactivated state (**Figure 9F** vs. **Figure 9E**). In conclusion, the variations in interaction energies and the involvement of specific VSD residues, as well as the differential interactions with the lipid membrane, contribute to the distinct binding characteristics of PTx2 to VSD IV in different states.

**d. MD Simulations with Activated VSD II and IV as a Part of the Full hNa$_V$1.7 Channel**

While the toxin binding regions remained largely stable, structural instability of the rest of the VSDs was observed in certain simulations, likely due to their detachment from the full ion channel's structure. Control MD simulations with the full hNa$_V$1.7 channel were conducted to evaluate the toxin – VSD complex stability and the subsequent effect on their binding. Given the unavailability of complete hNa$_V$1.7 or homologous ion channel structures with deactivated VSDs, our simulations were limited to activated VSD II and IV integrated into the full hNa$_V$1.7 channel.

As shown in **Supplementary Figure S8**, analysis of three 300-ns full-channel simulation replicas showed improved stability in VSDs, with backbone RMSDs relative to the starting structures of around 2-3 Å for activated VSD II and IV, versus 3-4 Å for isolated VSDs (see Figures 6F and 8F). The toxin's binding region RMSDs stabilized to around 2 Å (VSD II) or 1-1.5 Å (VSD IV) by the end of the MD simulations, comparable to simulations with isolated VSDs.

**Supplementary Figure S9** offers an analysis of PTx2's positional variations relative to the VSD in scenarios where the VSD was either isolated or part of the full channel. This was done by calculating the RMSD of PTx2's backbone atoms when the trajectory was aligned to the VSD. Overall, PTx2's binding to VSD II and IV in their activated state was characterized by greater fluctuations. This could be attributed to multiple factors. For instance, when PTx2 interacted with the activated VSD II, it engaged more with surrounding lipids than when binding to the deactivated state; the lipids' inherent mobility could contribute to PTx2's positional instability. In the case of PTx2 binding to the activated VSD IV, the toxin was observed to rise well above the membrane surface, resulting in fewer VSD contacts and increased exposure to the aqueous environment, which could explain the noted fluctuations. Despite these movements, the RMSD of PTx2 generally stabilized as the simulations progressed, although there were exceptions such as replica 1 of the deactivated VSD IV and replica 2 of the activated VSD IV – both involving isolated VSD systems. Nonetheless, these instances of fluctuation appeared not to compromise the toxin's binding, as the time series of interaction energies for these replicas showed minimal disruption, as shown in **Supplementary Figures S12C** and **S12D**.

**Supplementary Figures S9B** and **S9D** present a comparison of the RMSD time series data for PTx2 in its interaction with both isolated and non-isolated (attached to the full ion channel) activated VSD II and IV systems. When bound to the activated form of VSD II, PTx2 showed only slight positional deviations, regardless of whether the VSD was part of the full ion channel or isolated. In contrast, PTx2's binding to the activated VSD IV was more stable when the VSD was part of the full channel than in the isolated configuration. This finding is particularly notable given that the activated VSD IV generally displayed more structural rigidity due to its fewer interactions with the toxin in both scenarios. Nonetheless, in the isolated VSD scenario, PTx2 underwent a subtle rotational movement at the binding site (illustrated in **Figure 8D**), a phenomenon not observed when the VSD was integrated within the full channel (see **Supplementary Figure S8D**).

Further examination of the MD simulation trajectories revealed that when the activated VSD IV was isolated, the toxin's high exposure to the solvent and minimal VSD binding led to the toxin moving around and causing the isolated VSD's S3 and S4 helices to sway along (as

shown in Figure 8F, with an RMSD for the VSD of approximately 3 Å). Conversely, when the activated VSD IV was part of the complete channel, the movement of the VSD's S4 was restricted (as indicated in **Supplementary Figure S8F**, with an RMSD for the VSD of about 2 Å), resulting in the toxin maintaining its initial binding orientation without significant deviation.

**Supplementary Figures S10-S11** depict binding interactions between the toxin and the full-channel activated VSD II and IV during the latter half the simulations when the protein structures demonstrated stability. These results largely align with those from the isolated VSD simulations. Specifically, for VSD II, PTx2's K27 hydrogen bonded with VSD's F813 and A815, W5, V20, and K26 had hydrophobic interactions with VSD S3-S4 loop residues, and L23, W24 with VSD S1-S2 loop residues. PTx2's R22 and K28 formed salt bridges with VSD's D816 and E818, respectively. PTx2's interactions with lipids included T8 in hydrogen bonding, W30 in hydrophobic interactions, K4 and K27 in salt bridges, and W7 in cation-π interactions. In VSD IV, hydrogen bonds included K26-E1589 and K27-Y1591/V1593 pairs, hydrophobic interactions involved PTx2's W5, V20, K26, K27 with VSD's Y1591, F1592, and V1593, and a salt bridge formed between PTx2's K27 and VSD's E1589. W30 also interacted with lipids.

Minor differences were observed compared to the isolated VSD simulations. For example, in the VSD II case, the toxin's K26 formed a salt bridge with VSD II's D816, whereas R13 formed a salt bridge and W30 engaged in cation-π interactions with lipids - forming interactions not seen in the isolated VSD case. In the VSD IV case, in addition to W30, W5 also formed hydrophobic interactions with lipids, unlike in the isolated VSD scenario.

**Supplementary Figure S13** details the calculation of MM/PBSA binding energies for PTx2 with the activated VSD II and IV attached to the hNa$_V$1.7 channel. Even though the PTx2 – activated VSD II interaction showed a more favorable binding enthalpy, $\Delta H$, there was no notable distinction in the overall free energy of binding, $\Delta G$, when comparing the activated VSD II with the activated VSD IV. This aligns with the energy computations for the isolated VSD systems shown in Figure 10, which indicate similar binding free energies of the toxin for both VSD II and IV in the activated state.

In conclusion, integrating the VSDs into the full channel improved stability, although the toxin binding interaction outcomes largely mirrored those observed in isolated VSD MD simulations. As we do not have the full hNa$_V$1.7 structure with VSDs in the deactivated state, our study primarily used isolated VSDs, whose structures we have in both the activated and deactivated states, to allow for a consistent comparison.

## 3.4) Discussion

### a. PTx2 Interactions with hNa$_V$1.7 VSD II and IV in Different States

Using existing experimental data as a guide, we generated models of PTx2 bound to human Na$_V$1.7 VSD II and IV in both the activated and deactivated states. These models accurately reproduced the crucial toxin – ion channel's VSD II interactions, as previously elucidated through mutagenesis experiments and structural studies involving chimeric ion channels [17,22]. Our findings also shed new light on PTx2 interactions with VSD IV, unveiling substantial differences compared to its interactions with VSD II.

MD simulations can provide a dynamic perspective on toxin-receptor interactions, transcending the static results from experimental structures and docking. This is especially important when the binding site is surrounded by lipids, making the fluid nature of lipid membranes play a crucial role. As lipids are continually in motion, their interactions with the toxin can significantly influence and even shift the binding of the toxin to the receptor, a phenomenon not captured in static structures or even typical docking runs but evident in MD simulations. Indeed, we discovered several new insights into PTx2 – hNa$_V$1.7 interactions by performing the MD simulations. For example, initially, PTx2's residue K27 showed no interaction with VSD II in both states, and M6 interacted with hydrophobic VSD II residues. However, after 1 μs of the MD simulation, K27 extensively interacted with VSD II, while M6's interactions diminished. In the case of PTx2 binding to deactivated VSD IV, MD revealed PTx2's residue K28 as the dominant contact with VSD's D1586, overshadowing the initially observed PTx2's R22 interaction. For the activated VSD IV, interactions involving PTx2's K27 and V20 emerged only in MD simulations. Additionally, MD runs can detail the roles of individual residues in interaction energies, deepening our grasp of the binding

mechanics. This detailed perspective, highlighting the influence of dynamic lipids, is vital for understanding molecular interactions and subsequent therapeutic peptide design.

Based on our atomistic structural modeling and MD simulations we established that PTx2 primarily binds to VSD II in different conformational states through interactions with residues on the S3-S4 loop of the VSD II. In both states, PTx2's residue K27 formed hydrogen bonds with the channel's residues F813 and A815 located on the S3-S4 loop. In the deactivated state of VSD II, a hydrogen bond is established between residue K26 of PTx2 and residue E811 of VSD II. On the contrary, when VSD II is activated, K26 interacted with VSD II residue L812 instead, and this interaction was less frequent and more variable. Similarly, PTx2's R22 established a more robust and consistent salt bridge with D816 (on the S3-S4 loop) when bound to the deactivated VSD II state compared to the activated state, as it occasionally shifted its interaction to E753 (on the S1-S2 loop). Also, PTx2 exhibited a hydrophobic interaction via residue L29 with channel's residue L814 in the deactivated state of VSD II, an interaction not observed in the activated state.

PTx2 interacts with VSD IV primarily through residues on the VSD IV S3-S4 loop, but significant differences were observed in the toxin's binding to the deactivated and activated states of VSD IV. For example, PTx2's K26 formed a hydrogen bond with G1581 (located in the S3-S4 loop) in the deactivated state and with E1589 (also in the S3-S4 loop) in the activated state of VSD IV. In the activated state, K26's interaction with E1589 was more frequent and consistent, suggesting stronger binding. Another differential interaction can be seen with the toxin's residue K27. When binding to the activated VSD IV, K27 interacted with three different residues in the S3-S4 loop, Y1591, V1593, and F1592, while in the deactivated state it interacted mostly with F1583 and A1585 in S3-S4 loop only.

### b. PTx2 Interactions with Lipids when Binding to hNa$_V$1.7 VSD II and IV

A significant advancement highlighted in this study is the novel atomistic understanding of PTx2's interactions with surrounding POPC membrane lipids upon binding to the Na$_V$1.7 channel. Lipid interactions play a vital role in stabilizing the position and orientation of PTx2 for binding to VSD II and IV in different conformational states. These novel insights from

simulations could open a new avenue for enhancing PTx2 binding to the channel by optimizing toxin's interactions with nearby membrane lipids [26].

Upon PTx2's binding to the activated VSD II, PTx2's residue T8 formed a hydrogen bond while K4 and K27 formed salt bridges with nearby POPC $PO_4$ groups. W7, W24, and W30 engaged in hydrophobic interactions with the lipids' fatty acid tails, and W7 formed cation-π interactions with POPC choline group. When PTx2 binds to the deactivated VSD II, PTx2's W5 and W24 engaged in hydrophobic interactions while K4 and K27 formed salt bridges with POPC $PO_4$ groups. T8 still engaged in hydrogen bonding with nearby lipids, but no hydrophobic interactions originating from W7 and W30 were observed unlike the previous case. However, PTx2's R13 was observed to interact sporadically with lipids through salt bridge formation.

When PTx2 binds to the activated VSD IV, the only interaction with lipids was a hydrophobic one originating from W30, which was stronger and more consistent than when PTx2 binds to the deactivated state. However, when PTx2 binds to the deactivated VSD IV, salt bridges with POPC head groups were formed by toxin's residues K4 and K27 (and with R13 and K28 to a much lesser extent). It is noteworthy that K27 – POPC salt bridge interactions were also substantially more frequent and consistent in the deactivated VSD IV than in the deactivated VSD II. Hydrophobic interactions with POPC were observed not only from W30, like the activated VSD IV case, but also from W5 and W24, which are unique to the deactivated VSD IV case.

In summary, this first view into the lipid modulation of toxin binding may constitute a novel and emerging area of research for optimizing peptide toxin's specificity and selectivity.

**c. Energetics of PTx2 Binding**

Lastly, we complemented the data with specific binding energetic contributions from individual toxin's and ion channel's residues in different VSD conformational states, enabling the identification of potential mutation candidates to enhance or weaken toxin's state-specific binding. This information can support the design of toxin analogs with enhanced potency, holding promise as a way to identify candidates for the development of innovative

treatments for chronic pain. Our analysis of PTx2 binding energetics to different states of VSD II and IV indicates that PTx2 binds most favorably to the deactivated VSD II, followed by the activated VSD IV. A substantial difference in the free energy of PTx2 binding is observed when comparing the deactivated VSD II to both the activated VSD II and the deactivated VSD IV. PTx2 binding to and trapping VSD II in its deactivated state can prevent the opening of the hNa$_V$1.7 channel. Likewise, trapping VSD IV in its activated state can keep the channel in a non-conducting, inactivated state. The net effect is a decrease in the $Na^+$ current flowing through the Na$_V$1.7 channels when PTx2 is introduced, which aligns with experimental results[19]. From a therapeutic standpoint, optimizing PTx2's interactions with these states could enhance its effectiveness in pain therapy through hNa$_V$1.7 channel inhibition. Furthermore, experimental findings demonstrate that the efficacy of PTx2 as an inhibitor of hNa$_V$1.7 is significantly influenced by its binding to the lipid membrane[26]. Therefore, improving PTx2 interactions with cell membrane lipids could also augment PTx2's inhibition of hNa$_V$1.7.

Based on our analysis in **Figure 11A** and **Supplementary Figure S5**, we have identified certain residues of PTx2 that interact with either the activated and deactivated states of VSDs and lipids but exhibit relatively weak interaction energies. For residues that primarily interact with membrane lipids (T8, R13, E17), introducing mutations that enhance their hydrogen bonding interactions with lipids for T8, remove negative charge for E17 or diminish preferential state-specific salt bridge formation for the deactivated VSD IV system could be beneficial. V20 interacts exclusively with the VSDs through hydrophobic interactions but has remarkably low interaction energy despite its relatively long contact duration. Therefore, a non-conservative mutation could be considered. A recent mutagenesis study achieved a more potent and selective peptide by performing the V20R mutation[14].

A promising strategy entails fine-tuning state-specific interactions. Mutations could be made to increase PTx2's affinity for states that decrease $Na^+$ conduction by preventing channel activation or trapping it in an inactivated state (deactivated VSD II and activated VSD IV, respectively) compared to those that facilitate $Na^+$ conduction (activated VSD II and deactivated VSD IV). Residue L23 demonstrated hydrophobic interactions with both the VSDs and the lipid molecules in the membrane. However, it favors binding to the activated

VSD II and deactivated VSD IV which are both undesirable. In addition, K4, R13, K26, and K28 exhibit significantly higher binding energies upon binding to the deactivated VSD IV compared to the other states. This heightened affinity is undesirable as it impedes channel inactivation. Consequently, careful mutation of these residues could serve to increase PTx2's affinity for the desired states of the channel. As an example, a mutagenesis study discovered that K26R and K28E resulted in improved PTx2 potency[14]. Another experimental study also revealed that charge-reversing and charge-neutralizing mutations of PTx2's basic residues R22 and K26 lead to significant loss of potency[22], which was also corroborated by a computer simulation study of PTx2 mutant – hNa$_V$1.7-Na$_V$Ab deactivated VSD II system using free energy perturbation (FEP)[28]. However, this MD simulation method, more rigorous and computationally expensive than *a posteriori* interaction energy analysis used in this work, was unable to predict gain in potency for some charge-preserving K26 and R22 mutations (e.g., K26R and R22noR, where norR is norarginine)[28]. This suggests that using PTx2–hNa$_V$1.7 VSD II and IV systems in different conformational states as we explored in this work might be crucial for a more accurate estimates of peptide toxin mutation effects using FEP or other approaches as will be explored in our follow-up studies.

By strategically introducing mutations to enhance the interactions of these residues, there is a possibility of enhancing the affinity of PTx2 towards the VSDs, their different conformational states and/or lipids. These mutations can be designed to optimize specific interactions, such as salt bridges, hydrogen bonding, electrostatic, hydrophobic and/or other types of interactions, involved in PTx2's binding to the VSDs and lipids as indicated in the contact maps in **Supplementary Figures S1-S4**. Nonetheless, it is vital to balance this enhancement with the potential effects on state specificity and ion channel subtype selectivity for toxin – receptor binding, as well as the net effect on protein's overall structure and stability.

### c. Limitations and Future Directions

Using computational modeling and simulation tools, we can model the atomic-resolution structures and relative energetics of peptide toxin interactions with hNa$_V$1.7 in different conformational states. This allows for the accurate prediction of key binding interactions,

their subsequent experimental validation and design of more potent and selective peptides targeting hNa$_V$1.7 or other ion channels as we did recently[14].

Although the computational workflow in our study can provide valuable insights into the complexities of peptide – receptor interactions at the atomic resolution, several drawbacks are associated with each methodology. Given the absence of complete hNa$_V$1.7 channel structures with VSDs in the deactivated state, our modeling and simulations were limited to the isolated VSDs of the channel. This limitation could potentially introduce instabilities during the MD simulations. Nonetheless, we observed only minor displacements of the PTx2's binding regions during the majority of our simulations as demonstrated by the RMSD profiles shown in **Figures 6 and 8**. We also performed preliminary MD simulations of a full hNa$_V$1.7 channel with VSDs II and IV in the activated state as shown in **Supplementary Figures S8-S11 and S13** demonstrating fairly similar results to the isolated VSD simulations with somewhat higher stabilities (lower RMSDs) and a few additional interactions observed. In the future, we will expand our study by modeling and simulating full hNa$_V$1.7 channel structural models in different conformational states.

Peptide backbone flexibility, especially at the N- and C- termini, is difficult to capture during the protein-protein docking and, therefore, RosettaDock might not be able to sample the full range of possible toxin conformations. Additionally, Rosetta scoring uses an implicit membrane model that might not account for energetic contributions involving membrane lipids that could be important to stabilize the toxin's binding pose. Although a lipid membrane can be explicitly included in MD simulations, simulation accuracy often depends on the quality of the force field and parameters used in the simulation, and it can be difficult to determine the best set of parameters for any given system[62]. Fully atomistic molecular dynamics simulations are limited by their timescale (ns to μs), and hence enhanced sampling MD simulation techniques such as string method with the swarm of trajectories[63], weighted ensemble (WE) [64], Gaussian accelerated MD (GaMD)[65] or their combination[66] need to be employed to observe peptide – protein binding and large-scale protein conformational transitions.

For estimating the free energy of binding, MM/PBSA methodology may suffer from accuracy problems especially for predicting absolute binding free energies as was demonstrated by a study done by Sheng *et al.*[67]. The accuracy can be improved such as by damping the solvated and Coulombic terms for highly charged systems[67,68], using residue type-specific dielectric constants[69], and applying potentially more accurate entropy calculation methods[70]. Yet, MM/PBSA methodology employed here was successfully utilized in our recent study[71] to correctly predict conformation-dependent binding of a small molecule ligand norepinephrine and large stimulatory G protein to beta-2 adrenergic receptor, another integral membrane protein. In order to predict effect of toxin's mutations on state-specific ion channel binding, potentially more accurate but computationally expensive MD simulation technique, free energy perturbation (FEP) [72], can be used, which was recently employed to study relative energetics of PTx2 mutants – hNa$_V$1.7-Na$_V$Ab deactivated VSD II interactions [28]. A similar approach can be used in our future studies focused on optimizing specific toxin – ion channel interactions, while in this work we aimed to predict molecular determinants of those state-specific interactions using a consistent set of wild-type toxin – hNa$_V$1.7 VSD II and IV structural models and microsecond-long MD simulations. Other molecular simulation and analysis methodology advancements such as using enhanced sampling techniques GaMD, WE or their combination[64–66,73] will be explored in our follow-up studies to thoroughly explore both thermodynamics and kinetics of toxin – ion channel interactions with the current work providing a necessary framework and achieving a good agreement with several experimental findings[19,22].

In conclusion, our molecular modeling and simulation protocol can be a helpful guide for future designs of more potent therapeutic peptide variants with improved selectivity for different states of the hNa$_V$1.7 channel and other Na$_V$ channel subtypes as will be explored in follow-up studies. In addition, our results demonstrate that the interactions between PTx2 residues and lipids are important for state-specific binding to hNa$_V$1.7 VSDs and should be considered in future ion channel inhibitor studies as well. These findings open the possibility to improve the balance of VSD and lipid-binding capability of the toxin and, consequently, its efficacy as an ion channel inhibitor. We anticipate that this methodology can be readily expanded to study the interactions between other peptides and ion channels. Advancing our

understanding of the molecular mechanisms of ion channel modulation by peptide toxins will allow us to harness the full potential of these molecules for therapeutic applications.

## 3.5) Supplementary Information

*Amino acid sequences of the structures used in the study.*

*Multiple sequence alignment was done using CLUSTAL OMEGA (1.2.4)*

```
Protoxin2                       YCQKWMWTCDSERKCCEGMVCRLWCKKKLW


hNaV1.7_VSD_II_deactivated      CSPYWIKFKKCIYFIVMDPFVDLAITICIVLNTLFMAMEHHPMTEEFKNVLAIGNLVFTG        60
hNaV1.7_VSD_II_activated        CSPYWIKFKKCIYFIVMDPFVDLAITICIVLNTLFMAMEHHPMTEEFKNVLAIGNLVFTG        60
hNaV1.7_VSD_IV_deactivated      ------KIQGCIFDLVTNQAFDISIMVLICLNMVTMMVEKEGQSQHMTEVLYWINVVFII        54
hNaV1.7_VSD_IV_activated        ----GNKIQGCIFDLVTNQAFDISIMVLICLNMVTMMVEKEGQSQHMTEVLYWINVVFII        56
                                       *:: **: :* :  .*::* : * ** : * :*:.  ::.:.:**   *:**

hNaV1.7_VSD_II_deactivated      IFAAEMVLKLIAMDPYEYFQVGWNIFDSLIVTLSLVELFLADVEGLSVLR--SFRLLRVF        118
hNaV1.7_VSD_II_activated        IFAAEMVLKLIAMDPYEYFQVGWNIFDSLIVTLSLVELFLADVEGLSVLR--SFRLLRVF        118
hNaV1.7_VSD_IV_deactivated      LFTGECVLKLISLRHY-YFTVGWNIFDFVVVIISIVGMFLADLIETYFVSPTLFRVIRLA        113
hNaV1.7_VSD_IV_activated        LFTGECVLKLISLRHY-YFTVGWNIFDFVVVIISIVGMFLADLIETYFVSPTLFRVIRLA        115
                                  :*:.* *****::  * ** ******* ::* :*:* :****:    .:    **::*

hNaV1.7_VSD_II_deactivated      KLAKSW------  124
hNaV1.7_VSD_II_activated        KLAKSW------  124
hNaV1.7_VSD_IV_deactivated      RIGRILRLV---  122
hNaV1.7_VSD_IV_activated        RIGRILRLVKGA  127
                                ::::.:
```

## Table S1. *Criteria for Different Type of Non-bonded Interactions Used in Analyses*

| TYPES OF INTERACTION | DETECTION CRITERIA | FILTERING CRITERIA |
|---|---|---|
| **HYDROPHOBIC INTERACTIONS** | All pairs of hydrophobic atoms (i.e., carbons with only carbon or hydrogen atoms as neighbors) not involved in other specific types of interactions shown below within 4 Å. | Exclude hydrophobic interactions between rings that interact through π-stacking. |
| **HYDROGEN BONDS** | Distance between the hydrogen bond donor (D) and acceptor (A) should be less than 3.5 Å.<br>The angle D-H···A should be above 100°. | Exclude hydrogen bonds between atoms that already form a salt bridge. A hydrogen bond donor can participate in only one hydrogen bond, while acceptor atoms can form multiple hydrogen bonds. |
| **SALT BRIDGES** | Centers of oppositely charged groups that come within 5.5 Å. | |
| **CATION-Π** | Pairing of a positive charge and an aromatic ring if the distance between the charge center and the aromatic ring center is less than 6 Å. | |
| **Π-STACKING** | Centers of two aromatic rings within 5.5 Å. The angle between the rings should deviate no more than 30° from the optimal angle (90° for T-stacking, 180° for P-stacking). When projecting each ring center onto the opposite ring plane, the distance between the other ring center and the projected point (offset) should be less than 2 Å. | |

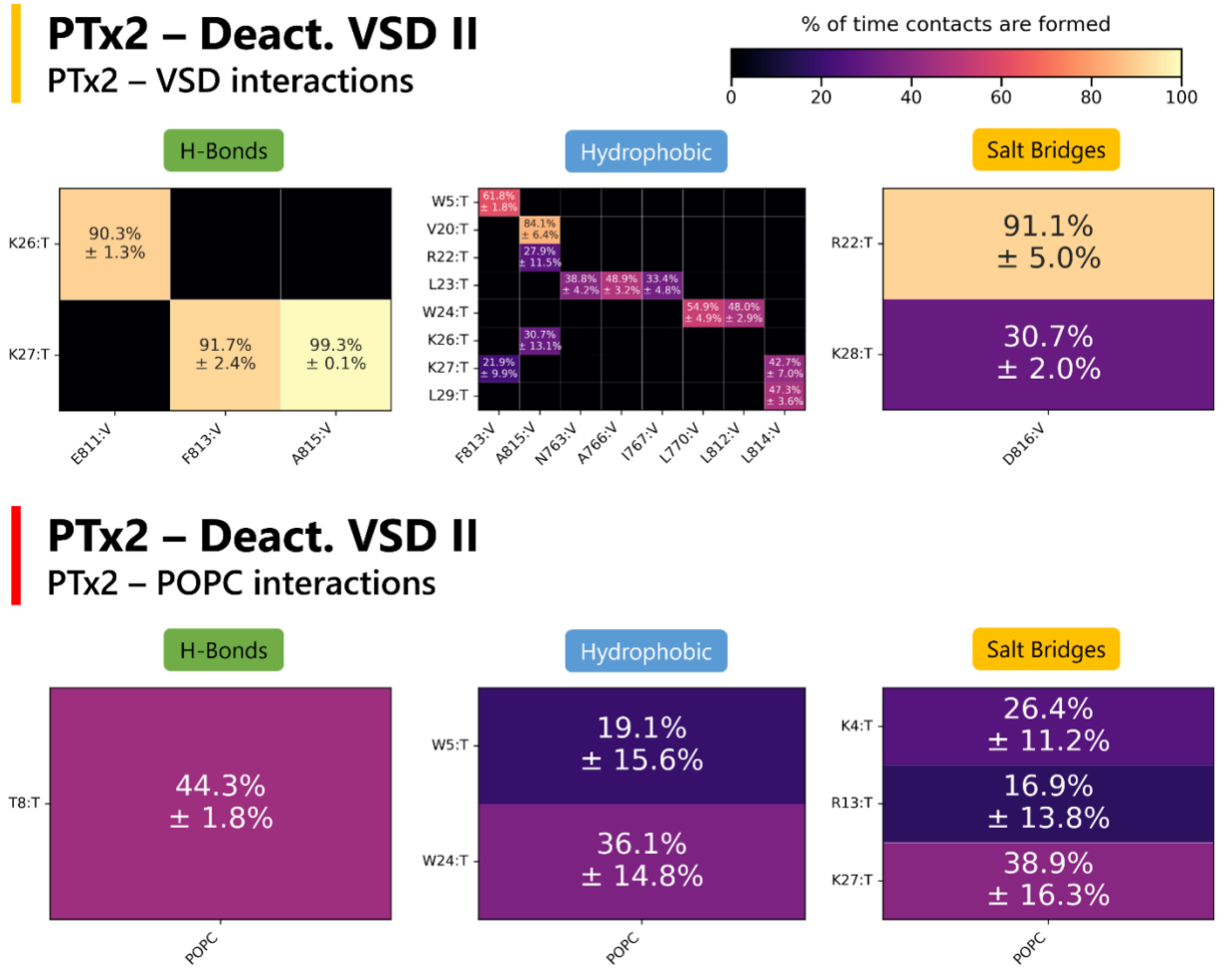



**Figure S1.** ***Contact maps showcasing PTx2 – VSD residue or POPC lipid interactions averaged over three MD simulation replicas of PTx2 – deactivated VSD II system.*** *PTx2 residues are denoted by the suffix ":T", while VSD residues are denoted by the suffix ":V" Error bars are standard errors of mean from 3 MD simulation replicas. Only contact pairs with duration above 5% (including standard errors) are displayed, and only the 2nd half of each MD simulation was used for the analysis.*

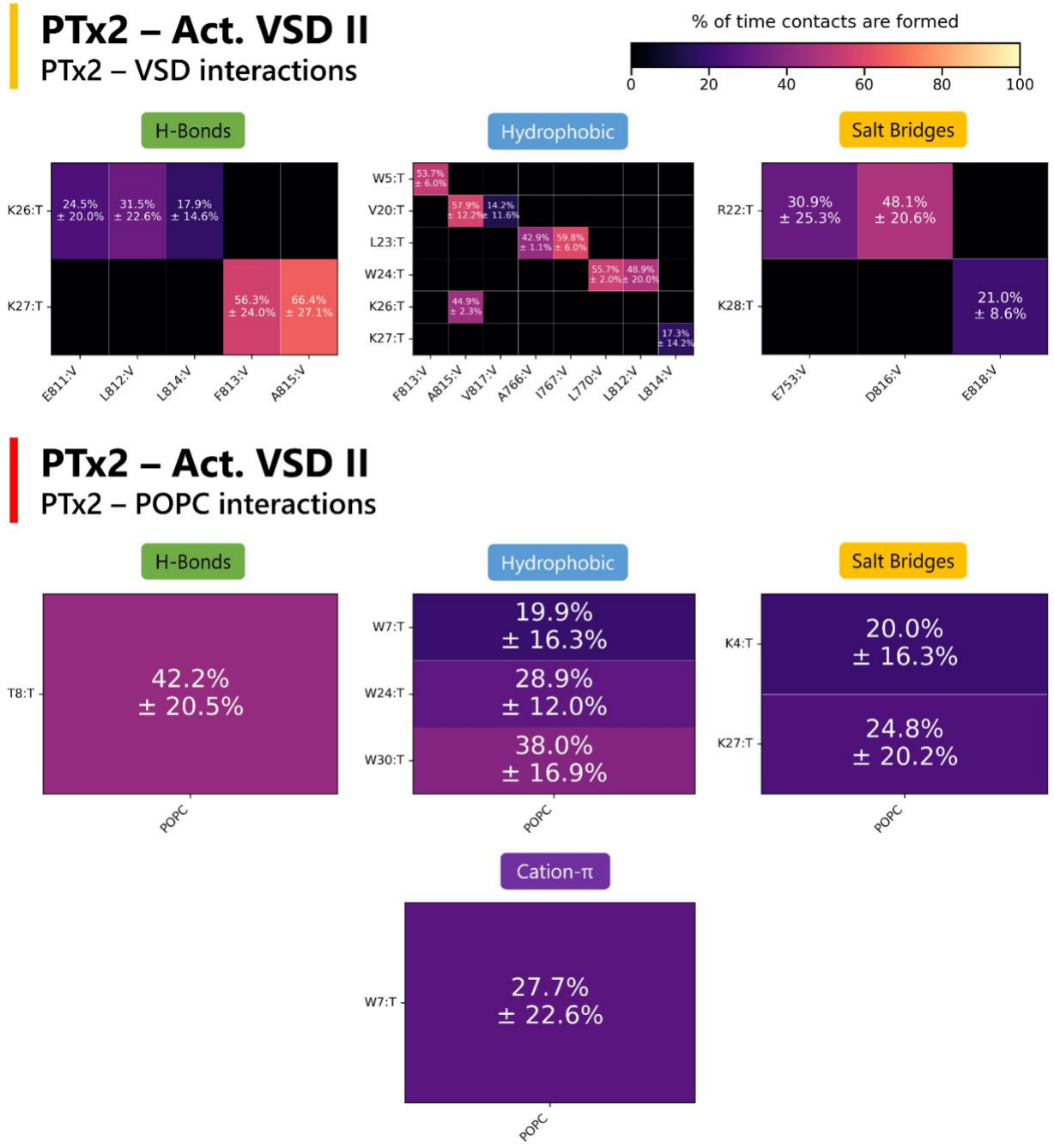



**Figure S2.** ***Contact maps showcasing PTx2 – VSD residue or POPC lipid interactions averaged over three MD simulation replicas of PTx2 – activated VSD II system.*** *PTx2 residues are denoted by the suffix ":T", while VSD residues are denoted by the suffix ":V". Error bars are standard errors of mean from 3 MD simulation replicas. Only contact pairs with*

*duration above 5% (including standard errors) are displayed, and only the 2nd half of each MD simulation was used for the analysis.*

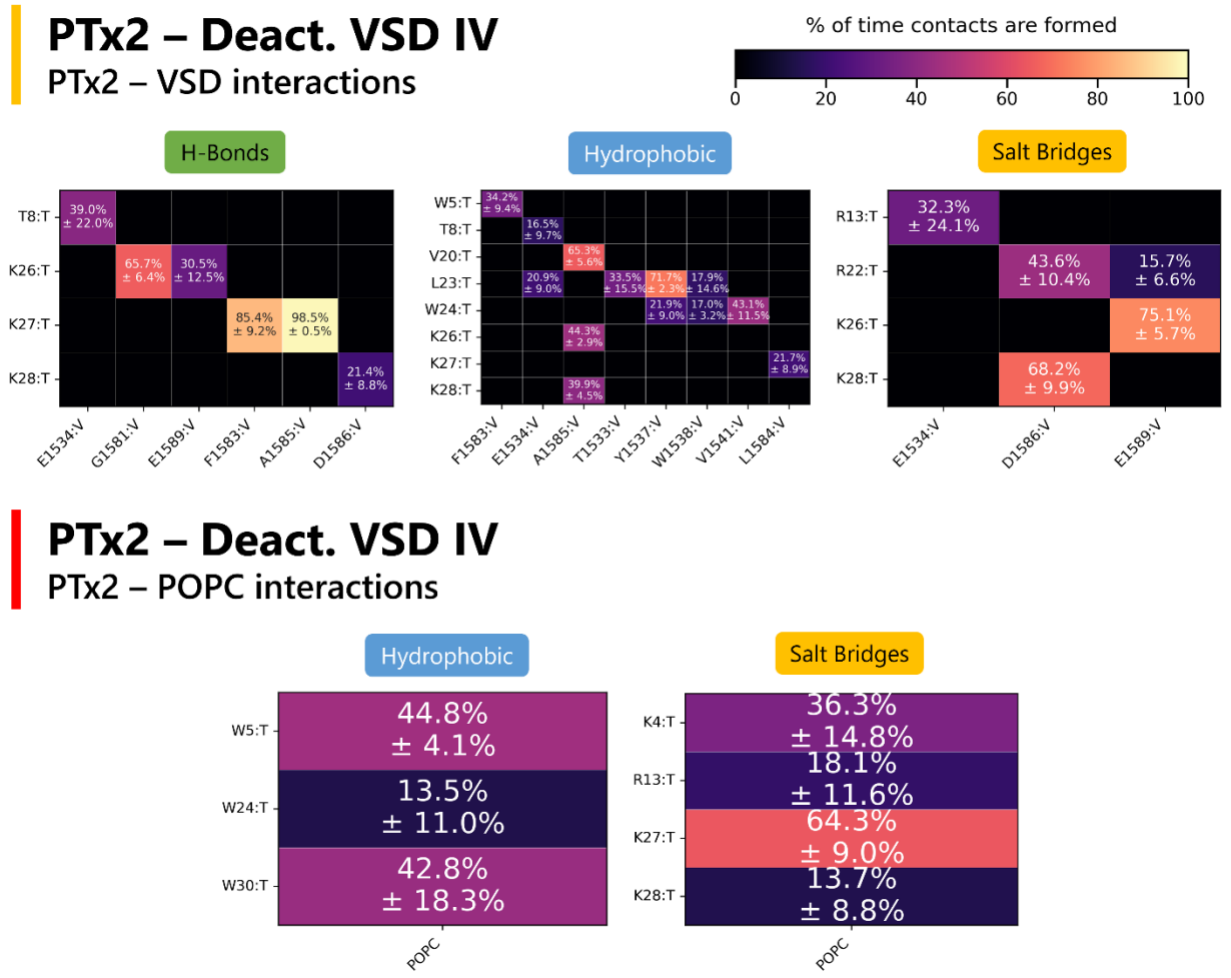



**Figure S3.** ***Contact maps showcasing PTx2 – VSD residue or POPC lipid interactions averaged over three MD simulation replicas of PTx2 – deactivated VSD IV system.*** *PTx2 residues are denoted by the suffix ":T", while VSD residues are denoted by the suffix ":V". Error bars are standard errors of mean from 3 MD simulation replicas. Only contact pairs with duration above 5% (including standard errors) are displayed, and only the 2nd half of each MD simulation was used for the analysis.*

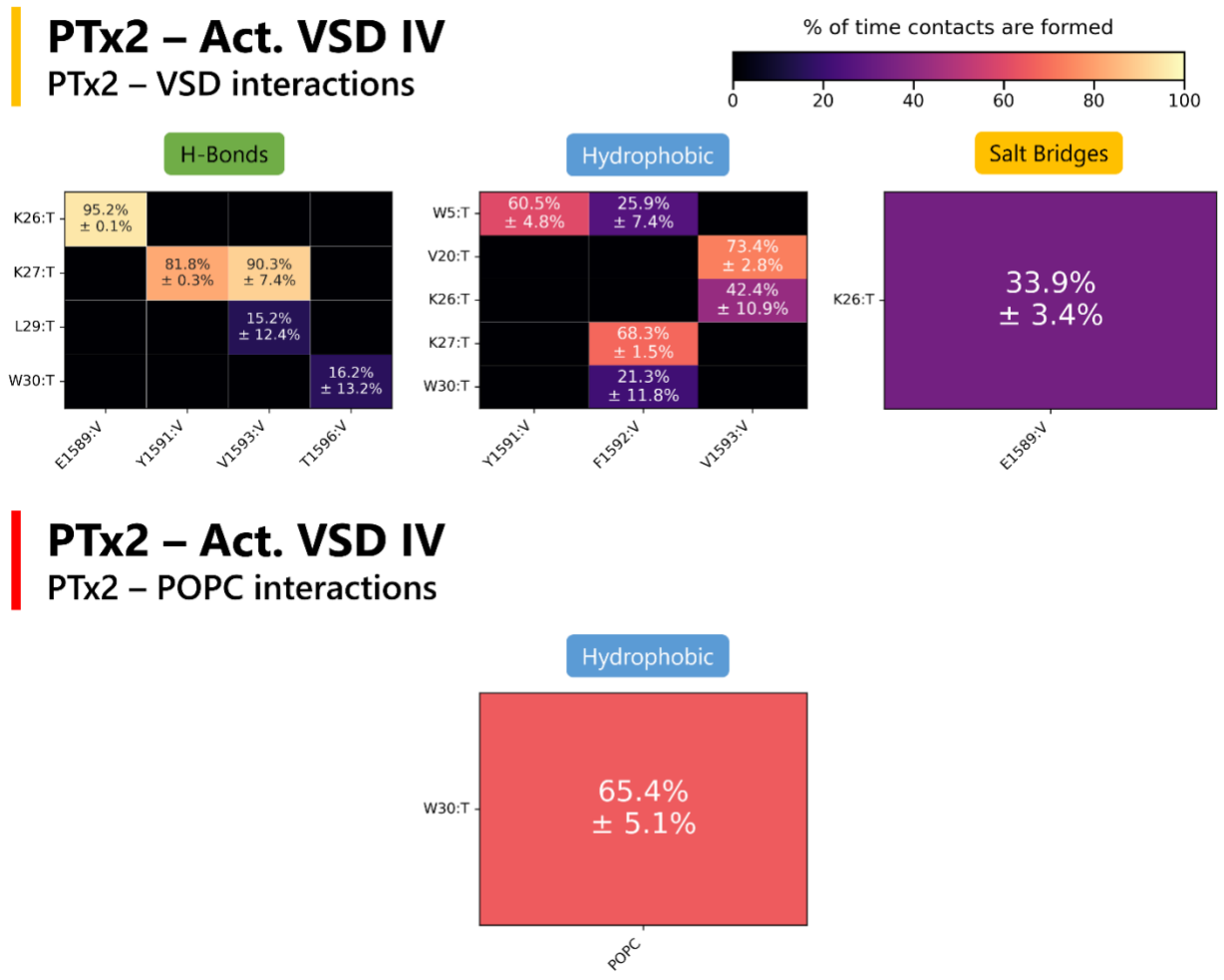



**Figure S4.** ***Contact maps showcasing PTx2 – VSD residue or POPC lipid interactions averaged over three MD simulation replicas of PTx2 – activated VSD IV system.*** *PTx2 residues are denoted by the suffix ":T", while VSD residues are denoted by the suffix ":V". Error bars are standard errors of mean from 3 MD simulation replicas. Only contact pairs with duration above 5% (including standard errors) are displayed, and only the 2nd half of each MD simulation was used for the analysis.*

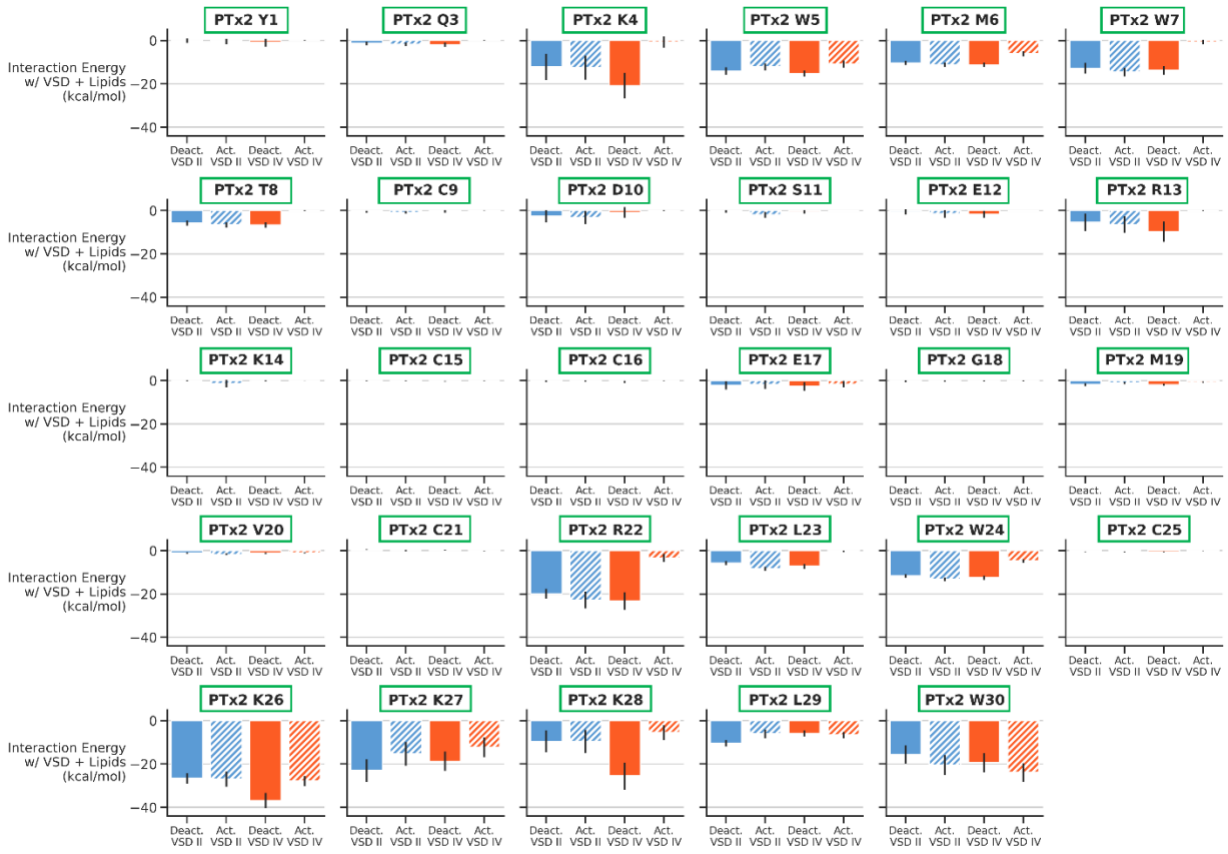



**Figure S5.** ***Interaction energies (van der Waals + electrostatic) contributed by each PTx2 residue toward the VSD and lipid binding process for different PTx2 - hNav1.7 VSD II/IV*** **MD simulation** ***systems.*** *Only residues with combined interaction energies across the whole system more favorable than –5 kcal/mol when averaged over all replicas are displayed,* and only the 2nd half of each MD simulation was used for the analysis.

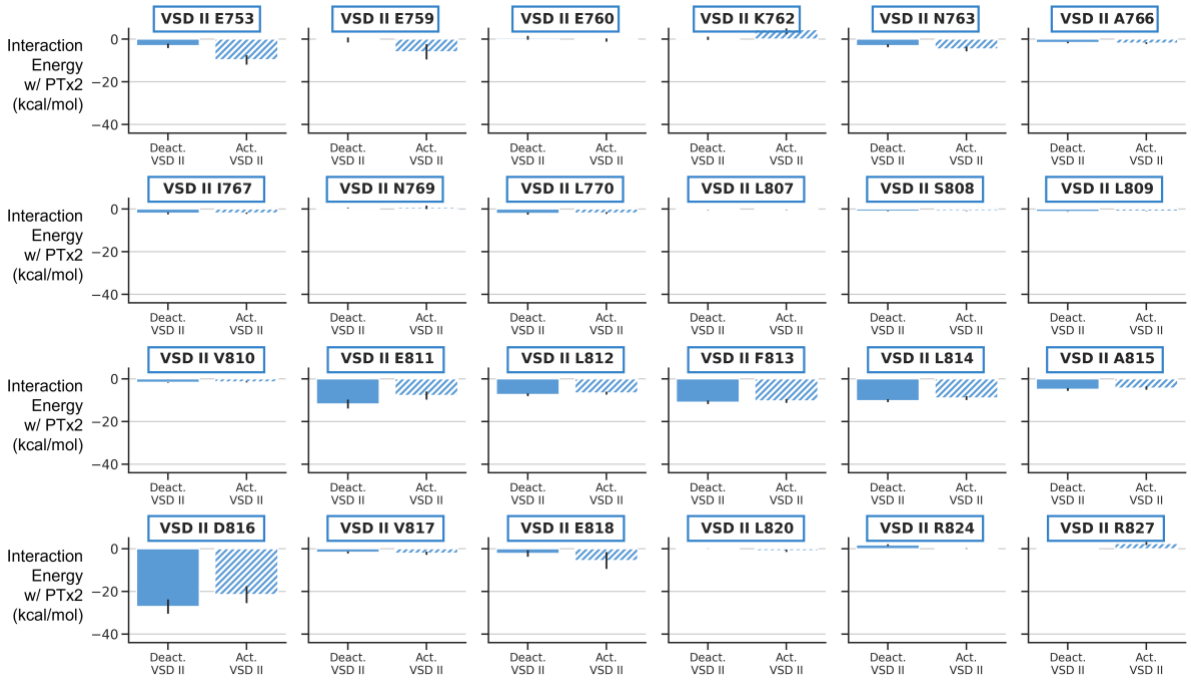



**Figure S6.** ***Interaction energies (van der Waals + electrostatic) contributed by each VSD II residue toward the PTx2 binding process for different PTx2 - hNav1.7 VSD II/IV MD simulation systems.*** *Only residues with combined interaction energies across the whole system more favorable than –0.5 kcal/mol when averaged overall replicas are displayed, and only the 2nd half of each MD simulation was used for the analysis.*

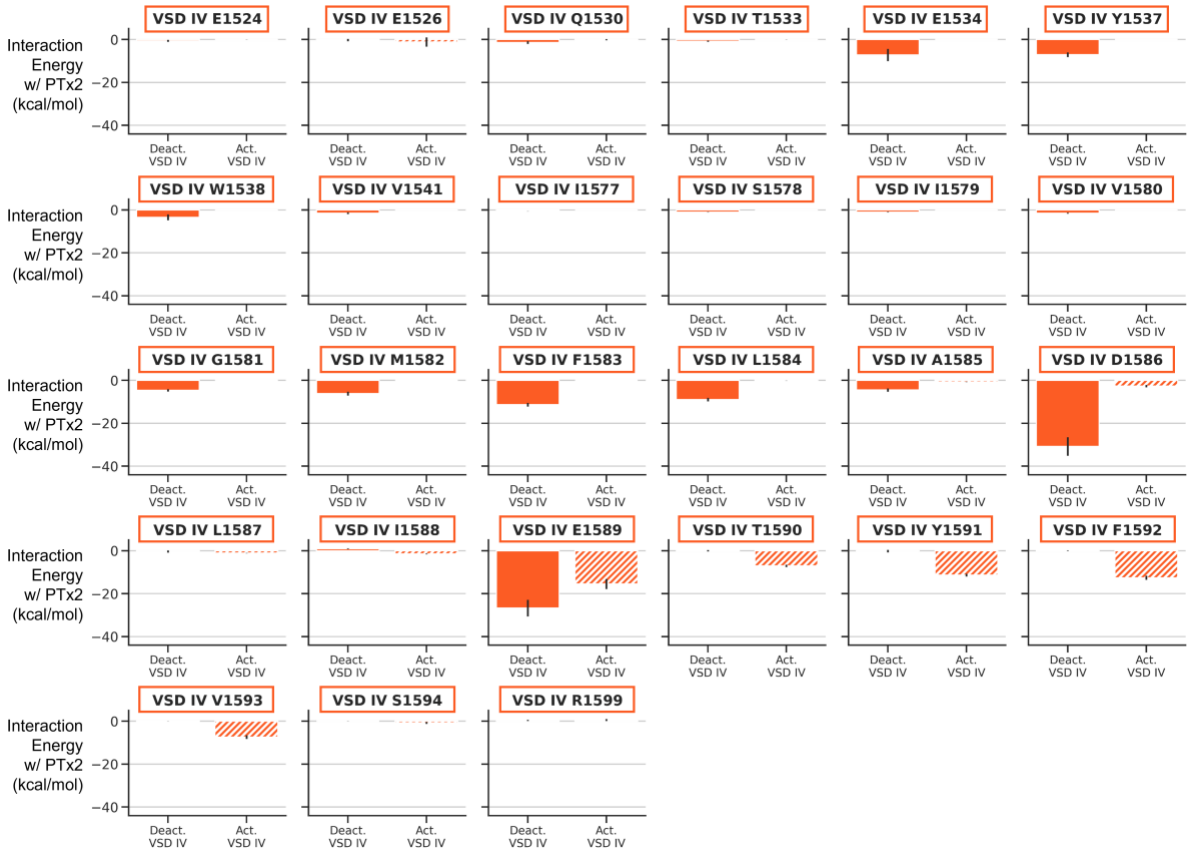



**Figure S7.** ***Interaction energies (van der Waals + electrostatic) contributed by each VSD IV residue toward the PTx2 binding process for different PTx2 - hNav1.7 VSD II/IV MD simulation systems.*** *Only residues with combined interaction energies across the whole system more favorable than −0.5 kcal/mol when averaged over all replicas are displayed, and only the 2nd half of each MD simulation was used for the analysis.*

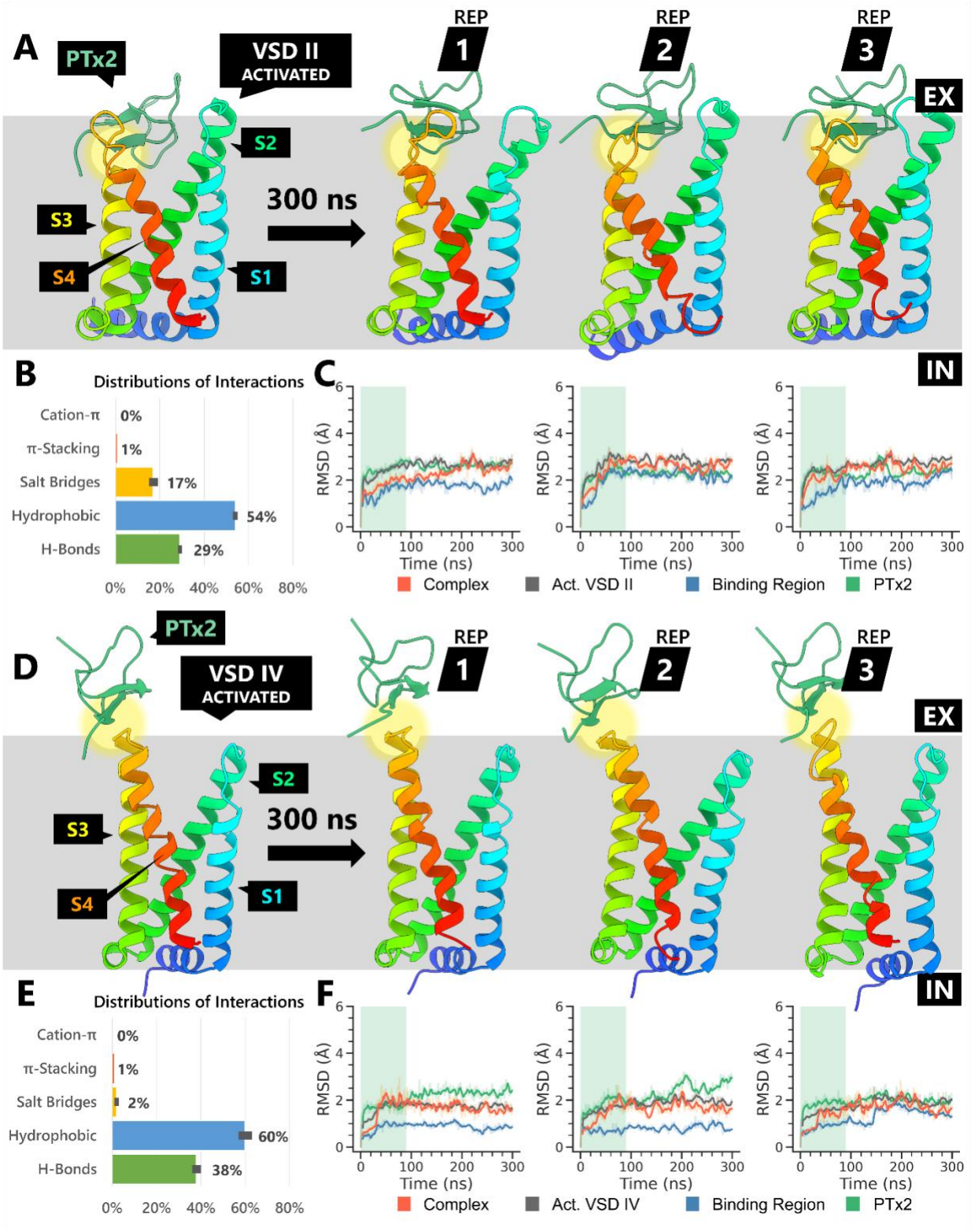



**Figure S8.** ***A general overview of the three replicas (REP) of 300-ns molecular dynamics simulations of PTx2 – hNav1.7 activated VSD II (A, B, C) or VSD IV (D, E, F) as a part of the***

***full hNav1.7 ion channel.*** *The lipid membrane is represented as a gray box, with the extracellular- and intracellular-facing sides labeled as "EX" and "IN" respectively. PTx2 – VSD binding sites are highlighted by yellow circles.* ***A, D)*** *Conformational changes of the toxin – VSD complex following the 300-ns-long molecular dynamics simulation for each replica.* ***B, E)*** *Distribution of all toxin – VSD interaction types encountered during the MD simulations, computed by dividing the sum of a specific interaction type by the sum of all interactions across all MD simulation frames.* ***C, F)*** *Time series of the root-mean-square deviation (RMSD) of PTx2 – VSD complex, VSD, toxin's binding region, and PTx2 backbone atoms compared to the initial structures. The plotted values represent the moving averages calculated over intervals of 5 ns each. The MD simulation equilibration period is shown as a light green box.*

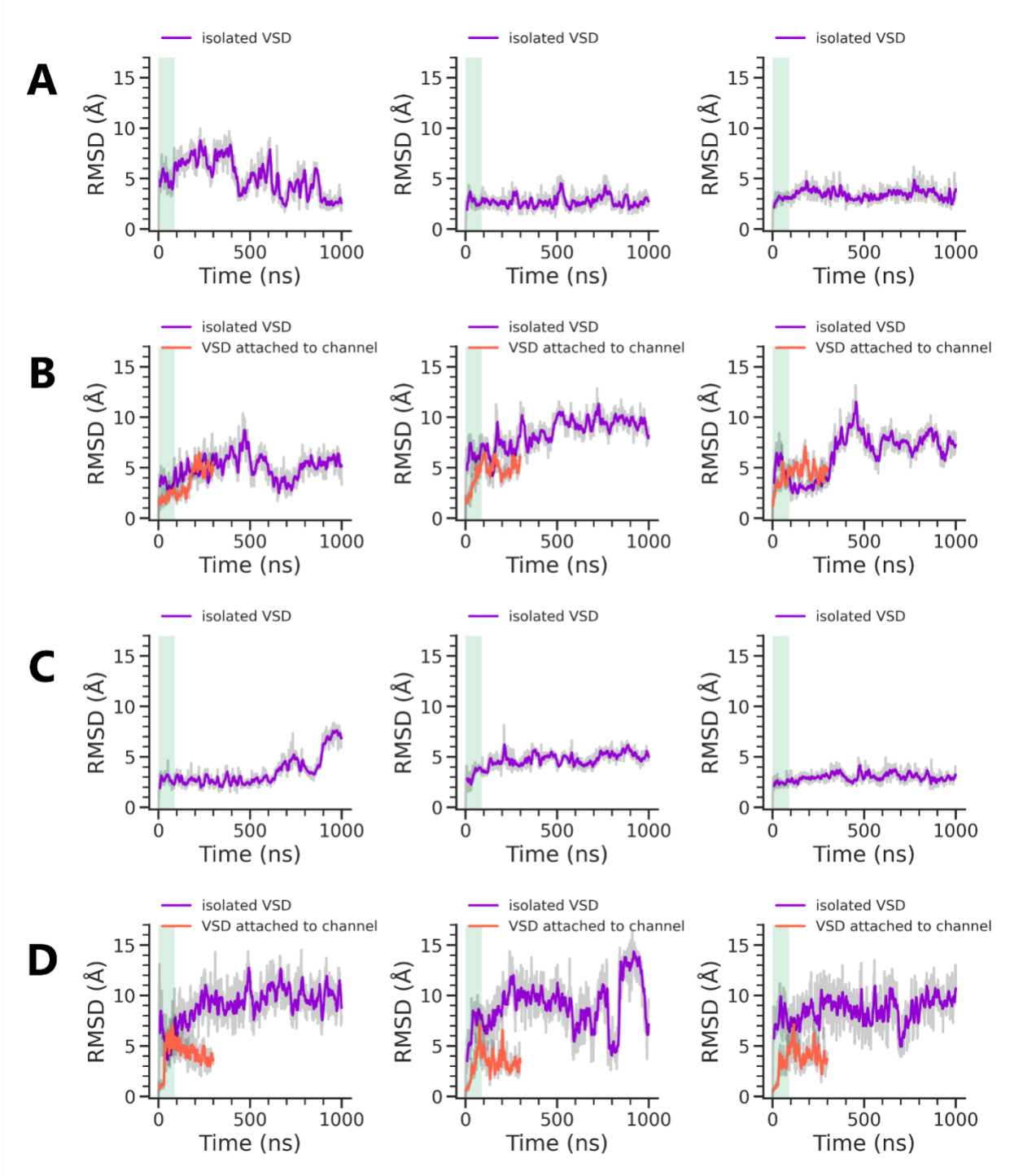



**Figure S9.** ***Time series of the root-mean-square deviation (RMSD) of PTx2 backbone atoms relative to the VSD compared to the initial structures.*** *RMSD plots from* ***A)*** *deactivated VSD II,* ***B)*** *activated VSD II, either isolated or attached to the channel,* ***C)*** *deactivated VSD IV, and* ***D)*** *activated VSD IV, either isolated or attached to the channel, are*

*displayed for three MD simulation replicas each. The colored bold lines represent the moving averages calculated over intervals of 10 ns each, while the gray lines represent the raw values. The MD simulation equilibration period is shown as a light green box.*

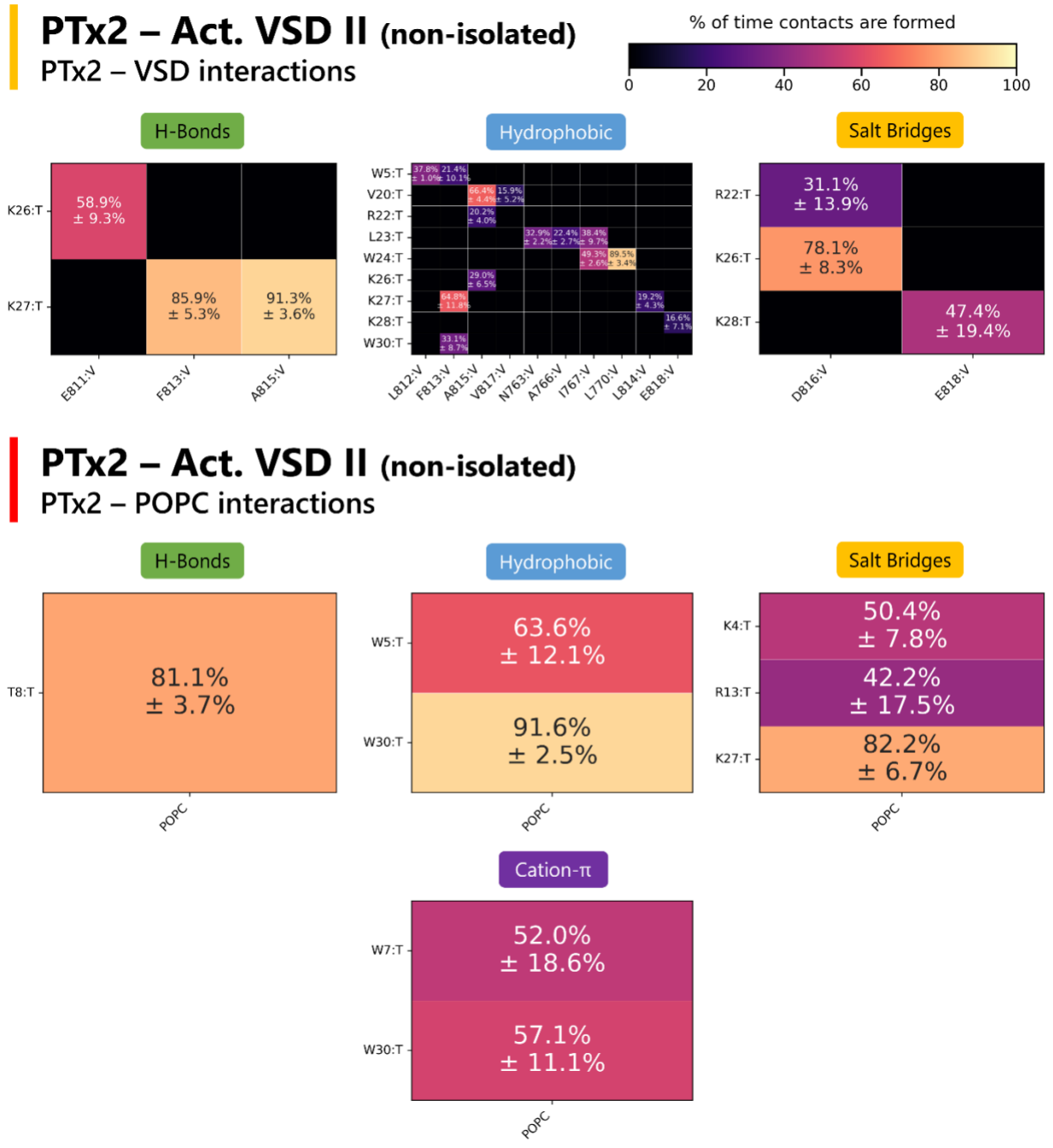



**Figure S10.** ***Contact maps showcasing PTx2 – VSD residue or POPC lipid interactions averaged over three MD simulation replicas of PTx2 – activated VSD II system (now as a***

***part of the full ion channel).*** *PTx2 residues are denoted by the suffix ":T", while VSD residues are denoted by the suffix ":V". Error bars are standard errors of mean from 3 MD simulation replicas. Only the 2nd half of each MD simulation was used for the analysis.*

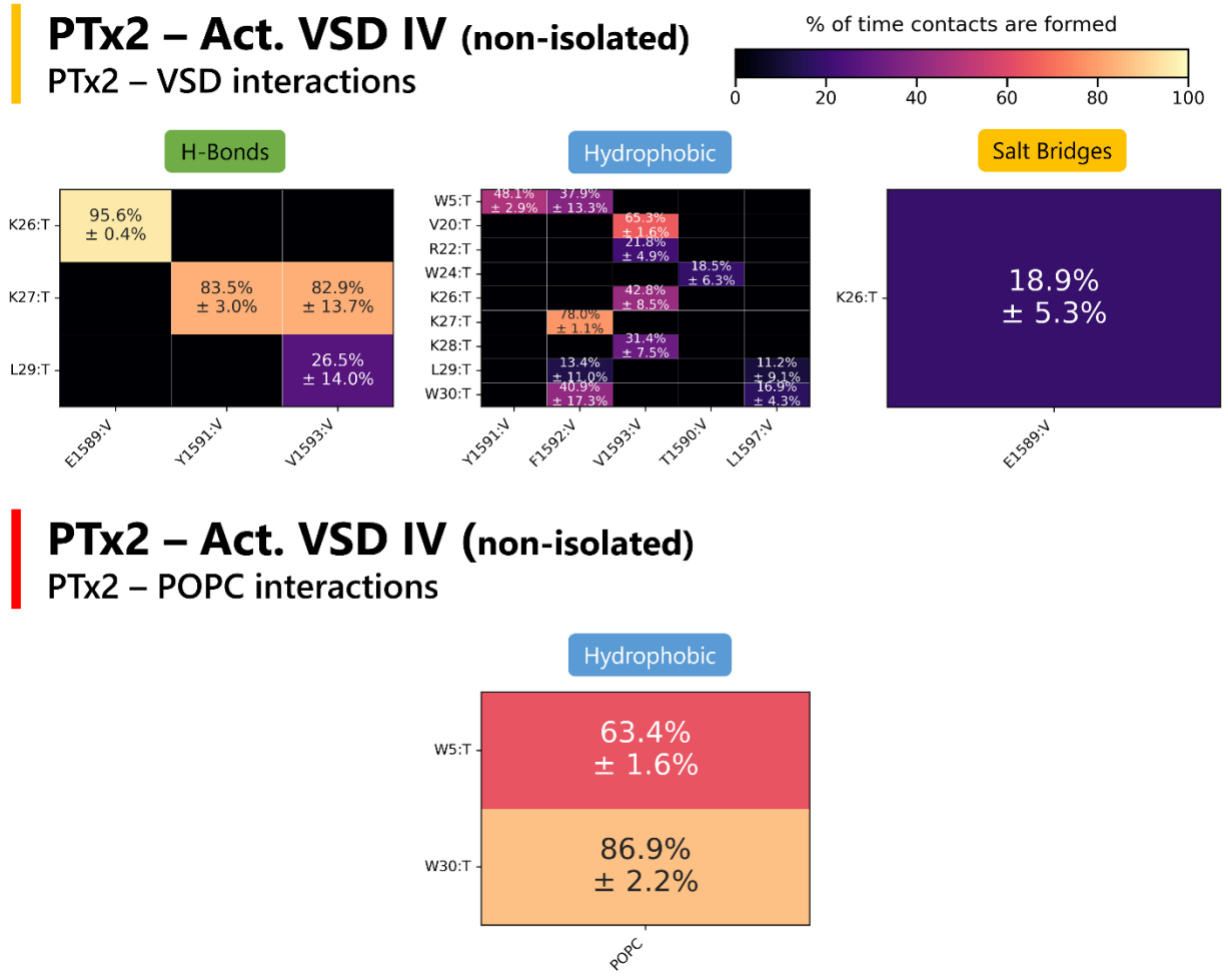



**Figure S11.** ***Contact maps showcasing PTx2 – VSD residue or POPC lipid interactions averaged over three MD simulation replicas of PTx2 – activated VSD IV (now as a part of the full ion channel).*** *PTx2 residues are denoted by the suffix ":T", while VSD residues are denoted by the suffix ":V". Error bars are standard errors of mean from 3 MD simulation replicas. Only the 2nd half of each MD simulation was used for the analysis.*

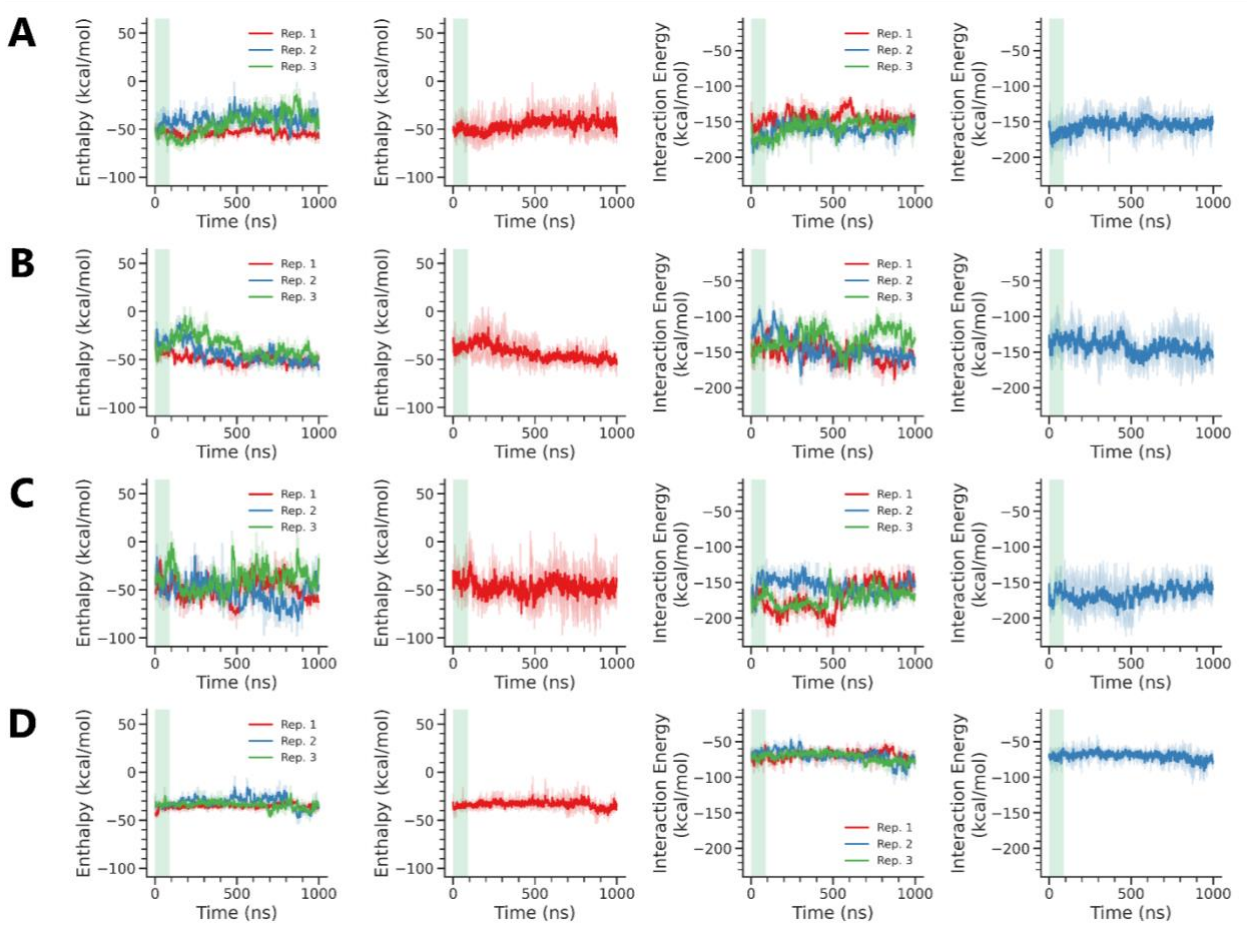



**Figure S12.** ***Time series of MM/PBSA computed enthalpy and interaction energies for PTx2 – hNav1.7 isolated VSD II and IV systems during the all-atom molecular dynamics simulations.*** *Energies from PTx2 interactions with* ***A)*** *deactivated VSD II,* ***B)*** *activated VSD II,* ***C)*** *deactivated VSD IV, and* ***D)*** *activated VSD IV are displayed. The first and third columns indicate individual energy values from different replicas. The plotted values represent the moving averages calculated over intervals of 5 ns each. The second and fourth columns indicate the averaged energy values across the replicas including standard errors of the mean shown by lighter colors. The MD simulation equilibration period is shown as a light green box.*

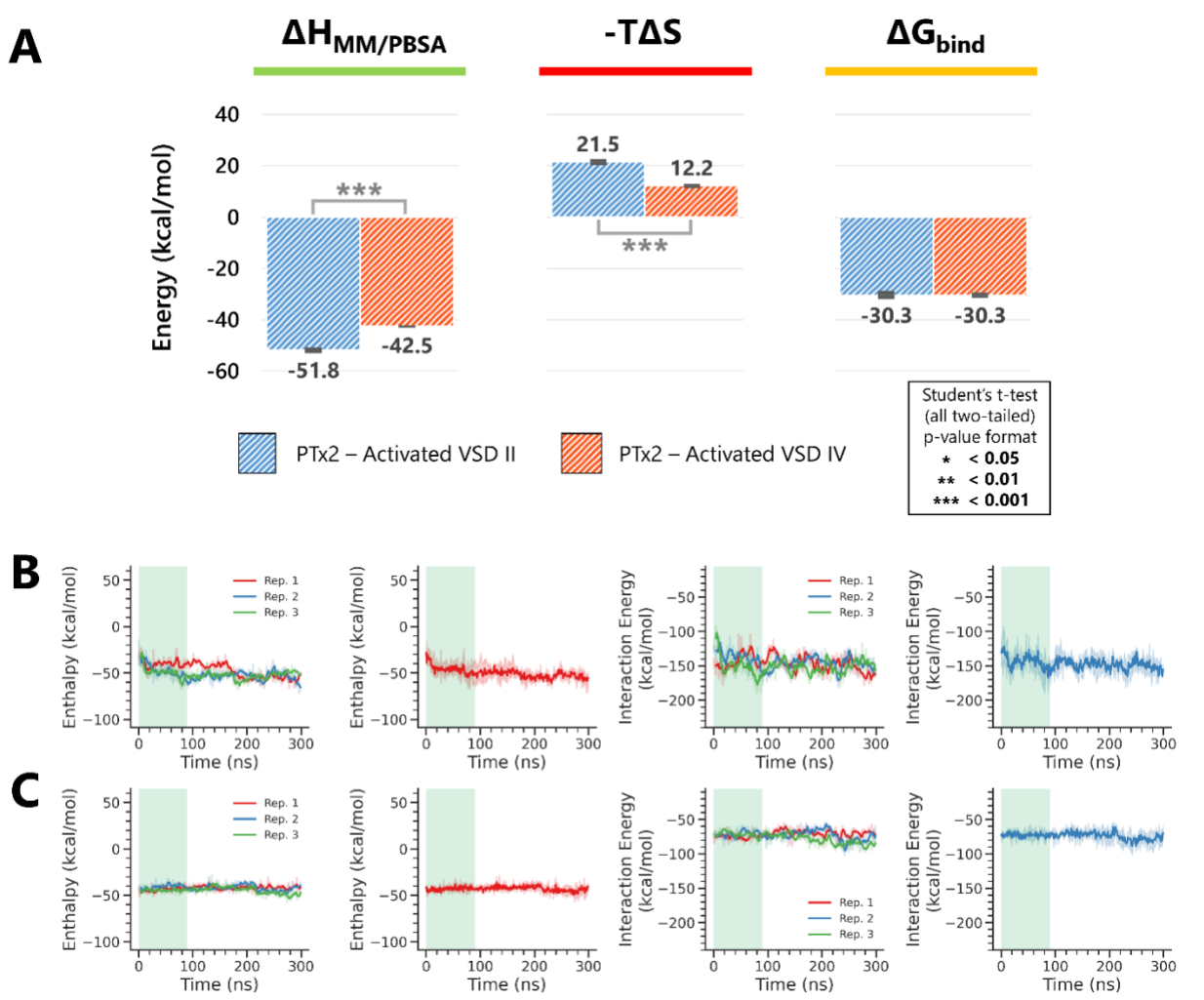



**Figure S13.** ***MM/PBSA binding energies in PTx2 – hNav1.7 activated VSD II and IV systems (full channel simulation).*** **A)** *Energies averaged over three replicas from 100 ns onward of the 300-ns-long all-atom molecular dynamics simulations.* ***B)*** *Time series of MM/PBSA computed enthalpy and interaction energies for PTx2 interactions with activated VSD II and* ***C)*** *activated VSD IV during the full channel simulations are displayed. The first and third columns indicate individual energy values from different replicas. The second and fourth columns indicate the averaged energy values across the replicas including standard errors of the mean shown by lighter colors. The MD simulation equilibration period is shown as a light green box.*

## 3.6) References


1. Ahern, C. A., Payandeh, J., Bosmans, F. & Chanda, B. The hitchhiker's guide to the voltage-gated sodium channel galaxy. *J. Gen. Physiol.* **147**, 1–24 (2016).
2. Bennett, D. L., Clark, A. J., Huang, J., Waxman, S. G. & Dib-Hajj, S. D. The Role of Voltage-Gated Sodium Channels in Pain Signaling. *Physiol. Rev.* **99**, 1079–1151 (2019).
3. Capes, D. L., Goldschen-Ohm, M. P., Arcisio-Miranda, M., Bezanilla, F. & Chanda, B. Domain IV voltage-sensor movement is both sufficient and rate limiting for fast inactivation in sodium channels. *J. Gen. Physiol.* **142**, 101–112 (2013).
4. Clairfeuille, T. *et al.* Structural basis of α-scorpion toxin action on Nav channels. *Science* **363**, (2019).
5. Chanda, B., Asamoah, O. K. & Bezanilla, F. Coupling interactions between voltage sensors of the sodium channel as revealed by site-specific measurements. *J. Gen. Physiol.* **123**, 217–230 (2004).
6. Chanda, B. & Bezanilla, F. Tracking voltage-dependent conformational changes in skeletal muscle sodium channel during activation. *J. Gen. Physiol.* **120**, 629–645 (2002).
7. Chowdhury, S. & Chanda, B. Sodium channels caught in the act. *Science* **363**, 1278–1279 (2019).
8. Nassar, M. A. *et al.* Nociceptor-specific gene deletion reveals a major role for Nav1.7 (PN1) in acute and inflammatory pain. *Proc. Natl. Acad. Sci.* **101**, 12706–12711 (2004).
9. Drenth, J. P. H. & Waxman, S. G. Mutations in sodium-channel gene SCN9A cause a spectrum of human genetic pain disorders. *J. Clin. Invest.* **117**, 3603–3609 (2007).

10. Golshani, A. E., Kamdar, A. A., Spence, S. C. & Beckmann, N. M. Congenital Indifference to Pain: An Illustrated Case Report and Literature Review. *J. Radiol. Case Rep.* **8**, 16–23 (2014).

11. Dib-Hajj, S. D., Yang, Y., Black, J. A. & Waxman, S. G. The NaV1.7 sodium channel: from molecule to man. *Nat. Rev. Neurosci.* **14**, 49–62 (2013).

12. Yang, X., Wei, X., Mu, Y., Li, Q. & Liu, J. A review of the mechanism of the central analgesic effect of lidocaine. *Medicine (Baltimore)* **99**, e19898 (2020).

13. Hsu, E. S. Medication Overuse in Chronic Pain. *Curr. Pain Headache Rep.* **21**, 2 (2017).

14. Nguyen, P. T. *et al.* Computational design of peptides to target NaV1.7 channel with high potency and selectivity for the treatment of pain. *eLife* **11**, e81727 (2022).

15. Nguyen, P. T. & Yarov-Yarovoy, V. Towards Structure-Guided Development of Pain Therapeutics Targeting Voltage-Gated Sodium Channels. *Front. Pharmacol.* **13**, 842032 (2022).

16. Tran, H. N. T. *et al.* Enzymatic Ligation of a Pore Blocker Toxin and a Gating Modifier Toxin: Creating Double-Knotted Peptides with Improved Sodium Channel NaV1.7 Inhibition. *Bioconjug. Chem.* **31**, 64–73 (2020).

17. Shen, H., Liu, D., Wu, K., Lei, J. & Yan, N. Structures of human Nav1.7 channel in complex with auxiliary subunits and animal toxins. *Science* **363**, 1303–1308 (2019).

18. Schmalhofer, W. A. *et al.* ProTx-II, a selective inhibitor of NaV1.7 sodium channels, blocks action potential propagation in nociceptors. *Mol. Pharmacol.* **74**, 1476–1484 (2008).

19. Xiao, Y., Blumenthal, K., Jackson, J. O., Liang, S. & Cummins, T. R. The tarantula toxins ProTx-II and huwentoxin-IV differentially interact with human Nav1.7 voltage sensors to inhibit channel activation and inactivation. *Mol. Pharmacol.* **78**, 1124–1134 (2010).

20. de Lera Ruiz, M. & Kraus, R. L. Voltage-Gated Sodium Channels: Structure, Function, Pharmacology, and Clinical Indications. *J. Med. Chem.* **58**, 7093–7118 (2015).

21. Wulff, H., Christophersen, P., Colussi, P., Chandy, K. G. & Yarov-Yarovoy, V. Antibodies and venom peptides: new modalities for ion channels. *Nat. Rev. Drug Discov.* **18**, 339–357 (2019).

22. Xu, H. *et al.* Structural Basis of Nav1.7 Inhibition by a Gating-Modifier Spider Toxin. *Cell* **176**, 702-715.e14 (2019).

23. Ahuja, S. *et al.* Structural basis of Nav1.7 inhibition by an isoform-selective small-molecule antagonist. *Science* **350**, aac5464 (2015).

24. Flood, E., Boiteux, C., Lev, B., Vorobyov, I. & Allen, T. W. Atomistic Simulations of Membrane Ion Channel Conduction, Gating, and Modulation. *Chem. Rev.* **119**, 7737–7832 (2019).

25. Deplazes, E. Molecular Simulations of Disulfide-Rich Venom Peptides with Ion Channels and Membranes. *Molecules* **22**, 362 (2017).

26. Henriques, S. T. *et al.* Interaction of Tarantula Venom Peptide ProTx-II with Lipid Membranes Is a Prerequisite for Its Inhibition of Human Voltage-gated Sodium Channel NaV1.7. *J. Biol. Chem.* **291**, 17049–17065 (2016).

27. Mateos, D. L. & Yarov-Yarovoy, V. Structural Modeling of Peptide Toxin - Ion Channel Interactions using RosettaDock. 2022.06.29.498146 Preprint at https://doi.org/10.1101/2022.06.29.498146 (2022).

28. Katz, D., Sindhikara, D., DiMattia, M. & Leffler, A. E. Potency-Enhancing Mutations of Gating Modifier Toxins for the Voltage-Gated Sodium Channel NaV1.7 Can Be Predicted Using Accurate Free-Energy Calculations. *Toxins* **13**, 193 (2021).

29. Angus, M. & Ruben, P. Voltage gated sodium channels in cancer and their potential mechanisms of action. *Channels* **13**, 400 (2019).

30. Bajaj, S. & Han, J. Venom-derived peptide modulators of cation-selective channels: friend, foe or frenemy. *Front. Pharmacol.* **10**, 58 (2019).

31. Borrego, J. *et al.* Peptide inhibitors of Kv1. 5: an option for the treatment of atrial fibrillation. *Pharmaceuticals* **14**, 1303 (2021).

32. Chow, C. Y. *et al.* Venom peptides with dual modulatory activity on the voltage-gated sodium channel NaV1. 1 provide novel leads for development of antiepileptic drugs. *ACS Pharmacol. Transl. Sci.* **3**, 119–134 (2019).

33. Zhang, F. *et al.* BmK AEP, an anti-epileptic peptide distinctly affects the gating of brain subtypes of voltage-gated sodium channels. *Int. J. Mol. Sci.* **20**, 729 (2019).

34. Rohl, C. A., Strauss, C. E. M., Misura, K. M. S. & Baker, D. Protein structure prediction using Rosetta. *Methods Enzymol.* **383**, 66–93 (2004).

35. Mandell, D. J., Coutsias, E. A. & Kortemme, T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nat. Methods* **6**, 551–552 (2009).

36. Stein, A. & Kortemme, T. Improvements to robotics-inspired conformational sampling in rosetta. *PloS One* **8**, e63090 (2013).

37. Gray, J. J. *et al.* Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **331**, 281–299 (2003).

38. Wang, C., Bradley, P. & Baker, D. Protein-protein docking with backbone flexibility. *J. Mol. Biol.* **373**, 503–519 (2007).

39. Alford, R. F., Fleming, P. J., Fleming, K. G. & Gray, J. J. Protein Structure Prediction and Design in a Biologically Realistic Implicit Membrane. *Biophys. J.* (2020) doi:10.1016/j.bpj.2020.03.006.

40. Kozma, D., Simon, I. & Tusnády, G. E. PDBTM: Protein Data Bank of transmembrane proteins after 8 years. *Nucleic Acids Res.* **41**, D524–D529 (2013).

41. Wang, R. Y.-R. *et al.* Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. *eLife* **5**, (2016).

42. Song, Y. *et al.* High-resolution comparative modeling with RosettaCM. *Struct. Lond. Engl. 1993* **21**, 1735–1742 (2013).

43. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

44. Salomon-Ferrer, R., Case, D. A. & Walker, R. C. An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* **3**, 198–210 (2013).

45. Huang, J. & MacKerell Jr, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).

46. Klauda, J. B. *et al.* Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).

47. Best, R. B. *et al.* Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ, ψ and Side-Chain χ1 and χ2 Dihedral Angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).

48. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

49. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

50. Ryckaert, J.-P., Ciccotti, G. & Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341 (1977).

51. Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F. & Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **43**, W443–W447 (2015).

52. Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate–DNA Helices. *J. Am. Chem. Soc.* **120**, 9401–9409 (1998).

53. Kollman, P. A. *et al.* Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **33**, 889–897 (2000).

54. Wang, C. *et al.* Calculating protein–ligand binding affinities with MMPBSA: Method and error analysis. *J. Comput. Chem.* **37**, 2436–2446 (2016).

55. Miller, B. R. I. *et al.* MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).

56. Shirts, M. R. *et al.* Lessons learned from comparing molecular dynamics engines on the SAMPL5 dataset. *J. Comput. Aided Mol. Des.* **31**, 147–161 (2017).

57. Lu, Q. & Luo, R. A Poisson–Boltzmann dynamics method with nonperiodic boundary condition. *J. Chem. Phys.* **119**, 11035–11047 (2003).

58. Duan, L., Liu, X. & Zhang, J. Z. H. Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein–Ligand Binding Free Energy. *J. Am. Chem. Soc.* **138**, 5722–5728 (2016).

59. Roe, D. R. & Cheatham, T. E. I. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).

60. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**, 70–82 (2021).

61. White, S. H. & Wimley, W. C. Hydrophobic interactions of peptides with membrane interfaces. *Biochim. Biophys. Acta* **1376**, 339–352 (1998).

62. Guvench, O. & MacKerell, A. D. Comparison of protein force fields for molecular dynamics simulations. *Methods Mol. Biol. Clifton NJ* **443**, 63–88 (2008).

63. Chen, H. *et al.* A companion guide to the string method with swarms of trajectories: Characterization, performance, and pitfalls. *J. Chem. Theory Comput.* **18**, 1406–1422 (2022).

64. Zuckerman, D. M. & Chong, L. T. Weighted ensemble simulation: review of methodology, applications, and software. *Annu. Rev. Biophys.* **46**, 43–57 (2017).

65. Wang, J. & Miao, Y. Peptide Gaussian accelerated molecular dynamics (Pep-GaMD): Enhanced sampling and free energy and kinetics calculations of peptide binding. *J. Chem. Phys.* **153**, (2020).

66. Ahn, S.-H., Ojha, A. A., Amaro, R. E. & McCammon, J. A. Gaussian-Accelerated Molecular Dynamics with the Weighted Ensemble Method: A Hybrid Method Improves Thermodynamic and Kinetic Sampling. *J. Chem. Theory Comput.* **17**, 7938–7951 (2021).

67. Sheng, Y., Yin, Y., Ma, Y. & Ding, H. Improving the Performance of MM/PBSA in Protein–Protein Interactions via the Screening Electrostatic Energy. *J. Chem. Inf. Model.* **61**, 2454–2462 (2021).

68. Spiliotopoulos, D. *et al.* dMM-PBSA: A New HADDOCK Scoring Function for Protein-Peptide Docking. *Front. Mol. Biosci.* **3**, (2016).

69. Liu, X., Peng, L. & Zhang, J. Z. H. Accurate and Efficient Calculation of Protein–Protein Binding Free Energy-Interaction Entropy with Residue Type-Specific Dielectric Constants. *J. Chem. Inf. Model.* **59**, 272–281 (2019).

70. Sun, H. *et al.* Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Phys. Chem. Chem. Phys.* **20**, 14450–14460 (2018).

71. Han, Y. *et al.* Elucidation of a dynamic interplay between a beta-2 adrenergic receptor, its agonist, and stimulatory G protein. *Proc. Natl. Acad. Sci.* **120**, e2215916120 (2023).

72. Kollman, P. Free energy calculations: applications to chemical and biochemical phenomena. *Chem. Rev.* **93**, 2395–2417 (1993).

73. Ahn, S.-H., Jagger, B. R. & Amaro, R. E. Ranking of Ligand Binding Kinetics Using a Weighted Ensemble Approach and Comparison with a Multiscale Milestoning Approach. *J. Chem. Inf. Model.* **60**, 5340–5352 (2020).

# CHAPTER 4:

# Leveraging Deep Learning to Predict Different Conformations of the hERG Potassium Channel – A Notorious Drug Anti-Target

Khoa Ngo*[1,2,3], Vladimir Yarov-Yarovoy[2,4], Colleen E. Clancy*[2,3,5], and Igor Vorobyov*[2,5]

[1]*Biophysics Graduate Group, UC Davis, CA*

[2]*Department of Physiology and Membrane Biology, UC Davis, CA*

[3]*Center for Precision Medicine and Data Science, UC Davis, CA*

[4]*Department of Anesthesiology and Pain Medicine, UC Davis, CA*

[5]*Department of Pharmacology, UC Davis, CA*

** These authors are co-corresponding authors*

**Author Contributions:**

I devised the approaches detailed in the manuscript and implemented them to generate hERG models in various conformational states. I developed scripts to perform all analyses described in the study and was responsible for creating all figures in both the main manuscript and the supplementary section.

*"To know, is to know that you know nothing. That is the meaning of true knowledge."*

– Socrates (470-399 BC)

## 4.1) Introduction

The human voltage-gated potassium channel $K_V11.1$, encoded by the KCNH2 or human Ether-à-go-go-Related Gene (hERG) gene, is a key player in cardiac electrophysiology, underpinning the rapid component of the delayed rectifier $K^+$ current ($I_{Kr}$) in cardiac myocytes[1]. This current plays a crucial role in the repolarization phase of the cardiac action potential, essential for the rhythmic and stable electrical propagation and consequent contraction[2]. Perturbation to hERG channel function, whether stemming from genetic anomalies or pharmacological interventions, can precipitate multiple arrhythmogenic disorders[2].

The distinctive pharmacological sensitivity of hERG channels, prone to blockade by a diverse array of drugs, has placed them at the forefront of cardiac safety pharmacology and the drug discovery process. Blockade of hERG by drugs can lead to QT interval prolongation known as an acquired long QT syndrome and escalate the risk of *torsades de pointes* (TdP), a potentially fatal arrhythmia[3]. This issue has prompted the withdrawal of various drugs from the market and underscored the necessity of incorporating hERG safety evaluations in the drug development pipeline[4–6].

The hERG channel is a homotetramer, with each subunit containing six membrane-spanning segments (S1-S6)[7]. The segments S5 and S6 along with the intervening loops and short pore helix form the channel pore domain (PD), crucial for potassium ion passage, while segments S1-S4 form the voltage sensing domain (VSD), responding to voltage changes across the cell membrane. Notably, the hERG channel also features specialized intracellular regions: the Per-Arnt-Sim (PAS) domain at the N-terminus and the cyclic nucleotide-binding domain (CNBD) at the C-terminus. The structural intricacies of the hERG channel, particularly its pore and voltage-sensing domains, render it sensitive to a variety of drugs, highlighting the imperative for intricate structural understanding of drug channel interactions to foster safer drug designs. This susceptibility is not uniform but varies depending on the channel conformational state, a phenomenon known as state-dependent drug block. Drugs may

preferentially bind to and block the channel in specific states (open, closed, or inactivated), which can differentially affect cardiac repolarization and rhythm[8,9].

Capturing the dynamic spectrum of hERG channel states poses a formidable challenge. While cryo-electron microscopy (cryo-EM) has offered invaluable insights into the open state of the channel[7,10], a comprehensive view of the closed and inactivated states has remained elusive. Thus, even as we embark on a scientific era of explosive growth fueled by the convergence of protein structure insights, computational capabilities, and AI based modeling and synthetic data, the next frontier is marked by the need to reveal all relevant conformational states of proteins. The existing knowledge gaps constrain both predictive capabilities regarding drug – protein interactions and the creation of therapies through drug discovery to find specific and selective drugs, or in the case of hERG, to minimize adverse interactions. For example, a recent study by Yang *et al.* introduced a multiscale model framework to forecast drug-induced cardiotoxicity at cellular and tissue levels, utilizing atomistic simulations of drug interactions with hERG[11]. However, the absence of hERG structural models in the inactivated and closed states limited the predictive potential of atomistic scale simulations of state-specific drug binding.

The emergence of AlphaFold2, a protein structure prediction tool driven by machine learning, has brought a paradigm shift in structural biology[12]. The proficiency of AlphaFold2 in accurately predicting protein structures presents a significant opportunity to model numerous proteins whose experimental structures remain unsolved. Nonetheless, the conventional use of AlphaFold2 tends to produce a single-state structure for the protein, which in the case of hERG corresponds to the open state. In this study, we presented a novel approach to guide AlphaFold2 to predict models of the hERG channel in the closed and inactivated states, thereby unveiling the elusive structural underpinnings of state-specific drug interactions and gating mechanisms of the channel.

We sought insights into the state-specific interactions of drugs with hERG channel, which are vital for revealing the mechanisms behind drug-induced cardiotoxicity and for designing safer pharmaceuticals. Our methodology blended advanced computational techniques to

yield an exhaustive understanding of the dynamics of the hERG channel at atomistic resolution. We commenced with structural modeling using AlphaFold2, followed by drug docking simulations via Rosetta GALigandDock[13] to evaluate drug – ion channel interactions. All-atom molecular dynamics (MD) simulations were employed to investigate ion conduction attributes and dynamic channel behavior across conformational states. We also present an interaction network analysis that revealed key protein residues and domains involved in state transitions, offering a molecular framework for deciphering hERG channel gating mechanisms.

## 4.2) Results and Discussion

### *a. Generating hERG channel conformational states*

The hERG channel undergoes transitions between closed, open, and inactivated conformational states. Experimental cryo-EM structures published so far[7,10] appeared to resolve only one of those states (open and conducting) based on previous molecular simulation studies[11,14,15]. To guide AlphaFold2 in generating diverse conformations pertinent to each conformational state, we integrated fragments from experimental structures, each characteristic of the discrete closed, open or inactivated state. We then utilized AlphaFold2 to explore a broad array of conformational states and generate the final structures. This approach ensured that the resulting conformations were diverse and relevant to each discrete conformational state. Inspired by the Meiler lab's successful use of AlphaFold2 for sampling alternative conformations of transporter and receptors[16–18], we expect that this method will be capable of yielding possible representations of the hERG channel in different states, offering an improvement over earlier computational methods that depended on homology modeling[19,20].

To construct the closed state of the hERG channel, we used regions of the voltage-sensing domains (VSDs) from the rat EAG channel cryo-EM structure in its closed conformation (PDB 8EP1[21], residues H208 – H343) as a structural template and combined them with the selectivity filter (SF) region from an available hERG cryo-EM structure (PDB 5VA2[7], residues I607 – T634), as depicted in **Figure 1a**. Recognizing that the physiological closed state of the

hERG channel might only partially resemble these structural elements, we sought to introduce a degree of flexibility in our model sampling process.

Due to the time-intensive process of creating multiple sequence alignments (MSA) for AlphaFold2, the ColabFold[22] software was made to streamline protein structure prediction by combining MMseqs2[23] sequence search toolkit with AlphaFold2, enhancing runtime efficiency while preserving high prediction accuracy. First, we provided this hybrid structure as a template for ColabFold. ColabFold utilized only a portion of the MSA in the modeling process to reduce memory usage. We adjusted the limit to 256 cluster centers and 512 additional sequences, a reduction from the standard 512 and 1024, respectively. Altering the random seed can lead to variance in cluster centers and, consequently, in structure predictions. Therefore, we set AlphaFold2 to cycle through 20 random seeds (each seed was used to produce 5 models) to foster diversity in the predicted conformational states. Additionally, we enabled the model stochastic (dropout) component during inference to sample from an ensemble of models for the more ambiguous segments of the structure prediction. Past research has indicated that these settings may permit the sampling of alternative protein conformations[16–18,22]. In total, we generated 100 models for further analysis. The N-terminal PAS domain (residues M1 – R397) was not included in the modeling due to graphics card memory limitation making the resultant model (W398 – R863) resemble hERG 1b isoform[24].

For the open hERG channel state, we utilized the existing cryo-EM structure of hERG (PDB 5VA27), albeit with missing extracellular loops. We then reconstructed the extracellular loops using the Rosetta LoopRemodel mover[25] with the results shown in **Figure 1b**. This reconstructed model served as a base for molecular dynamics and drug docking simulations. However, we also wanted to test the potential for AlphaFold2 to emulate this structure. The modeling procedure was initiated with constraints akin to those used for the closed state. This time, specific regions of the putative open state hERG model (PDB 5VA27) — namely the VSD (W398 – V549), the SF with adjacent pore helix (I607 – T634), a part of S6 and the cytosolic domain (S660 – R863) — were provided to ColabFold as a structural guide. Then, we generated 100 diverse models for the detailed cluster analysis.

The task of modeling hERG in an inactivated state was more complex due to the absence of direct experimental references. As illustrated in **Figure 1c**, we extracted a region of the S6 and the entire cytosolic domain (S660 – R863) from the open-state hERG channel model and employed ColabFold under similar conditions to those described above for closed and open states. Our method involved letting AlphaFold2 reconstruct the transmembrane domain from the ground up under conditions that facilitated the sampling of varied conformations[17]. In the resulting models, a subset of these models exhibited SFs (residues S624 – G628) akin to the non-conductive variant observed under low extracellular $K^+$ conditions, as described in a preprint study by Lau *et al*.[26,27] This non-conductive SF is primarily identified by a distinct lateral flip of the backbone carbonyl oxygens in residue V625 away from hERG central axis, creating a potential barrier that may prevent $K^+$ from crossing from S3 to S2 ion binding sites. We extracted this non-conductive SF and the adjacent pore helix (residues Y607 – T634) structures from our array of models and merged it with the activated VSDs (W398 – V549) and the cytosolic domain (S660 – R863) from the open-state hERG structure (PDB 5VA2[7]) to use as structural template. Employing the same parameters in ColabFold as described above, we then generated 100 more models for further analysis.

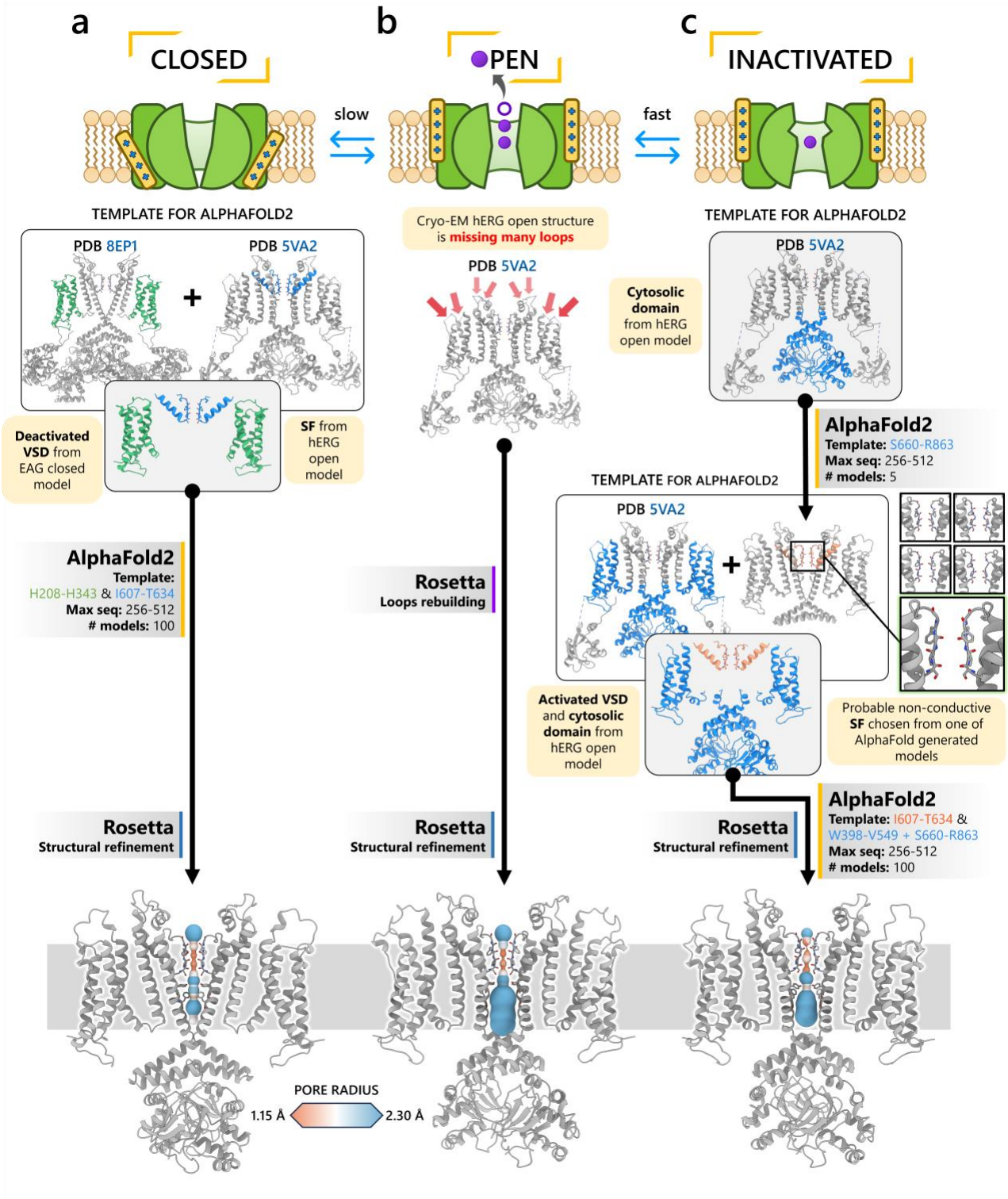



**Figure 1. Generation of hERG channel models in the closed (a), open (b), and inactivated (c) states.** *The lower limit of the pore radius color profile (1.15 Å) indicates the*

*minimum radius to accommodate a water molecule, and the upper limit (2.30 Å) indicates sufficient space to fit two water molecules side by side.*

### *b. Clustering AlphaFold2-generated hERG models*

Post-prediction, the resulting 100 models for each structural state were categorized into clusters based on all-atom root-mean-square deviation (RMSD) between those models, as shown in **Figure 2**. Closed-state models were clustered with a threshold of 0.75 Å across the entire channel, whereas inactivated and open states (non-empirical, AlphaFold-predicted) models focused on the SF (residues S624 – G628), with a more stringent threshold of 0.35 Å. We ranked the models in the cluster by their average per-residue confidence metric (predicted local distance difference test, or pLDDT), which assesses the likelihood that the predicted structure aligns with an experimentally determined structure[12]. pLDDT above 90 are considered to be highly reliable and those between 70 and 90 as reliable with generally good protein backbone structure prediction[12]. Lower scores indicate regions of lower confidence and may be unstructured. Cluster 1 is defined in this study to be the cluster with highest average pLDDT among its members.

Closed-state clusters: The analysis of closed-state clusters revealed very minor discrepancies when comparing RMSD and pLDDT values across those clusters. Comparing the top-ranked model in each cluster, the all-atom RMSD between cluster 1 (cluster with highest confidence) with cluster 2 is 0.36 Å, and with the outlier cluster is 0.95 Å (clusters with fewer than 3 members are classified as outliers). Excluding the outlier cluster, the average pLDDT scores between cluster 1 and cluster 2 are very similar. The predicted low RMSD between the models suggests a convergence of the prediction to a singular conformation, with minor differences likely arising from slight variations in the positioning of intracellular loop regions. Therefore, the top-ranked model in cluster 1 was chosen for subsequent simulations.

Inactivated-state clusters: Compared to the previous case, we observed more differences among the putative inactivated-state clusters. For this situation, we used a more precise metric for clustering by focusing exclusively on the SF (S624 – G628). Following this, we

ranked the clusters according to the average pLDDT of these specific residues. This method led to the identification of four main clusters and one outlier.

Cluster 1, which has the highest certainty (pLDDT = 90.3 ± 2.3), consists of models where the carbonyl of SF V625 alternates between turning inward and outward, i.e. toward and away from the central axis, among different subunits. Cluster 2 (pLDDT = 85.2 ± 1.6) is fairly similar, but the carbonyl of SF V625 is now flipped outward from the central axis in all subunits in most instances, and there is a noticeable constriction between G626 carbonyls. Cluster 3 (pLDDT = 77.5 ± 2.5) features the carbonyl reorientation of SF G626 and occasionally F627. In most models from these groups, the S6 helix rotates to varying extents, causing Y652 and particularly F656 side chains to rotate along and extending further into the hERG inner cavity. Both cluster 4 and the outlier cluster have low average pLDDTs (<70), rendering them unreliable for our analysis.

Intriguingly, while modeling hERG inactivation, we found that the SF conformations predicted by AlphaFold coincide with proposed hERG C-type inactivation mechanisms as highlighted in other experimental and computational studies[15,26,27]. This correlation is particularly notable considering the structures in the mentioned studies have not been published in any publicly accessible databases, to the best of our knowledge. Specifically, the behavior observed in clusters 1 and 2, namely the flipping of V625 and the constriction at G626 carbonyls, was previously reported in a preprint study by Lau *et al.*[26,27] Moreover, Li *et al.*'s computational work revealed an asymmetric SF conformation, where two opposing subunits exhibited similar V625 flipping and G626 narrowing as seen in our AlphaFold-predicted clusters 1 and 2, while the other two subunits displayed the G626 and F627 carbonyl reorientation characteristic of our cluster 3[15]. For our investigation, we chose to focus on the conformations seen in clusters 1 and 2, as this behavior was observed in both previous studies. We proceeded with further analysis and simulations using the best model from cluster 2, where all subunits consistently show the flipping of V625 carbonyls and pinching between G626 carbonyls.

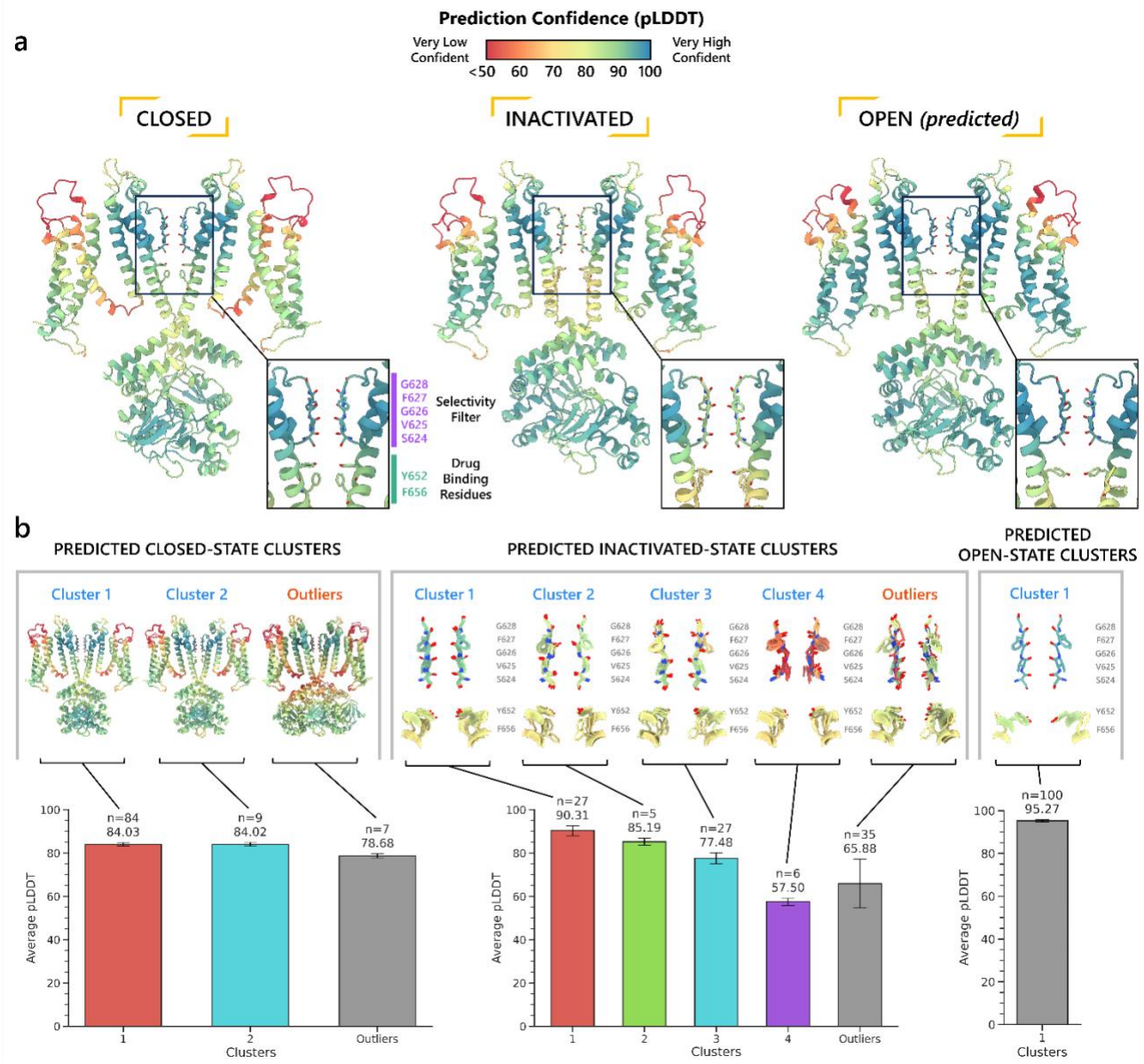



**Figure 2. Clustering of AlphaFold2-predicted hERG channel models.** ***a)*** *The model chosen for subsequent analysis.* ***b)*** *Clusters created from 100 models predicted for each state. Each structure visualized is colored according to the per-residue confidence metric (pLDDT). The closed-state models are clustered based on the RMSD of the entire protein models. The inactivated- and open-state models are clustered based on the RMSD of the selectivity filter (S624 – G628). To represent each cluster, the top 5 models ranked by average pLDDT are shown. The bar graphs display the mean pLDDT values for the clustered segment across all models*

*within each cluster, accompanied by the standard deviation. Clusters containing less than three models are categorized as outliers.*

Open-state clusters: In a control study, we used the conductive SF from the presumed open-state cryo-EM hERG structure (PDB 5VA2[7]), combined with the VSD and cytosolic domain, as the structural template for AlphaFold. Remarkably, the generated models converged almost uniformly into a single cluster, exhibiting negligible variance in both pLDDT and RMSD values across the models. The highest scoring model closely mirrors the experimental structure with reconstructed loops, showing only 0.5 Å all-atom RMSD due to minor differences in the loop conformations.

### *c. Comparison of state-specific hERG channel models*

We refined the preliminary atomistic structural models putatively representing each functional state of the hERG channel (open, inactivated and closed) using Rosetta[25,28] with an implicit membrane to optimize each residue orientation and resolve any steric clashes. The resulting models are compared in **Figure 3**. In **Figure 3a, b**, the closed-state model displays the most constricted channel pore, followed by the inactivated-state and then the open-state model. In the pore-lining S6 helix, the drug-binding residue Y652 maintains a consistent position in both closed and open hERG channel models. However, in the inactivated-state model, Y652 and its aromatic ring move slightly due to the rotation of the S6 helix, and together with the rotation of S624 sidechain hydroxyl oxygen away from the pore help form a shallow hydrophobic pocket above Y652 rings and beneath the SF. The rotation and movement of the pore-lining S6 helix also influences the position of residue F656 across the models, with its side chain protruding most into hERG inner cavity in the closed state, followed by the inactivated state and least in the open state.

Shown in **Figure S1a, b**, the SFs of the open- and closed-state models display similar arrangements, with carbonyl oxygens along the ion path all oriented inward to the central axis as in other $K^+$ channel structures, e.g., KcsA and $K_V1.2$, enabling efficient knock-on $K^+$ conduction[29,30]. In contrast, the inactivated-state model SF is distinct, marked by the lateral rotation of the V625 backbone carbonyl away from the central axis (**Figure S1c**), thereby

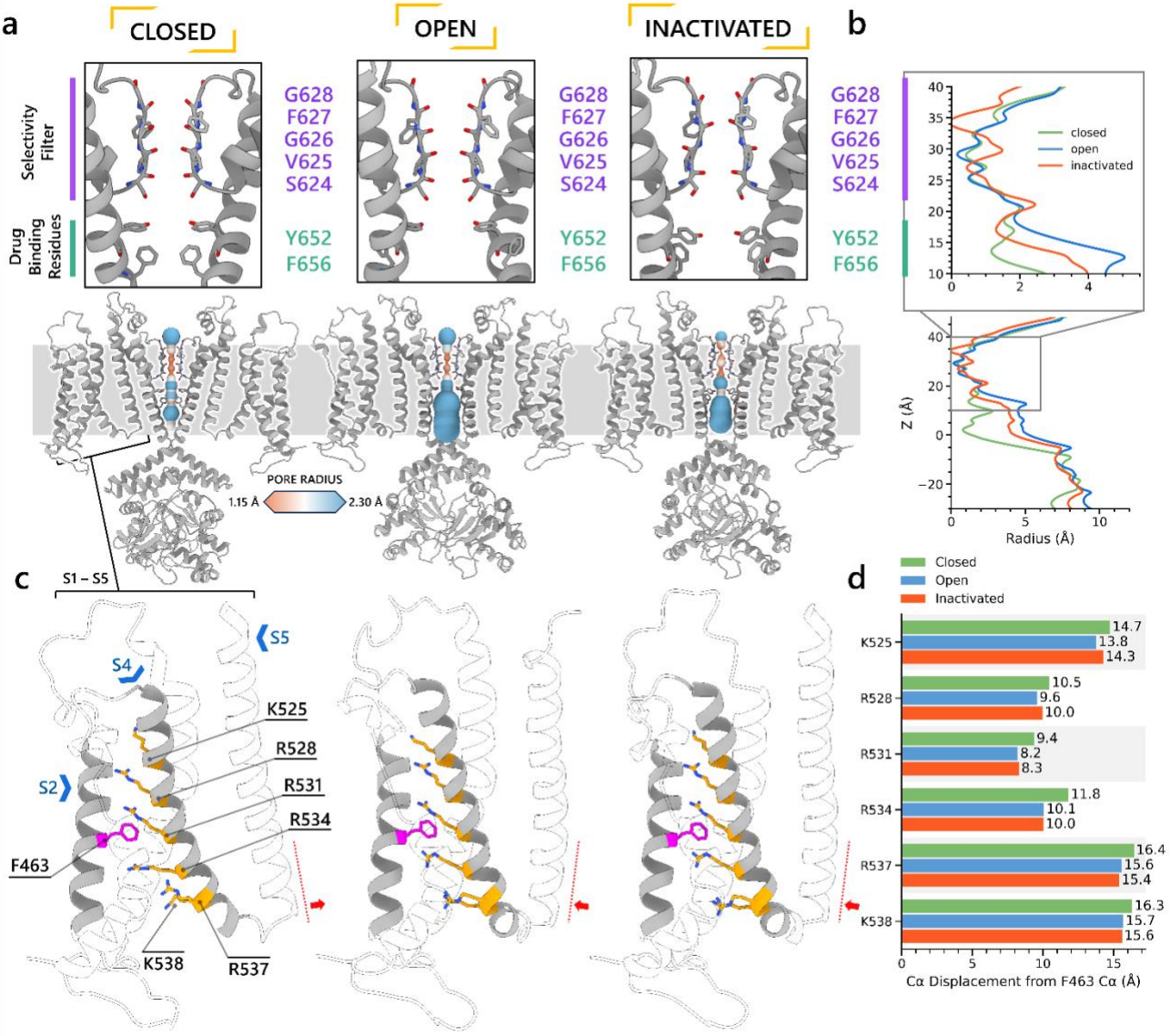
creating a potential barrier preventing ion crossing from the S3 to S2 binding sites. Additionally, we noted a constriction between the G626 backbone carbonyls and a repositioning of the S624 sidechain hydroxyl oxygen. In the model representing the inactivated state, the carbonyl oxygens of G628 and F627 exhibit an upward shift relative to their positions in the open-state model. **Figure S1d-f** presents the SF of all three models from an extracellular perspective. In both the closed- and inactivated-state models, the F627 side chain undergoes a clockwise rotation when contrasted with its orientation in the open state. This rotational behavior aligns with findings from prior simulation study[14] where it was noted in a metastable non-conductive state. The loop that links the upper SF to the S6 helix rotates anti-clockwise relative to its position in the open-state model, consequently narrowing the upper part of the SF.

Echoing our observations, Lau *et al*., in a preprint, have suggested that the observed flipping of V625 carbonyls might be a potential C-type inactivation mechanism for the hERG channel[26,27]. Their study involved altering the external $K^+$ concentration to predispose the channels towards a non-conductive, inactivated-like state[26,27]. Li *et al.* made a similar observation in their study, demonstrating the constriction between G626 carbonyls which occurs asymmetrically in only two opposing subunits and not in the others to be a potential C-type inactivation mechanism[15]. Indeed, this constriction was accompanied by the flipping of V625 carbonyls in the same subunits where the constriction was observed[15]. Both structural features, which are also replicated in the AlphaFold-predicted model, might collectively constitute components of the inactivation mechanism.

Furthermore, Cuello *et al.* in their study of KcsA channel identified a similar constriction at G77 within the SF and a corresponding reorientation of the V76 carbonyl, resulting in a dilation in the SF at this location and correspondingly the loss of the S2 and S3 ion binding sites[31]. Consequently, they suggested this backbone rearrangement as a fundamental molecular mechanism underlying C-type inactivation in $K^+$ channels[31]. In other studies on Shaker and $K_V1.3$ channels, dilation in the upper SF has been proposed to be a potential C-type inactivation mechanism but without notable constriction at the SF glycine[32–35]. The inactivation processes across various $K^+$ channels may exhibit similarities to some degree,

but possess distinct differences that could account for the observed variability in inactivation rates among different $K^+$ channels[1].

**Figure 3. Structural comparison of different hERG channel state models.** ***a)*** *Visual*




*comparison of the closed, open-, and inactivated-state models.* ***b)*** *Pore radius for the SF and drug binding region (upper) and for the entire pore (lower).* ***c)*** *Comparison of the VSD conformation in each model, showcasing the positively charged Arg and Lys gating-charge residues, located on the S4 helix, and the gating charge transfer center residue, F463, on the S2 helix.* ***d)*** *Measurement of the distance between the $C_\alpha$ atom of F463 to the $C_\alpha$ atom of each of the gating-charge residues.*

For the VSD, we measured the distances between the backbone $C_\alpha$ atoms of the gating charge residues (K525, R528, R531, R534, R537, K538) on the S4 helix and the gating charge transfer center F463 on the S2 helix, as shown in **Figure 3c**. Although we observed an increased distance between the gating charge residues and the charge transfer center residue in the closed state, this separation was not due to a straight downward movement of the S4 helix. Instead, the closed-state model S4 exhibited a minor kink around residue R531 and lateral movement toward the center of the channel, impacting the S4 – S5 linker and consequently nudging the S5 helix inward, effectively narrowing the pore. It should be emphasized that the predicted closed-state model exhibits lower confidence levels for the S4 helix and S4-S5 linker residues (pLDDT ≤ 75) when compared to their counterparts in models of other states, necessitating caution in interpreting the physiological implications of this observation. Conversely, the pore region, which demonstrates closure, is characterized by a higher prediction confidence (pLDDT ≥ 75), suggests a more robust and reliable representation in this part of the model.

In **Figure 4a** and **b**, we analyzed interactions of extracellular S5-P linker (I583 – Q592) along with SF and surrounding SF residues (S620 – N633) through heatmaps detailing hydrogen bonding, π stacking, cation-π, and salt bridge formation similarities and differences (**Table S1** shows detection criteria; distance-based contact maps[36] for all residues are shown in **Figure S2**). Distinct interaction patterns between open- and inactivated-state models were observed in these regions (**Figure 4b, c**). N633 atop the SF forms hydrogen bonds with S5-P linker G584 and Q592 from an adjacent subunit, while N629 forms hydrogen bonds with adjacent S631. Additionally, SF G626 forms a hydrogen bond with S620 behind the SF. Within the S5-P linker, N588 and D591 also display hydrogen bonding. However, these stabilizing interactions in the open-state model SF region are absent in the inactivated-state model,

where only intra-subunit hydrogen bonds between I583 (S5-P linker) and N633 (SF) occur, along with V630 bonding with N629 atop the SF.

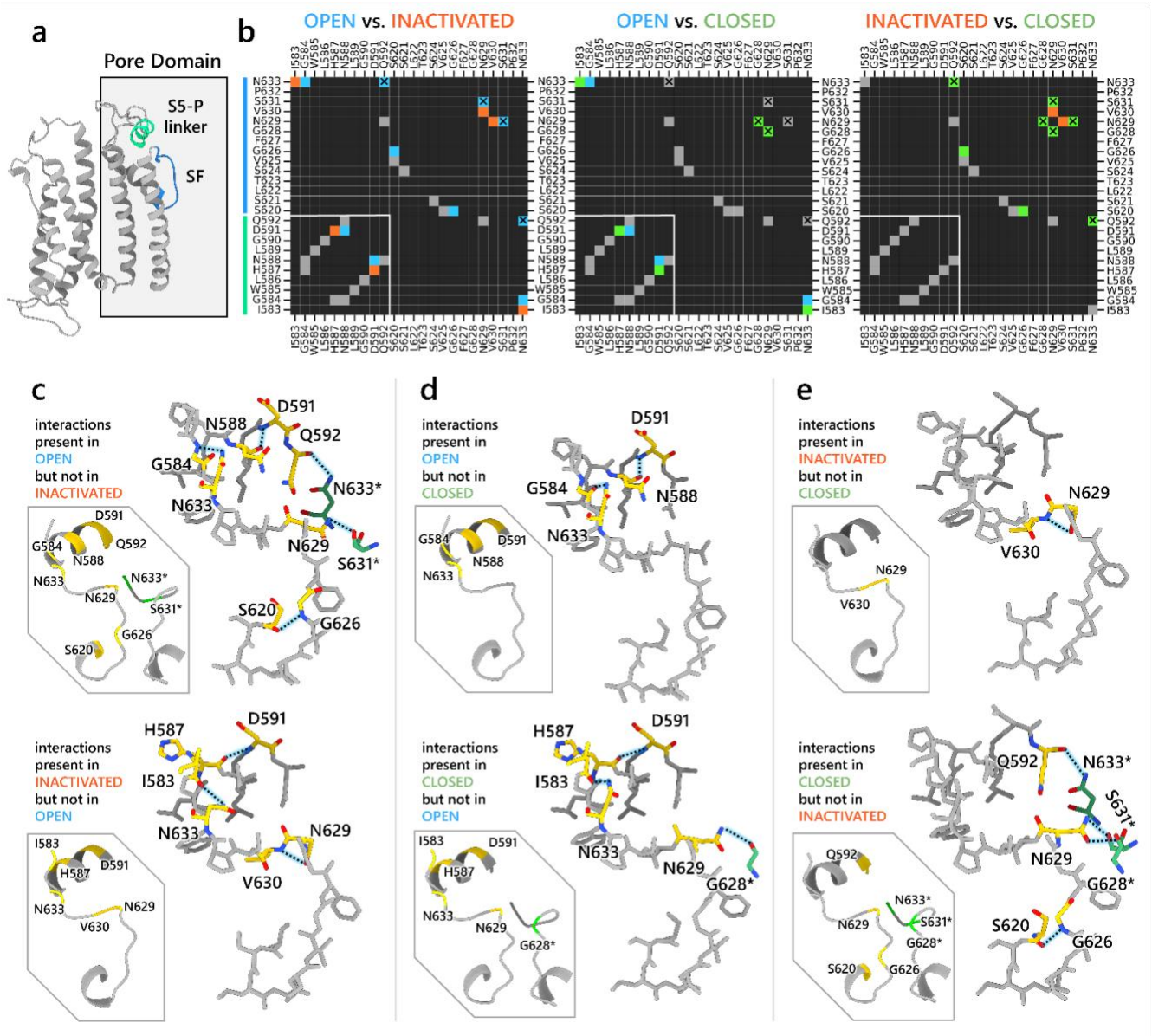



**Figure 4. Interaction network analysis showcasing residue-residue interactions in the S5-P linker (I583 – Q592) and region surrounding the SF (S620 – N633).** ***a)*** *An image of a hERG channel subunit with the analyzed regions colored in cyan and blue.* ***b)*** *Heatmaps showing intrasubunit and intersubunit (marked by X) interactions between each residue in the analyzed regions. The interactions analyzed are hydrogen bonding, π stacking, cation-π, and salt bridges. Black cells indicate no interactions. Gray cells indicate an interaction is present in*

*both states. Blue, orange, and green colored cells indicate the interaction is present only in the open, inactivated, or closed state, respectively, but not in the other state being compared in the map. White lines are added to separate S5-P linker residues from the SF region residues.* ***c, d, e)*** *Visualization of the interactions being present in one state but not the other. Gold-colored residues are involved in the interactions. Green-colored residues, named with an asterisk at the end, are from an adjacent chain but are interacting with gold-colored residues. Dashed lines represent hydrogen bonds.*

To corroborate our findings, mutations involving the residues above have been shown to impact hERG inactivation as evidenced in numerous clinical and experimental studies[14,37–44]. For instance, the S631A mutation, linked to short QT syndrome, causes a positive shift in hERG half-inactivation voltage[38], while S631C speeds up fast inactivation[39]. N629D, N629S, and N633S mutations, found in long QT syndrome patients, disrupt inactivation and $K^+$ selectivity [14,40]. N588K/E and Q592K modulate rapid inactivation[41,42], G584S leads to inactivation gating defects[45], and S620T abolishes hERG inactivation[43]. V630L caused a negative shift in the voltage dependence of steady-state inactivation[44], while H587P/K disrupts the C-type inactivation process and $K^+$ selectivity[37]. N588C and I583C mutations have high and intermediate impact of hERG inactivation[46]. Lastly, the D591R/Q592R double mutation inhibits inactivation[42].

There are fewer differences between the open- and closed-state models (**Figure 4b, d**). In the open-state model, D591 from the S5-P linker forms hydrogen bonds with N588, and G584 bonds with N633 at the top of the SF. These interactions are absent in the closed-state model, where H587 (instead of N588) bonds with D591 within the S5-P linker, and I583 (replacing G584) interacts with N633 at the SF top. Additionally, G628 from an adjacent subunit forms a hydrogen bond with N629 atop the SF. Analyzing the differences between the inactivated- and closed-state model (**Figure 4e**), the inactivated-state model uniquely features a V630-N629 hydrogen bond, whereas the closed-state model exhibits intersubunit hydrogen bonds

between N633 and Q592, and between S631/G628 and N629. Furthermore, in the closed-state model, S620 forms a hydrogen bond with G626, stabilizing the SF conformation.

In **Figure S3**, we compared the S6 pore-lining helix orientation across various models. The closed-state model features a straight S6 helix with a minor kink. On the contrary, both the open- and inactivated-state models exhibit a pronounced kink around I655, as identified in a prior study[47], which facilitates pore opening and distinguishes the inactivated-state from the closed-state model. Notably, a slight rotation differentiates the S6 helix in the open- and inactivated-state models, altering the alignment of drug-binding residues Y652 and F656. The interaction network analysis results from **Figure 4b, c** suggest that alterations in the hydrogen bond network around the SF region, during the transition from open to inactivated state, might pull on the S6 helix and influence its orientation (**Figure S1e, f**) – a subtle yet potentially impactful change for drug binding which we will explore later. In agreement with our observations, a study by Helliwell *et al.* also suggested that a slight clockwise rotation of the S6 helix in the hERG open-state cryo-EM structure[7] could align the S6 aromatic side chains, particularly F656, into a configuration to enable interactions with inactivation-dependent blockers that more accurately reflects experimental data[48].

### *d. Molecular dynamics simulations to assess ion conductivity*

We performed all-atom molecular dynamics (MD) simulations on two hERG channel models described above, one in the open state and the other in the predicted inactivated state, to evaluate their ion conduction capabilities. Unlike the closed-state model, both open- and inactivated-state models should allow ions and water to enter and traverse the channel pore reaching the SF region. However, only the open state-model is expected to facilitate ion conduction through its SF.

To investigate ion conduction in the selectivity filter, we considered two conditions, as shown in

**Figure 5a**: one in which the SF initially contained only $K^+$ ions and another in which both ions and water were present to test previously proposed direct (or Coulombic) and water-mediated $K^+$ conduction knock-on mechanisms[49,50] as in a previous study[14]. In the direct

knock-on (ions-only) scenario, we manually positioned $K^+$ atoms in the putative $K^+$ binding sites of S0, S2, S3, and S4 within the SF. For water-mediated knock-on (the alternating ions and water molecules) scenario, $K^+$ ions were placed in the S1, S3, and $S_{cav}$ positions, while water molecules were inserted into the S0, S2, and S4 positions. These models were incorporated into phospholipid bilayers consisting of *1-palmitoyl-2-oleoylphosphatidylcholine* (POPC) molecules, as depicted in

**Figure 5b**. Subsequently, we conducted simulations for each case under three membrane voltage conditions: 0 mV, 500 mV, and 750 mV, each lasting 1 μs. This resulted in a total of six MD simulations for each model.

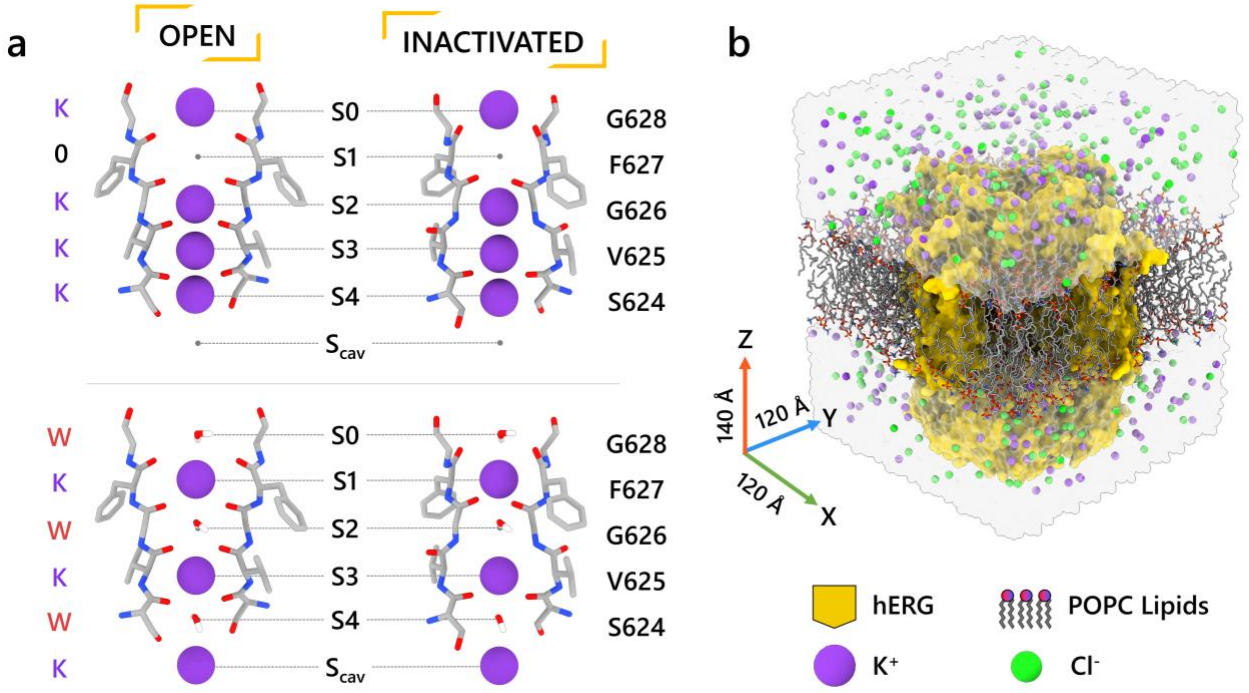



**Figure 5. Setup of MD simulations to assess ion conduction in the open and inactivated models.** ***a)*** *Initial configuration of the SF, set to fill with either all $K^+$ ions, or alternating $K^+$ and water molecules.* ***b)*** *An example MD simulation box showing a hERG channel model (shown in yellow surface representation) embedded in POPC lipid bilayer (shown as sticks) and solvated by an aqueous 0.3 M KCl solution (shown as a transparent surface with $K^+$ and $Cl^-$ ions shown as purple and green balls, respectively).*

For all instances where a non-zero membrane voltage was applied, we observed $K^+$ conduction after equilibration in the open-state model (**Figure 6a, c** and **Figure S4a, c**), whereas such conduction was not observed in the inactivated-state model (**Figure 6b, d** and **Figure S4b, d**). For ions-only initial arrangement we observed that all $K^+$ ions initially located in the SF went across during 1 μs MD runs under applied 750 and 500 mV membrane voltage (**Figure 6a** and **Figure S4a**), whereas for the alternating water-ion initial SF configuration we observed conduction of SF ions as well as additional $K^+$ moving all the way across the channel pore (**Figure 6c** and **Figure S4c**). In both cases we saw combination of direct and water-mediated knock-on mechanisms as in our previous hERG channel MD simulations[11,14]. Control MD simulations conducted under zero voltage conditions revealed a single $K^+$ conduction event for the open-state model (**Figure S4g**) when the SF was initially filled with water molecules and ions, while no conduction events were observed in the remaining cases (**Figure S4e, f, h**).

Subsequently, we conducted an analysis of pore radius changes throughout the MD simulations (full results in **Figure S5**). In

**Figure 7a**, under zero voltage conditions, we observed consistent and distinct pore radius profiles across all simulations within their respective models. Specifically, MD simulations featuring the inactivated-state model consistently displayed a narrower pore radius when compared to simulations involving the open-state model. However, when subjected to high voltage conditions, the open-state model exhibited a shift towards an inactivated-like state, leading to a reduction in the pore width, which is consistent with an increased hERG channel inactivation propensity at more depolarized voltages[1]. Although we did not observe the outward flipping of the V625 backbone carbonyl oxygens in the SF during the 1 μs simulations of the open-state model, we did observe the flipping of the F627 backbone carbonyl oxygens as shown in

**Figure 7b**. Interestingly, this SF conformation with flipped F627 but not V625 backbone carbonyls is also observed in cluster 3 of the AlphaFold-predicted models shown in **Figure 2**. In the inactivated-state model MD simulations, V625 backbone carbonyl oxygens flipped

inward into ion conducting conformation more often with increasing voltage, but as long as one out of four of them remained outward oriented, no $K^+$ could pass through the SF.

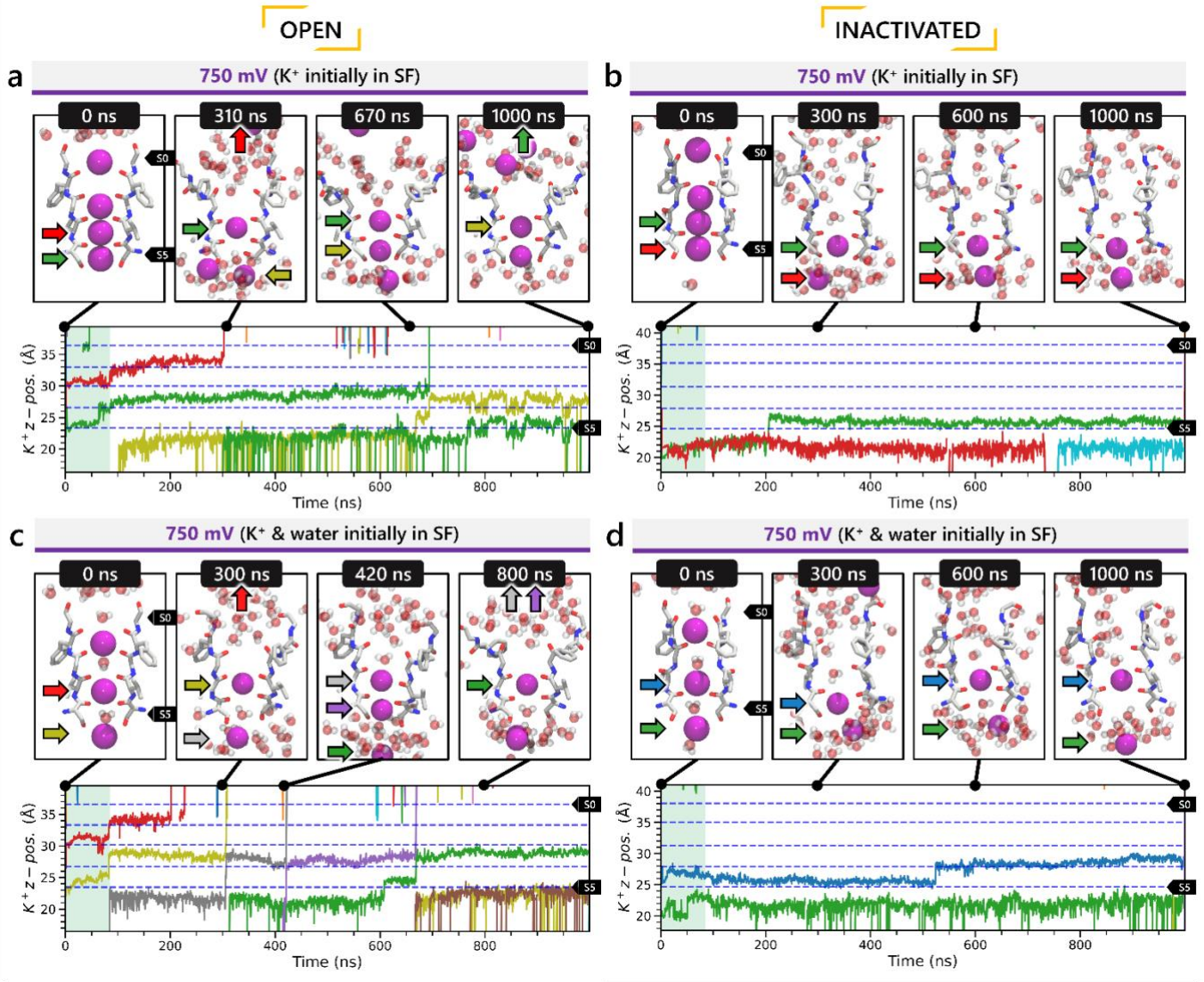



**Figure 6. Movement of $K^+$ ions through hERG SF.** *The z coordinates of $K^+$ ions are tracked as they traverse from the pore of the channel (lower y-axis limit) to the extracellular space (upper y-axis limit) under membrane voltage of 750 mV. Putative $K^+$ binding sites in the SF (S0 – S5) are marked using blue dashed lines in the plots.* ***a, c)*** *Results from MD simulations on the open model with the SF initially configured to have only $K^+$ ions or alternating $K^+$ / water*

*molecules, respectively.* ***b, d)*** *Results from MD simulations on the inactivated model with the SF initially configured to have only K⁺ ions or alternating K⁺ / water molecules, respectively.*

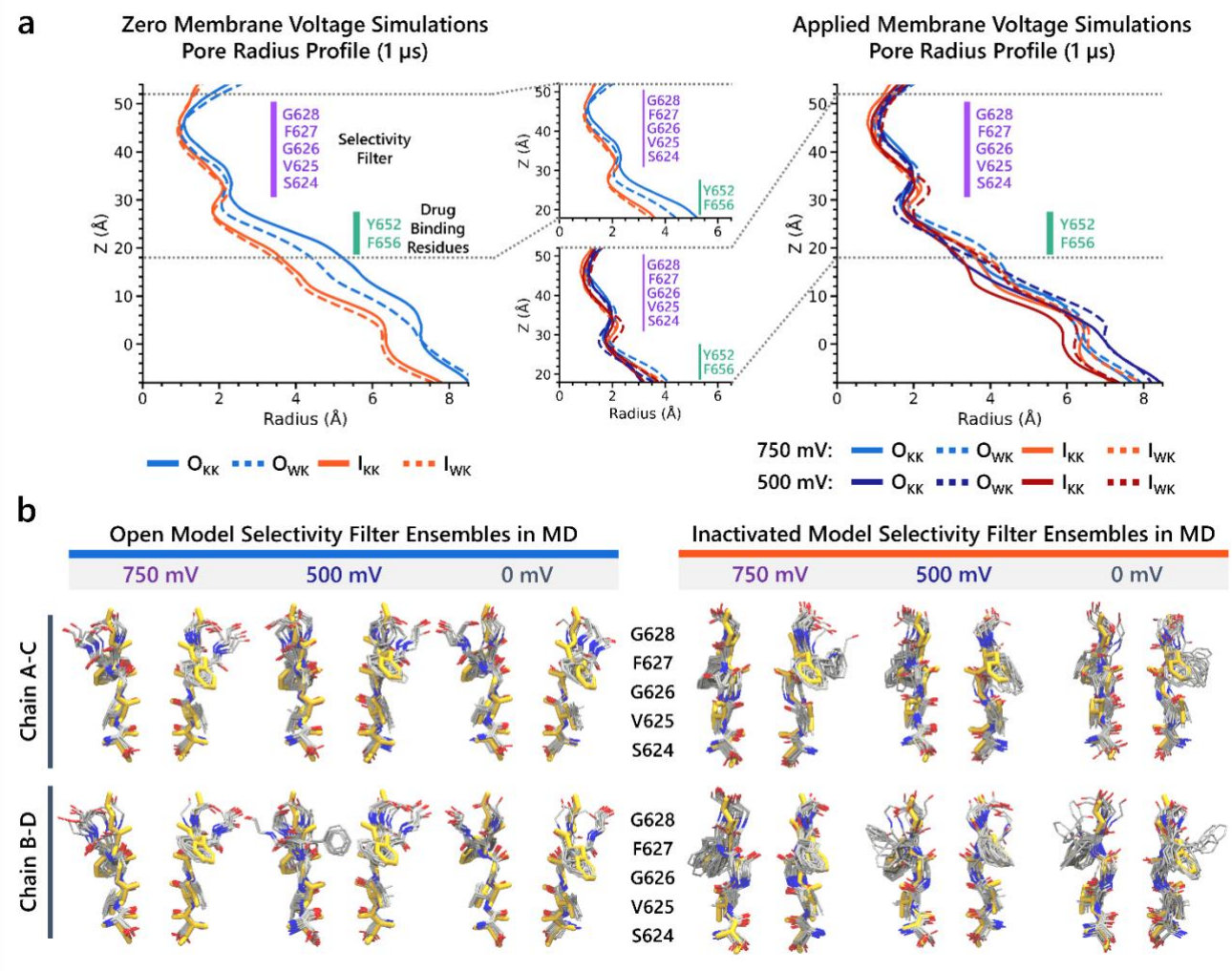



**Figure 7. Analysis of dynamics of the SF and pore conformations over the course of the 1 μs MD simulations.** ***a)*** *Pore radius averaged over each 1 μs long MD simulations with or without applied membrane voltage. Open- and inactivated-state model MD simulations are notated as O and I, respectively, with the subscripts KK and WK denoting whether the SF initially configured to have only K⁺ ions or alternating K⁺ / water molecules, respectively.* ***b)***

*Ensembles of SF conformation over the course of each MD simulation superimposed over each other. The gold-colored conformation indicates the initial conformation.*

### *e. Assessing state-dependent drug block using GALigandDock*

We utilized Rosetta GALigandDock software[13] to dock 19 drugs from different classes, including alternative protonation states, into our hERG state-specific channel models. This process aimed to evaluate and corroborate their state-dependent binding interactions with experimental studies, specifically in terms of relative binding affinities. **Figure 8** presents these findings in the form of Rosetta GALigandDock[13] binding energies (lower values mean more favorable binding). Consistent with published studies, most drugs showed stronger binding to the inactivated-state hERG channel model, including astemizole[8], terfenadine[8], cisapride[8], d/l-sotalol[8], dofetilide[8], and E-4031[51,52]. Drugs including moxifloxacin[53], quinidine[8], and perhexiline[8] did not show strong preference for the inactivated model, aligning with findings from hERG experimental studies using inactivation-deficient mutants[8] or "step-ramp" voltage protocol[53].

In our GALigandDock docking results, most drugs exhibited increased binding affinity to the closed-state hERG model compared to the open-state hERG model. Drugs are unable to bind to the closed state from the intracellular space because the pore is closed. However, they can become trapped if they are already bound when the channel transitions from open to closed state, as shown in experiments for dofetilide[54], cisapride[54], terfenadine[54,55], E-4031[55], nifekalant[56], etc. To model drug trapping, we placed the drug in a pocket beneath the selectivity filter in the closed pore configuration before docking. However, this method does not consider how the conformational shift from the open to the closed state might influence drug binding. Under physiological conditions, the pore gating motion from open to closed might expel drugs from the pore instead of pushing them deeper. This limitation might account for some inconsistencies noted in our docking study, particularly regarding the apparent trapping of drugs such as amiodarone and haloperidol, which is at odds with experimental results[57]. However, these preliminary results could pave the way for more

thorough investigations, employing advanced computational techniques to delve deeper into the dynamics of drug trapping[58,59].


**Figure 8. GALigandDock drug docking results to different hERG channel models.** *Each bar plot represents the estimated energy of binding in Rosetta energy units (R.E.U.) for the named drug to the open-, inactivated-, and closed-state hERG channel models. Lower values mean more favorable binding. 25,000 docking poses were generated for each drug/channel model pairing. The top 50 poses were clustered, and the plotted energy represents the average*

*free energy of the top cluster along with the standard deviation. The labels (0), (+), (±) after the drug name indicate its neutral, cationic, or zwitterionic form, respectively.*

In most cases, the increased drug binding affinity to the inactivated-state hERG model can be attributed to the protrusion of residue F656 aromatic side chain into hERG central cavity (**Figure 3a**). This protrusion has the potential to strengthen drug binding through providing extra π stacking or hydrophobic interactions. Furthermore, the side chains of F656 could act as a barrier, blocking drugs that are bound deeper within the pore from escaping even when they do not have direct interactions with the drugs. Similar observations were made for the closed-state model, where the pore closure also positions the F656 side chain inward. In addition, the slanted orientation of the Y652 aromatic ring and repositioning of the S624 sidechain hydroxyl oxygen in the inactivated model enabled the formation of a pocket beneath the SF (**Figure 3b**), enabling drugs to reach deeper into the pore.

As an example,
**Figure 9** focuses closely on the binding of astemizole, dofetilide, and quinidine in their cationic forms (the predominant species at the physiological pH 7.4) to different hERG channel models. Experimental evidence[8] suggests that dofetilide and astemizole bind more potently to the inactivated state, while quinidine does not exhibit this preference.

*Astemizole:* In the open-state model, astemizole engages in π stacking with the Y652 aromatic side chain via its benzimidazole group and forms hydrophobic contacts with adjacent residues (A653, F656, and Y652). Its binding occurs just below S6 helix residue Y652, not deeply within the pore, resulting in a binding pocket that remains partially open to external solvent. In contrast, in the inactivated-state model, astemizole demonstrates an increase in interactions and becomes fully enclosed within the binding site deeper into the pore, engaging in hydrophobic interactions with S6 residues S649, Y652, F656, I655 and S5 residue F557. Astemizole phenyl ether group engages in the π stacking with hERG Y652 and F656 side chains from adjacent subunits. At the other end, the aromatic ring of Y652 phenol side chain also engages in π stacking, while its oxygen atom acts as a non-ideal hydrogen bond acceptor (i.e., donor angle of ≤ 45° and bonding distance of ≤ 3.8 Å[60]) to astemizole

nitrogen. Furthermore, the SF residue S624 side chain participates as a hydrogen bond acceptor with the piperidine ring of the drug. The binding pattern in the closed-state model is similar to that in the inactivated-state model, but it includes additional hydrophobic interaction with S660 and an extra hydrogen bond provided by the phenol oxygen of Y652.

In 2021, a study by Asai *et al.* solved the cryo-EM hERG structure in complex with astemizole[10]. Although the structures deposited to the PDB (7CN0[10] and 7CN1[10]) did not depict any bound ligand, a possible astemizole binding pose was suggested based on residual EM density. The density map showed a π-π interaction of astemizole with Y652 and a hydrogen bond with S624[10]. In addition, electrophysiology study by Saxena *et al.* has identified F557 and Y652 as crucial for astemizole binding, as determined through mutagenesis[61]. In 2022, a study by Nano Imaging Services unveiled a cryo-EM structure of hERG with astemizole at a refined 2.7 Å resolution[62], showcasing a binding pose remarkably similar to what we observed in our study for the hERG inactivated state.

*Dofetilide:* In the open-state hERG model, dofetilide interacts with multiple residues of the pore lining S6 helices. For instance, it forms hydrogen bonds with two S660 residues from opposite subunits and two A653 residues from adjacent subunits. Additionally, there are hydrophobic interactions with Y652 and G657. When bound to the inactivated-state model, dofetilide settles deeper into the pore, extending horizontally above the Y652 residue, where it forms π stacking and cation-π interactions, as well as hydrogen bonds with the Y652 residues of all subunits. The backbone carbonyl oxygen of residue S649 also engages in hydrogen bonding with the nitrogen in the sulfonamide group of dofetilide. There are also hydrophobic contacts with hERG S6 residues S649, F656, and I655 as well as residue F557 on the S5 helix. The closed-state model pore provides a substantially narrower space, leading dofetilide to bind vertically along the pore axis, forming hydrophobic contacts with S624, S649, Y652, F656, and S660. Here, the oxygen and nitrogen of its sulfonamide group create a series of hydrogen bonds with hERG residues T623 and Y652. The central cationic ammonium group of the ligand forms cation-π interactions with two F656 residues from adjacent subunits. Lastly, the oxygens in the remaining sulfonamide group are involved in hydrogen bonding with the S660 residues from every subunit. In the inactivated-state and

closed-state models, the binding pocket is completely enclosed, notably unlike the open-state model.

In line with our results, mutations affecting A653 have demonstrated an influence on hERG channel blocking by dofetilide[63]. Additionally, electrophysiology study by Saxena *et al.* revealed that mutants Y652A and F557 altered dofetilide binding[61]. Simulation studies have highlighted interactions involving non-polar residues A653, Y652, F656, as well as polar residues S624, T623, S649, and S660 with dofetilide[64].

*Quinidine:* Unlike the previous two drugs, quinidine does not exhibit significant state-specific binding based on our docking predictions (**Figure 8**). In the open-state hERG model, its small size allows for complete enclosure within the binding pocket, where it forms non-ideal hydrogen bonds with residue Y652 and hydrophobic contacts with another Y652 from an adjacent chain as well as A653 and S660. In the inactivated-state model, quinidine binds lower in the pore, resulting in a partially open binding pocket. Here, its quinoline rings engage in π-stacking interactions with the F656 side chain above and hydrophobic contacts with another F656 from the opposite subunit. The hydroxyl oxygen forms a non-ideal hydrogen bond with residue S660. In the closed-state model, quinidine becomes trapped deeper in the pore and is fully enclosed. Its quinoline rings π stack with multiple F656 side chains below and the Y652 side chain above drug binding position. The hydroxyl oxygen now establishes a non-ideal hydrogen bond with Y652, with additional hydrophobic contacts observed with Y652 side chains and the T623 backbone. Corroborating our results, Sănchez-Chapula *et al.* previously found that mutating key aromatic residues (Y652, F656) in the S6 domain reduced the efficacy of quinidine by more than 100-fold[65].

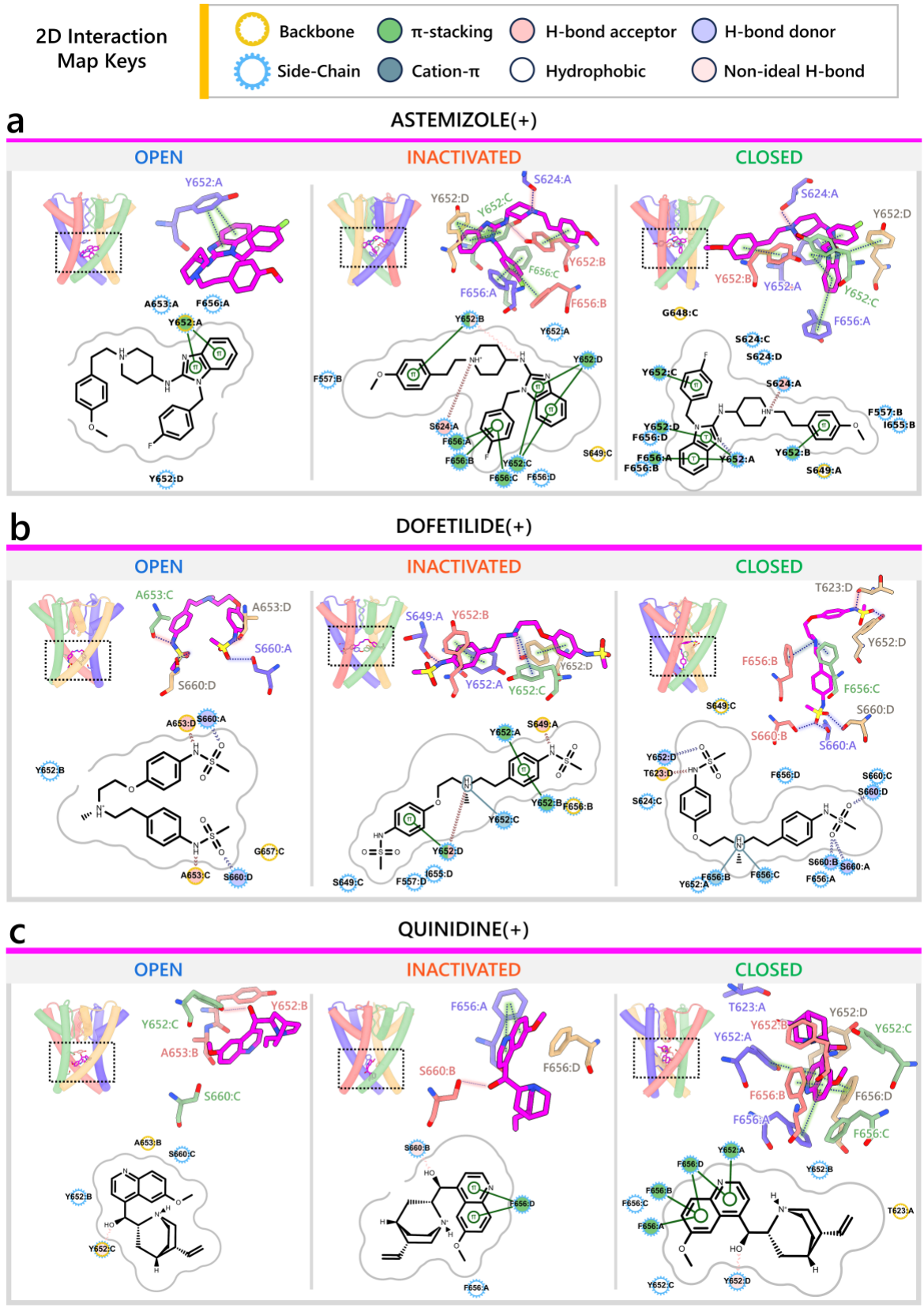

2D Interaction Map Keys
Backbone
π-stacking
H-bond acceptor
H-bond donor
Side-Chain
Cation-π
Hydrophobic
Non-ideal H-bond
a
ASTEMIZOLE(+)
OPEN
INACTIVATED
CLOSED
b
DOFETILIDE(+)
OPEN
INACTIVATED
CLOSED
c
QUINIDINE(+)
OPEN
INACTIVATED
CLOSED

**Figure 9. Visualization of interactions for cationic astemizole (a), dofetilide (b), and quinidine (c) to different hERG channel models.** *Each panel includes 3 subpanels showcasing drug interactions to the open-, inactivated, and closed-state hERG channel models. In each subpanel, an overview of where the drug binds within the channel pore is shown on the upper left, a 3D visualization of interactions between each channel residue (blue, red, green, and tan residues are from the subunit A, B, C, or D, respectively) to the drug (magenta) is shown on the upper right, and a 2D ligand – protein interaction map is shown at the bottom. A continuous gray line depicts the contour of the protein binding site, and any breaks in this line signal areas where the ligand is exposed to the solvent.*

### 4.3) Conclusions

Employing AlphaFold2 under default parameters typically leads to the prediction of the hERG channel in a single state conformation. In our study, we presented an approach to guide AlphaFold2 to model the hERG channel in various conformational states. We further validated the physiological significance of these states using computational simulations. This methodology has provided new insights into the structural and functional behavior of the hERG channel, including state-specific drug interaction mechanisms.

The predicted models, which putatively represent the closed, open and inactivated states of the hERG channel, have revealed detailed structural aspects that were previously elusive, particularly in the pore region where drug binding occurs. The closed-state hERG channel model demonstrates a constricted pore and deactivated voltage-sensing domains, suggesting a structural basis for the prevention of ion passage. The open-state hERG channel model, conversely, illustrates a pore conformation accessible to ion conduction, aligning with the physiological role of the channel in cardiac repolarization. The inactivated-state hERG channel model, a prediction of which greatly contributes to our understanding to hERG channel gating mechanism, shows a non-conductive selectivity filter conformation with significant changes to the pore region that could be a key in understanding state-specific drug-induced channel blockade. The non-conductive SF predicted by AlphaFold2, characterized by flipping of SF V625 carbonyl oxygens away from the central axis and minor constriction at the level of G626, aligns with prior computational analyses and preliminary

experimental reports as a possible hERG C-type inactivation mechanism[15,26,27]. The structures referenced in these studies have not been published in any publicly available databases and used for the training of AlphaFold2 to our knowledge. Our approach thus underscores the substantial yet untapped potential of AlphaFold2 to predict *de novo* channel conformations with physiological relevance.

Molecular dynamics simulations were crucial in confirming these model functional properties. The observed differences in ion conduction between the open and inactivated hERG channel states illustrate the impact of structural changes on ion channel activity. In addition, the drug docking analysis matched expectations, with significant preference for the inactivated-state model across a variety of hERG blockers. In the model representing the inactivated state, a subtle rotation of the S6 pore-lining helix from the open state results in a more prominent protrusion of the F656 side chains into the central cavity and widens the inner cavity below the SF due to the repositioning of Y652 side chains. These structural modifications facilitate deeper binding of certain drugs in the pore, increase interactions between drugs and the channel, impede ion movement, and obstruct the exit routes for drugs that are bound within the pore. Although prior studies have proposed potential molecular mechanisms for hERG C-type inactivation[15,26,27], they have primarily concentrated on conformational shifts leading to a non-conductive selectivity filter. Our research, however, also sheds light on how these alterations extend to the pore region and subsequently impact drug binding, an issue of significant relevance in safety pharmacology and drug development.

Moreover, the study highlights the potential for deep-learning based computational methods like AlphaFold2 in structural biology, especially for complex proteins like ion channels and a simultaneous prediction of multiple conformations putatively representing different channel states. This is a challenging task for experimental structure determination as it requires changes in experimental conditions and/or classification of multiple cryo-EM images. Our proposed AlphaFold2 approach using multiple templates and fine-tuned structural prediction parameters allowed us to overcome this challenge and predict atomistic structural models of the hERG channel in the putative closed, open, and inactivated

states. These models were further validated using all-atom molecular dynamics simulations under applied voltage to test ion conduction through model channel pores as well as molecular docking of multiple drugs to test their state-dependent binding preferences compared to known experimental data. The intersection of structural modeling, molecular docking and molecular dynamics simulations, and supporting experimental data offers a comprehensive approach to understanding protein structures and their links to biological functions.

Despite the promising results, our study is not without limitations. While AlphaFold has demonstrated remarkable accuracy in numerous instances[66–68], it is important to note that the predicted models, particularly in regions where the model has lower confidence, may not always be reliably accurate to predict drug binding poses[69]. Moreover, hERG channel inactivation and closure might encompass a range of states, and the conformations we have identified could potentially represent just a few possibilities within this broad spectrum. The results from GALigandDock drug docking, while providing valuable insights, should be regarded as provisional due to inherent limitations of this approach, including limited accuracy and transferability[19]. Additionally, knowing the computational binding free energies of a drug is not the complete story; binding rates such as $k_{on}$ (association rate) and $k_{off}$ (dissociation rate) are also crucial for a quantitative evaluation of drug binding to the channel. The results herein primarily serve as a basis for further detailed investigation using methods like molecular dynamics simulations linked to multiscale functional modeling, as demonstrated in previous studies[11,70].

In conclusion, this study not only advances our knowledge of the hERG channel structural dynamics but also sets a precedent for using AlphaFold2 or similar computational methods, such as RoseTTAFold[71], to study other ion channels and membrane proteins. The insights gained from this study are expected to have a lasting impact on the fields of ion channel physiology and pharmacology including cardiac drug safety, providing a foundation for future research aimed at developing safer and more effective therapeutic agents.

## 4.4) Methods

### *a. hERG channel model generation and refinement*

Structural data for the Eag $K_V$ channel in its closed conformation (PDB 8EP1[21]) and the hERG channel in its open state (PDB 5VA2[7]) were sourced from the Protein Data Bank (PDB[72]). To accurately represent the open state of hERG, we integrated known cryo-EM structures (PDB 5VA2[7]) with modeled segments of absent loops, employing Rosetta[28] software suite LoopRemodel mover in a symmetrical configuration. The refinement of all final models was performed using Rosetta FastRelax protocol, which iterated 15 times and included a simulated POPC membrane environment. For each finalized model, 10 separate relaxation runs were executed, and the highest-scoring model from these runs was selected for further simulation and analysis.

Cloud computational resources, specifically NVIDIA A10 GPUs (24 GB of VRAM) provided by Oracle, facilitated the modeling processes in ColabFold[22]. To enhance sampling diversity, in each modeling session, ColabFold was configured to utilize 256 sequence clusters and 512 extra sequences (*max_msa = "256:512"*) under the MSA option and have dropouts enabled (*use_dropout* = True). The prediction phase was further diversified by running 20 iterations with different random seeds (*num_seeds* = 20), each producing 5 models and resulting in 100 total models per run. An exception to this approach was the initial phase of inactivated-state hERG model creation, where a single seed number was used to generate 5 models. The number of iteration cycles was set to 20, with an early stopping tolerance of 0.5 based on per-residue confidence metric pLDDT deviation from the previous cycle.

### *b. Atomistic MD simulations of hERG channel conduction*

The CHARMM-GUI web server[73] was employed to embed hERG channel structural models within tetragonal patches of phospholipid bilayers, each comprising approximately 230 to 240 POPC lipid molecules. The resulting assemblies were immersed in a 0.3 M KCl aqueous solution, culminating in molecular systems with an approximate total of 144,000 atoms.

Residue protonation states were determined assuming a neutral pH environment, and each channel subunit was terminated with standard charged N- and C-termini.

MD simulations were conducted using the Amber22[74] software suite. The simulations utilized standard all-atom CHARMM force fields specific to proteins (CHARMM36m[75]), lipids (C36[76]), and ions, in conjunction with the TIP3P water model[77]. The systems were maintained at 310.15 K and 1 atm pressure in the isobaric-isothermal (*NPT*) ensemble, facilitated by Langevin temperature equilibration and the Berendsen barostat. Non-bonded interactions were calculated with a cutoff of 12 Å. Long-range electrostatic forces were addressed using the Particle Mesh Ewald (PME) method[78], and van der Waals interactions were not subjected to long-range correction as per recommendations for the C36 lipid force field. All hydrogen-involved covalent bonds were constrained using the SHAKE algorithm to enable a 2-fs MD simulation time step.

Equilibration began with harmonic restraints imposed on protein atoms and lipid dihedral angles, initially set at 20 kcal/mol/Å$^2$ and reduced to 2.5 kcal/mol/Å$^2$ over 6 ns. A subsequent 90 ns equilibration phase further decreased the restraints to 0.1 kcal/mol/Å$^2$, initially encompassing all atoms in the protein and eventually focusing solely on the backbone atoms of pore-lining residues (G546 to F720). In selected MD simulations, an electric field was applied along the Z direction to mimic membrane voltage, increasing linearly over the final 10 ns of equilibration to reach either 500 or 750 mV. This setup prefaced a production phase lasting 910 ns, totaling 1 μs of total simulation time per each case.

### *c. Docking of drugs to hERG channel models*

Ligand structures (i.e., pharmaceutical compounds) were retrieved from the PubChem[79] and ZINC[80] databases. Afterward, each ligand underwent modification processes including bond corrections and setting of protonation states using the Avogadro software[81]. In this study, we considered the protonation state of the top two most dominant species at the physiological pH of 7.4, computed using the Henderson–Hasselbalch equation. After these

initial modifications, each ligand partial atomic charges, as well as atom and bond types, were refined using AM1-BCC correction via the Antechamber module in AmberTools22[74,82].

Each ligand was individually positioned within the pore of the hERG channel, specifically between the key drug-binding residues, Y652 and F656, located on the pore lining S6 helices. This manual placement was a preparatory step for the docking process. Docking was executed using the GALigandDock[13] Rosetta mover, a component of the Rosetta software suite[28]. The *DockFlex* mode was utilized for this purpose with a spatial padding of 5 Å. For every individual hERG channel – ligand pair, a total of 20 parallel runs were conducted. Each run generated 1,250 unique docking poses, cumulatively providing a substantial collection of 25,000 poses for each ligand – hERG channel complex. This extensive array of poses was intended to ensure a comprehensive exploration of potential ligand binding configurations and orientations within the hERG channel pore.

Utilizing in-house Python scripts, we selected the top 50 docking poses for each ligand based on the lowest free energy of binding as calculated by GALigandDock. These selected poses were then clustered based on ligand all-atom RMSD, factoring in the symmetry of the hERG channel to ensure accurate grouping. The clustering was executed with an RMSD threshold of 2.5 Å, and only clusters containing a minimum of 8 poses were deemed significant; those with fewer were classified as outliers. Within the cluster with the most favorable average binding free energy, we selected the top scoring pose as the representative pose for further analysis. This pose was then subjected to a detailed analysis of protein-ligand interactions utilizing the Grapheme Toolkit from the OpenEye software suite[83]. The criteria for detecting interactions are outlined in the OEChem Toolkit manual[60], with two modifications: the *MaxContactFraction* is set to 1 (default: 1.2), and the *MaxCationPiAngle* is adjusted to 30° (default: 40°). The interaction patterns and binding sites were subsequently rendered as a two-dimensional image for comprehensive visual interpretation. Additionally, for a more detailed understanding of the spatial arrangement, three-dimensional visualization of the protein-ligand complexes was conducted using the ChimeraX[20] software.

### *d. Molecular graphics and interaction analysis*

Molecular graphics visualization was performed using ChimeraX[84]. MD trajectory and simulation images were visualized using VMD[85]. Interaction network analysis was performed using the Protein-Ligand Interaction Profiler[86] with criteria outlined in **Table S1**.

## 4.5) Supplementary Information

**Table S1. Criteria for Different Type of Non-bonded Interactions Used in Analyses**

| TYPES OF INTERACTION | DETECTION CRITERIA | FILTERING CRITERIA |
|---|---|---|
| **Hydrogen Bonds** | Distance between the hydrogen bond donor (D) and acceptor (A) should be less than 3.8 Å.<br>The angle D-H···A should be above 110°. | Exclude hydrogen bonds between atoms that already form a salt bridge.<br>A hydrogen bond donor can participate in only one hydrogen bond, while acceptor atoms can form multiple hydrogen bonds. |
| **Salt Bridges** | Geometric centers of oppositely charged groups that come within 4.5 Å. | |
| **Cation-π** | Pairing of a positive charge and an aromatic ring if the distance between the charge center and the aromatic ring center is less than 6 Å. | |
| **π-Stacking** | Geometric centers of two aromatic rings within 5.5 Å. The angle between the rings should deviate no more than 30° from the optimal angle (90° for T-stacking, 180° for P-stacking). When projecting each ring center onto the opposite ring plane, the distance between the other ring center and the projected point (offset) should be less than 2 Å. | |

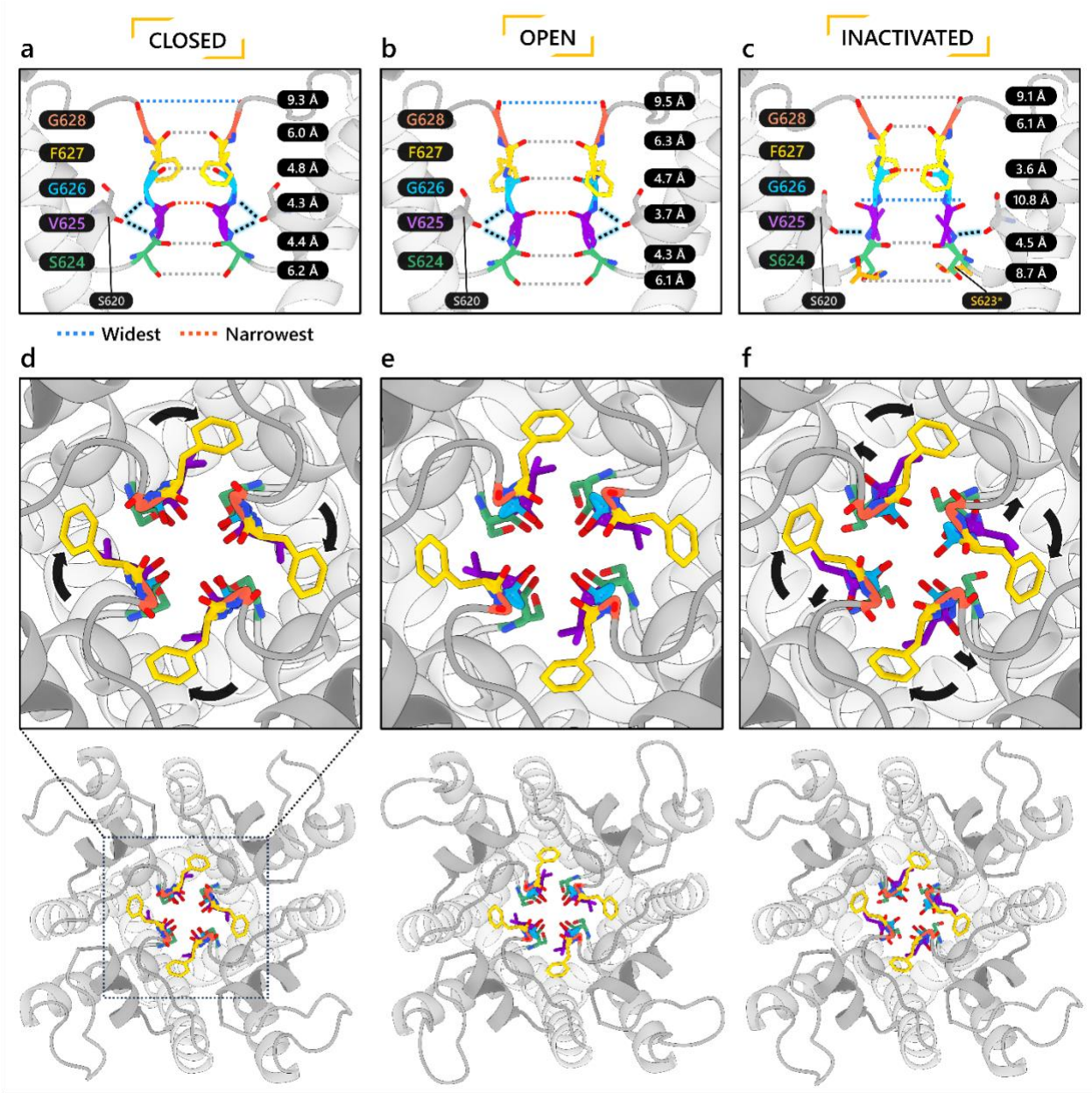



**Figure S10. Comparison of the SF in hERG closed- (a, d), open- (b, e), and inactivated-state (c, f) models.** ***a, b, c)*** *Measurement of the distances between each carbonyl oxygen lining the conduction pathway in the SF. In the open- and closed-state models, S620 carbonyl interacts with G626 and S624 backbone amides. In the inactivated-state model, the hydrogen bond between S620 and G626 is absent due to a reorientation of V625 backbone. However, at the bottom of the SF, S624 sidechain interacts with S623 carbonyl from an adjacent subunit.* ***d, e, f)*** *View of the SF from the extracellular side. Large arrows indicate the rotation of the F627 side*

*chain, while small arrows show the rotation of the loop that connects the upper SF to the S6 helix, all relative to the equivalent in the open-state model.*

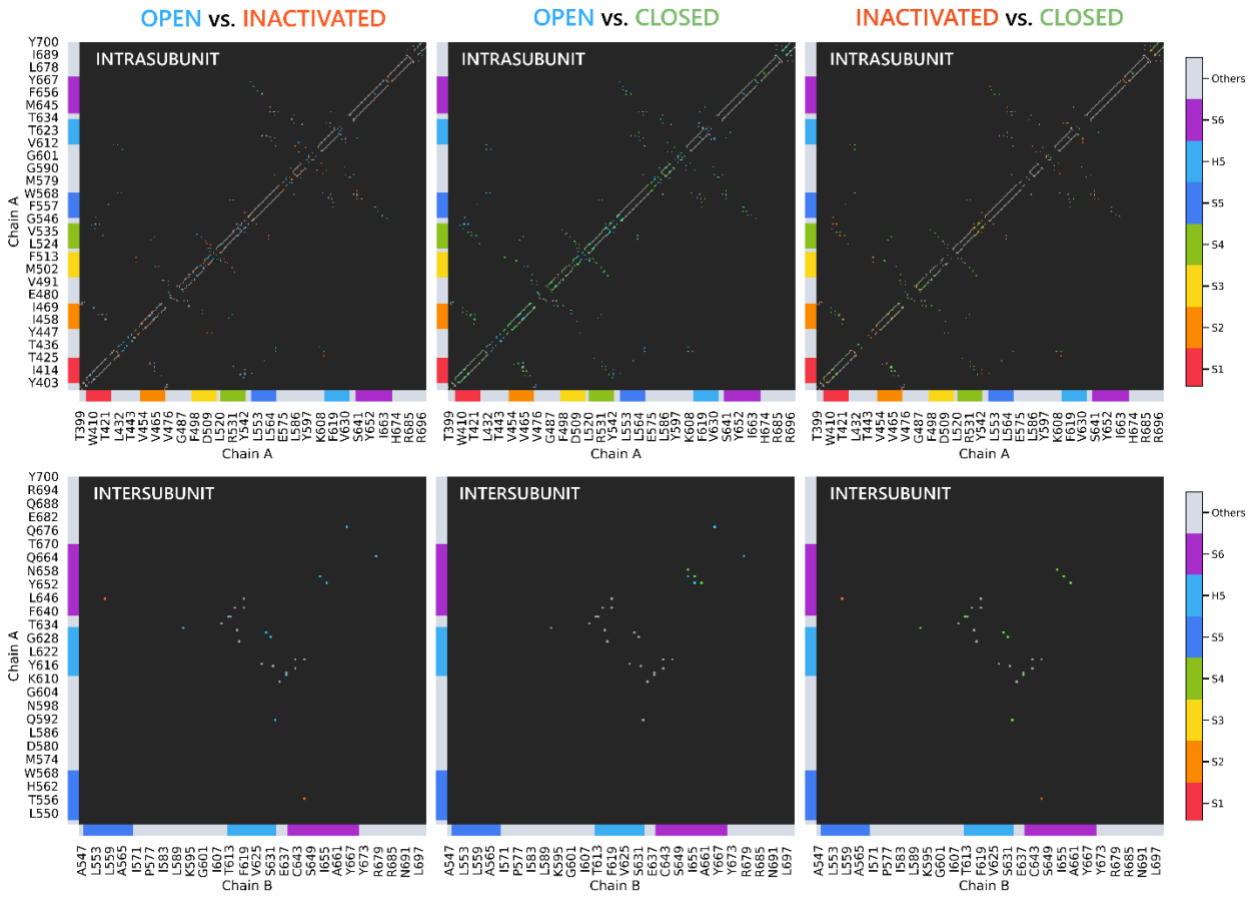



**Figure S11. Distance-based contact maps comparing intra- and intersubunit contacts between each model.** *Two residues whose $C_\alpha$ atoms are within 6 Å of each other are considered to be in contact, provided there are no $C_\alpha$ atoms belonging to a third residue in between. Black cells indicate no contacts. Gray cells indicate a contact is present in both states. Blue, orange, and green colored cells indicate the interaction is present only in the open, inactivated, or closed state, respectively, but not in the other state being compared in the map. Colored topology labels are included along the left and bottom edges of the maps showing the specific regions of hERG to which the residues correspond.*

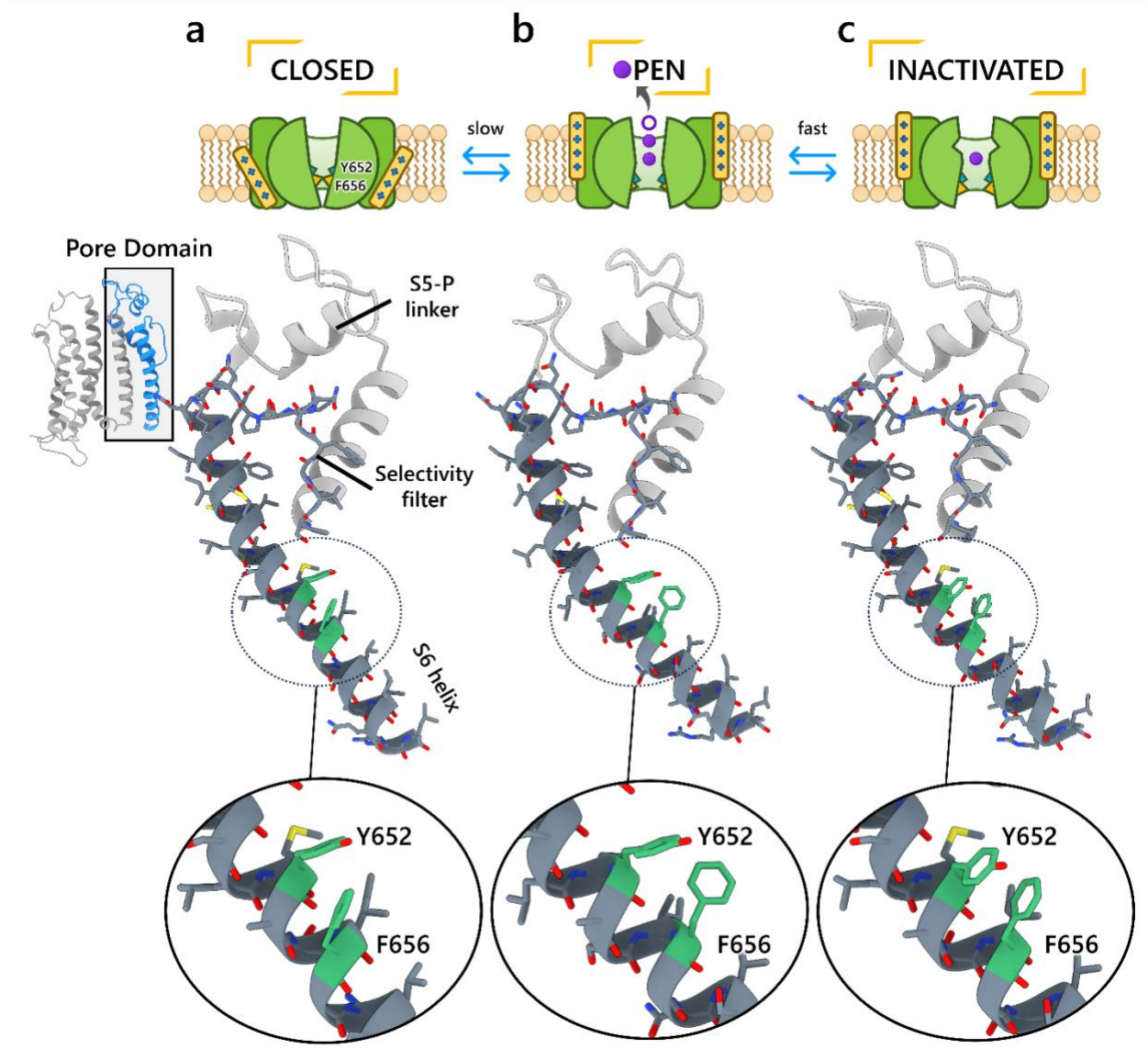



**Figure S12. Comparison of the S6 helix conformation for the hERG closed- (a), open- (b), and inactivated-state (c) models.** *Residues E575 – L666 from the pore domain are visualized. Selectivity filter (SF) residues and those on the S6 helix are shown with their side chains displayed as colored sticks. C atoms are gray, O are red, N are blue, S are yellow, H are not shown. The drug binding residues Y652 and F656 are highlighted in green.*

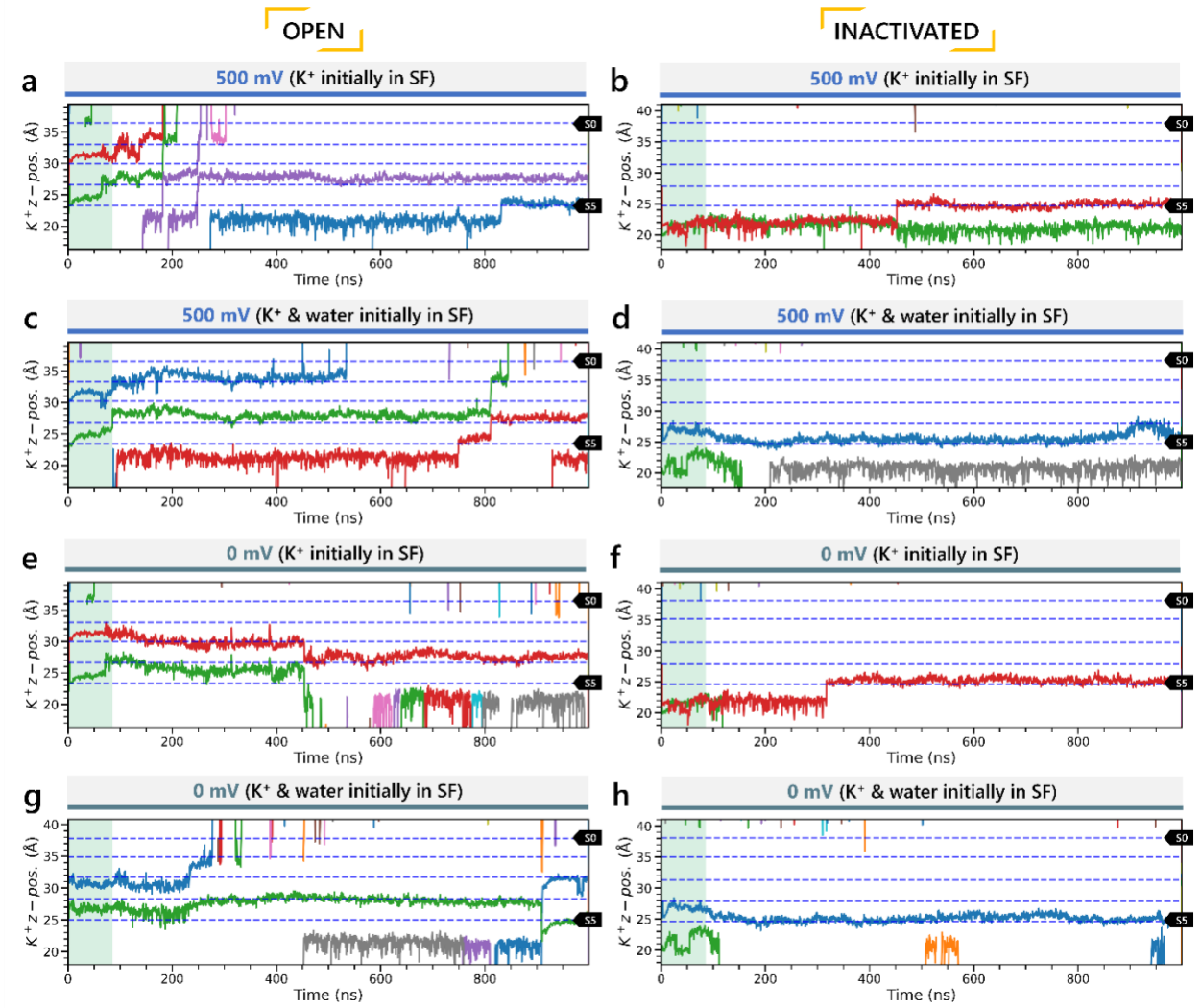



**Figure S13. Movement of K⁺ ions through hERG selectivity filter (SF).** *The z coordinates of K⁺ ions are tracked as they traverse from the pore of the channel (lower y-axis limit) to the extracellular space (upper y-axis limit). Putative K⁺ binding sites in the SF (S0 – S5) are marked using blue dashed lines in the plots.* ***a, c)*** *Molecular dynamics (MD) simulations with the applied 500 mV membrane voltage of the open-state model with the SF initially configured to have only K⁺ ions or alternating K⁺ / water molecules, respectively.* ***b, d)*** *MD simulations with the applied 500 mV membrane voltage of the inactivated-state model with the SF initially configured to have only K⁺ ions or alternating K+ / water molecules, respectively.* ***e, g)*** *MD simulations without applied membrane voltage of the open-state model with the SF initially configured to have only K⁺ ions or alternating K+ / water molecules, respectively.* ***f, h)*** *MD simulations*

*without applied membrane voltage of the inactivated-state model with the SF initially configured to have only K⁺ ions or alternating K⁺ / water molecules, respectively.*

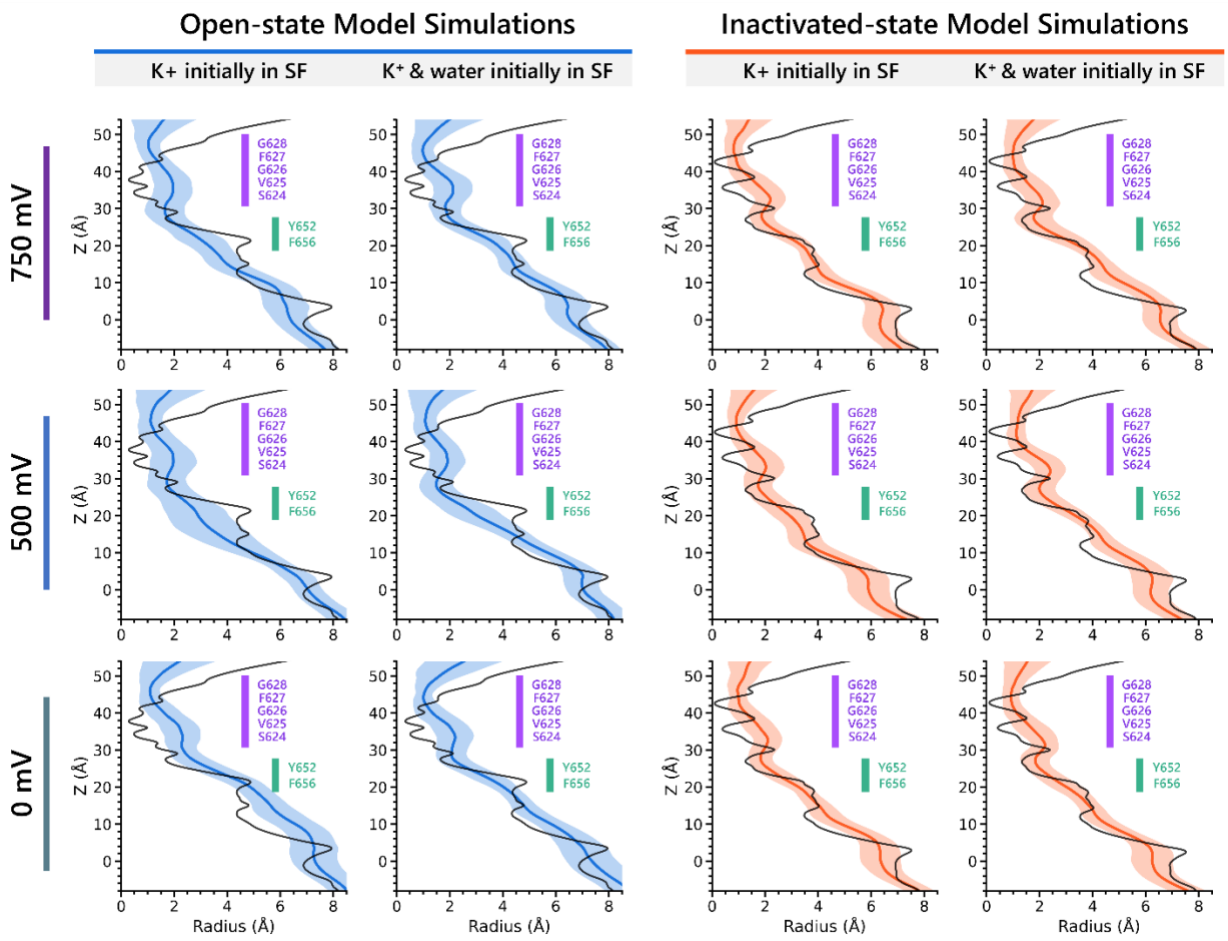



**Figure S14. Analysis of modulations of the selectivity filter (SF) and pore radii over the course of the 1 μs long molecular dynamics (MD) simulations.** *The blue/orange-colored lines represent the average pore radius and the shaded regions represent the standard deviation measured in MD simulations for a given Z value. The black lines represent the initial pore radii. The label on the left indicates the voltage of the simulations in each row.*

## 4.6) References


1. Vandenberg, J. I. *et al.* hERG K(+) channels: structure, function, and clinical significance. *Physiol. Rev.* **92**, 1393–1478 (2012).
2. Sanguinetti, M. C. & Tristani-Firouzi, M. hERG potassium channels and cardiac arrhythmia. *Nature* **440**, 463–469 (2006).
3. Li, M. & Ramos, L. G. Drug-Induced QT Prolongation And Torsades de Pointes. *Pharm. Ther.* **42**, 473–477 (2017).
4. Ferri, N. *et al.* Drug attrition during pre-clinical and clinical development: understanding and managing drug-induced cardiotoxicity. *Pharmacol. Ther.* **138**, 470–484 (2013).
5. Kocadal, K., Saygi, S., Alkas, F. B. & Sardas, S. Drug-associated cardiovascular risks: A retrospective evaluation of withdrawn drugs. *North. Clin. Istanb.* **6**, 196–202 (2018).
6. Waldo, A. L. *et al.* Effect of d-sotalol on mortality in patients with left ventricular dysfunction after recent and remote myocardial infarction. *The Lancet* **348**, 7–12 (1996).
7. Wang, W. & MacKinnon, R. Cryo-EM Structure of the Open Human Ether-à-go-go-Related K+ Channel hERG. *Cell* **169**, 422-430.e10 (2017).
8. Perrin, M. J., Kuchel, P. W., Campbell, T. J. & Vandenberg, J. I. Drug Binding to the Inactivated State Is Necessary but Not Sufficient for High-Affinity Binding to Human Ether-à-go-go-Related Gene Channels. *Mol. Pharmacol.* **74**, 1443–1452 (2008).
9. Priest, B., Bell, I. M. & Garcia, M. Role of hERG potassium channel assays in drug development. *Channels* **2**, 87–93 (2008).
10. Asai, T. *et al.* Cryo-EM Structure of K+-Bound hERG Channel Complexed with the Blocker Astemizole. *Struct. Lond. Engl. 1993* **29**, 203-212.e4 (2021).
11. Yang, P.-C. *et al.* A Computational Pipeline to Predict Cardiotoxicity: From the Atom to the Rhythm. *Circ. Res.* **126**, 947–964 (2020).

12. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

13. Park, H., Zhou, G., Baek, M., Baker, D. & DiMaio, F. Force Field Optimization Guided by Small Molecule Crystal Lattice Data Enables Consistent Sub-Angstrom Protein-Ligand Docking. *J. Chem. Theory Comput.* **17**, 2000–2010 (2021).

14. Miranda, W. E. *et al.* Selectivity filter modalities and rapid inactivation of the hERG1 channel. *Proc. Natl. Acad. Sci.* **117**, 2795–2804 (2020).

15. Li, J., Shen, R., Reddy, B., Perozo, E. & Roux, B. Mechanism of C-type inactivation in the hERG potassium channel. *Sci. Adv.* **7**, eabd6203 (2021).

16. Sala, D., Hildebrand, P. W. & Meiler, J. Biasing AlphaFold2 to predict GPCRs and kinases with user-defined functional or structural properties. *Front. Mol. Biosci.* **10**, 1121962 (2023).

17. del Alamo, D., Sala, D., Mchaourab, H. S. & Meiler, J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife* **11**, e75751 (2022).

18. Brown, B. P., Stein, R. A., Meiler, J. & Mchaourab, H. S. Approximating Projections of Conformational Boltzmann Distributions with AlphaFold2 Predictions: Opportunities and Limitations. *J. Chem. Theory Comput.* (2024) doi:10.1021/acs.jctc.3c01081.

19. Maly, J. *et al.* Structural modeling of the hERG potassium channel and associated drug interactions. *Front. Pharmacol.* **13**, 966463 (2022).

20. Emigh Cortez, A. M. *et al.* Structural modeling of hERG channel-drug interactions using Rosetta. *Front. Pharmacol.* **14**, 1244166 (2023).

21. Mandala, V. S. & MacKinnon, R. Voltage-sensor movements in the Eag Kv channel under an applied electric field. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2214151119 (2022).

22. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

23. Mirdita, M., Steinegger, M. & Söding, J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinforma. Oxf. Engl.* **35**, 2856–2858 (2019).

24. Phartiyal, P., Jones, E. M. C. & Robertson, G. A. Heteromeric Assembly of Human Ether-à-go-go-related Gene (hERG) 1a/1b Channels Occurs Cotranslationally via N-terminal Interactions*. *J. Biol. Chem.* **282**, 9874–9882 (2007).

25. Fleishman, S. J. *et al.* RosettaScripts: A Scripting Language Interface to the Rosetta Macromolecular Modeling Suite. *PLOS ONE* **6**, e20161 (2011).

26. Carus H.Y. Lau *et al.* Potassium dependent structural changes in the selectivity filter of HERG potassium channels. *bioRxiv* 2023.12.14.571769 (2023) doi:10.1101/2023.12.14.571769.

27. Vandenberg, J. *et al.* Structural Basis for Rapid Voltage Dependent Inactivation of HERG Potassium Channels. *ResearchSquare* (2021) doi:10.21203/rs.3.rs-1105661/v1.

28. Leman, J. K. *et al.* Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods* **17**, 665–680 (2020).

29. Doyle, D. A. *et al.* The structure of the potassium channel: molecular basis of K+ conduction and selectivity. *Science* **280**, 69–77 (1998).

30. Long, S. B., Campbell, E. B. & Mackinnon, R. Crystal structure of a mammalian voltage-dependent Shaker family K+ channel. *Science* **309**, 897–903 (2005).

31. Cuello, L. G., Jogini, V., Cortes, D. M. & Perozo, E. Structural mechanism of C-type inactivation in K+ channels. *Nature* **466**, 203–208 (2010).

32. Tan, X.-F. *et al.* Structure of the Shaker Kv channel and mechanism of slow C-type inactivation. *Sci. Adv.* **8**, eabm7814.

33. Tyagi, A. *et al.* Rearrangement of a unique Kv1.3 selectivity filter conformation upon binding of a drug. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2113536119 (2022).

34. Chandy, K. G., Sanches, K. & Norton, R. S. Structure of the voltage-gated potassium channel KV1.3: Insights into the inactivated conformation and binding to therapeutic leads. *Channels* **17**, 2253104.

35. Selvakumar, P. *et al.* Structures of the T cell potassium channel Kv1.3 with immunoglobulin modulators. *Nat. Commun.* **13**, 3854 (2022).

36. Noel, J. K., Whitford, P. C. & Onuchic, J. N. The Shadow Map: A General Contact Definition for Capturing the Dynamics of Biomolecular Folding and Function. *J. Phys. Chem. B* **116**, 8692–8702 (2012).

37. Dun, W., Jiang, M. & Tseng, G.-N. Allosteric effects of mutations in the extracellular S5-P loop on the gating and ion permeation properties of the hERG potassium channel. *Pflüg. Arch.* **439**, 141–149 (1999).

38. Butler, A., Zhang, Y., Stuart, A. G., Dempsey, C. E. & Hancox, J. C. Action potential clamp characterization of the S631A hERG mutation associated with short QT syndrome. *Physiol. Rep.* **6**, e13845 (2018).

39. Fan, J. S., Jiang, M., Dun, W., McDonald, T. V. & Tseng, G. N. Effects of outer mouth mutations on hERG channel function: a comparison with similar mutations in the Shaker channel. *Biophys. J.* **76**, 3128–3140 (1999).

40. Satler, C. A., Vesely, M. R., Duggal, P., Ginsburg, G. S. & Beggs, A. H. Multiple different missense mutations in the pore region of HERG in patients with long QT syndrome. *Hum. Genet.* **102**, 265–272 (1998).

41. Cordeiro, J. M., Brugada, R., Wu, Y. S., Hong, K. & Dumaine, R. Modulation of IKr inactivation by mutation N588K in KCNH2: A link to arrhythmogenesis in short QT syndrome. *Cardiovasc. Res.* **67**, 498–509 (2005).

42. Clarke, C. E. *et al.* Effect of S5P α-helix charge mutants on inactivation of hERG K+ channels. *J. Physiol.* **573**, 291–304 (2006).

43. Ficker, E., Jarolimek, W., Kiehn, J., Baumann, A. & Brown, A. M. Molecular determinants of dofetilide block of HERG K+ channels. *Circ. Res.* **82**, 386–395 (1998).

44. Nakajima, T. *et al.* Novel Mechanism of HERG Current Suppression in LQT2. *Circ. Res.* **83**, 415–422 (1998).

45. Zhao, J. T. *et al.* Not all hERG pore domain mutations have a severe phenotype: G584S has an inactivation gating defect with mild phenotype compared to G572S, which has a dominant negative trafficking defect and a severe phenotype. *J. Cardiovasc. Electrophysiol.* **20**, 923–930 (2009).

46. Liu, J., Zhang, M., Jiang, M. & Tseng, G.-N. Structural and Functional Role of the Extracellular S5-P Linker in the HERG Potassium Channel. *J. Gen. Physiol.* **120**, 723–737 (2002).

47. Thouta, S. *et al.* Proline Scan of the hERG Channel S6 Helix Reveals the Location of the Intracellular Pore Gate. *Biophys. J.* **106**, 1057–1069 (2014).

48. Helliwell, M. V. *et al.* Structural implications of hERG K+ channel block by a high-affinity minimally structured blocker. *J. Biol. Chem.* **293**, 7040–7057 (2018).

49. Lam, C. K. & de Groot, B. L. Ion Conduction Mechanisms in Potassium Channels Revealed by Permeation Cycles. *J. Chem. Theory Comput.* **19**, 2574–2589 (2023).

50. Roux, B. Ion channels and ion selectivity. *Essays Biochem.* **61**, 201–209 (2017).

51. Numaguchi, H. *et al.* Probing the Interaction Between Inactivation Gating and dd-Sotalol Block of HERG. *Circ. Res.* **87**, 1012–1018 (2000).

52. Wang, S., Morales, M. J., Liu, S., Strauss, H. C. & Rasmusson, R. L. Modulation of HERG affinity for E-4031 by [K+]o and C-type inactivation. *FEBS Lett.* **417**, 43–47 (1997).

53. Alexandrou, A. J. *et al.* Mechanism of hERG K+ channel blockade by the fluoroquinolone antibiotic moxifloxacin. *Br. J. Pharmacol.* **147**, 905–916 (2006).

54. Windley, M. J. *et al.* Measuring kinetics and potency of hERG block for CiPA. *J. Pharmacol. Toxicol. Methods* **87**, 99–107 (2017).

55. Windisch, A. *et al.* Trapping and dissociation of propafenone derivatives in HERG channels. *Br. J. Pharmacol.* **162**, 1542 (2011).

56. Kamiya, K., Niwa, R., Mitcheson, J. S. & Sanguinetti, M. C. Molecular Determinants of hERG Channel Block. *Mol. Pharmacol.* **69**, 1709–1716 (2006).

57. Stork, D. *et al.* State dependent dissociation of HERG channel inhibitors. *Br. J. Pharmacol.* **151**, 1368–1376 (2007).

58. Miao, Y., Bhattarai, A. & Wang, J. Ligand Gaussian accelerated molecular dynamics (LiGaMD): Characterization of ligand binding thermodynamics and kinetics. *J. Chem. Theory Comput.* **16**, 5526–5547 (2020).

59. Branduardi, D. & Faraldo-Gómez, J. D. String method for calculation of minimum free-energy paths in Cartesian space in freely-tumbling systems. *J. Chem. Theory Comput.* **9**, 4140–4154 (2013).

60. OEPerceiveInteractionOptions — Toolkits -- Python. https://docs.eyesopen.com/toolkits/python/oechemtk/OEBioClasses/OEPerceiveInteractionOptions.html.

61. Saxena, P. *et al.* New potential binding determinant for hERG channel inhibitors. *Sci. Rep.* **6**, 24182 (2016).

62. Solving the hERG anti-target with cryo-EM | Nano Imaging. https://www.nanoimagingservices.com/blog/characterization-of-the-herg-anti-target.

63. Stepanovic, S. Z. *et al.* The evolutionarily conserved residue A653 plays a key role in HERG channel closing. *J. Physiol.* **587**, 2555–2566 (2009).

64. Wang, Y. *et al.* Role of the pH in state-dependent blockade of hERG currents. *Sci. Rep.* **6**, 32536 (2016).

65. Sănchez-Chapula, J. A., Ferrer, T., Navarro-Polanco, R. A. & Sanguinetti, M. C. Voltage-dependent profile of human ether-a-go-go-related gene channel block is influenced by a single residue in the S6 transmembrane domain. *Mol. Pharmacol.* **63**, 1051–1058 (2003).

66. AlQuraishi, M. AlphaFold at CASP13. *Bioinforma. Oxf. Engl.* **35**, 4862–4865 (2019).

67. Jumper, J. *et al.* Applying and improving AlphaFold at CASP14. *Proteins* **89**, 1711–1721 (2021).

68. Varadi, M. & Velankar, S. The impact of AlphaFold Protein Structure Database on the fields of life sciences. *Proteomics* **23**, e2200128 (2023).

69. Karelina, M., Noh, J. J. & Dror, R. O. How accurately can one predict drug binding modes using AlphaFold models? *eLife* **12**, RP89386 (2023).

70. DeMarco, K. R. *et al.* Molecular determinants of pro-arrhythmia proclivity of d- and l-sotalol via a multi-scale modeling pipeline. *J. Mol. Cell. Cardiol.* **158**, 163–177 (2021).

71. Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

72. Berman, H. M. *et al.* The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

73. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).

74. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

75. Huang, J. & MacKerell Jr, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J. Comput. Chem.* **34**, 2135–2145 (2013).

76. Klauda, J. B. *et al.* Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B* **114**, 7830–7843 (2010).

77. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

78. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).

79. Kim, S. *et al.* PubChem Substance and Compound databases. *Nucleic Acids Res.* **44**, D1202-1213 (2016).

80. Tingle, B. I. *et al.* ZINC-22–A Free Multi-Billion-Scale Database of Tangible Compounds for Ligand Discovery. *J. Chem. Inf. Model.* **63**, 1166–1176 (2023).

81. Hanwell, M. D. *et al.* Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics* **4**, 17 (2012).

82. Case, D. A. *et al.* AmberTools. *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).

83. Molecular Modeling Software | OpenEye Scientific. https://www.eyesopen.com.

84. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci. Publ. Protein Soc.* **30**, 70–82 (2021).

85. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38, 27–28 (1996).

86. Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F. & Schroeder, M. PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res.* **43**, W443–W447 (2015).

# CHAPTER 5:

# Lipid Regulation of SK2 Channels – A Therapeutic Target for Cardiac Arrhythmia

Ryan L. Woltz[1], Woori Choi[1], Khoa Ngo[2], Brandon J. Harris[2], Yanxiao Han[2], Kyle C. Rouen[2], Pauline Trinh[1], Yang Zhang[1], Zhong Jian[3], Ye Chen-Izu[3], Eamonn Dickson[2], Ebenezer N. Yamoah[4], Vladimir Yarov-Yarovoy[2,5], Igor Vorobyov*[2], Xiao-Dong Zhang*[1], Nipavan Chiamvimonvat*[1,3,6]

[1]*Department of Internal Medicine, Division of Cardiovascular Medicine, UC Davis, CA*

[2]*Department of Physiology and Membrane Biology, UC Davis, CA*

[3]*Department of Pharmacology, UC Davis, CA*

[4]*Department of Physiology and Cell Biology, University of Nevada, Reno, NV*

[5]*Department of Anesthesiology and Pain Medicine, UC Davis, CA*

[6]*Department of Veterans Affairs, Northern California Health Care System, Mather, CA*

** These authors are co-corresponding authors*

**Author Contributions:**

I developed the computational programs that were used to set up and analyze the molecular dynamics simulations investigating the interactions between PIP2 and SK2 channels. Specifically, my scripts were instrumental in generating the plots shown in figures 3c, 4e/g, 5c-e, and 6g. The insights derived from these analyses played a crucial role in shaping the major findings of the study. Although I provided the scripts for running and analyzing the simulations, I did not directly execute these simulations or carry out the experiments described in the study.

*"Knowing is not enough; we must apply.*
*Willing is not enough; we must do."* – Johann Wolfgang von Goethe (1749 – 1832)

## 5.1) Introduction

Atrial fibrillation is the most common cardiac arrhythmia seen in clinical practice and is associated with increased morbidity and mortality[1,2]. SK2 channels are pivotal in linking changes in intracellular calcium to the heart's electrophysiological properties, making their regulation critical for developing effective treatments[3].

SK2 channels belong to the family of small-conductance $Ca^{2+}$-activated $K^+$ channels[3]. These channels are unique because they are gated exclusively by intracellular calcium levels, without direct voltage sensitivity[3]. This makes them crucial for connecting calcium signaling to the electrical responsiveness of cardiac cells, making them key targets for arrhythmia treatments. Their importance is further underscored by their higher expression in atrial myocytes compared to ventricular ones, highlighting their significant role in atrial electrophysiology[3].

Historically, the modulation of SK2 channels has been of great interest, particularly their potential as therapeutic targets for atrial fibrillation[4]. However, the specific regulatory mechanisms, especially how lipid signaling molecules like PIP2 influence their function, have remained less understood. Previous studies have shown that PIP2 interacts with these channels, affecting their activity in heterologous expression systems, but detailed understanding of this interaction in native cardiac cells has been lacking[5].

Our study aims to dissect the interactions between PIP2 and SK2 channels using a combination of experimental and computational techniques. By resolving the atomic-level dynamics and functional consequences of PIP2 binding to SK2 channels, we seek to provide new insights into the modulation of cardiac excitability and arrhythmogenesis. These insights are crucial not only for understanding cardiac ion channel physiology but also for developing novel therapeutic strategies for managing atrial fibrillation and other cardiac arrhythmias involving SK2 channels[3].

## 5.2) Methodology and Results

### a. Optogenetic Analysis of PIP2's Role in Regulating $I_{K,Ca}$ in Ventricular Myocytes

To investigate the regulation of apamin-sensitive current ($I_{K,Ca}$) in ventricular myocytes by PIP2, we employed optogenetic constructs to deplete PIP2 in cells expressing hSK2 channels, shown in **Figure 1A**. Initial tests were conducted using Chinese hamster ovary (CHO) cells, where $I_{K,Ca}$ was recorded using a voltage ramp protocol. Blue light exposure initiated PIP2 depletion by bringing 5-Ptase to the membrane, catalyzing the conversion of PIP2 to PIP. During blue light exposure, there was a significant reduction in $I_{K,Ca}$, which recovered when the light was turned off, demonstrating the sensitivity of the SK2 channels to PIP2 levels **(Figure 1B)**. Apamin was used at the end of the experiments to confirm the apamin sensitivity of the recorded currents.

To extend these findings to cardiomyocytes, we combined functional analyses with optogenetics in rabbit ventricular myocytes **(Figure 1C, D)**. These cells were transfected with optogenetic constructs containing either active protein phosphatase (5P) or an inactive variant (5P-Dead). Using a two-pulse protocol to elicit $I_{K,Ca}$, we recorded the currents before and after blue light exposure. The results showed a significant inhibition of $I_{K,Ca}$ after blue light exposure in cells with active 5P compared to controls, indicating that PIP2 depletion leads to a marked reduction in SK2 channel activity **(Figure 1E, F)**.

These findings confirm the crucial role of PIP2 in the activation of SK2 channels in cardiomyocytes. Despite this, the exact atomistic mechanisms underlying this regulation remain unclear. Further exploration using atomistic structural modeling and simulations is necessary to fully understand how PIP2 regulates SK2 channel activity at the molecular level.

### b. Development hSK2 Models Using Molecular Modeling and Molecular Dynamics (MD) Simulations

Next, we developed homology models of the hSK2 channel using the cryo-EM structures of the hSK4 channel complexed with CaM in closed, intermediate, and open states (PDB IDs: 6CNM, 6CNN, and 6CNO, respectively[6]).

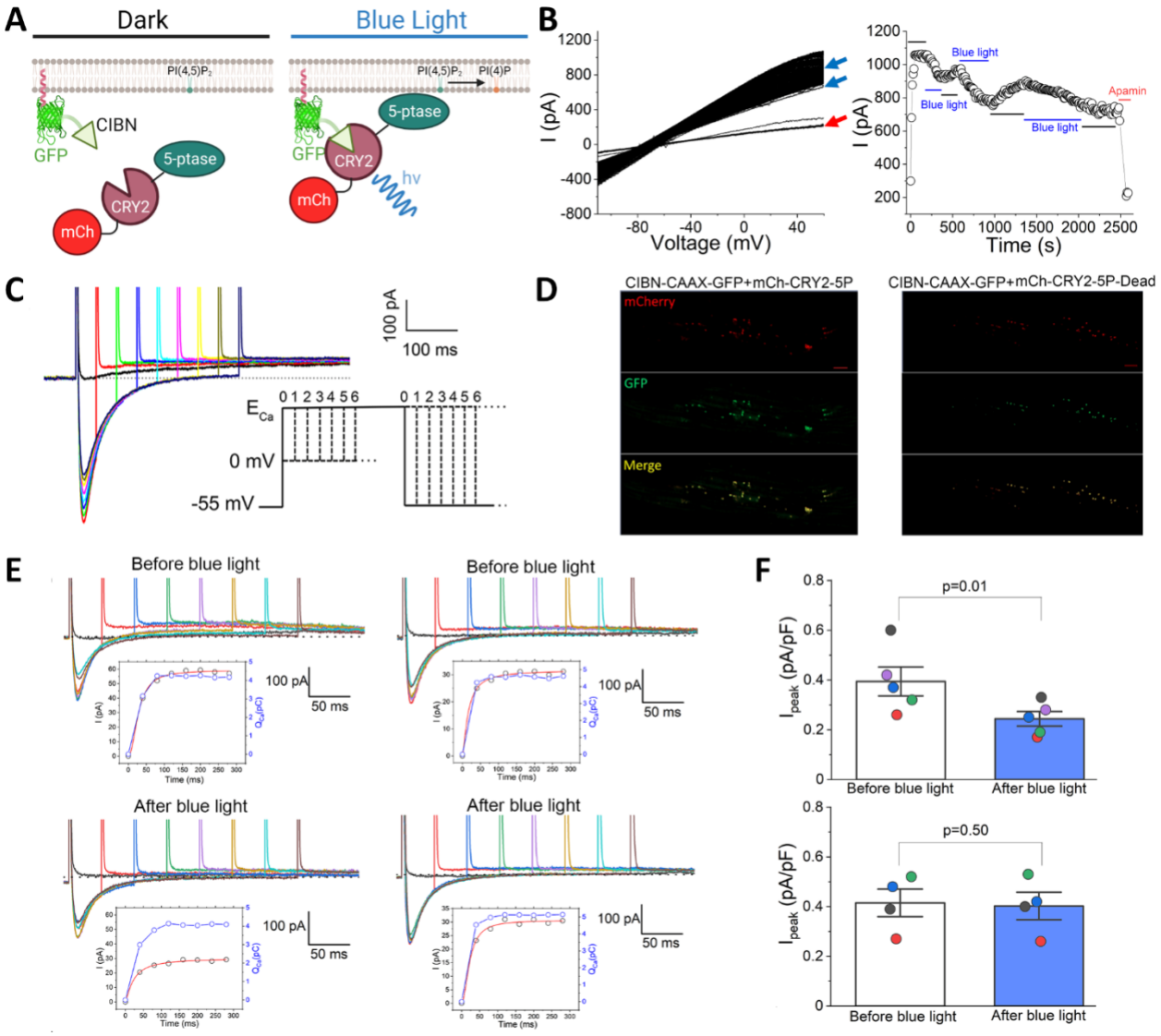



**Figure 1.** ***A.*** *Schematic illustrating the dimerization of the CRY2 and CIBN optogenetic constructs at the plasma membrane and dephosphorylation of PIP2 via inositol-5-phosphatase activity to produce phosphatidylinositol 4-phosphate (PIP).* ***B.*** *$I_{K,Ca}$ recorded in CHO cells using a voltage-ramp protocol prior to and during blue light exposure (blue arrow in the left panel and black lines in the right panel).* ***C.*** *Current traces and the two-pulse voltage-clamp protocol used to isolate apamin-sensitive current activated by $Ca^{2+}$ influx through L-type $Ca^{2+}$ channels (LTCCs). The voltage-clamp protocol uses a prepulse to progressively induce $Ca^{2+}$ influx via LTCCs from a holding potential of -55 mV, followed by a test pulse to monitor the apamin-sensitive $I_{K,Ca}$. We recorded $Ca^{2+}$ current ($I_{Ca}$) for each cell to measure the reversal potential ($E_{Ca}$). The test pulse was stepped to the observed $E_{Ca}$ to minimize inward $I_{Ca}$. Specific inhibitors*

*for transient outward $K^+$ current ($I_{to}$), rapidly activating ($I_{Kr}$) and slowly activating delayed rectifier $K^+$ currents ($I_{Ks}$), inward rectifier $K^+$ currents ($I_{K1}$), and $Cl^-$ currents were applied. The instantaneous outward $K^+$ currents progressively increased depending on the $Ca^{2+}$ influx.* ***D.*** *Rabbit ventricular myocytes were transfected with optogenetic constructs containing protein phosphatase 5 (5P, Left Panel) compared to control (inactive phosphatase, 5P-Dead, Right Panel) using magnetic nanoparticles.* ***E.*** *The activation kinetics of $I_{K,Ca}$ was quantified compared to the total charge entered through LTCC during the prepulse ($Q_{Ca}$ in pC). Current traces together with insets showing activation kinetics of $I_{K,Ca}$ (red line) compared to the total charge entered through LTCCs during the prepulse ($Q_{Ca}$ in pC, blue line), before and after the blue light for rabbit ventricular myocytes transfected with optogenetic constructs containing phosphatase (5P, Left Panel) compared to control (inactive phosphatase, 5P-Dead, Right Panel).* ***F.*** *$I_{K,Ca}$ was significantly blocked after blue light exposure compared to control cells. *P=0.01.*

We compared the sequence of hSK2 (UniProt ID: Q9H2S1) with the cryo-EM structure of hSK4 used as a template. The sequence identity between hSK4 and hSK2 is 46%, with a sequence similarity of 63%. For the CaM binding domain (CaMBD) in the C-terminal domain of hSK2 (residues 412-488), the sequence identity is 47%, with a sequence similarity of 68%. Given that sequence identities above 30-40% are considered sufficient for reliable modeling, hSK4 structures were deemed appropriate templates for generating hSK2 models.

Using the Rosetta Comparative Modeling (RosettaCM) protocol[7], we generated hSK2 homology models through two main steps: cryo-EM map refinement and homology modeling. During this process, additional restraints were imposed to maintain the integrity of the selectivity filter (SF) and CaMBD. Models without CaM exhibited large movements in the CaMBD due to the flexible hinge loop connecting the S6 transmembrane domain to the CaMBD **(Figure 2A)**. Conversely, including CaM resulted in convergence and improved agreement among the top hSK2-CaM models for the CaMBD **(Figure 2B-D)**.

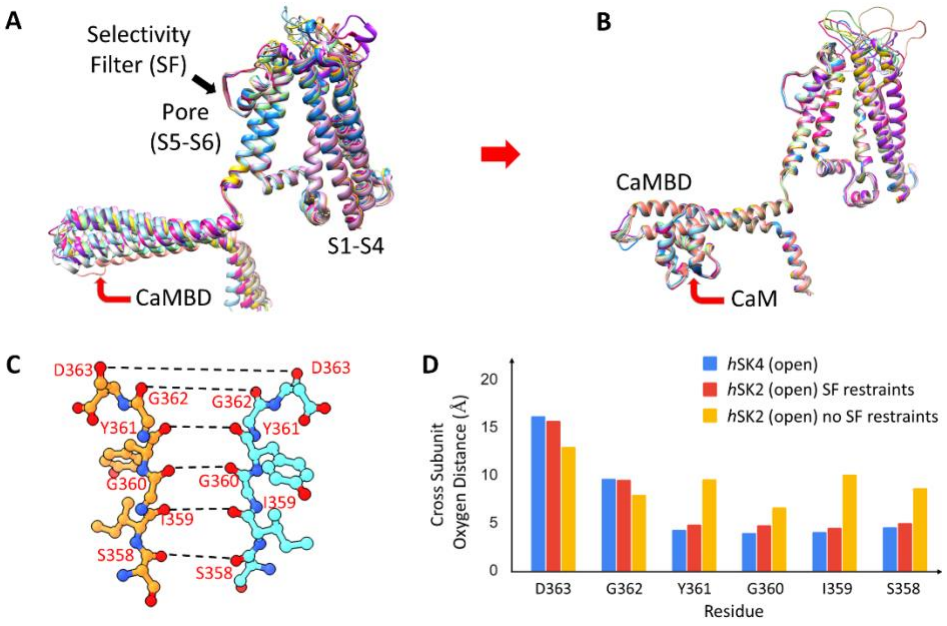



**Figure 2. Homology models of hSK2-CaM complexes based on the open state of hSK4-CaM cryo-EM structure (PDB: 6CNO) using Rosetta structural modeling.** ***A****) Top 10 homology models of hSK2 channel without CaM show large movements in the CaMBD in the C terminus of hSK2 channels, due to the flexible loop connecting pore-lining transmembrane helix S6 to the CaMBD.* ***B****) Inclusion of CaM in homology modeling shows convergence among the top 10 models for the CaMBD.* ***C****) Selectivity filter (SF) of hSK2 channel shown using ball-and-stick representation with the distances of backbone carbonyl oxygen atoms (red balls) shown as dashed black lines (amino acid residue numbering based on hSK2 channel). Nitrogen atoms are shown as blue balls.* ***D****) Cross-subunit backbone carbonyl oxygen distances for amino acid residues for the open state of hSK4 cryo-EM structure (blue bars), hSK2-CaM homology models with 100 kcal/mol/Å² restraints being applied to the backbone $C_\alpha$ atoms (red bars), and hSK2-CaM homology models without restraints (yellow bars).*

The hSK2-CaM complex models were embedded into a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) lipid membrane, solvated by 0.15 M KCl aqueous solution using CHARMM-GUI[8] **(Figure 3A)**, and subjected to MD simulations. We conducted 18 MD simulations on Anton 2 with varying PIP2 concentrations (2.5%, 5%, and 10%). An extended equilibration protocol was performed on the high-performance computer Expanse using Amber18[9] **(Figure 3B)**. After 100 ns of equilibration, the simulations were transferred to

Anton 2 for production runs extending up to 5 μs **(Figure 3C)**. To ensure system stability, position restraints were applied to the hSK2-CaM complex or its components at each step of the equilibration MD runs, with weak restraints on the SF backbone atoms maintained throughout the production runs. The MD simulation trajectories, analyzed via root-mean-square deviation (RMSD) calculations, demonstrated stable hSK2-CaM complexes and selectivity filters throughout the 5-μs simulations.

### c. Pore Mapping and Conductivity Validation through HOLE Analyses

To further validate the homology models for the hSK2-CaM complex, central ion channel pore mapping was performed using the HOLE program[10] to confirm the closed, intermediate, and open states of the models. **Figure 4A-C** highlight the radius profiles across the entire channel pore and the intracellular hydrophobic gate region at the V390 residues in the pore-lining S6 helices of the hSK2-CaM tetramer. These profiles are shown for the closed, intermediate, and open states, demonstrating clear distinctions in pore dimensions at each state. Time series profiles of the cross-subunit distances between V390 residues, depicted in **Figure 4D and 4E**, show consistent distances of approximately 6 Å in the closed and intermediate state models and about 14 Å in the open state models throughout the 5-μs-long MD simulations.

To validate the conductivity of the SK2-CaM complex homology models in the open state, we tracked the trajectory of $K^+$ ions over a 5-μs MD simulation, as shown in **Figure 4F and 4G**. Using an applied electric field approach with a 750 mV transmembrane voltage, 150 mM of $K^+$ ions in the aqueous solution, and 1 M $K^+$ ions seeded in the pore, we observed multiple $K^+$ pore permeation events. These events corresponded to a single-channel current of approximately 2.47 pA, closely matching the experimental values for SK2 channels.

These results confirm that the homology models accurately represent the different conformational states of the hSK2-CaM complex. The observed $K^+$ ion conduction in the open state model validates its functional integrity, reinforcing the reliability of these homology models in replicating the known behaviors of SK2 channels.

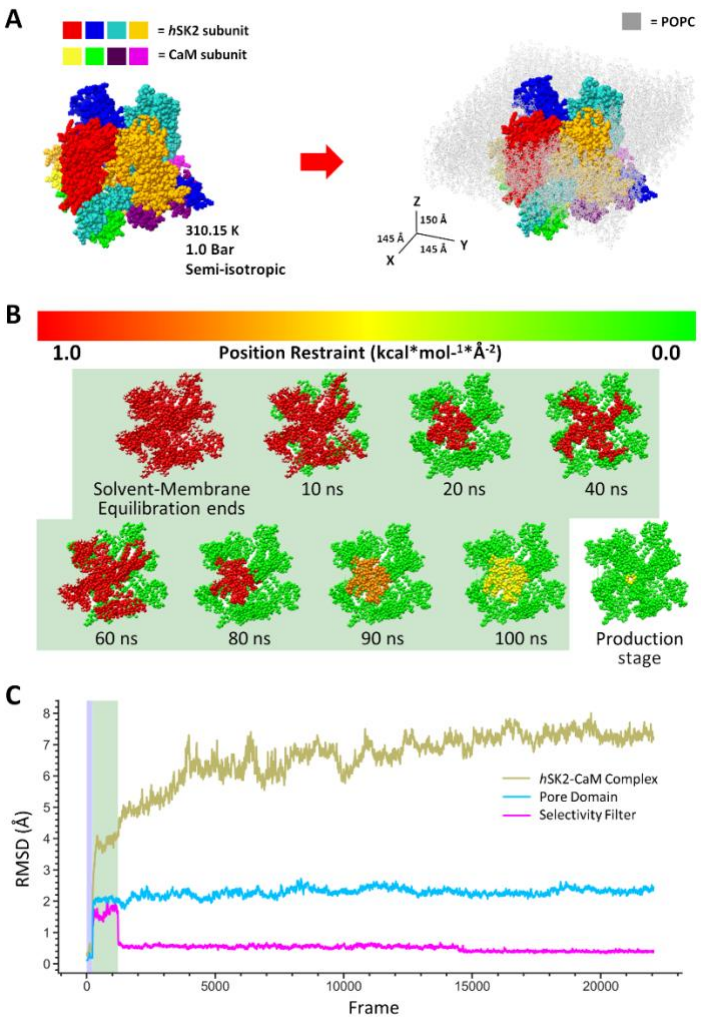



**Figure 3. MD simulation parameters for a *h*SK2-CaM complex.** ***A)*** *Visualization of the four subunits of hSK2-CaM complex before (left) and after (right) embedding into a POPC bilayer (light gray). PIP2 molecules are included in the lower bilayer leaflet only but are not visualized. MD simulation conditions and box size dimensions in X, Y and Z are shown.* ***B)*** *An extended equilibration protocol for 100 ns using Amber18 on the high-performance computer cluster (HPC) Expanse provides stable MD simulations by varying the location and strength of the restraints applied to the hSK2-CaM tetramer up to each time point indicated in ns and shaded in light green. It was followed by mostly unrestrained production run on Anton 2 supercomputer.* ***C)*** *Root-mean-square deviation (RMSD) profiles for the backbone $C_\alpha$ hSK2-CaM tetramer and the selectivity filter (SF) during MD simulations equilibrated for 100 ns using Amber18 (light green shaded area corresponding to the shaded area in panel* ***B****), followed by production runs of 5,000 ns on Anton 2.*

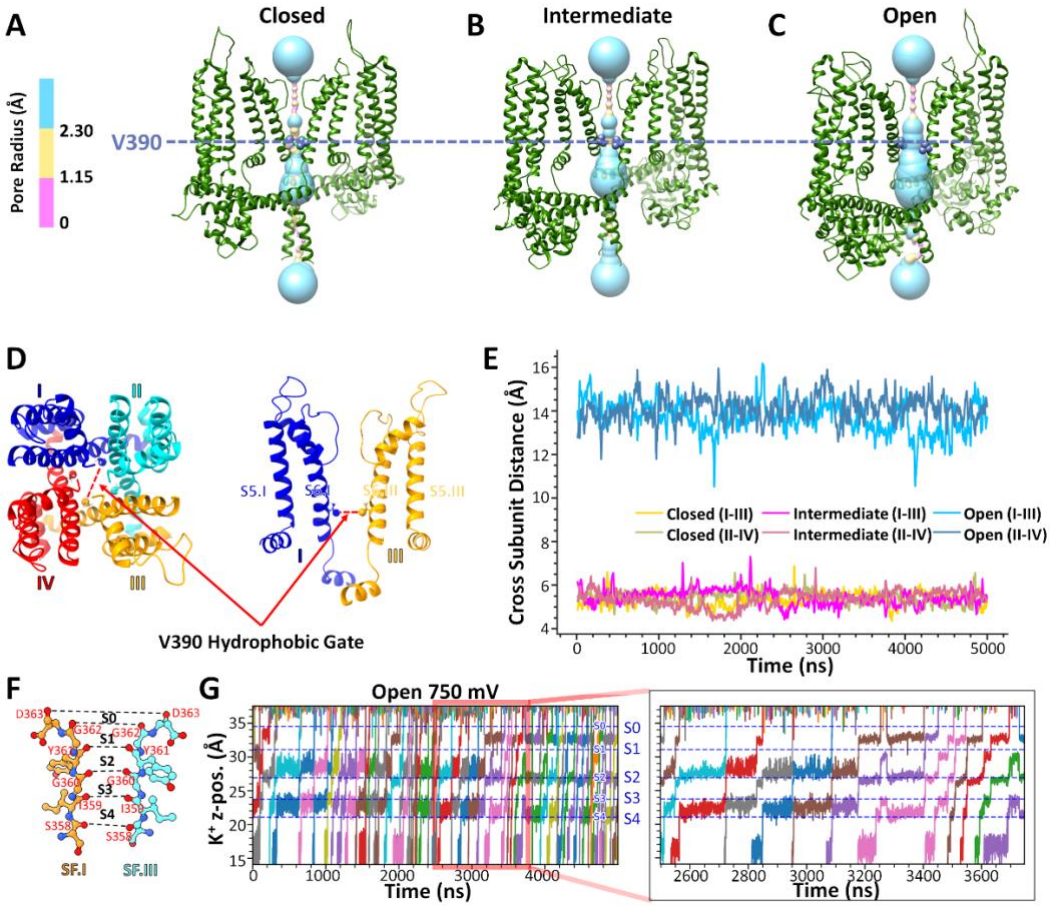



**Figure 4. Intracellular hydrophobic gate V390 residue cross-subunit distances and K⁺ ion conduction for homology models of the hSK2-CaM tetramer.** ***A-C)*** *location of the V390 residue in the S6 transmembrane helix of the hSK2-CaM tetramer in closed, intermediate, and open conformational states. Pore dimensions were computed using HOLE program and are shown using surface representation with coloring corresponding to the pore radii.* ***D-E)*** *Quantification of the cross-subunit distances between backbone $C_\alpha$ atoms of V390 residues (open = blue, intermediate = magenta, and closed = yellow).* ***F-G.*** *Illustration of ion conduction in the open conductive homology model of hSK2-CaM tetramer. The time series of K⁺ ion z-positions (K⁺ z-pos.) during 5-ms MD simulations with an applied voltage of 750 mV is depicted as they pass through the selectivity filter (SF). Five distinct positions of K⁺ in the SF, sites S0-S4, defined by the backbone carbonyl oxygen atoms (colored in red) of SF amino acid residues (S358-D363) have been labeled and separated by horizontal dashed lines. The inset on the right panel shows K⁺ ion conduction at an expanded time scale.*

### d. Dynamics of PIP2 Binding to the hSK2-CaM Complex in MD Simulations

To understand the dynamics and binding sites of PIP2 on SK2 channels, PIP2 molecules were randomly placed in the lower leaflet of lipid bilayers at concentrations ranging from 2.5-10% (physiologically 1-3%) during MD simulations. Our setup aimed to capture unbiased PIP2 binding sites and to provide insights into how PIP2 interacts with the basic residues of SK2 channels. A new analysis protocol was designed, focusing on salt bridges formed between the hSK2 channel and PIP2, using a 5.5 Å cutoff from the center of the charges.

The analysis enabled tracking the lateral movement of PIP2 molecules within the lipid membrane as they interacted with different amino acid residues. **Figure 5A-B** show the movement of PIP2 into the binding pockets of the SK2-CaM complex over 5-μs MD simulations from the membrane periphery. Initially, PIP2 binds to residues depicted in cyan in **Figure 5A-B**, and then sequentially transitions to other binding sites depicted in green, yellow, orange, red, and finally dark red over the course of the simulations. Two main PIP2 binding pockets were identified: the transient site and the activation site. The transition from the transient to activation sites occurs via the transfer residue R418, as shown in **Figure 5C-E**.

In conclusion, the data reveal that PIP2 binds to specific transient and activation sites on the hSK2-CaM complex, moving through a defined pathway that involves key residues, with the residue R418 playing a crucial role in transferring PIP2 from the transient to the activation site. This detailed dynamic interaction underscores the importance of PIP2 in SK2 channel regulation.

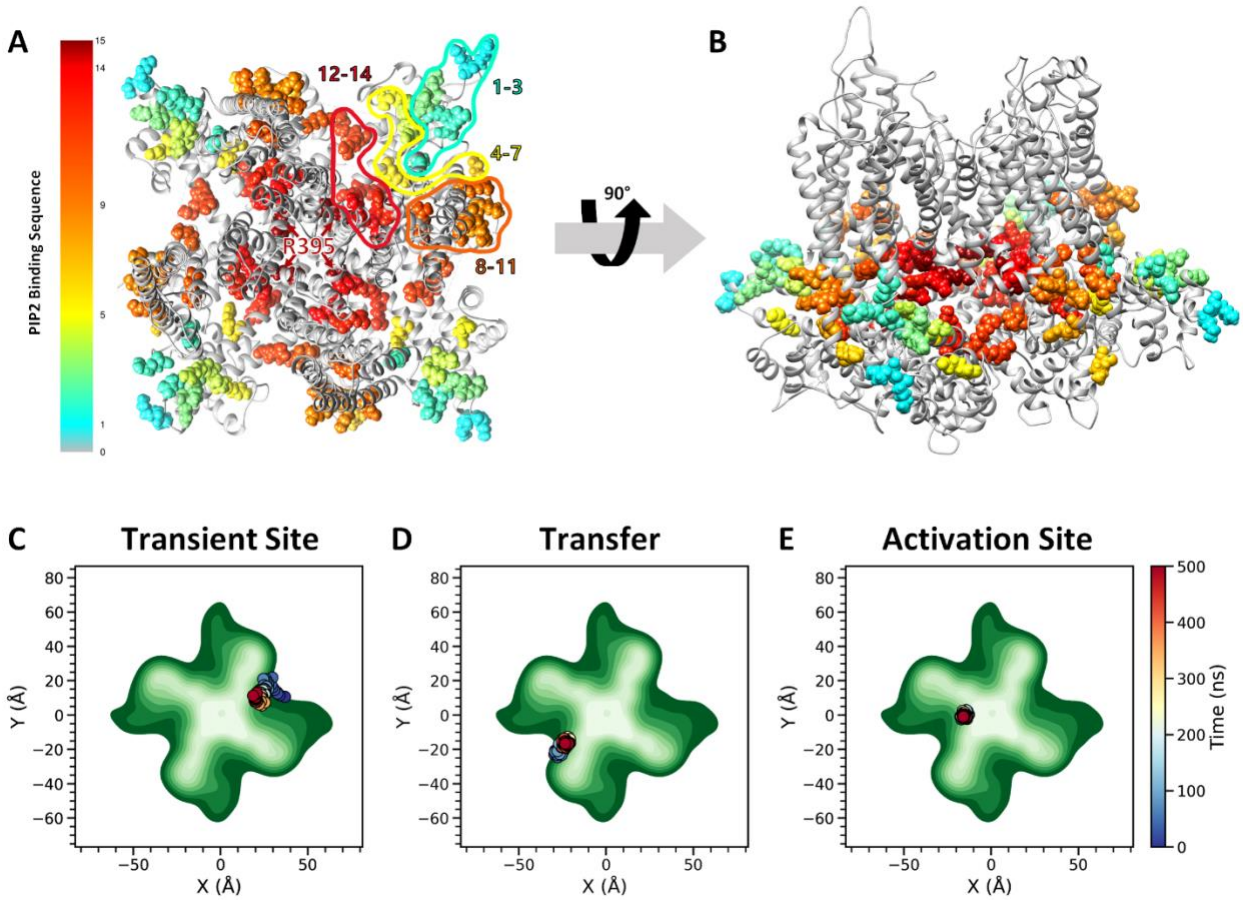



**Figure 5. Tracking of PIP2 in the membrane *XY* plane over time identifies two PIP2 binding pockets, denoted transient and activation binding sites.** ***A, B)*** *Movement of PIP2 molecules into the binding pockets of hSK2-CaM complex during MD simulations. The color scale provides time progression over the 5-μs MD simulations from cyan to green, yellow, orange, red, and dark brown. The hSK2-CaM complex is shown with side and top views in panels* ***A*** *and* ***B****, respectively.* ***C-E)*** *PIP2 molecule 1, 2 and 3 locations at the transient and activation binding sites during 5 ms of the MD simulations. PIP2 location is shown by a colored dot corresponding to the simulation time, whereas hSK2-CaM complex atomic density is shown in different shades of green with darker corresponding to higher density.* ***C)*** *PIP2 molecule 1 at the transient binding site, which includes amino acid residues R271, K278, and R286.* ***D)*** *Transfer of PIP2 molecule 2 from the transient to the activation binding site.* ***E)*** *PIP2 molecule 3 at the activation binding site, which includes residues R299, R395, K396, and K471.*

### e. Salt Bridge Formation between PIP2 and hSK2-CaM Amino Acid Residues

MD simulations provide significant insights into how the binding of PIP2 within the activation site results in hSK2 channel activation. To determine if PIP2 binding disrupts any stabilizing salt bridges in the channel, we analyzed intra- and inter-subunit interactions within the hSK2 structure. Remarkably, our analyses revealed that only one salt bridge is routinely disrupted by PIP2 binding: the cross-subunit R395:E398 salt bridge, located close to the hydrophobic gate at the S6 residue V390 of two adjacent subunits.

**Figure 6A-B** highlight the formation of the R395:E398 salt bridge between adjacent hSK2 subunits. **Figure 6C-F** focus on clustering MD simulation poses with and without the PIP2:R395 salt bridge, showing PIP2 interactions with other basic residues of hSK2. **Figure 6G** demonstrates the dynamics of salt bridge formation and disruption between PIP2 and these hSK2 residues. In **Figure 6C-F** and the top panels of **Figure 6G**, representative poses from MD simulations that form the PIP2:R395 salt bridge are shown in darker shades of pink and green, while poses that do not form this salt bridge are shown in lighter shades.

As shown in **Figure 6C, G**, the amino acid residues in the activation site that form salt bridges with PIP2 include R299, R395, K396, and K471. Interestingly, the formation of the PIP2:R395 salt bridge appears to correlate with the breaking of the PIP2:R299 salt bridge, while interactions with K396 and K471 are not affected. The R395 residue shows a significant change in side chain orientation when it forms a salt bridge with either E398 or PIP2, as exemplified in **Figure 6E-F**.

To summarize, the MD simulations reveal that PIP2 binding within the activation site primarily disrupts the R395:E398 salt bridge, which is crucial for hSK2 channel activation. The formation and disruption of specific salt bridges, especially involving R395, play a key role in the dynamic regulation of the channel, with distinct changes in side chain orientations reflecting these interactions.

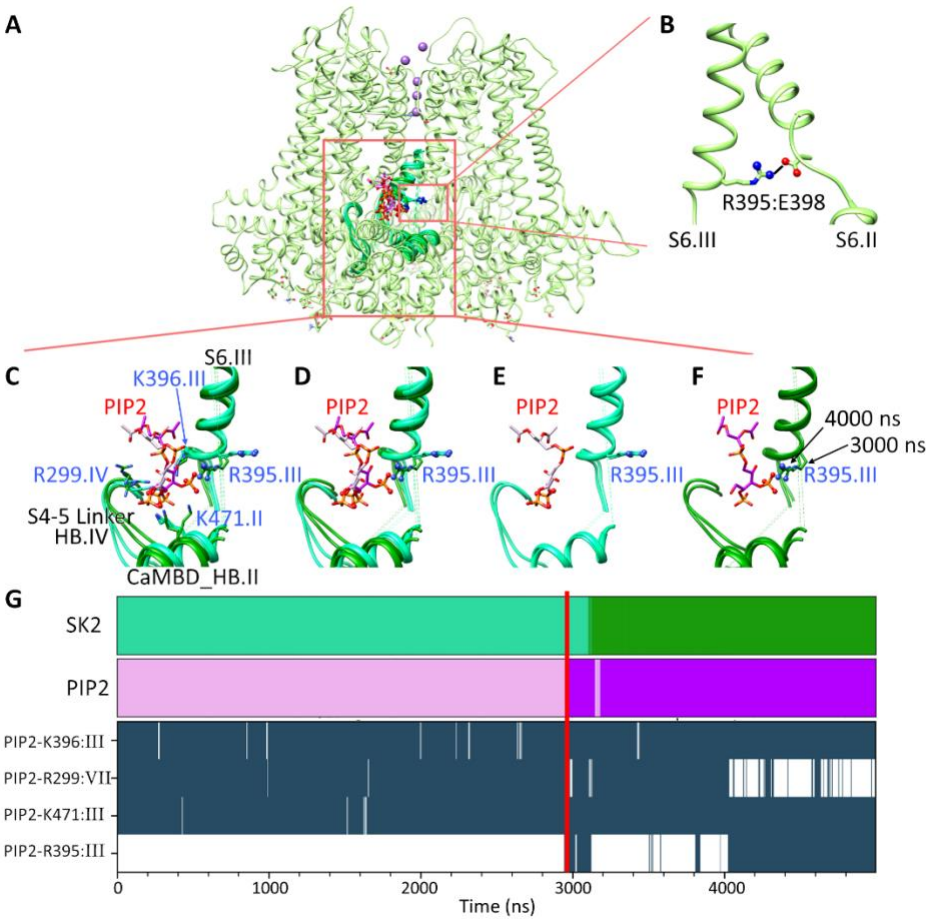



**Figure 6. Formation of salt bridges between PIP2 and amino acid residues from the hSK2-CaM tetramer.** ***A****) hSK2-CaM complex (shown as green ribbons with 70% transparency) to highlight the location of PIP2 binding shown as a wireframe. Only two subunits are shown for clarity. The red box highlights the region of zoom-in panels in* ***B-F****.* ***B****) Zoom-in view of the inter-subunit salt bridge between R395 of subunit III and E398 of subunit II located on S6 helices shown as ribbons.* ***C-F****) PIP2 carbon atoms are shown in pink and hSK2 backbone is shown in green. Phosphorus (Orange), oxygen (Red) and nitrogen (Blue) atoms involved in PIP2:R395 salt bridges are shown in ball-and-stick representation. Representative poses of hSK2-CaM from clustering of MD simulation data that are able to form the PIP2:R395 salt bridge are shown in darker shades and poses unable to form this salt bridge are shown in lighter shades.* ***C****) Poses with all major amino acid residues forming salt bridges with PIP2 in the activation site, R299, R395, K396, and K471.* ***D****) R395 side chain orientations shown to exemplify differences when it forms a salt bridge with E398 vs. PIP2.* ***E****) Only hSK2-CaM poses*

*that do not form the PIP2:R395 salt bridge.* ***F****) Only hSK2-CaM poses that do form the PIP2:R395 salt bridge.* ***G****) Two top graphs show clustering of MD simulation results based on time showing hSK2-CaM (SK2) and PIP2 poses, which are able to form the PIP2:R395 salt bridge in darker shades of green and pink, respectively. The four lower graphs show time series of salt bridge formation (dark-blue) and breaking (white) between several hSK2-CaM residues and PIP2 using a 3.6 Å cutoff. The red vertical line across all the graphs shows time point at which R395:PIP2 salt bridge was first detected.*

### f. Solvent Accessible Surface Area (SASA) Analysis Supporting PIP2:R395 Salt Bridge Formation

To understand how PIP2 binding within the activation site results in hSK2 channel activation, we performed solvent-accessible surface area (SASA) analyses. These analyses quantified the solvent exposure of each amino acid in the PIP2 binding sites across different channel states (closed, intermediate, and open). We compared the SASA for each amino acid residue in the four subunits, focusing on whether the residue was bound or unbound by PIP2, and tracked changes as residues formed new salt bridges with PIP2.

The SASA analyses revealed that PIP2 binding significantly affects the solvent exposure of specific residues. Notably, the R395 residue exhibited a significant increase in SASA upon PIP2 binding in the intermediate state, indicating its crucial role in the transition to the open state. Time series data showed that the R395 residue transitions from forming a salt bridge with E398 to binding PIP2, which correlates with changes in SASA and the stability of the salt bridge formation **(Figure 6 C-F)**.

Overall, the findings suggest that PIP2 binding disrupts the R395:E398 salt bridge, facilitating the conformational transition of the hSK2 channel to the open state. The consistent changes in SASA values for residues like R395, K278, R299, and R418 upon PIP2 binding further highlight their importance in channel activation. This dynamic interaction underscores the regulatory role of PIP2 in SK2 channel function.

### g. Proposed Structural Mechanism of SK2 Channel Activation by PIP2

**Figure 7A-B** summarize our proposed atomistic structural and dynamic mechanisms for how PIP2 works in concert with CaM to activate the hSK2 channel. The hSK2-CaM complexes are depicted in closed, intermediate, and open states, showing only two subunits for clarity. In the closed state, the CaMBD at the C-termini of the hSK2 channel interacts with the C-lobes of apo-CaM. Upon $Ca^{2+}$ binding to CaM in the intermediate state, CaM extends and interacts with the S4-S5 linker, which pulls the linker down and out, enabling PIP2 to bind to residue R395. This interaction disrupts the R395:E398 salt bridge, leading to the widening of the pore domain in the open state.

**Figure 7C-D** illustrate that multiple PIP2 molecules reside within the inner leaflet of the membrane in the hSK2-CaM intermediate state, with $Ca^{2+}$ bound to CaM. The MD simulations show PIP2 molecules sequentially traversing through the basic amino acid residues in CaMBD, similar to objects on a conveyor belt. When PIP2 binds to residue R395, it disrupts the R395:E398 salt bridge between adjacent hSK2-CaM subunits. This disruption triggers the expansion of the channel's pore gate, facilitating $K^+$ passage through the channel.

## 5.3) Discussion and Conclusion

In this study, we present the first report on the regulation of apamin-sensitive SK currents in cardiomyocytes by PIP2, leveraging optogenetics and magnetic nanoparticles combined with structural analyses. We used homology modeling and all-atom MD simulations to uncover the atomistic structural and dynamic mechanisms of how PIP2, in concert with $Ca^{2+}$-CaM, activates the hSK2 channel. Previous studies had identified PIP2 binding sites but lacked a detailed understanding of the mechanisms underlying hSK2 channel activation by PIP2[11]. Our computational study aimed to identify all key PIP2 binding sites in an unbiased manner by randomly placing PIP2 in the lipid membrane during simulations. We found that the amino acid residue R395, located close to the intracellular hydrophobic gate, is critical for channel activation. PIP2 binding to R395 disrupts the R395:E398 salt bridge, increases the flexibility of the S6 transmembrane segment, and activates the channel. These findings offer a new platform for future structural-based drug designs targeting SK channels.

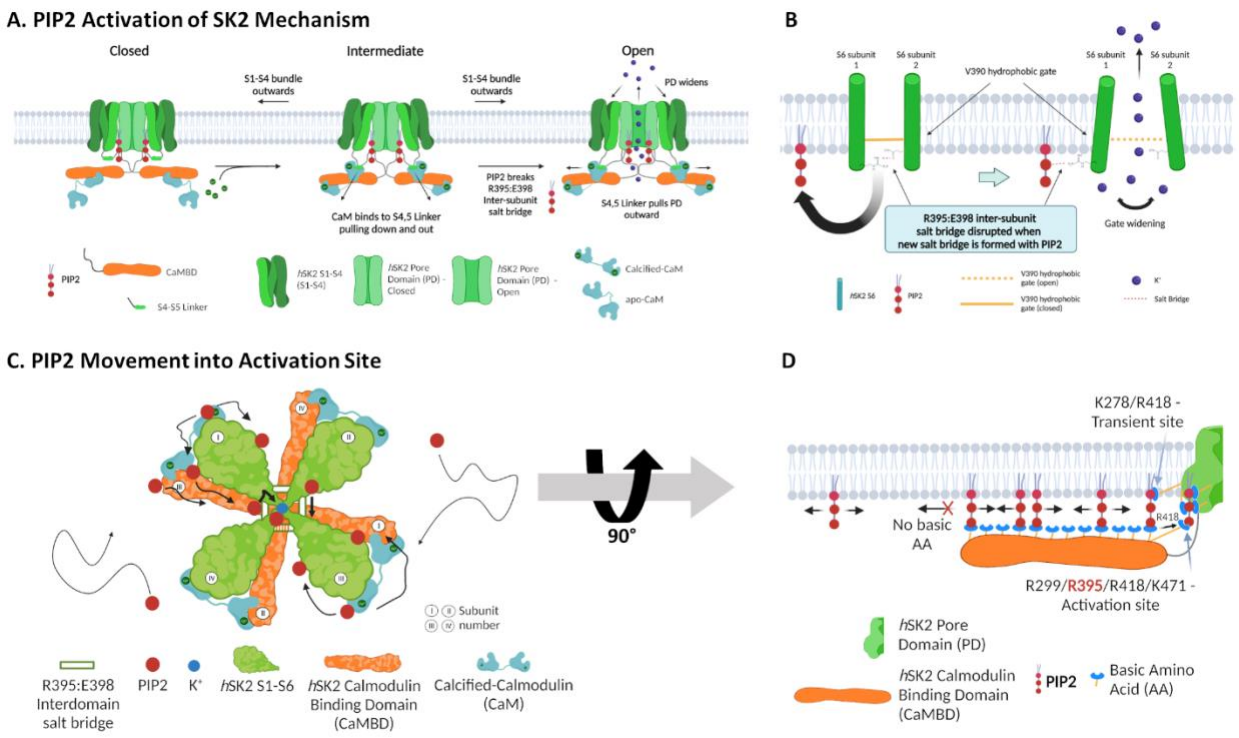



**Figure 7. Diagrams depicting how PIP2 works in concert with $Ca^{2+}$ and calmodulin (CaM) in the proposed activation mechanism of the SK2 channel.** ***A****) The conformational state transition mechanism of the hSK2-CaM complex.* ***B****) The widening of the SK2 intracellular gate through the disruption of the cross-subunit R395:E398 salt-bridge by PIP2.* ***C****) The movement of PIP2 into the SK2 activation site.* ***D****) Depiction of the conveyor belt-like movement of PIP2 by binding to multiple basic residues in the CaMBD of SK2 channels.*

Alternative mechanisms of SK channel activation by PIP2 have been explored in prior studies, including a crystal structure of the hSK2 channel containing CaMBD, CaM, and PIP2[11]. Our current study, using cryo-EM structures of hSK4 complexed with CaM and built using Rosetta, showed differences from previous models. Notably, mutations in the S6-CaMBD linker region disrupt multiple salt bridge networks, affecting the overall conformation of the channel but not directly impacting PIP2 binding. Prior studies also identified different PIP2 binding pockets involving residues from CaM and the S6-CaMBD linker. Our study found that PIP2 binds initially to R299 before transitioning to the activation site, highlighting differences in PIP2 binding dynamics between hSK2 and hSK4 channels.

MD simulations revealed the dynamic mechanisms of PIP2 binding and its role in channel activation under more physiologically relevant conditions, with PIP2 embedded in the lipid bilayer. We found that PIP2 did not bind to previously predicted CaM residues but did bind to residue K116 while trafficking to the transient site. Our new analysis protocol focused on salt bridges between hSK2-CaM and PIP2, showing that PIP2 binding disrupts the cross-subunit R395:E398 salt bridge, crucial for channel activation. The R395:E398 salt bridge's presence varied across channel states, being most frequent in the closed state and least frequent in the open state. The increase in SASA for R395 upon PIP2 binding supports its role in the transition to the open state.

Regulation of cardiac ion channels by PIP2 represents a novel mechanism for lipid regulation of cardiac function. PIP2 has been shown to regulate multiple cardiac ion channels and may play critical roles in arrhythmia initiation and maintenance. PIP2 dysregulation in heart failure, mediated by sustained phospholipase C activation, significantly affects cardiac ion channel activities. SK channels integrate beat-to-beat changes in intracellular $Ca^{2+}$ with membrane potentials, playing essential roles in cardiac excitability and arrhythmogenesis. New insights into the dynamic regulation of cardiac SK channels by PIP2 are crucial for developing therapeutic strategies for cardiac arrhythmias.

The current study provides new insights into the atomistic mechanisms of how PIP2 works with $Ca^{2+}$-CaM to activate hSK2 channels. The critical role of R395 is supported by its location near the hydrophobic gate and its dynamic salt bridge formation with E398 and PIP2. Future mutational studies will further validate these findings. Additional computational studies are needed to determine if PIP2 binds R395 in closed or open states and to quantify the increased flexibility near E398 when the R395 salt bridge is disrupted. Further MD simulations of SK1-3 and SK4 channels will provide insights into distinct PIP2 activation mechanisms. These studies will enhance our understanding of the atomistic-level mechanisms for channel function modulation and support future structure-based drug designs for therapeutic inhibitors and activators of the SK channel family.

## 5.4) References


1. Wyndham, C. R. Atrial fibrillation: the most common arrhythmia. *Tex. Heart Inst. J.* **27**, 257–267 (2000).
2. Nesheiwat, Z., Goyal, A. & Jagtap, M. Atrial Fibrillation. in *StatPearls* (StatPearls Publishing, Treasure Island (FL), 2024).
3. Zhang, X.-D., Lieu, D. K. & Chiamvimonvat, N. Small Conductance Ca2+-Activated K+ Channels and Cardiac Arrhythmias. *Heart Rhythm Off. J. Heart Rhythm Soc.* **12**, 1845–1851 (2015).
4. Diness, J. G., Bentzen, B. H., Sørensen, U. S. & Grunnet, M. Role of Calcium-activated Potassium Channels in Atrial Fibrillation Pathophysiology and Therapy. *J. Cardiovasc. Pharmacol.* **66**, 441–448 (2015).
5. Zhang, M., Meng, X.-Y., Zhang, J., Cui, M. & Logothetis, D. E. Molecular overlap in the regulation of SK channels by small molecules and phosphoinositides. *Sci. Adv.* **1**, (2015).
6. Lee, C.-H. & MacKinnon, R. Activation mechanism of a human SK-calmodulin channel complex elucidated by cryo-EM structures. *Science* **360**, 508–513 (2018).
7. Leman, J. K. *et al.* Macromolecular modeling and design in Rosetta: recent methods and frameworks. *Nat. Methods* **17**, 665–680 (2020).
8. Jo, S., Kim, T., Iyer, V. G. & Im, W. CHARMM-GUI: A web-based graphical user interface for CHARMM. *J. Comput. Chem.* **29**, 1859–1865 (2008).
9. Case, D. A. *et al.* The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

10. Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A. & Sansom, M. S. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graph.* **14**, 354–360, 376 (1996).

11. Zhang, M. *et al.* Selective phosphorylation modulates the PIP2 sensitivity of the CaM-SK channel complex. *Nat. Chem. Biol.* **10**, 753–759 (2014).

# CHAPTER 6:

# Shaping the World of Tomorrow: Innovations and Future Directions in Computational Physiology

*"In my view, all that is necessary for faith is the belief that by doing our best we shall succeed in our aims: the improvement of mankind."* – Rosalind Franklin (1920 – 1958)

---

## 6.1) Dissertation Summary

The research presented in this dissertation has significantly advanced our understanding of ion channels and their roles in physiological processes, focusing on the human (denoted by the prefix "h") voltage-gated sodium channel hNa$_V$1.7, voltage-gated potassium channels hK$_V$11.1 also known as hERG (human *Ether-à-go-go*-Related Gene), and small conductance calcium activated potassium channel hSK2. Utilizing state-of-the-art computational techniques, we have modeled conformational state-specific structures of these channels, studied their structural dynamics, ion conduction as well as their modulations by lipids, small-molecule and peptide ligands, revealing critical insights into their functions and drug interactions.

Our research on the hNa$_V$1.7 channel provides a detailed understanding of its interaction with the protoxin-II peptide, offering a basis for the design of novel pain therapeutics[2]. The knowledge of state-specific toxin binding to hNa$_V$1.7 informs the design of therapeutic treatments for chronic pain, demonstrating the importance of understanding the structural nuances of ion channels and their ligand interactions[3,4].

Our findings on the hERG channel elucidate the atomic-level mechanisms underlying drug-induced arrhythmias, emphasizing the importance of the inactivated state of this channel for pro-arrhythmic drug binding and highlighting potential pathways for safer cardiac medications. Studies have shown that understanding the inactivated state of the hERG channel is crucial for developing drugs that minimize the risk of arrhythmias[1]. Despite these advances, experimental approaches for determining different ion channel states still have

limitations. Therefore, we leveraged deep learning-based tools like AlphaFold2[5] to predict various hERG channel conformational states, providing models that align closely with experimental data[5–7]. These computational predictions offer a powerful complement to traditional experimental methods, enabling more precise and detailed insights into ion channel behavior.

Additionally, our combined experimental and molecular simulation research on the hSK2 channel modulation by phosphatidylinositol 4,5-bisphosphate (PIP2) lipid has identified the critical role of the channel's R395 residue in the S6 transmembrane segment, which was hard to predict using experiments alone. This finding enhances our understanding of hSK2 channel regulation and provides a new platform for the structure-based design of therapeutic inhibitors and activators, particularly relevant for ongoing clinical trials targeting cardiac arrhythmias[8].

Overall, this dissertation underscores the power of integrative computational approaches in predicting drug interactions, designing targeted therapies, and uncover fundamental mechanisms of ion channel gating. By combining protein structural modeling, molecular dynamics simulations, deep learning predictions, and experimental validations, we have made significant strides in understanding the complex behavior of ion channels. These insights are crucial for developing safer and more effective treatments for cardiac and neurological disorders.

## 6.2) Outlook

The future of computational physiology holds immense promise as technological advancements continue to expand the capabilities of simulation and modeling techniques. These advancements are particularly transformative when combined with artificial intelligence and machine learning, which are reshaping the landscape of biological research. Tools like AlphaFold[5,9] have revolutionized protein structure prediction by providing highly accurate models, which were previously unattainable[10,11]. The continuous improvement of these artificial-intelligence-driven tools is expected to lead to even more accurate and

efficient modeling of complex biological systems, thereby enhancing our understanding of protein functions and their implications in health and disease.

In the context of ion channels, which are crucial for numerous physiological processes and disease mechanisms, future research should focus on integrating multi-scale models. These models would bridge the gap between molecular interactions and whole-organism physiology[12]. Multi-scale modeling is a powerful approach that connects detailed molecular dynamics simulations with higher-level physiological models, such as those of cells, tissues and organs. This integration can elucidate how molecular-level drug interactions affect overall physiological outcomes. For example, coupling molecular simulations of ion channel function with cardiac electrophysiology models can provide insights into how specific drug interactions can lead to arrhythmias, thus informing the development of safer cardiac drugs[13]. Moreover, incorporating patient-specific data into these models could revolutionize personalized medicine by enabling treatments tailored to individual genetic and physiological profiles. This patient-specific approach is essential for optimizing therapeutic efficacy and minimizing adverse effects, leading to a new era of precision medicine.

Furthermore, advancements in high-performance computing and quantum computing are set to further enhance the resolution and accuracy of molecular dynamics simulations. High-performance computing allows for the simulation of larger and more complex biological systems over longer timescales, which is critical for studying the dynamic behavior of ion channels and their interactions with various biomolecules. Quantum computing, although still in its early stages, holds the potential to solve complex problems that are currently beyond the reach of classical computers. By enabling simulations that can model the quantum nature of molecular interactions, quantum computing could provide unprecedented insights into ion channel function and drug interactions.

Another critical area of development is the integration of artificial intelligence with experimental data to refine and validate computational models. By training artificial intelligence algorithms on large datasets from experimental studies, researchers can develop models that predict biological behaviors with high accuracy. The resulting models can identify patterns and relationships that might be missed by traditional approaches, thus

accelerating the discovery of new drug targets and therapeutic strategies[14]. For instance, using machine learning to analyze ion channel mutations and their effects on channel function can help in the design of targeted therapies for diseases such as epilepsy, cystic fibrosis, and cardiac arrhythmias[15,16].

In addition to these technological advancements, there is a growing emphasis on the need for collaborative efforts across disciplines. The complexity of biological systems and the challenges of modeling them accurately necessitate collaboration between computational biologists, experimental scientists, clinicians, and data scientists. Such interdisciplinary efforts are essential for translating computational predictions into clinically relevant outcomes and ensuring that new therapies are both effective and safe[17].

By harnessing the power of advanced computational tools, artificial intelligence, multi-scale modeling, patient-specific data, and cutting-edge computing technologies, we can accelerate the pace of drug discovery and development, improve the safety and efficacy of treatments, and move closer to the realization of precision medicine. The journey ahead is filled with opportunities for groundbreaking discoveries that will transform our approach to understanding and treating human diseases. These advancements will not only enhance our scientific knowledge but also have a profound impact on patient care and public health globally.

## 6.3) References


1. Perrin, M. J., Kuchel, P. W., Campbell, T. J. & Vandenberg, J. I. Drug Binding to the Inactivated State Is Necessary but Not Sufficient for High-Affinity Binding to Human Ether-à-go-go-Related Gene Channels. *Mol. Pharmacol.* **74**, 1443–1452 (2008).
2. Ngo, K. *et al.* Elucidating molecular mechanisms of protoxin-II state-specific binding to the human NaV1.7 channel. *J. Gen. Physiol.* **156**, e202313368 (2024).
3. Bennett, D. L., Clark, A. J., Huang, J., Waxman, S. G. & Dib-Hajj, S. D. The Role of Voltage-Gated Sodium Channels in Pain Signaling. *Physiol. Rev.* **99**, 1079–1151 (2019).

4. Xu, H. *et al.* Structural Basis of Nav1.7 Inhibition by a Gating-Modifier Spider Toxin. *Cell* **176**, 702-715.e14 (2019).
5. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
6. Wang, W. & MacKinnon, R. Cryo-EM Structure of the Open Human Ether-à-go-go-Related K+ Channel hERG. *Cell* **169**, 422-430.e10 (2017).
7. Carus H.Y. Lau *et al.* Potassium dependent structural changes in the selectivity filter of HERG potassium channels. *bioRxiv* 2023.12.14.571769 (2023) doi:10.1101/2023.12.14.571769.
8. Zhang, X.-D., Lieu, D. K. & Chiamvimonvat, N. Small Conductance Ca2+-Activated K+ Channels and Cardiac Arrhythmias. *Heart Rhythm Off. J. Heart Rhythm Soc.* **12**, 1845–1851 (2015).
9. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 1–3 (2024) doi:10.1038/s41586-024-07487-w.
10. AlQuraishi, M. AlphaFold at CASP13. *Bioinforma. Oxf. Engl.* **35**, 4862–4865 (2019).
11. Jumper, J. *et al.* Applying and improving AlphaFold at CASP14. *Proteins* **89**, 1711–1721 (2021).
12. Clancy, C. E. *et al.* Multiscale Modeling in the Clinic: Drug Design and Development. *Ann. Biomed. Eng.* **44**, 2591–2610 (2016).
13. Yang, P.-C. *et al.* A Computational Pipeline to Predict Cardiotoxicity: From the Atom to the Rhythm. *Circ. Res.* **126**, 947–964 (2020).
14. Paul, D. *et al.* Artificial intelligence in drug discovery and development. *Drug Discov. Today* **26**, 80–93 (2021).
15. Boßelmann, C. M., Hedrich, U. B. S., Lerche, H. & Pfeifer, N. Predicting functional effects of ion channel variants using new phenotypic machine learning methods. *PLoS Comput. Biol.* **19**, e1010959 (2023).

16. Ashrafuzzaman, M. Artificial Intelligence, Machine Learning and Deep Learning in Ion Channel Bioinformatics. *Membranes* **11**, 672 (2021).

17. Smye, S. W. & Frangi, A. F. Interdisciplinary research: shaping the healthcare of the future. *Future Healthc. J.* **8**, e218 (2021).