COMPUTATIONAL STUDIES OF PROTEIN STRUCTURE, DYNAMICS, AND FUNCTION IN NATIVE-LIKE ENVIRONMENTS

BY

HUAN RUI

Submitted to the graduate degree program in Bioinformatics
and the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy.

Chairperson: Wonpil Im, Ph. D.

Eric Deeds, Ph. D.

John Karanicolas, Ph. D.

Mario Rivera, Ph. D.

Ilya Vakser, Ph. D.

Date Defended: 04/12/2013

The Dissertation Committee for Huan Rui
certifies that this is the approved version of the following dissertation:

COMPUTATIONAL STUDIES OF PROTEIN STRUCTURE, DYNAMICS, AND FUNCTION IN NATIVE-LIKE ENVIRONMENTS

Chairperson: Wonpil Im, Ph. D.

Date Approved: 04/12/2013

Abstract

Proteins are among the four unique organic constituents of cells. They are responsible for a variety of important cell functions ranging from providing structural support to catalyzing biological reactions. They vary in shape, dynamic behavior, and localization. All of these together determine the specificity in their functions, but the question is how. The ultimate goal of the research conducted in this thesis is to answer this question. Two types of proteins are of particular interest. They include transmembrane proteins and protein assemblies. Using computer simulations with available experimental data to validate the simulation results, the research described here aims to reveal the structure and dynamics of proteins in their native-like environment and the indication on the mechanism of their functions.

The first part of the thesis focuses on studying the structure and functions of transmembrane proteins. These proteins are consisted of transmembrane α-helices or β-strands, and each of the secondary structure elements adopts a unique orientation in the membrane following its local interactions. The structure of the entire protein is a collection of the orientations of these elements and their relative positions with respect to one another. These two basic aspects of membrane protein structure are studied in Chapter II and III. In Chapter II, efforts are given to determine the favorable orientation of a β-hairpin peptide, protegrin-1, in different lipid bilayers. The orientational preference results from the interplay between the protein and the surrounding lipid molecules. Chapter III is centered on revealing the structure and dynamics of caveolin-1 in DMPC bilayers. Caveolin-1 forms a re-entrant helix-turn-helix structure with two α-helices embedded in the membrane bilayer. The study shows that caveolin-1 monomer is rather

dynamic and maintains its inserted conformation via both specific and non-specific protein-lipid interactions. To investigate the structural and dynamic impact on the function of a membrane protein, molecular dynamics simulations of the voltage-dependent anion channel are performed and the results are presented in Chapter IV. It is found in this chapter that the electrostatic interactions between charged residues on the channel wall facing the lumen are responsible for retarding the cation current, therefore giving the channel its anion selectivity.

The second category of protein that is of interest in this thesis is the assembled protein complex, especially the ones that are highly symmetric. Actually, many membrane proteins belong to this category as well, but the study presented here in Chapter V involves simulations performed on a soluble protein complex, bacterioferritin B from *Pseudomonas Aeruginosa*. It is revealed by the simulations that the dynamic behavior of the protein is magnified by the symmetry and is tightly associated to its function.

Acknowledgements

There are many people I want to express my gratitude to for their kindness and support during my study in graduate school. I am where I am today all because of you.

First and foremost, I would like to thank my advisor, Professor Wonpil Im. It is my great pleasure to spend my years in graduate school under his mentorship. His expertise in the field of computational biophysics has helped me acquire the knowledge and skills necessary for conducting research as an independent researcher. Besides giving me guidance in scientific issues, he is also an inspiring figure to me. It is from him that I learned what is of the most importance in a scientific research project, and I deeply appreciate the opportunity and encouragement he gives me to learn and apply new skills to my research.

I would also like to thank Dr. Ilya Vakser, Dr. John Karanicolas, Dr. Eric Deeds, and Dr. Mario Rivera for being on my thesis committee, reading my thesis, and providing me insightful comments. I had the chance to discuss many of my research problems with them while I pursue my doctoral degree and have since learned a lot from them.

I want to devote my thanks to all the past and current Im lab members as well. They have helped me tremendously on many technical issues I encountered in my research. I am especially thankful to Dr. Jinhyuk Lee, who was a postdoc when I first joined the lab and had shown me how the helix and hairpin restraint potential works, Dr. Taehoon Kim, who was the most senior graduate student at the time and had helped me set up my first molecular dynamics simulation, Sunhwan Jo, who is a guru and my idol in programing and all technology related issues, Dr. Soohyung Park, who is a current

postdoc and had taught me how to use the algorithm he developed to conduct the window-exchange umbrella sampling simulations, and many others.

I am also glad that The University of Kansas in Lawrence is my first stop in my journey of learning and exploring in the United States and I have since made many life-long friends who keep me company during my years in Lawrence. I would like to express my thanks to the Dye family. Brian and Judy Dye are my adopted parents and they have provided me food, shelter, and emotional support when I need the most. My thanks also goes to all my best friends in Lawrence: Rolf Petermann for making me pizza and providing me great music, Beth Ruhl for listening to me complaining about hardship and telling me to be positive, and Connita Fiorella Sardon-Quiroz for being so cheerful and awesome all the time.

At last, I want to say thank you to my beloved parents. I could have not been here at this moment of my life if it were not for them. They shower me with unconditional love when I was growing up and gave me the courage to go abroad and pursue a career in scientific research. They have done everything they can to support me both in my career and in my personal life. My gratitude towards them is beyond words.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction

1.1. An overview of protein structure, dynamics, environment, and function

Proteins serve as the bridge between the genotypes and phenotypes of organisms. The genetic materials stored in DNA and RNA are constantly being retrieved and translated to proteins, which perform a variety of cellular functions such as enzyme catalysis, gene regulation, energy production, and signal transduction. The functional specificity of a protein usually comes from its distinctive structure and/or the embedded dynamic properties. For example, membrane channels are usually cylindrical tubes with pathways connecting both sides of a membrane. Such architecture facilitates communication across the membrane. Another example is the molecular switches. These proteins change their conformations upon certain stimuli, which can be the binding of a ligand or a change in pH, and either activate or inhibit certain biological pathways. The structure and dynamics of a protein is affected by its environment as in the case of the molecular switch mentioned above. To study the mechanism of action in proteins, all the above factors have to be included.

1.2. Membrane proteins

Membrane proteins belong to one of the three major protein classes. Approximately 20-30% of genes in most known genomes encode membrane proteins [1]. Besides their abundance, they are indispensible for cells to function properly. They mediate signal transduction across the membrane, facilitate the traffic ions and metabolites into and out of cells and cellular organelles, catalyze reactions, and promote cell recognition (Figure 1-1). Dysfunction and misregulation of them can lead to severe diseased conditions. In fact, ~50% of current medicinal drugs target membrane proteins [2]. To better understand

how membrane proteins function and to further use this information for the development of novel therapeutic tools, thorough investigation of membrane protein structure, dynamics and its interplay with the lipid bilayers is required.

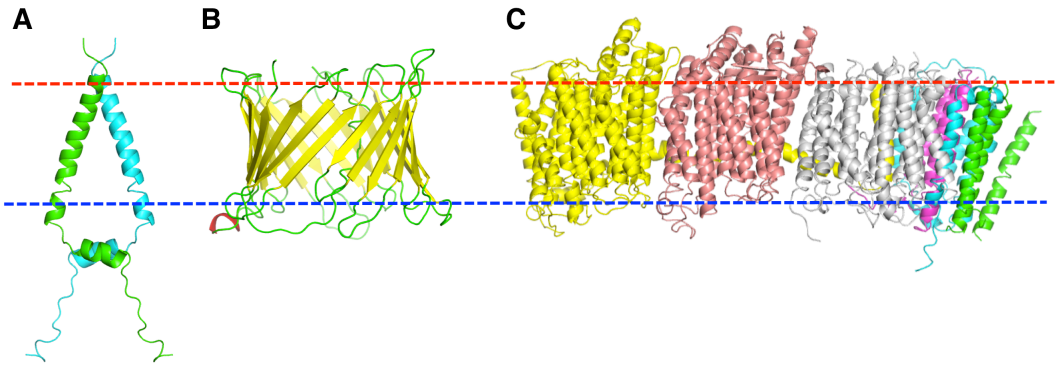


Figure 1.1. Membrane protein examples. (A) Epidermal growth factor receptor, erbB-1 homodimer (PDB ID:2M20). (B) Voltage-dependent anion channel (PDB ID:2K4T). (C) Respiratory complex I (PDB ID:3RKO).

The study of membrane protein structure has been on the radar of biophysicists ever since the mid-20th century [3, 4]. In 1985, Johann Deisenhofer and Hartmut Michel determined the atomic-resolution structure of the *Rhodopseudomonas viridis* photosynthetic reaction center [5]. This is a cornerstone in membrane protein biophysics history and marks the start of a modern era in membrane protein research. The following three decades have witnessed an exponential growth of high-resolution unique membrane protein structures [6], the number of which has registered 385 as of February 2013 [7]. Most of them are composed of α-helices or β-strands or both. These two secondary structures are the common structural motifs in membrane proteins [8]. When present in transmembrane proteins, these structural elements have specific orientations with regard

to a membrane bilayer. It is their orientations together that determine the overall structure of a transmembrane protein.

The access to membrane protein structures leads to another question: what are the dynamic behaviors of these proteins? It is becoming more evident that determining the structures is only the first step, and the function of membrane proteins also depends on their dynamics. Many membrane proteins undergo conformational changes to perform their function. For example, glutamate transporters undergo conformational transition between inward and outward facing states to enable the uptake of the neurotransmitter glutamate [9], and ligand binding and pH change control the transition between the open and closed states of the pentameric ligand-gated channels [10, 11]. Besides these large conformational transitions, there also exist local fluctuations of protein backbone. The role of such fluctuations is usually to help achieve an intermediate state in between two metastable structural states so that large conformational change can occur.

The structure and dynamics of membrane proteins are only part of the story. The proper function of these proteins requires the presence of their native-like environment, the membranes. There is emerging evidence that membrane protein structure and dynamics are affected by their immediate environment [12]. The indication of this is that the membrane environment is an indispensible factor in studying membrane protein functions. This adds another layer of complexity to the problem, because it is well established that different cells have varied lipid compositions [13], and the even distribution of lipids in a single patch of membrane can be very heterogeneous [14]. Depending on the composition, membranes are also different in their physical properties including hydrophobic thickness, intrinsic curvature, fluidity, and so on. Change in these

properties would affect the equilibrium distribution of different membrane protein conformations and also bias the transition paths between them. It is likely that cells utilize this interplay between protein-membrane interactions and the structure and dynamics of the protein to fine-tune the membrane protein function.

A large variety of experimental techniques have been employed to investigate membrane protein structure, dynamics, and interactions and their implications on the function. Structures of membrane proteins are usually determined using X-ray crystallography and NMR spectroscopy. However, these classical experimental techniques have their limitations. In X-ray crystallography, the hurdle faced by the community is the difficulty in obtaining well-ordered crystals. Growing membrane protein crystals is extremely tricky as it involves endless rounds of detergent condition optimization. In NMR experiments, the two major problems are the sample purification and signal overlaps [15], which limit the application to many membrane proteins. Most of the previous solution NMR studies focused on β-barrels and a few low molecular weight α-helical transmembrane peptides, however, new methods have been developed in recent years to allow a wider spectrum of target proteins to be studied [16]. For example, progress in solid-state NMR spectroscopy shows that it is now possible to extract both the structure and dynamics information of some large membrane protein complexes in lipid bilayers [17, 18]. Also, by using cryo-electron tomography combined with techniques for image averaging and 3D classification a handful of near atomic-resolution structures have been retrieved [19-22]. Among them are the nuclear pore complex [23], and the flagellar motor of *Treponema primitia* [24], both of which are transmembrane proteins. In addition to all these advances, the use of interdisciplinary approaches in studying membrane

protein function is gradually gaining traction. By incorporating site-directed spin labeling (SDSL)-EPR measurements and available NMR data, Han, et al. determined the binding mode of hemagluttinin in lipid bilayers and the structural change that happens at different pH [25]. The latest functional model of EGFR is the result of a synergetic work of NMR spectroscopy and microsecond time scale molecular dynamics simulations [26, 27]. The field of membrane protein function studies is at a point where the integration of experimental and theoretical approaches is inevitable, and such integration effort will benefit future membrane protein function research tremendously.

1.3. Supramolecular protein assemblies

Many fundamental biological functions are carried out by large protein assemblies [28]. This is exemplified by the participation of ribosomes in protein production and the involvement of nuclear pore complexes in communication between the nucleus and the rest of the cell. As a matter of fact, most proteins have quaternary structures and they assemble to form oligomeric complexes [29]. Within these proteins, 50-70% are composed of two or more copies of identical subunit and are symmetric [30, 31]. Large protein assemblies such as the microtubules, the ferritins, some membrane channels, and viral capsids all belong to this category (Figure 1.2). Because of their universality and their functional role in biological context, there is a pressing need to study their structure, dynamics, and interactions, which would open doors to the understanding of the ways these complexes function. A thorough examination of how these systems work will also aid designing and synthesizing molecular machines that can perform useful functions [32].

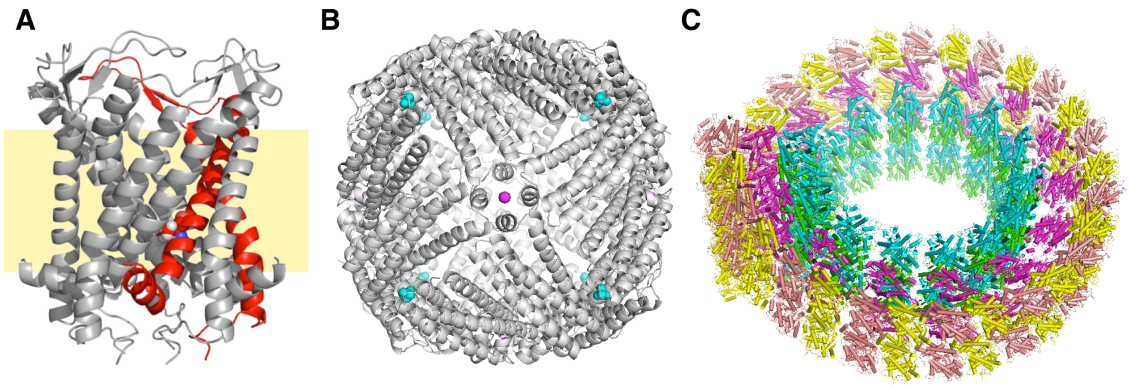


Figure 1.2. Examples of some large protein assemblies. (A) Mechanosensitive channel of large conductance (PDB ID:2OAR). (B) Bacterioferritin (PDB ID:3ISF). (C) Kinesin-microtubule ring complex (PDB ID:3EDL)**.**

To gain a mechanistic view of protein complex functions, the first step is to characterize their structures and dynamics [33]. High-resolution approaches like the X-ray crystallography and NMR spectroscopy have been shown to be able to successfully reveal protein complex structures in atomic detail. In 1999, the crystal structure of the RNA polymerase core complex from *Thermus aquaticus* was resolved at 3.3 Å resolution [34]. Almost 10 years later, Yusupov and coworkers reported the 3.0 Å crystal structure of an eukaryotic ribosome [35]. The progress in determining atomic-resolution complex structures is compromised due to technical difficulties associated with protein expression and crystallization. The dynamics of these large complexes can be probed by tools including single molecule fluorescence spectroscopy and magical angle spinning NMR as illustrated by the studies of myosin dynamics in complex with actin filaments [36] and HIV-1 capsid protein CA assembly [37]. Besides these experimental methods, computational endeavors have also made considerable contribution to the structure and dynamics research of large protein assemblies. They include knowledge-based structural

modeling [38], normal mode analysis [39, 40], and molecular dynamics simulations [41, 42], just to list a few. These techniques are commonly used and have shown promising results.

1.4. Molecular dynamics simulations

Molecular dynamics (MD) simulation is one of the frequently used computational modeling approaches that can be used in conjunction with experimental techniques to enhance our comprehension of the biological world. It has the ability to examine the structure and dynamics of any molecule in its native-like environment at the atomic level. It predicts the position evolution of all the atoms in a given system over a course of time following Newton's classical equation of motion:

$$\mathbf{F}_\alpha = m_\alpha \ddot{\mathbf{r}}_\alpha = -\frac{\partial U\left(\left\{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right\}\right)}{\partial \mathbf{r}_\alpha}. \tag{1-1}$$

In this equation, $\mathbf{F}_\alpha$ is the force exerted on atom α with mass m_α, and $\mathbf{r}_\alpha$ is its position. The potential energy function of the system is given by $U\left(\left\{\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right\}\right)$ and involves the positions of all the atoms in the system. For most biologically relevant systems, it is expressed in the form of a summation over bonded and non-bonded energy terms:

$$\begin{aligned} U\left(\left\{r_1, \mathbf{r}_2, \cdots, \mathbf{r}_N\right\}\right) = & \sum_{\text{bonds},i} \frac{1}{2} k_{b,i}\left(b_i - b_i^0\right)^2 + \sum_{\text{angles},i} \frac{1}{2} k_{\theta,i}\left(\theta_i - \theta_i^0\right)^2 \\ & + \sum_{\text{torsion},i} k_{\phi,i}\left[1 + \cos\left(n_i \phi_i - \delta_i\right)\right] \\ & + \sum_i \sum_{j, j \neq i} \left\{ \varepsilon_{ij} \left[\left(\frac{r_{ij}^{\min}}{r_{ij}} \right)^{12} - 2\left(\frac{r_{ij}^{\min}}{r_{ij}} \right)^{6} \right] + \frac{q_i q_j}{\varepsilon r_{ij}} \right\} \end{aligned} \tag{1-2}$$

The items on the right side of the equation describe bond, angle, torsion, van der Waals, and electrostatic energies, respectively [43]. Because of the complexity of the problem, no simple analytical solution to equation 1-1 is available. The propagation of a system has to be carried out under numerical integration algorithms. The velocity Verlet algorithm is one among them [44]. This algorithm evolves both the position and velocity of atoms in a system and is implemented in majority of the available MD simulation packages [45-48]. The scheme it follows to propagate the position is given below:

$$\mathbf{r}\left(t+\delta t\right)=\mathbf{r}\left(t\right)+\dot{\mathbf{r}}\left(t\right)\delta t+\frac{1}{2}\ddot{\mathbf{r}}\left(t\right)\delta t^{2}, \tag{1-3}$$

where **r**(t) is position of an atom at time t, and δt is the time step for the integration. At the same time, the velocity of the atom is also updated:

$$\dot{\mathbf{r}}\left(t+\delta t\right)=\dot{\mathbf{r}}\left(t\right)+\ddot{\mathbf{r}}\left(t\right)\frac{\delta t}{2}+\ddot{\mathbf{r}}\left(t+\delta t\right)\frac{\delta t}{2}. \tag{1-4}$$

The first two terms on the right-hand side give the mid-point velocity, which is then used to update to the final velocity at time $t + \delta t$. Usually, with all the bonds involving hydrogen atoms having their lengths fixed through the SHAKE algorithm [49, 50], the time step for the propagation is 2 fs.

The first MD simulation was performed for a hard sphere system in 1957 [51]. It was not until 20 years later that this approach was first applied to study the dynamics of biomolecules. In 1977, Karplus and coworkers examined the internal motions of the bovine pancreatic trypsin inhibitor using MD simulations and pointed out that the fluid-like nature of the protein interior is likely involved in its biological function [52]. Since the publication of this study, the number of MD simulation studies on biomolecules has

skyrocketed, benefiting from the exponential growth of the available structures and the computational power. Nowadays, MD simulations of systems with hundreds of thousands of atoms at microsecond time scales are becoming increasingly feasible [53]. A good exemplar is the MD simulation study on drug binding in the G-protein coupled receptors, which presents the binding pathways in atomic detail and provides future possibilities in drug design and the study of protein allosteric effects [54]. Since current MD simulations can achieve similar time scales to many biologically relevant events, they are becoming a valuable alternative to study the mechanisms of these events [55, 56].

Chapter II

Protegrin-1 Orientation and Physicochemical Properties in Membrane Bilayers Studied by Potential of Mean Force Calculations[1]

[1] Reused from *Journal of Computational Chemistry*, Huan Rui and Wonpil Im, **31**. 2010. pp 2859-2867. Copyright (2010). With permission from Wiley Periodicals, Inc.

Summary

Protegrin-1 (PG-1) belongs to the family of antimicrobial peptides. It interacts specifically with the membrane of a pathogen and kills the pathogen by releasing its cellular contents. To fully understand the energetics governing the orientation of PG-1 in different membrane environments and its effects on the physicochemical properties of the peptide and membrane bilayers, we have performed the potential of mean force (PMF) calculations as a function of its tilt angle at four distinct rotation angles in explicit membranes composed of either DLPC (1,2-dilauroylphosphatidylcholine) or POPC (1-palmitoyl-2-oleoylphosphatidylcholine) lipid molecules. The resulting PMFs in explicit lipid bilayers were then used to search for the optimal hydrophobic thickness of the EEF1/IMM1 implicit membrane model, in which a two-dimensional PMF in the tilt and rotation space was calculated. The PMFs in explicit membrane systems clearly reveal that the energetically favorable tilt angle is affected by both the membrane hydrophobic thickness and the PG-1 rotation angle. Local thinning of the membrane around PG-1 is observed upon PG-1 tilting. The thinning is caused by both hydrophobic mismatch and arginine-lipid head group interactions. The two-dimensional PMF in the implicit membrane is in good accordance with those from the explicit membrane simulations. The ensemble-averaged Val16 ^{15}N and ^{13}CO chemical shifts weighted by the two-dimensional PMF agree fairly well with the experimental values, suggesting the importance of peptide dynamics in calculating such ensemble properties for direct comparison with experimental observables.

2.1. Introduction

Antimicrobial peptides are crucial components of the mammalian innate immune system [57]. When a pathogen is present, they can self-insert into the pathogen membrane and destroy the pathogen by depolarizing its membrane. Due to their ability to kill a broad range of pathogens like bacteria, fungi, and viruses, they are often considered potential antibiotics [58]. The secondary structure of the antimicrobial peptide varies. Most of them are unstructured in solution and form either α-helices or β-strands upon partitioning into membranes. The rest are β-hairpins reinforced by inter-strand disulfide bonds [59], represented by a family of antimicrobial peptides including protegrin-1 (PG-1), horseshoe crab tachyplesins, and mammalian defensin [60]. PG-1 is a short peptide with 18 amino acids (RGGRL CYCRR RFCVC VGR). It was first discovered and purified from porcine leukocytes [61]. The solution NMR experiment reveals that the structure is composed of two anti-parallel β-strands connected by a short turn [62]. Four cysteine residues are present and form two inter-strand disulfide bonds (Cys6-Cys15, Cys8-Cys13), which are both structurally and functionally essential for PG-1 [63].

The antimicrobial activity of PG-1 is highly dependent on the lipid composition. Efforts have been made to study the orientation and interactions of PG-1 in different membrane environments both experimentally [64-66] and computationally [67-69]. Based on the solid-state NMR measurements, Hong and coworkers suggested that PG-1 stays as a transmembrane monomer tilted at 55º ± 5º from the membrane normal in a 1,2-dilauroylphosphatidylcholine (DLPC) bilayer [66], while it shows concentration-dependent dimerization in a 1-palmitoyl-2-oleoylphosphatidylcholine (POPC) bilayer. In anionic lipid bilayers representing the bacteria outer membrane, an oligomeric pore

structure was proposed [70]. Nonetheless, the orientation of PG-1 monomer in either the dimer or the oligomer with respect to the membrane remains elusive. In particular, to uncover the mechanisms underlying the polymerization and the membrane disruption activities of PG-1, the following questions need to be addressed. What are the dominant orientation of PG-1 in different membrane bilayers and the underlying energetics? Does the PG-1 tilting affect the thickness of the membrane? Moreover, what mechanism does PG-1 utilize to disturb the lipid bilayers? Unfortunately, these questions are difficult to be addressed by experimental techniques alone due to their limited resolution and the experimental difficulties associated with membrane proteins/peptides.

Potential of mean force (PMF) calculations offer a means to determine the complicated energetics and thermodynamic properties along the chosen reaction coordinates [71]. Although calculating the PMF from straightforward molecular dynamics (MD) simulations seems impractical, special sampling techniques can be applied to enhance the conformational sampling along the reaction coordinates. In the present study, we employed the umbrella sampling technique [72] with the weighted histogram analysis method (WHAM) [73] to construct the PMFs as a function of PG-1 tilt and rotation angles in both explicit and implicit membrane bilayers. Two pure lipid bilayers, DLPC and POPC, were considered in the explicit membrane simulations. The rotation angle of PG-1 was held at different values to examine the influence of PG-1 rotation on its tilting in the bilayers. Our recently developed β-hairpin tilt and rotation restraint potentials [74] were used to restrain the tilt and rotation angles of PG-1. The resulting PMFs in the explicit DLPC bilayers were then compared with those generated from the EEF1/IMM1 [75, 76] implicit membrane simulations at various membrane

hydrophobic thicknesses in order to find an optimal hydrophobic thickness of the implicit membrane model. Afterwards, a two-dimensional PMF calculation as a function of PG-1 tilt and rotation angles was carried out in the implicit membrane with the optimal hydrophobic thickness. At the end, the application of the two-dimensional PMF is illustrated by the calculation of solid-state NMR observables.

2.2. Computational Methods

2.2.1. Defining the tilt and rotation angles

The orientation of PG-1 is defined by its tilt (τ) and rotation (ρ) angles (Figure 2.1). With the *Z*-axis parallel to the membrane normal, τ is defined as the angle between the hairpin principal axis and the unit vector along the Z-axis. To define ρ, both the internal and external references have to be specified. The internal reference is given by the vector pointing from a point on the hairpin axis to atom N on Cys6, while the unit vector along the Z-axis is chosen as the external reference. One can then define ρ as the angle between the projections of such two vectors on the plane made by the second and third principal axes of the hairpin. Detailed expressions can be found in the work of Lee et al[74, 77].

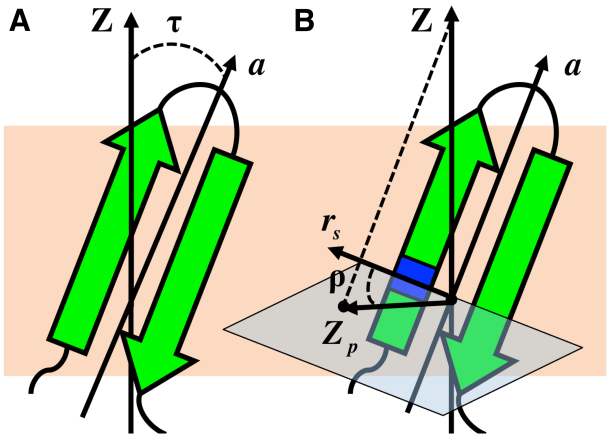


Figure 2.1. Definitions of (A) tilt angle (τ) and (B) rotation angle (ρ). The hairpin axis (**a**) is the eigenvector that corresponds to the smallest eigenvalue of the hairpin inertia tensor. $\mathbf{Z_p}$ is the projection of the *Z*-axis on the plane (light blue) made by the second and third principal axes. $\mathbf{r_s}$ is the principal plane projection of the vector pointing from **a** to the reference atom N on Cys6 (blue).

2.2.2. Umbrella sampling simulations in explicit membranes

The structure of PG-1 was obtained by taking chain A from PDB:1ZY6 [65] (solid-state NMR structure) using the PDB Reader module in CHARMM-GUI (http://www.charmm-gui.org) [78] with disulfide bonds and amidated C-terminus. With the hairpin axis of PG-1 parallel to the membrane normal (i.e., the *Z*-axis) ($\tau = 0^o$) and its center at $Z = 0$, the peptide/membrane systems with either DLPC or POPC lipid were generated by the CHARMM-GUI Membrane Builder module [79, 80], bathed in 0.2 M KCl solution. Each system was then subjected to 750 ps equilibration, in which the last 350 ps was performed with the $P2_1$ periodic boundary condition [81] to allow flipping of lipid molecules between the top and bottom leaflets. The peptide/membrane systems in DLPC and POPC bilayers after equilibration are shown in Figure 2.2. To investigate the influence of PG-1 rotation and membrane tension on the PG-1 tilting energetics, we constructed a total of eight systems with different PG-1 rotation angles and bilayer tensions. The system information is given in Table 2.1. In each system, the initial structures for umbrella sampling simulations were generated by tilting the peptide from 0^o to the specified maximum tilt angle (τ_{max}) in Table 2.1 every 2^o with a 100 ps time interval, followed by a 16 ns production for each window, resulting in a total of 3.648 μs umbrella sampling simulations. We used a force constant of 10,000 kcal/(mol·rad^2) for the tilt angle restraint potential [74] and a rotation angle force constant of 1,000 kcal/(mol·rad^2) for the systems with rotation restraints (ρ_{cons}) in Table 2.1. All calculations were performed in the NPγT ensemble [82, 83] at 303.15 K using the biomolecular simulation program CHARMM [45] with the all-atom parameter set PARAM22 [84] including the dihedral cross-term corrections (CMAP) [85] and recently optimized lipid parameters [86].

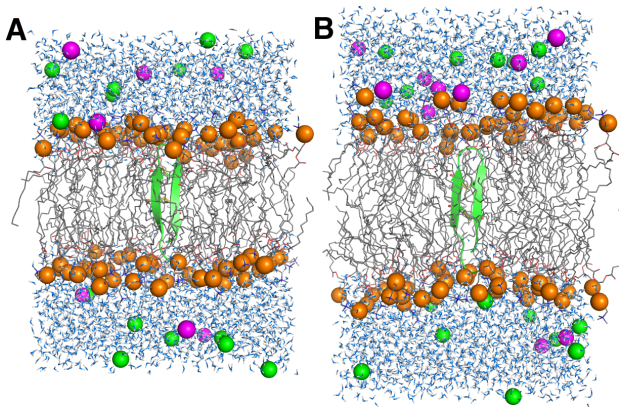


Figure 2.2. The peptide/membrane systems in (A) DLPC and (B) POPC bilayers. PG-1 is centered at the middle of the membrane with its principal axis parallel to the membrane normal. The secondary structure of PG-1 (green) is shown. Lipid molecules (grey) are illustrated by lines. Phosphate atoms (orange) are shown in spheres to illustrate the hydrophobic/hydrophilic interface. Water molecules (blue) are shown in lines. 0.2 M K^+ (magenta) and Cl^- (green) are also presented with spheres. The figure is produced with PyMOL [87].

Table 2.1. System information of umbrella sampling simulations in explicit membranes.

Systems	Lipids (top/bottom)	τ_{max}	ρ_{cons}	γ (dyne/cm)	Number of ions (K^+/Cl^-)
DLPC_S1	DLPC(41/41)	70º	-	20	7/14
DLPC_S2	DLPC(41/41)	50º	0º	20	7/14
DLPC_S3	DLPC(41/41)	40º	90º	20	7/14
DLPC_S4	DLPC(41/41)	50º	180º	20	7/14
DLPC_S5	DLPC(41/41)	70º	-	0	7/14
DLPC_S6	DLPC(41/41)	70º	-	10	7/14
POPC_S7	POPC(38/38)	50º	-	20	10/17
POPC_S2	POPC(38/38)	40º	90º	20	10/17

2.2.3. Umbrella sampling simulations in implicit membrane

The EEF1/IMM1 [75, 76] implicit membrane model in CHARMM [45] was employed for the simulations. The planar implicit membrane was centered at $Z = 0$ with the

membrane normal parallel to the *Z*-axis. The EEF1/IMM1 calculations were prepared using the Implicit Solvent Modeler module in CHARMM-GUI [78]. The simulations were conducted at 298 K using Langevin dynamics with a 2 fs time-step. All bond lengths involving hydrogen were fixed with the SHAKE algorithm [88]. In the one-dimensional PMF calculations, the initial systems for umbrella sampling simulations were generated by tilting PG-1 every 2º from 0º to 70º, resulting in a total of 36 windows, at each of the four rotation angles used in the explicit DLPC membrane PMF calculations (Table 2.1). Then, each window was subjected to a 200 ps equilibration and a 4 ns production. We repeated the PMF calculations with a different hydrophobic thickness of the implicit membrane from 9 Å to 12 Å in an increment of 0.5 Å. The optimal hydrophobic thickness was determined to be 10 Å, which showed the maximum agreement with the PMFs calculated in explicit membranes (see 2.3. Results and Discussion). In the two-dimensional PMF calculation as a function of PG-1 τ and ρ, the initial structures were created by tilting PG-1 every 2º from 0º to 70º, followed by rotating it every 5º from -180º to 180º at each τ, resulting in a total of 2,628 windows. A 200 ps equilibration followed by a 10 ns production was performed at each window. In the implicit membrane umbrella sampling simulations, we used the same force constants for the hairpin tilt and rotation restraint potentials as those employed in the explicit membrane PMF calculations.

2.2.4. WHAM calculations

The biased τ and ρ distribution from the umbrella sampling simulations was unbiased by the WHAM [73] analysis to construct the PMF. In the one-dimensional PMF calculation,

the free energy was calculated every 0.1º, with the energy tolerance of 10^{-5} kcal/mol. To calculate the PMF in the τ and ρ space, the interval of free energy calculation was set to 1.0º along the τ and ρ directions. The energy tolerance for the WHAM iteration was 10^{-5} kcal/mol.

2.3. Results and Discussion

2.3.1. PMFs as a function of tilt angle in explicit lipid bilayers

To examine the convergence of the PMFs, the trajectory in each simulation was sequentially divided into ten parts with equal time duration and the PMF was calculated from each of the sub-trajectories. The averaged PMFs from the last eight sub-trajectories (i.e., after 3.1 ns) are shown in Figure 2.3. The largest PMF fluctuations appear at the two extremes along τ, i.e., either at small or large tilt angles. Even in these energetically prohibitive regions, the PMF fluctuation does not exceed ±1.46 kcal/mol, suggesting the calculated PMFs are well converged. Therefore, all the averaged properties in this study were calculated using trajectories after 3.1 ns. The free energy minimum tilt angle (τ_{min}) and the free energy change from τ = 0º to τ_{min} ($\Delta G_{0\rightarrow\tau_{min}}$) in each system are summarized in Table 2.2.

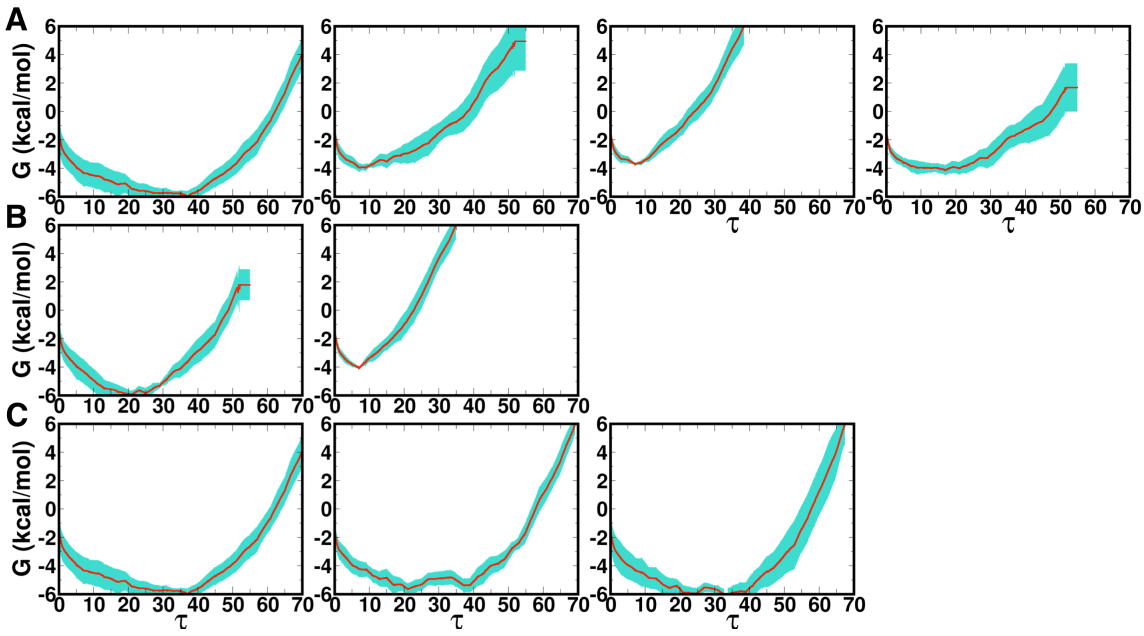


Figure 2.3. The averages (red) and fluctuations (cyan) of the PMFs calculated from the explicit membrane systems with (A) ρ ≈ -90º, 0º, 90º, and 180º in DLPC bilayers (from

left to right); (B) ρ ≈ -90º and 90º in POPC bilayers (from left to right); (C) γ = 20, 10 and 0 dyne/cm in DLPC bilayers (from left to right).

Table 2.2. Free energy minimum tilt angle (τ_{min}) and associated free energy changes ($\Delta G_{0\to\tau_{min}}$).

Systems	ρ	τ_{min}	$\Delta G_{0\to\tau_{min}}$ (kcal/mol)
DLPC_S1	~ -90º	37.0º	-6.12 ± 0.40
DLPC_S2	0º	8.7º	-4.05 ± 0.21
DLPC_S3	90º	7.1º	-3.67 ± 0.12
DLPC_S4	180º	17.1º	-4.20 ± 0.34
DLPC_S5	~ -90º	33.3º	-6.35 ± 0.39
DLPC_S6	~ -90º	21.1º	-5.81 ± 0.40
POPC_S1	~ -90º	20.9º	-5.87 ± 0.23
POPC_S2	90º	7.0º	-4.04 ± 0.12

The PG-1 conformational stability is measured in terms of its backbone H-bonding fraction. According to the solid-state NMR experiment [89], there are six possible backbone H-bonds: Leu5-Val16, Tyr7-Val14, and Arg9-Phe12. In the current study, we define the H-bond (D−H···A) by an H···A distance < 2.8 Å and a D−H···A angle > 120º. The H-bond fraction at a specified PG-1 tilt angle was calculated as the time average of the instantaneous fractions in each window, which is given by $f_{\mathrm{HBOND}} = \left\langle N_{\mathrm{calc}} / N_{\mathrm{total}} \right\rangle_{\tau}$ where $N_{\mathrm{total}} = 6$ and N_{calc} is the number of instant H-bonds in PG-1. As shown in Figure 2.4, except at energetically unfavorable large tilt angles, the β-hairpin structures are well preserved with $f_{\mathrm{HBOND}} > 0.75$, corresponding to 4~5 H-bonds maintained throughout the simulations. Therefore, the PG-1 β-hairpin structure appears to be stable in membrane environments.

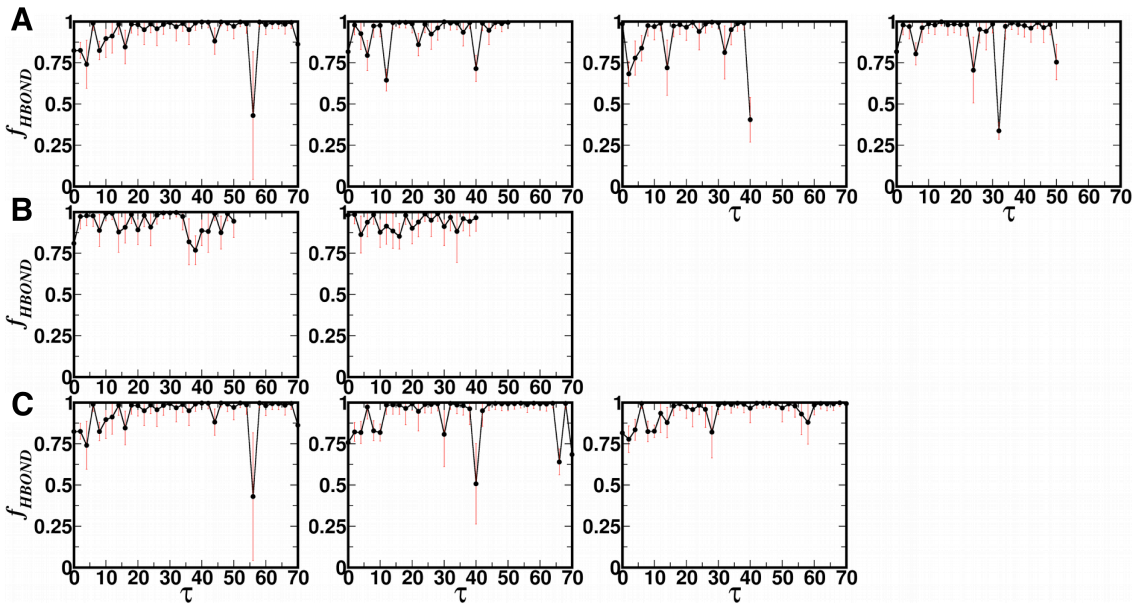


Figure 2.4. The average H-bond frequencies (solid black) of PG-1 at each τ in explicit membrane simulations with (A) ρ ~ -90º, ρ = 0º, 90º, and 180º in DLPC bilayers (from left to right); (B) ρ ~ -90º and ρ =90º in POPC bilayers (from left to right); (C) γ = 20, 10, and 0 dyne/cm in DLPC bilayers (from left to right). The standard errors are shown in red lines.

The resulting PMFs at different PG-1 rotation angles in two lipid bilayers (Figure 2.3A and B and Table 2.2) reveal that the PG-1 tilting is largely affected by two factors: the PG-1 rotation and the membrane hydrophobic thickness. The rotation angle of PG-1 is the dominant factor that determines the thermally-accessible tilt angle in a bilayer with a given lipid type. As shown in Table 2.1, τ_{min} varies significantly with ρ: 37.0º (ρ ≈ -90º), 8.7º (ρ ≈ 0º), 7.1º (ρ ≈ 90º), and 17.7º (ρ ≈ 180º) in DLPC bilayers, and 20.9º (ρ ≈ -90º) and 7.0º (ρ ≈ 90º) in POPC bilayers. The same rotational preference was also observed in our previous MD simulations [69]. In addition to τ_{min}, $\Delta G_{0\to\tau_{min}}$ varies with ρ. The largest $\Delta G_{0\to\tau_{min}}$ in DLPC and POPC bilayers are -6.12 ± 0.40 kcal/mol and -5.87 ± 0.23 kcal/mol, respectively, when ρ ≈ -90º.

The bilayer hydrophobic thickness (L_H) is another factor that affects the tilting behavior of PG-1 in the membrane environments. With the same ρ, PG-1 appears to have reduced τ_{min} in the POPC bilayers ($L_H \approx 28$ Å [90]) compared to that in the DLPC bilayers ($L_H \approx 21$ Å [91]) (Figure 2.3A and B and Table 2.2). Such difference was also observed in our previous MD studies [69], which can be explained as the result of the hydrophobic mismatch between PG-1 and the lipid bilayers. The same argument holds for single-pass TM-helices as they tilt more in membranes with smaller hydrophobic thickness [92]. We also examined the influence of membrane surface tension on PG-1 tilting. Within the same lipid type, the systems with different surface tensions show similar PMF profiles and τ_{min} (Figure 2.3C and Table 2.2), which is discussed in the next section in detail.

What are the microscopic interactions that determine such an optimal orientation of PG-1 in membranes? In the previous MD studies of PG-1, we found that the ρ dependence of the thermally-accessible τ arises from the sidechain position of the aromatic residues [69]. In addition, the PMFs as a function of τ at specific ρ provide another route to explore other factors such as the asymmetric hydrophobicity of the hairpin. As shown in Figure 2.5, the PG-1 structure can be simplified as a rectangular box with four surrounding facets denoted by F_S, F_N, F_R, and F_C. Tilting in the direction of each facet corresponds to PG-1 tilting at four distinct rotation angles, i.e., -90º, 0º, 90º and 180º, respectively. Among the four faces, F_S (ρ = -90º) has the highest hydrophobicity according to the hydrophobic scale proposed by Janin [93], followed by F_C (ρ = 180º), F_N (ρ = 0º), and F_R (ρ = 90º). Similarly, the free energy minimum tilt angles are τ_{min} (ρ ≈ -90º) > τ_{min} (ρ ≈ 180º) > τ_{min} (ρ ≈ 0º) > τ_{min} (ρ ≈ 90º) in DLPC

bilayers, and τ_{min} ($\rho \approx -90^{o}$) > τ_{min} ($\rho \approx 90^{o}$) in POPC bilayers. Therefore, the rotational preference in PG-1 tilting appears to be governed by its anisotropic hydrophobicity.

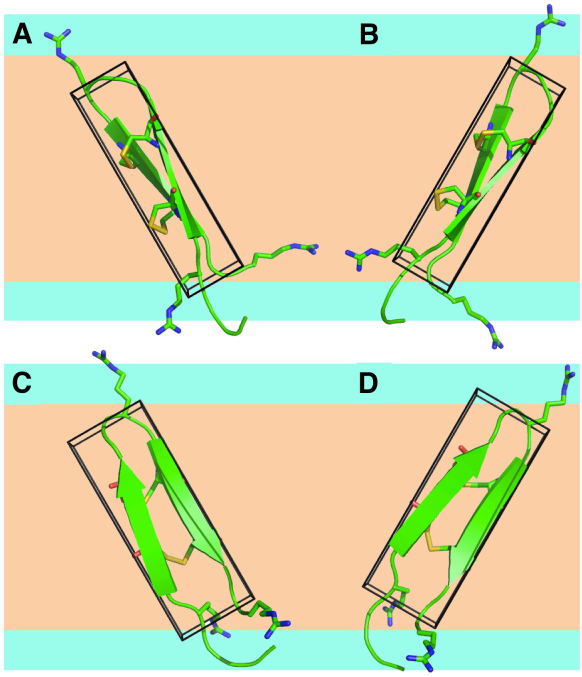


Figure 2.5. The tilting of PG-1 in the direction of different facet with (A) F_S ($\rho = -90^{o}$), (B) F_R ($\rho = 90^{o}$), (C) F_C ($\rho = 180^{o}$), and (D) F_N ($\rho = 0^{o}$). The figures are produced by PyMol[87]. The hydrophobic region (orange) and head group region (cyan) of the membrane are shown by colored bars. Arg4, Arg11 and Arg18 are highlighted in stick representation to illustrate their positions with regard to the membrane.

2.3.2. Physicochemical properties of the PG-1 systems

The physicochemical properties of PG-1 and the lipid bilayers depend largely on the configuration of the peptide in the membrane environments. Some properties, such as the solid-state NMR chemical shifts, are measured based on a large structural set in a sample and the observed value represents an ensemble average. In order to calculate the statistically significant chemical shifts, structural ensembles at different peptide orientations should be included. The chemical shifts of Val16 ^{15}N and ^{13}CO were

calculated as a function of τ in DLPC_S1 (Figure 2.6A), which is closely related to the solid-state NMR experiment setup [66]. The chemical shift σ is calculated by

$$\sigma = \left\langle \hat{Z} \cdot \left[\sum_{i=1}^{3} \hat{e}_i \sigma_{ii} \hat{e}_i \right] \cdot \hat{Z} \right\rangle_{\text{ens}} \tag{2-1}$$

where σ_{ii} and $\hat{e}_i$ are the magnitude and unit vector of the instantaneous chemical shift tensor for ^{15}N or ^{13}CO [94]. In the present calculations, we used the same chemical shift tensors as those in the solid-state NMR study [66]. The difference between the calculated chemical shifts and the experimental values is denoted by the root mean-square deviation δ_σ,

$$\delta_\sigma = \sqrt{\frac{(\sigma_N^{(\text{calc})} - \sigma_N^{(\text{exp})})^2 + (\sigma_C^{(\text{calc})} - \sigma_C^{(\text{exp})})^2}{2}} \tag{2-2}$$

where $\sigma_N^{(\text{calc})}$ and $\sigma_C^{(\text{calc})}$ are the calculated ^{15}N and ^{13}CO chemical shifts, while the corresponding experimental values are $\sigma_N^{(\text{exp})}$ (143 ± 2 ppm) and $\sigma_C^{(\text{exp})}$ (216 ± 5 ppm) [66], respectively. The resulting δ_σ as a function of τ in Figure 2.6B spans a wide range from 23.9 ± 12.7 ppm at τ = 42º to 66.6 ± 13.4 ppm at τ = 2º. As shown in Figure 2.3A, a rough correspondence is found between the τ regions with low δ_σ and thermal accessibility, i.e., 20º ≤ τ ≤ 41º. However, the magnitude of δ_σ is still considerably large, such as 45.5 ± 8.9 ppm at τ = 30º, which is probably due to limited sampling, the uniformly weighted averaging, and the single rigid tensor approximation. In order to examine these issues, we performed a two-dimensional PMF calculation as a function of PG-1 τ and ρ in the EEF1/IMM1 implicit membrane [75, 76] and calculated the ensemble-averaged chemical shifts, which is discussed in the next section.

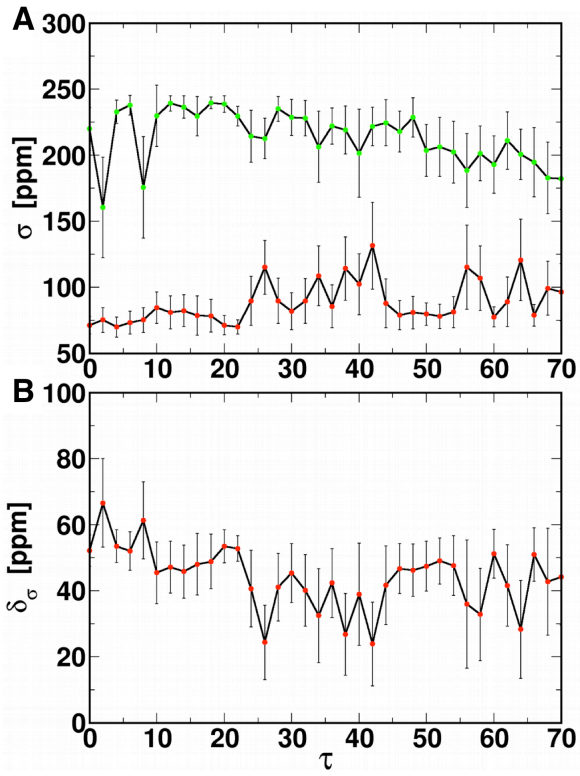


Figure 2.6. (A) The chemical shifts of Val16 ^{15}N (red) and ^{13}CO (green) as a function of τ, calculated from system DLPC_S1. (B) The root-mean-square deviation (δ_σ) between the calculated and the experimental chemical shifts as a function of PG-1 tilt angle in system DLPC_S1. The standard deviations are shown in thin black lines.

Another important physical property in membrane systems is the local membrane adjustment upon the peptide insertion. To minimize the energy penalty of exposing the nonpolar residues to aqueous solution, the lipids in the vicinity of an integral membrane protein would change their hydrophobic length to maximize the hydrophobic match [95]. Given the hydrophobic length of PG-1 (~30 Å) [62] and the suggested tilt angle of 55º [66], the previous solid-state NMR study proposed a local membrane thinning effect of PG-1 inside a DLPC bilayer [66]. However, no direct measurements are available. In order to investigate the membrane thinning as a function of PG-1 tilting in the lipid

bilayers, we calculated the membrane thickness as the average distance between the C1 carbon atoms of the lipid acyl chains in the top/bottom leaflets. The local thinning of the membrane (ΔL_H) is defined by the thickness difference between the bulk and contact lipid molecules, which are distinguished by a lipid-peptide heavy atom distance of 4 Å. Figure 2.7 shows the averaged membrane thinning as a function of τ in DLPC_S1 and POPC_S1. The thinning is around 2 Å in DLPC_S1 and 5 Å in POPC_S1 with a deviation of about ±1 Å when τ is in the thermally-accessible region. Similar amplitudes of local membrane thinning were also observed in the previous PG-1 MD studies [68, 96, 97].

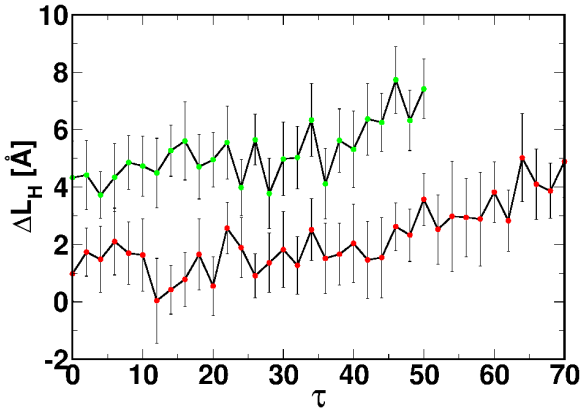


Figure 2.7. Local membrane thinning as a function of PG-1 tilt angle in system DLPC_S1 (red) and POPC_S1 (green).

Besides the apparent impact from membrane hydrophobic thickness, the influence of the membrane tension on the thinning effect is also worth mentioning. As shown in Figure 2.8, decreasing the membrane tension induces the increase of the bulk lipid bilayer hydrophobic thickness, which consequently magnifies the thinning effect (Figure 2.9).

Nonetheless, the hydrophobic thickness of the lipids close to the peptide is not significantly changed, which is the reason why the PMFs are similar regardless of the applied membrane tensions (Figure 2.3C). Together with the PMF of PG-1 in the POPC bilayers, it becomes clear that the hydrophobic mismatch can be accommodated by the local lipid adaptations when the extent of a hydrophobic mismatch is moderate. In this case, the thermally favorable orientation of the peptide is not affected. However, if such extent exceeds some threshold (e.g., DLPC to POPC), the local bilayer adaptations can no longer compensate the energy penalty associated with a mismatch. Instead, an orientational change of the peptide (mostly in tilt angle) occurs as seen in the POPC bilayers. In addition to the lipid type and the membrane tension, the thinning effect also depends on the tilt angle of PG-1 (Figure 2.7). As the peptide tilts, its effective hydrophobic length decreases, leading to the increase of the negative hydrophobic mismatch, in which the effective hydrophobic length of PG-1 is smaller than the membrane hydrophobic thickness. As a result, the local membrane hydrophobic thickness shrinks to maximize the hydrophobic match. Besides the hydrophobic interactions between the hydrophobic regions of the peptide and the lipid tails [95], the membrane thinning around PG-1 is also attributed to the specific electrostatic interactions between the arginine guanidinium groups and the phosphate groups in the lipid head groups [69].

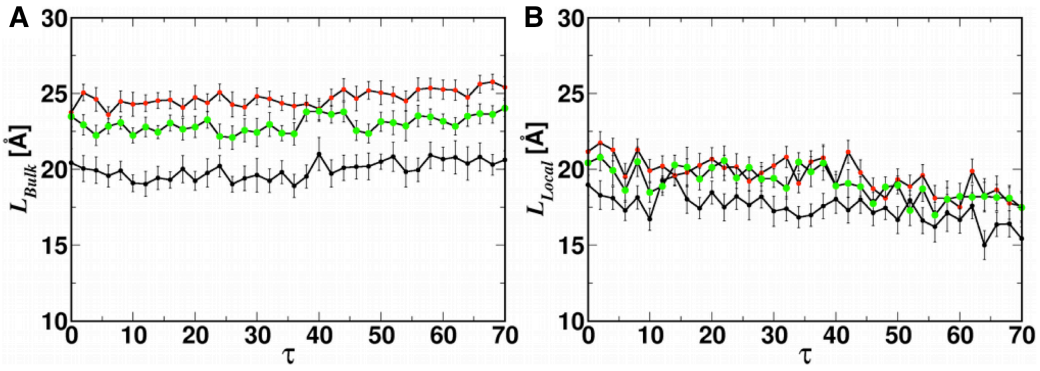


Figure 2.8. The hydrophobic thickness of (A) the bulk lipid bilayer (L_{Bulk}) and (B) the local lipid bilayer (L_{Local}) as a function of tilt angle (τ) in the DLPC systems with different membrane tensions γ = 0 (red), 10 (green), and 20 (black) dyne/cm.

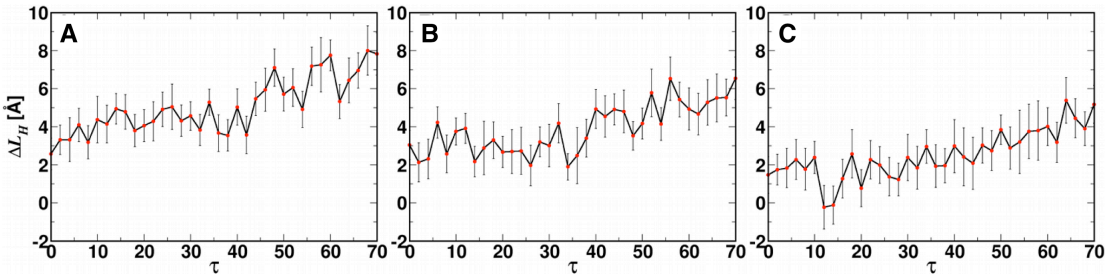


Figure 2.9. The local membrane thinning (ΔL_H) as a function of tilt angle (τ) in the DLPC systems with γ = (A) 0, (B) 10, and (C) 20 dyne/cm.

There are six arginines in PG-1 at position 1, 4, 9, 10, 11, and 18. The influence of different arginines on the membrane thinning can be probed by the correlation coefficient ($C_{Z-\Delta L}$) between the guanidinium group (heavy atoms) positions along the *Z*-axis (Z_{gua}) and the corresponding membrane hydrophobic thickness change (ΔL_{H}) at

different ρ (Figure 2.10); $C_{Z-\Delta L} = \mathrm{cov}\left(\Delta L_{\mathrm{H}}, Z_{\mathrm{gua}}\right) / \sigma_{\Delta L_{\mathrm{H}}} \sigma_{Z_{\mathrm{gua}}}$ where $\sigma_{\Delta L_{\mathrm{H}}}$ and $\sigma_{Z_{\mathrm{gua}}}$ are the standard deviations of ΔL_{H} and Z_{gua} of each Arg residue. The value of $C_{Z-\Delta L}$ is between 0 and 1 with a higher value corresponding to a better correlation between Z_{gua} and ΔL_{H}. The $C_{Z-\Delta L}$ of each arginine is summarized in Table 2.3 and then mapped onto the PG-1 structure (Figure 2.11). Different arginine guanidinium groups are responsible for the thinning effect at different ρ as PG-1 tilts in the bilayer. When ρ ≈ -90º, the $C_{Z-\Delta L}$ of Arg11 is the highest among all the arginines. As PG-1 is rotated 90º, nevertheless, the strongest correlations appear to be at Arg4 and Arg9. In systems with ρ = 0º and 180º, Arg9/18 and Arg4/11/18 are mostly correlated with the thinning effect. Given the most probable orientation of PG-1 in the membrane bilayers, i.e., τ = 37.0º, ρ ≈ -122.5º in DLPC bilayers; τ = 20.9º, ρ ≈ -92.5º in POPC bilayers, the electrostatic interactions between the lipid head groups and the guanidinium groups of Arg4, Arg11, and Arg18 appear to be the determinant of the membrane thinning (Figure 2.5).

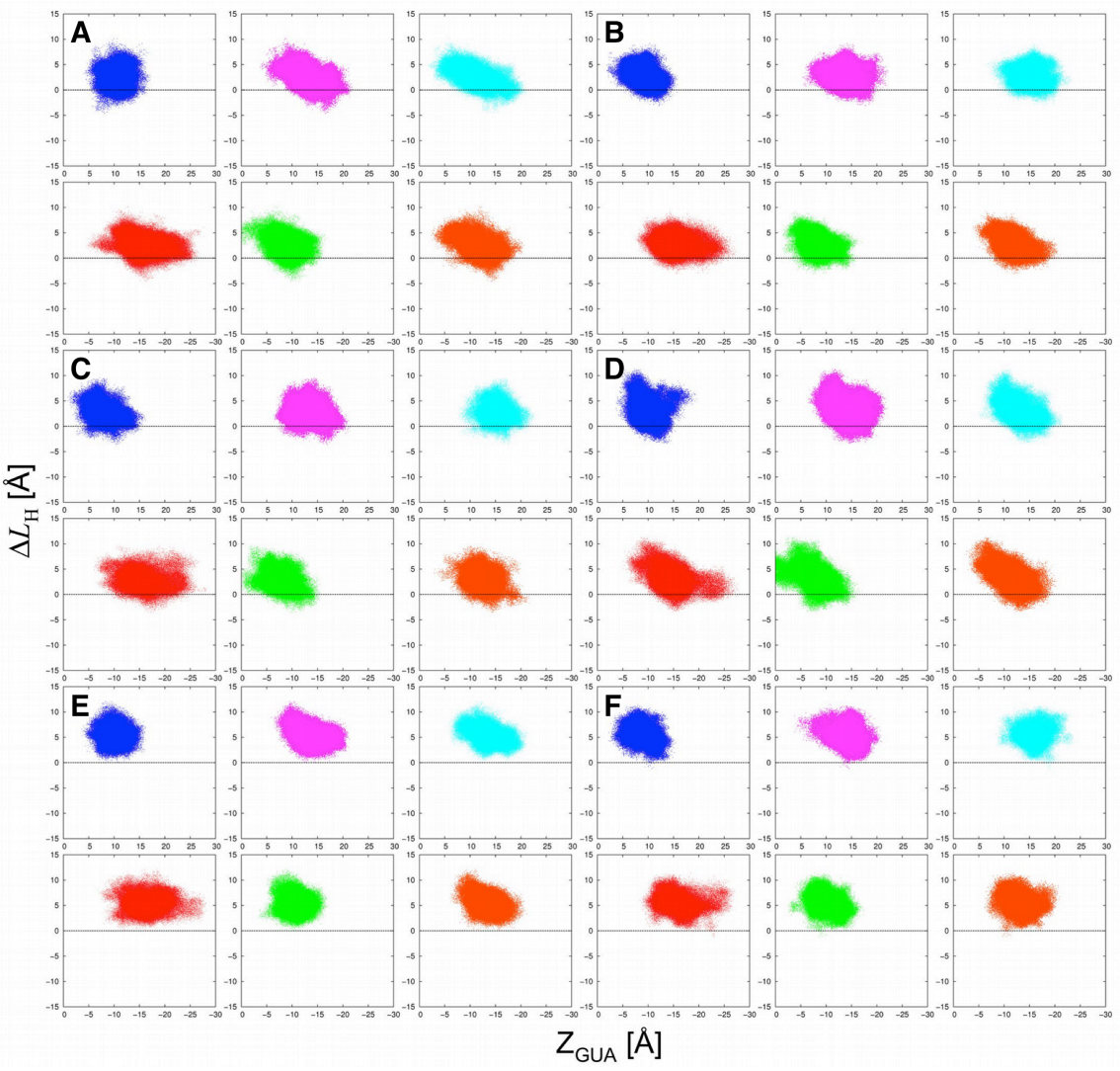


Figure 2.10. The local membrane thinning (ΔL_{H}) versus the *Z*-coordinate of guanidinium group (Z_{gua}) for Arg9 (blue), Arg10 (magenta), Arg11 (cyan), Arg1 (red), Arg4 (green), and Arg18 (orange) in (A) DLPC_S1, (B) DLPC_S2, (C) DLPC_S3, (D) DLPC_S4, (E) POPC_S1, and (F) POPC_S2 (see **Table 2.1** for system names).

Table 2.3. The correlation coefficients ($C_{Z-\Delta L}$) between the Z positions of guanidinium groups (Z_{gua}) and the local membrane thinning (ΔL_H) for all the arginine residues.

	DLPC_S1	POPC_S1	DLPC_S3	POPC_S2	DLPC_S2	DLPC_S4
Arg1	0.25	0.01	0.14	0.21	0.16	0.42
Arg4	0.32	0.09	0.30	0.29	0.30	0.47
Arg9	0.03	0.08	0.34	0.28	0.37	0.19
Arg10	0.46	0.41	0.13	0.17	0.25	0.32
Arg11	0.55	0.41	0.12	0.06	0.08	0.48
Arg18	0.48	0.36	0.22	0.23	0.41	0.55

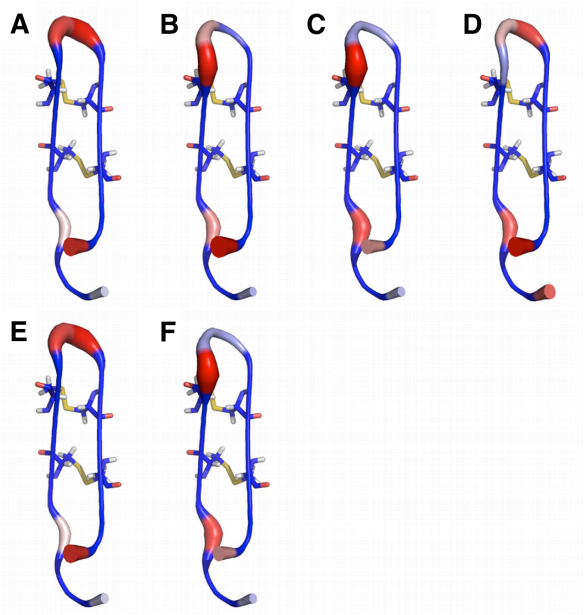


Figure 2.11. The correlation coefficients ($0 \leq C_{Z-\Delta L} \leq 1$) between the Z positions of guanidinium groups (Z_{gua}) and the local membrane thinning (ΔL_H) are mapped onto the PG-1 structure for (A) DLPC_S1, (B) DLPC_S2, (C) DLPC_S3, (D) DLPC_S4, (E) POPC_S1, and (F) POPC_S2 (see **Table 2.1** for system names). PG-1 is displayed with a tube representation. The radius (small to big) and color (blue to red) of the tube denote the magnitude of $C_{Z-\Delta L}$ (small to big). The figure is produced with PyMol [87].

2.3.3. Two-dimensional PMF as a function of both tilt and rotation angles in implicit membrane

In umbrella sampling simulations, increasing the number of reaction coordinates will intensify the computational cost drastically, especially in all-atom systems. To calculate the orientation free energy profile (τ and ρ) of PG-1 with both efficiency and acceptable level of accuracy, we employed the EEF1/IMM1 implicit membrane model [75, 76] in which the membrane is simplified to a continuum low-dielectric slab for the effective solvation energy calculation. Before calculating the two-dimensional free energy profile, however, the hydrophobic thickness of the implicit membrane must be carefully adjusted to well represent the characteristics of the explicit lipid bilayers (i.e., DLPC in the present study), including the local thinning effect upon the PG-1 insertion. For this purpose, we performed four sets (ρ = -90º, 0º, 90º, and 180º) of umbrella sampling simulations with different membrane hydrophobic thickness in the implicit membrane model. For each hydrophobic thickness, the PMF as a function of τ at a specified ρ was calculated and compared with the corresponding PMF from the explicit membrane simulations. As shown in Table 2.4, the implicit-membrane PMFs with a hydrophobic thickness of 10 Å best reproduce the explicit-membrane PMFs despite a slightly large deviation of 3.26 kcal/mol at ρ = 90º. Such large discrepancy at ρ = 90º, however, can be negligible, because the prohibitive free energy of PG-1 at those rotation angles would result in low probability in finding such a configuration in a natural system. One can also compare the physicochemical properties of PG-1 obtained from the implicit membrane simulations with those from the all-atom simulations in DLPC bilayers. The H-bonding fractions and Val16 $^{15}N/^{13}CO$ chemical shifts were computed for the assessment. According to the

good agreement between the properties calculated from both implicit and explicit membrane simulations (Figures 2.12 and 2.13), the optimized hydrophobic thickness of 10 Å was used for the two-dimensional PMF calculations as a function of PG-1 τ and ρ in the implicit membrane.

Table 2.4. The minimum mean-square deviation ($\delta_{E,\min}$) between the PMFs calculated from the implicit membrane simulations with various hydrophobic thickness and the all-atom simulations.

L_H(Å) / ρ (°)	9.5	10.0	10.5	11.0	11.5	12.0
-90°	0.91	0.96	1.36	1.52	1.44	1.95
0°	9.76	0.91	1.10	0.65	0.97	2.10
90°	3.11	3.26	3.17	3.02	3.06	2.82
180°	0.75	0.42	0.49	0.51	0.77	0.96

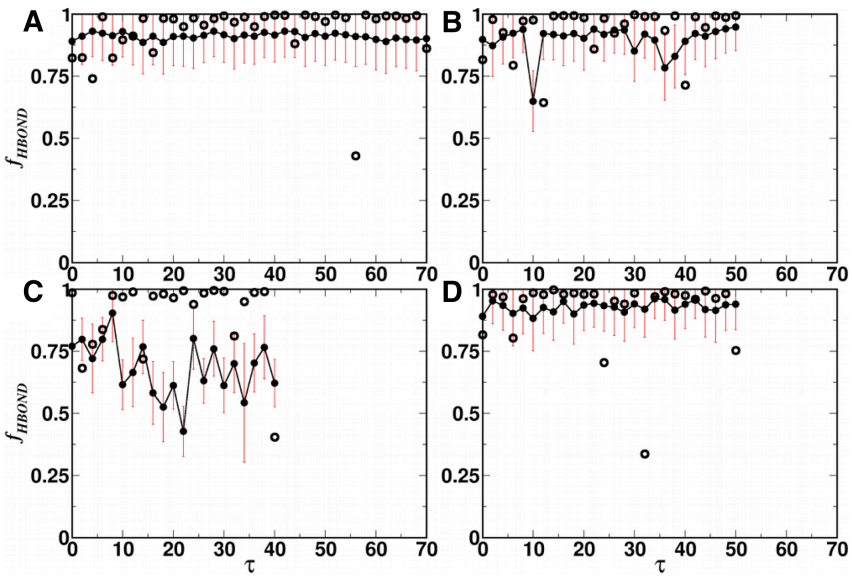


Figure 2.12. The average H-bonding frequencies (solid black) at each τ calculated in the implicit membrane with a hydrophobic thickness of 10 Å. The rotation angle restraints are ρ = (A) -90°, (B) 0°, (C) 90°, and (D) 180°. For the purpose of comparison, the H-bonding frequency at each τ calculated from the explicit membrane simulations (empty

black) are also shown.

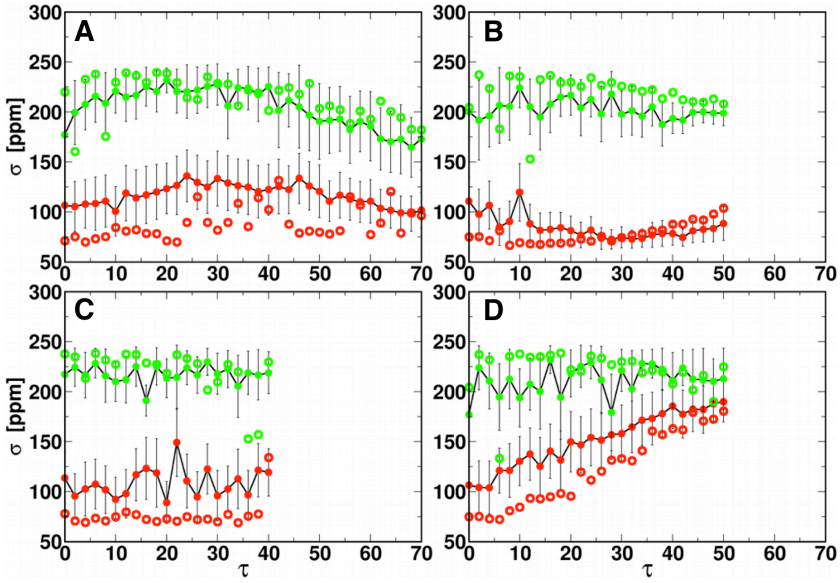


Figure 2.13. The average chemical shifts of Val16 ^{15}N (solid red) and ^{13}CO (solid green) at each τ calculated in the implicit membrane with a hydrophobic thickness of 10 Å. The rotation angle restraints are ρ = (A) -90º, (B) 0º, (C) 90º, and (D) 180º. For the purpose of comparison, the average chemical shifts of Val16 ^{15}N (empty red) and ^{13}CO (empty green) at each τ calculated from the explicit membrane simulations are also shown.

The resulting two-dimensional PMF in PG-1 τ and ρ space is shown in Figure 2.14. The free energy minimum is found at (29º, -93º), which corresponds fairly well to the minimum (37º, -122.5º) obtained from the simulations in the explicit DLPC bilayers. Although there is an 8º difference in τ, which arises possibly from the estimation of solvation energy in the implicit membrane model [75, 76], the thermally-accessible τ and ρ ranges (22º ≤ τ ≤ 32º, -122º ≤ ρ ≤ -75º) is similar to those (20.6º ≤ τ ≤ 40.1º, ρ ≈ -122.5º) from the all-atom simulations in the DLPC bilayers (Figure 2.3A). The thermally-accessible orientations are the ones that have free energies within 0.6 kcal/mol from the

minimum free energy (i.e., the first basin in the free energy contour). Following Equation 2-1, we calculated the averaged chemical shifts of Val16 ^{15}N and ^{13}CO of PG-1 at each pair of τ and ρ. The resulting chemical shifts are compared with the experiment observations in terms of the root mean-square deviation δ_σ (Equation 2-2). The orientations that best satisfy the experiment values ($\delta_\sigma \leq 4$ ppm) are then mapped as τ/ρ pairs on the free energy surface in Figure 2.14. It is clear that the free energies at these orientations are relatively low, yet not necessarily the lowest. The orientations are mostly scattered into three regions centered at (20º, -175º), (35º, -113º), and (52º, -113º) with the last one roughly corresponding to an experimentally-suggested orientation based on a rigid-body geometrical best-fit [66]. Interestingly, the calculated chemical shifts even in these orientational disparities are similar to the experiment measurements. Such correspondence appears to arise from the weak constraints of only two solid-state NMR observables as well as the dynamic nature of the peptide in a given orientation. This result indicates the difficulties in determining an orientation of a flexible peptide with limited experimental observables. When only few experimental observables are available, the calculated values from a single static structure may not reflect experimental measurement correctly. Instead, an ensemble of structures with proper weightings should be used in computing the observables.

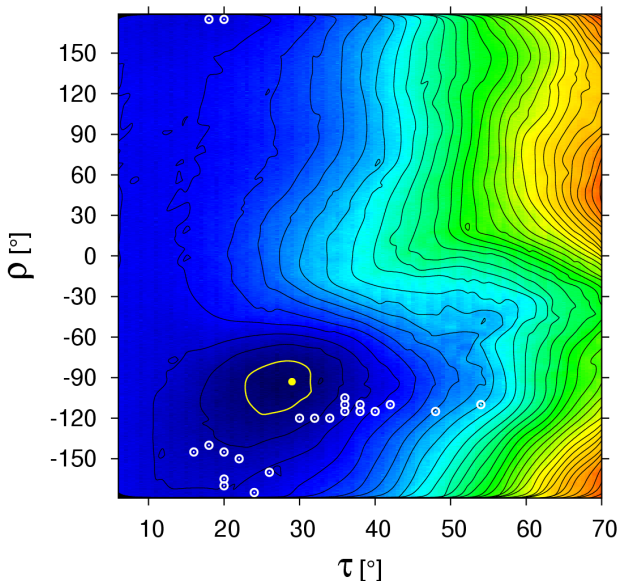


Figure 2.14. The two-dimensional PMF as a function of PG-1 tilt and rotation angles calculated in the implicit membrane with a hydrophobic thickness of 10 Å. The blue to red color change denotes an increase of 16.71 kcal/mol. The orientations with the minimum free energy (solid yellow) and chemical shift root mean-square deviation (δ_σ) < 4 ppm (empty white) are mapped onto the PMF surface. The first blue basin outlined in yellow lines shows the thermally- accessible orientation of PG-1. The contour lines are drawn every 0.6 kcal/mol from blue to red.

The weighting factors can be obtained from the two-dimensional PMF as a function of PG-1 τ and ρ. The ensemble-averaged properties like chemical shifts can then be calculated as a weighted sum from various free energy states because the experimental observations of such properties are the summation of the contributions from all the configurational states that the system has visited. In the current study, the configuration states of interest are τ and ρ of PG-1. Given the probability $P(\tau,\rho)$ of having a certain configuration at (τ,ρ), a physicochemical property $\langle X_{\text{obs}}\rangle_{\text{ens}}$ can be calculated by

$$\langle X_{\text{obs}}\rangle_{\text{ens}} = \int \langle X_{\text{obs}}\rangle_{\tau,\rho} P(\tau,\rho)\, d\tau d\rho \qquad (2\text{-}3)$$

where $X(\tau,\rho)$ is the instantaneous value of such property at (τ,ρ). In an equilibrium system with temperature T, the probability $P(\tau,\rho)$ is given by [98]

$$P(\tau',\rho') = \frac{e^{-W(\tau,\rho)/k_B T}}{\int_{\tau,\rho} d\tau\, d\rho\, e^{-W(\tau'\rho)/k_B T}} \tag{2-4}$$

The ensemble-averaged chemical shifts based on Equation 2-3 correspond moderately to the experiment observations. The results are 118.8 ppm (^{15}N) and 215.1 ppm (^{13}CO) with deviations of 24.2 ppm and 0.9 ppm from the experimental measurements [66].

The relatively large deviation of the ^{15}N chemical shift from the experimental value is likely due to the rigid tensor approximation of ^{15}N, which is known to be dependent on various factors, such as the residue type, the neighboring residues, the secondary structure, and the local interactions [99]. To illustrate the sensitivity of the calculated ^{15}N chemical shift to the choice of the tensors, we computed the ensemble-averaged ^{15}N chemical shift based on different tensors in the literature and the results are summarized in Table 2.5. Depending on different ^{15}N tensors, the calculated chemical shift varies from 118.8 ppm to 130.6 ppm. The deviation from the experimental measurements is reduced from 24.2 ppm to 12.4 ppm, when the ^{15}N tensors measured in N-acetyl-^{15}N-Val were used. Considering the dynamic aspect of the ^{15}N tensors and its impact on the calculated chemical shifts, one has to be careful in choosing an appropriate chemical tensor. To calculate the ensemble-averaged chemical shifts accurately, multiple tensor definitions need to be considered also. Since the chemical shifts are weighted with the probabilities of all the states (based on Equation 2-3) and averaged over the whole configurational space, rather than simply calculated from a singe rigid structure, they can

be directly compared with the experimental observables. The structural dynamics, which is important in calculating the thermodynamic properties like the chemical shifts, is included in the weighted-average. Therefore, the structural ensemble containing all the structures from the thermally-accessible orientations would best represent a system with given experimental observables. The ensemble structures with the most probable orientations offer a dynamic perspective in structure/orientation determination, which would otherwise be difficult to illustrate when a single structure with an arbitrarily chosen conformation is used.

Table 2.5. Alternative ^{15}N chemical shift tensor definitions and calculated ensemble-averaged ^{15}N chemical shift.

	σ_{11}*	σ_{22}	σ_{33}	$\beta^{\dagger}$	$\alpha^{\#}$	$\sigma_{N,calc}$ [§]
Ala-^{15}N-Leu[‡]	64	77	217	17º	25º	118.8
N-acetyl-^{15}N-Val[¥]	59.6	80.5	235.3	21º ± 2º	20º ± 15º	130.5
	59.6	80.5	235.3	21º ± 2º	5º	130.6
	59.6	80.5	235.3	21º ± 2º	35º	129.3
N-acetyl-^{15}N-Val-Leu[¥]	60.2	87.1	230.1	20º ± 2º	34º ± 12º	130.5
N-acetyl-^{15}N-Val-Leu[¥]	60.2	87.1	230.1	19º ± 2º	44º ± 13º	128.9
Boc-Gly-^{15}N-ValGlyAla-OPac[≠]	61.8	79.8	218.8	10.5º	0º	119.8°

*All the magnitude of tensors and chemical shifts are in ppm.

[†]The angle (β) between σ_{33} and the N-H bond vector.

[#]The angle (α) between σ_{11} and the projection of N-H bond vector on the plane made by σ_{11} and σ_{22}.

[§]The ensemble-averaged chemical shift (in ppm) calculated based on Equation 2-4. The experimental value is 143 ± 2 ppm.

[‡]From reference [100].

[¥]From reference [101].

[≠]From reference [102].

2.4. Summary and Conclusions

The PMFs as a function of PG-1 tilt angle in DLPC and POPC bilayers show that the most favorable orientations are at (37.0°, ~90.0°) and (20.9°, ~90.0°) respectively. The PMF-minimum tilt angle (τ_{min}) is dependent on both the PG-1 rotation and the hydrophobic thickness of the membrane environment (Figure 2.3 and Table 2.2). The ρ dependency stems from the anisotropic hydrophobicity of PG-1, which can be represented as a rectangular box with four surrounding faces of different hydrophobicity scales (Figure 2.5). With a more hydrophobic side facing down, PG-1 exhibits an increased τ_{min} and vice versa. Another determinant of τ_{min} is the hydrophobic mismatch [95] between the hydrophobic regions of the membrane and the peptide. In a membrane bilayer with smaller hydrophobic thickness (e.g., the DLPC bilayer), PG-1 tilts more to minimize the energy penalty associated with exposing the hydrophobic sidechains to the aqueous environment. The dependence of PG-1 orientation on the lipid types may as well result in the membrane specific PG-1 polymerization [70]. PG-1 prefers to stay in a small tilt angle orientation in membranes with greater hydrophobic thicknesses (e.g., POPC bilayers). Such orientation is likely to favor the interface formation of PG-1 dimers and oligomers. Unexpectedly, the variation in surface tension shows little impact on the PG-1 tilting PMF (Figure 2.3C). While smaller surface tension thickens the bulk membrane, the thickness of the lipid bilayer adjacent to the peptide turns out to be similar in all three cases (i.e., γ = 20, 10, and 0 dyne/cm) (Figures 2.8 and 2.9). This result suggests that the peptide-lipid interface can be maintained when the hydrophobic thickness of the bulk lipid region has a modest change. Excessive alteration in the bulk lipid bilayer thickness

(e.g., from DLPC to POPC), however, results in the change of the thermally favorable peptide orientation to minimize the hydrophobic mismatch (Figure 2.3C, 2.8, and 2.9).

PG-1 disrupts the membranes by inducing local membrane thinning of the bilayer upon its insertion (Figure 2.7). The magnitude of the thinning effect is around 2 Å in DLPC bilayers and 5 Å in POPC bilayers. Two mechanisms are used by PG-1 to reduce the hydrophobic thickness of the surrounding membrane. They are the nonspecific hydrophobic interactions and the specific electrostatic interactions between PG-1 and lipid molecules. The primary contribution of the former is the hydrophobic interactions between the peptide and lipid bilayers while the latter is mainly the electrostatic interactions between the arginine guanidinium groups of PG-1 and the lipid phosphate groups. The correlation analysis between the *Z* coordinate of each guanidinium group and the instantaneous membrane thinning shows that the membrane thinning around PG-1 arises mostly from the interactions between the lipid phosphate groups and Arg4, 11 and 18 (Figures 2.10 and 2.11, and Table 2.3). In anionic lipid bilayers representing bacteria outer membranes, we speculate that such electrostatic interactions may take the major role in disturbing the membranes around PG-1.

The solid-state NMR chemical shifts has been used to characterize the orientation of PG-1 in membrane environments [66] via a rigid-body orientational search. As shown in the PMFs (Figures 2.3 and 2.14), however, PG-1 has a wide range of thermally-accessible orientations, suggesting that one has to take such conformational/configurational dynamics into account in the calculation of the chemical shift values. The importance of the dynamic feature is shown by using two different approaches in calculating the chemical shifts of Val16 $^{15}N/^{13}CO$. First, the chemical shifts were calculated as a time-

average along PG-1 tilt angle using explicit membrane umbrella sampling trajectories. While the PG-1 structures in the thermally-accessible region produce the chemical shifts similar to the experimental values, the deviations from the experimental measurements were considerably large (Figure 2.6). However, when we considered the entire orientational space (tilt and rotation angles) and the PMF-weighted ensemble average (Equation 2-3 and Figure 2.14), the calculated chemical shifts are 118.8 ppm (^{15}N) and 215.1 ppm (^{13}CO) with deviations of 24.2 ppm and 0.9 ppm from the experimental measurements [66]. We further illustrate that the large discrepancy in ^{15}N chemical shift can be reduced with different sets of ^{15}N chemical shift tensor definitions (Table 2.5), demonstrating the importance of the choice in chemical shift tensor definitions for the chemical shift calculations in a membrane system. To increase the accuracy, one may also need to consider multiple tensor definitions in the chemical shift calculations. By incorporating the ensemble of PG-1 structures with thermally-accessible conformations (based on Equation 2-3) and proper chemical shift tensors, the present study highlights the importance of structural dynamics in protein/peptide conformation/configuration determinations.

Chapter III

Probing Caveolin-1 Structure and Dynamics in Membrane Bilayers Using Molecular Dynamics Simulations

Summary

Caveolin induces membrane curvature and drives the formation of caveolae that participate in many crucial cell functions such as endocytosis. Caveolin contains two membrane-embedded helices (H1 and H2) connected by a three-residue linker with both N- and C-tfermini exposed to the cytoplasm. Although a U-shape configuration is assumed based on its inaccessibility from the extracellular matrix, caveolin structure in a bilayer remains elusive. This work aims to characterize the structure and dynamics of caveolin-1 (D82 to S136; $Cav1_{82-136}$) in a DMPC bilayer using NMR, fluorescence emission measurements, and molecular dynamics (MD) simulations. The topology of $Cav1_{82-136}$ from NMR chemical shift indexing analysis serves as guideline for generating its structural models. 50 independent MD simulations (80 ns each) are performed to identify its favorable conformation and orientation in the bilayer. A system with the preferred configuration is then chosen and simulated for 1 μs to further explore its stability and dynamics. The results of these simulations mirror those from the tryptophan fluorescence measurements (i.e., its insertion depth in the bilayer) and corroborate that $Cav1_{82-136}$ inserts in the membrane with U-shape conformations. The preferred crossing angle range between H1 and H2 is from 45° to 64° and the insertion angle is ~70°. The majority of the conformations have R101 and Y118 at or near the H1-H2 interface, which are responsible for causing membrane deformation near $Cav1_{82-136}$.

3.1 Introduction

Caveolae are plasma membrane invaginations rich in sphingomyelin and cholesterol. They are usually 500-1000 Å in diameter and bulge inwards to the cytoplasm [103-105]. Many mammalian cells contain a large number of caveolae. These cells are usually highly differentiated. Examples include cardiac myocytes, endothelial cells, fibroblasts, macrophages, and smooth muscle cells [106, 107]. Besides their high occurrence in certain cells, caveolae are also considered multifunctional organelles that are of biological importance. They participate in endocytosis, signal transduction, membrane trafficking and protection, and many other vital cellular processes [105, 108-111]. Misregulation and dysfunction of caveolae have been linked to numerous human diseases such as infection, muscular dystrophy, cardiac disease, Alzheimer's disease, and cancer [112-115].

One characteristic protein in caveolae is caveolin, which is necessary for caveolae formation [116, 117]. There are three caveolin isoforms, caveolin-1, 2, and 3, among which caveolin-1 is the most ubiquitous [107]. Caveolin-1 is a membrane protein that adopts an unusual topology with both N- and C-termini exposed to the cytoplasm [118, 119]. The N-terminal region (residues 1-101) is the least conserved and varies among species. The last 20 residues (residues 82-101, Figure 3.1) of this region are termed the caveolin scaffold domain (CSD), which has been implicated in caveolin oligomerization, membrane binding, cholesterol binding, and interactions with other proteins. The C-terminal domain contains 44 amino acids in all mammalian caveolins and its functions are thought to include membrane attachment and protein interactions [120, 121]. The N- and C-terminal regions are bridged by the hydrophobic transmembrane domain (TMD,

residues 102-134, Figure 3.1), which has been proposed to adopt and intra-membrane loop structure [119]. The U-shaped TMD inserts in the membrane and is thought to be responsible for inducing membrane curvature and the formation of caveolae [122].

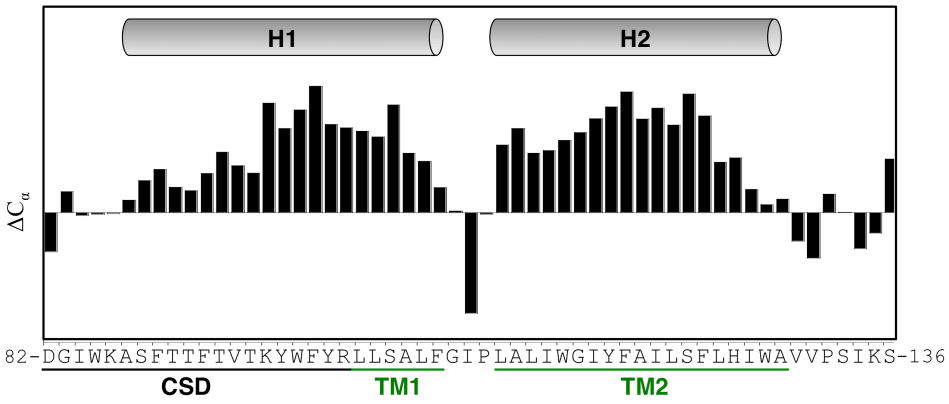


Figure 3.1. Chemical shift index plot of $Cav1_{82\text{-}136}$. Helix 1 (H1) spans residues 87-107 and Helix 2 (H2) spans residues 111-129. The putative turn between H1 and H2 includes residues 108-110. The regions of the CSD domain, TM1, and TM2 are indicated on the sequence of $Cav1_{82\text{-}136}$.

Efforts have been devoted to uncovering the structures of caveolins and how they trigger caveolae formation. The TMD is thought to adopt a U-shape re-entrant helical structure [123]. Such a TMD topology is supported by data from NMR and circular dichroism spectroscopy measurements revealing a helix-break-helix (TM1-turn-TM2) motif [124]. Unlike the TMD, the secondary structure of the CSD is still under debate. While some studies suggest a β-strand [125, 126], others show evidence of an amphipathic α-helix [122, 127]. The unique topology and structure of caveolin-1 is likely essential for its mechanism of action to drive caveolae formation [123, 128]. Despite considerable efforts, however, the atomic structure of caveolin-1 remains unknown and an understanding of how caveolin induces caveolae formation at the molecular level is far

from complete. There are still many questions that need to be addressed. In particular, what is the membrane bound structure of caveolin-1 and what is the binding mode, i.e., how does it orient with respect to the membrane bilayer? In addition, are there any specific caveolin-lipid interactions and what kind of indications on caveolin oligomerization and caveolae formation can be made from these local interactions?

This study aims to address the above questions using molecular dynamics (MD) simulations guided and supported by NMR and tryptophan fluorescence emission data. At first, a series of NMR experiments are carried out for a caveolin-1 construct containing residues 82 to 136 ($Cav1_{82-136}$). The secondary structure of $Cav1_{82-136}$ is then revealed by the chemical shift indexing analysis. With this information on the secondary structure, 50 different initial $cav1_{82-136}$ model structures are generated and subjected to MD simulations (80 ns each) with the goal of finding the conformational and orientational preference of caveolin-1 in the membrane. In addition, a longer simulation (1 μs) is conducted on the system that contains a favorable $cav1_{82-136}$ configuration to further explore its stability and dynamics as well as its interactions with the lipid bilayers. The simulation results show that $cav1_{82-136}$ is embedded in the bilayer and its U-shape conformation is maintained. The results also reveal a preference for residues R101 and Y118 to be close to or at the H1-H2 (Figure 3.1) interface. It is likely that the interactions between these residues and the surrounding lipid molecules help to maintain the U-shape conformation. The inserted caveolin-1 orientation is largely attributed to the anchorage of H1 at the top and bottom leaflets. Residues in both ends of H1 can interact with lipid headgroups and these interactions are also responsible for causing membrane deformation.

3.2. Computational Methods

3.2.1. Preparation of caveolin-1 constructs

Wild type caveolin-1 (residues 82-136; $Cav1_{82-136}$) was purified and expressed according to the protocol established by Diefenderfer *et al*. [129]. Full-length caveolin-1 contains four tryptophan residues (W85, W98, W115, and W128). Four single tryptophan mutant constructs of caveolin-1 residues 62-178 ($Cav1_{62-178}$) were prepared using the Agilent QuikChange site-directed mutagenesis kit (Santa Clara, CA). In each mutant, one of the four native tryptophan residues was retained and the other three were mutated to phenylalanine. These constructs were expressed and purified according to the aforementioned protocol. For compatibility with cyanogen bromide cleavage, M111, which is not strictly conserved, was mutated to leucine based on sequence homology to caveolin-2 and caveolin-3. Also, C133 in $Cav1_{82-136}$ and C133, C143, and C156 for $Cav1_{62-178}$, which are sites of palmitoylation, were mutated to serines. These mutations were made to avoid unwanted and biologically irrelevant disulfide bonding. This is considered to be a mild mutation as palmitoylation was found to be unnecessary for the correct trafficking of caveolin-1 to the plasma membrane.

Expression of uniformly ^{2}H, ^{15}N, and ^{13}C labeled $Cav1_{82-136}$ was done according to the procedure of Marley *et al*. with slight modification in that the growth was harvested 8 hours after induction [130]. Specific amino acid labeling was performed as described by Truhlar et al. with slight modification [131]. The cell growth was done on a 500 mL scale using M9 minimal media supplemented with 50 mg of the desired ^{15}N amino acid and 250 mg each of the other 19 ^{14}N amino acids. Expression was induced at

an optical density (600 nm) of 0.6 using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and grown for 6–8 hours at 310K.

3.2.2. Two- and Three-dimensional NMR spectroscopy

Lyophilized WT $Cav1_{82-136}$ or single tryptophan mutants of $Cav1_{82-136}$ were reconstituted into buffer (100 mM LMPG; 100 mM NaCl; 20 mM phosphate pH 7; 10% D_2O) to a concentration of approximately 1.0 mM. Vigorous vortexing of this mixture resulted in a clear homogeneous solution. Next, the sample was passed through a 0.2 μm regenerated cellulose spin filter.

All NMR spectra were acquired at 310K using a 600 MHz Bruker Avance II spectrometer (Billerica, MA) equipped with a cryoprobe. For analysis and backbone assignments, the following TROSY-based pulse sequences were employed: HSQC, HNCA, HNCACB, and HN(CO)CA. The spectra were processed using NMRPipe and Sparky [132, 133]. For chemical shift indexing, observed Cα chemical shifts were subtracted from random coil chemical shifts as described by Wishart *et al*. [134]. To aid backbone assignments, specific amino acid labeling was employed (Gly, Phe, Tyr, Leu, Ile, and Val).

3.2.3. Tryptophan fluorescence measurements

0.288 mg of a single tryptophan construct and 26.57 mg of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) were co-dissolved into 720 μL of 1,1,1,3,3,3-hexafluoroisopropanol. Next, 480 μL of deionized water was added. The solution was

vortexed, snap frozen in liquid nitrogen, and lyophilized for 24 hours. The protein-DMPC mixture was then taken up into 2.962 mL of 21 mM phosphate pH 7.0 with three minutes of vortex mixing. Next, 138 μL of 25% (w/w) 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) was added to the lipid-protein suspension. Three minutes of vortexing resulted in a clear solution. Bubbles were removed by centrifugation at 17,000 *x g*. The final protein concentration was approximately 5 μM.

Fluorescence emission spectra were acquired at 310 K using an Agilent Eclipse fluorometer (Santa Clara, CA). A standard 1 cm × 1 cm quartz cuvette was used. The excitation wavelength used was 295 nm to avoid unwanted tyrosine excitation [135]. Both the excitation and emission slit widths were set to 5 nm. The emission spectra were measured from 305-430 nm with a scan speed of 4 nm/s and 0.5 nm data point increments. The emission spectra of three separately prepared samples were recorded, and the λ_{max} values were averaged. To obtain λ_{max} values, as well as intensity values, the spectra were fit to a log-normal distribution using the software Igor Pro 6.22A (Portland, OR) [136].

3.2.4. Multiple caveolin-1 simulations in DMPC bilayers

An *in silico* $Cav1_{82-136}$ model was generated by the IC BUILD command in CHARMM [45] with its sequence shown in Figure 3.1. Based on the chemical shift indexing analysis (Figure 3.1), Residues A87-F107 (H1) and L111-A129 (H2) were modeled as ideal α-helices with their ϕ and ψ angles set to -57.8° and -47°, respectively; the other amino acid residues were modeled with their ϕ and ψ angles in the CHARMM residue topology file.

By randomly changing the ϕ and ψ angles of G108 in the linker (G108-P110), initial structures of caveolin-1 with different H1-H2 crossing angles (θ) were obtained. These structures were then placed into five categories ($\theta_{initial}$ = 45°, 55°, 65°, 75°, and 85°), each including 10 replicas, all of which have θ within 5° to $\theta_{initial}$. Before inserting each initial caveolin-1 model into membrane bilayers, it was reoriented so that its axis, defined by the vector sum of the principal axes of H1 and H2, coincided with the *Z*-axis, the membrane normal. Since no prior knowledge on the insertion depth of caveolin-1 was available, the model in each replica system was shifted along the *Z*-axis to randomly place the linker residues' centers of mass (COM) between Z = -5 Å and 5 Å.

These initial models with different configurations and orientations were embedded in DMPC bilayers and solvated by 0.15 M KCl solutions. This resulted in systems with approximately 52,000 atoms and 75 × 75 × 90 $Å^3$ in size. Each system was individually constructed using the *Membrane Builder* module [79, 137] in CHARMM-GUI [78] with the center of the membrane at the system origin and the membrane normal parallel to the *Z*-axis. The name of each system was given in the form of "cav1_P_Q" with P and Q indicating the value of $\theta_{initial}$ and the replica index. For example, "cav1_45_5" points to the fifth replica with $\theta_{initial}$ = 45°. Following the assembly of each system was a brief equilibration of 225 ps mainly to relax the initially uncorrelated system components. An 80-ns molecular dynamics simulation was performed afterwards. In the first 50 ns of each simulation, dihedral restraints were applied to H1 and H2 residues to maintain their secondary structure according to the chemical shift index results (Figure 3.1). After 50 ns, these restraints were removed and each system was simulated for another 30 ns. The total simulation time for all these systems was 4.0 μs.

The first 20 ns were designated to allow each system to reach equilibrium and therefore were not used in the trajectory analyses.

All calculations were performed in the NPT (constant particle number, pressure, and temperature) ensemble [138] at 310 K using CHARMM [45]. The CHARMM all-atom force field [84] with a modified version of dihedral cross-term correction [85] was used for the protein and the C36 lipid force field [139] was used for DMPC. The TIP3 water model [140] was employed for water molecules. A time step of 2 fs was enabled with the SHAKE algorithm [50]. Because of the U-shape conformation of caveolin-1 and the fact that it does not span the entire lipid bilayer, the protein cross sectional area in one lipid leaflet is larger than that in the opposite leaflet. Therefore, the $P2_1$ image transformation [81] was applied to allow the variation in the number of lipids at the top and bottom leaflets during the simulation. The nonbonded and dynamics options were kept the same as in the *Membrane Builder* input; the van der Waals interactions were smoothly switched off at 10–12 Å by a force-switching function [141] and the electrostatic interactions were calculated using the particle-mesh Ewald method [142] with a mesh size of ~1 Å for fast Fourier transformation, $\kappa = 0.34$ Å^{-1}, and a sixth-order B-spline interpolation.

3.2.5. 1-μs Anton simulation starting from a favorable caveolin-1 configuration

A snapshot of system cav1_65_3 at 35 ns was taken and used to initiate a 1-μs simulation. This snapshot was chosen because it contains a representative caveolin-1 configuration based on the structural analyses of the multiple MD simulations. The simulation was carried out on Anton [143], which is a special-purpose supercomputer

designed for long time scale MD simulations. The isothermal (NVT) ensemble was employed with the Nose-Hoover temperature coupling scheme [144]. The temperature (310 K) was set to the same as those used in the multiple 80-ns MD simulations described above. The lengths of all bonds involving hydrogen atoms were constrained using M-SHAKE [145]. The cutoff of the van der Waals and short-range electrostatic interactions was set to 10.06 Å. Long-range electrostatic interactions were evaluated with the k-space Gaussian Split Ewald method [146] and a 64 × 64 × 64 mesh. The simulation time step was 2 fs. The r-RESPA integration method [147] was employed and long-range electrostatics were evaluated every 6 fs.

3.2.6. Defining caveolin-1 structure and orientation in a membrane

Quantitative characterization of caveolin-1 structure in a membrane requires its internal configuration and its orientation with regard to the bilayer. The internal configuration of caveolin-1 can be defined by four parameters: the crossing angle (θ) between the H1 and H2 helical axes, the rotation angles of H1 (ρ_1) and H2 (ρ_2), and the pair of residues that are in close contact (Res_{ij}) (Figure 3.2A). For each helix, the angle between the projections of two vectors on the plane perpendicular to its helical axis is used to define the rotation. The two vectors include (i) the reference vector connecting the centers of the two helices and (ii) the vector pointing from the center of the helix to a reference atom. In H1, this atom is the Cα of T91 and in H2 it is the Cα of W115.

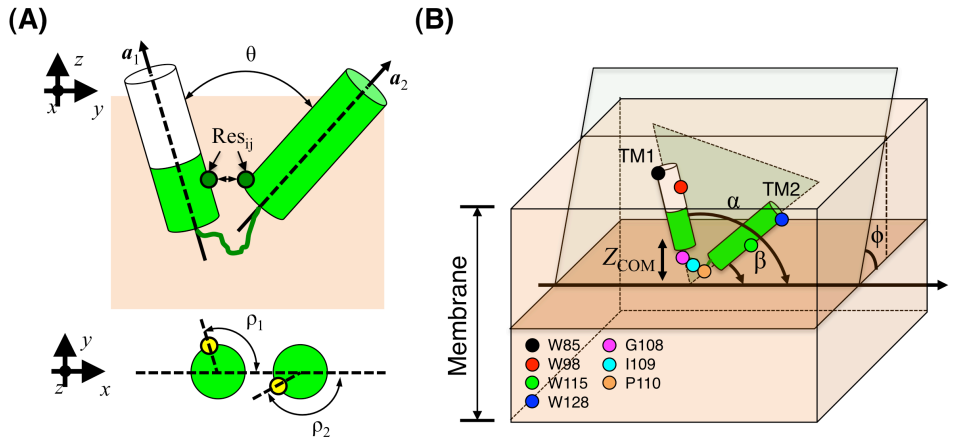


Figure 3.2. (A) Degrees of freedom used to define the internal conformation of caveolin-1. They include θ, ρ_1, ρ_2, and the contacting residues pairs (Res_{ij}). The yellow spheres indicate the reference atoms (Cα of T91 on H1 and Cα of W115 on H2) employed in ρ_1 and ρ_2 calculation. (B) Degrees of freedom (ϕ, α, β, and Z) describing the overall orientation of caveolin-1 with respect to a membrane bilayer. The membrane is centered at Z = 0 Å. Residues of interest are shown in colored spheres. The CSD domain (D82-R101) is colored in white.

To describe caveolin-1 orientation with reference to the membrane bilayer, four additional variables are needed. These variables are the insertion angle (ϕ), which is the angle between the molecular plane of caveolin-1 and the membrane, the tilt angles of H1 (α) and H2 (β) on the caveolin molecular plane, and the insertion depth (Z_{COM}), i.e., the Z-position of the linker residues' COM (Figure 3.2B). The caveolin molecular plane is the least-squares plane through all the Cα atoms defining the H1 and H2 helices. It is worth pointing out that both θ and the tilt angle pair α and β are necessary for an unambiguous description of caveolin-1 configuration. The values of α and β are indicative of θ in some cases, especially when both H1 and H2 lie in the molecular plane of caveolin-1. However, caveolin configurations with the helices deviating from the molecular plane do exist during the simulations. In this case, α and β alone do not provide enough information on caveolin configuration/orientation.

3.3. Results and Discussion

3.3.1. NMR peak assignments and chemical shift index data

The chemical shifts for $Cav1_{82-136}$ are obtained from its assigned ^{1}H-^{15}N-TROSY spectrum. They are well dispersed, indicating that $Cav1_{82-136}$ is structured. To assign the backbone resonances, a combination of HSQC, HNCA, HNCACB, HNCO, and HN(CO)CA experiments in conjunction with selective amino acid labeling were employed. Using these methods 96% percent of the residues were assigned and the chemical shifts of all of the Cα were obtained.

The core hydrophobic stretch of caveolin is from D82 to S136 and its chemical shift indexing is shown in Figure 3.1. Residues 87-107 (H1) and 111-129 (H2) have consistently positive ΔCα's, which is indicative of an α-helical structure [134]. The region between the two helices, residues 108-110, is the site of the putative intramembrane turn, which returns the polypeptide chain to the same side of the membrane [124]. Recent studies have shown that the residues in this region are critical for the structure of the caveolin protein [124]. Residues 82-86 and 130-136 do not have consistently positive or negative ΔCα's, indicating that these regions are unstructured or dynamic. From the ΔCα data in Figure 3.1, it is apparent that the values for the first half of H1 (residues A87 to T95) show distinctly lower ΔCα values than the end of the helix. These results indicate that the intramembrane domain of caveolin-1 is composed primarily of two helices that are roughly equal in length (17-21 amino acids each).

This result is in contrast to the model put forth by Parton *et al*., which is based on primary sequence analysis, and predicts three distinct helical regions in the

intramembrane domain (81-92 – break – 97-107 – break – 112-128) [122]. Clearly our experimental data add an important enhancement to this model by showing that residues 87-107 form one long continuous α-helix. Therefore our NMR data lends a significant refinement to the model of caveolin-1 topology.

3.3.2. Structural preference of caveolin-1 revealed by multiple MD simulations

Figure 3.3A shows the crossing angle distributions (θ) for all of the simulations. Clearly, most of the θ values are less than 90°. The average θ in the five simulation sets are 39°±12° (cav1_45), 45°±13° (cav1_55), 56°±15° (cav1_65), 57°±8° (cav1_75), and 71°±27° (cav1_85), respectively. This data shows that the U-shape conformation of caveolin-1 is largely maintained during all of the simulations. Closer inspection reveals that these systems visit similar range of θ values, mostly from 40° to 70° (Figure 3.3A and Table 3.1). Structures with θ below 40° are also accessible in nearly all five of the simulations sets, as shown in cav1_45_(3, 4, 7, and 10), cav1_55_(2, 6, 7, and 10), cav1_65_9, and cav1_85_9. In some of these systems, caveolin-1 forms a compact structure with more profound packing than those observed in other systems with larger θ. An opposite scenario is seen in system cav1_85_5, in which the caveolin-1 molecule opens up (θ = 144°±9°) and moves to the hydrophobic and hydrophilic interface of the top leaflet. Nevertheless, the plateau in the average distribution of θ in all the simulations shown in Figure 3.3A implies that the favorable range for θ is from 45° to 64°.

Table 3.1. Average θ in each of the 50 $Cav1_{82\text{-}136}$ MD simulation systems including the averages and standard errors of the ten replicas in each simulation set.

	45	55	65	75	85
1	46°±6°	48°±14°	65°±6°	69°±7°	73°±7°
2	44°±15°	35°±7°	71°±5°	57°±7°	57 °±7°
3	34°±3°	53°±6°	49°±13°	54°±6°	66°±10°
4	15°±4°	59°±6°	84°±8°	59°±7°	74°±8°
5	44°±5°	49°±10°	46°±6°	41°±7°	144°±9°
6	60°±6°	26°±6°	49°±6°	64°±7°	78°±5°
7	30°±4°	39°±6°	64°±5°	47°±7°	60°±5°
8	44°±4°	60°±8°	42°±6 °	50°±6°	69°±6°
9	44°±8°	56°±6°	29°±4°	64°±7°	37°±6°
10	30°±2°	23°±4°	60°±7°	62°±17°	47°±7°
Avg.± S.E.	39°±12°	45°±13°	56°±15°	57°±8°	71°±27°

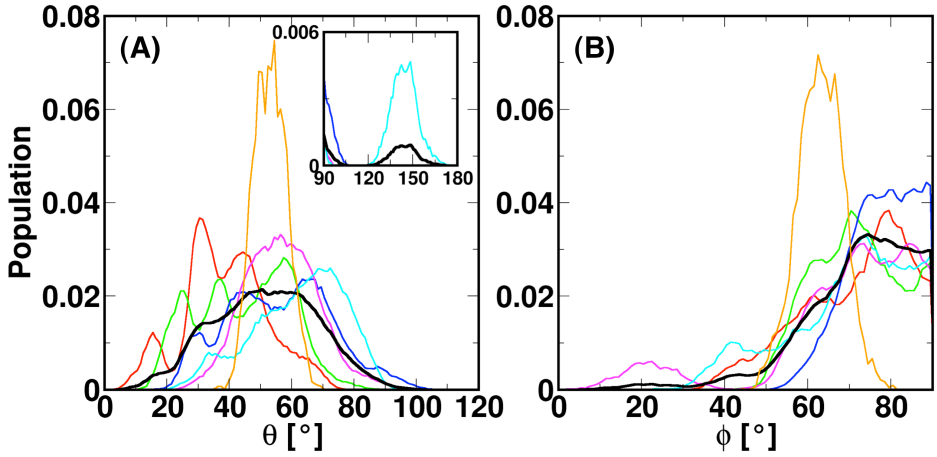


Figure 3.3. The population distributions of (A) θ and (B) ϕ in all the systems. Different sets of simulations are distinguished by colors with cav1_45 in red, cav1_55 in green, cav1_65 in blue, cav1_75 in magenta, cav1_85 in cyan, and the 1-μs simulation results in orange. The averaged distributions from all the 80-ns systems are shown in black. All the population histograms are calculated with a bin size of 1°.

It is also intriguing that θ is very flexible within the 45° to 64° range. The typical standard deviation observed for θ in any given system is 7° (Table 3.1), meaning a variation of θ as large as 14° from structures in one simulation run can be expected. By

examining the time evolutions of θ, fluctuations between the upper and lower bounds of θ are frequently seen and they happen on the order of tens of nanoseconds as exemplified in systems cav1_65_5 and cav1_75_5 (Figure 3.4). The oscillations of θ are likely associated with the way lipid molecules pack around caveolin-1 (Figure 3.5*A*-*C*). Both lipid tails and head groups can fill in the space between H1 and H2, and θ change is often accompanied by a corresponding change in the number of lipid tails (N_{tail}) and/or head groups (N_{head}) in between H1 and H2. It appears that the lipid molecules between H1 and H2 provide temporary structural support for the U-shape conformation of caveolin-1. However, since the lipids are also very dynamic, they can engage in or withdraw from the caveolin-1 structure fairly easily, therefore all the systems share similar pattern of θ fluctuations.

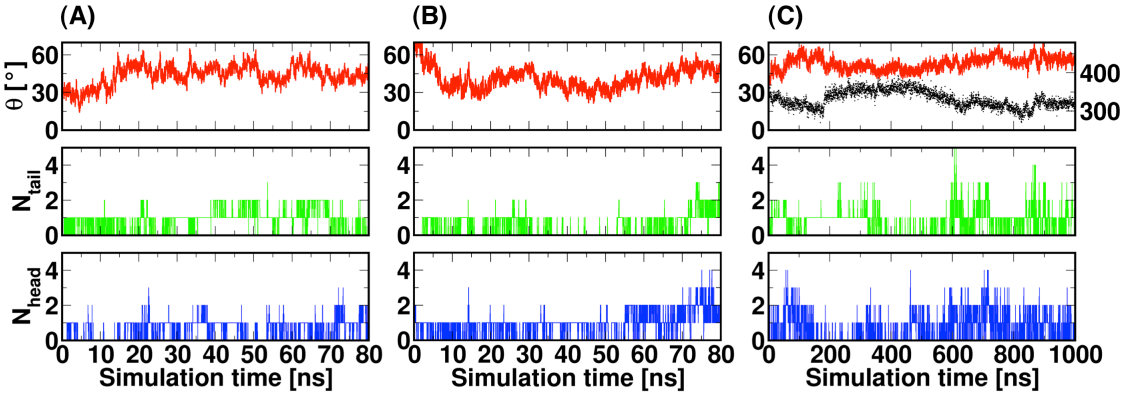


Figure 3.4. Time evolutions of θ (red), N_{tail} (green), and N_{head} (blue) in (A) cav1_65_5, (B) cav1_75_5, and (C) the 1-μs simulation. The top panel in (C) also plots the time evolution of ρ_1 (black) in the 1-μs simulation.

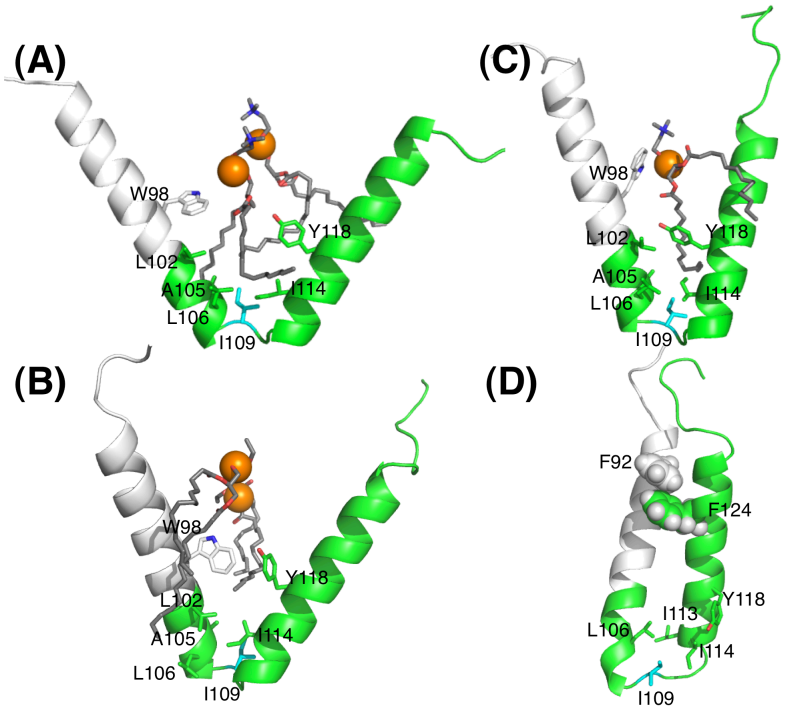


Figure 3.5. The crossing angle (θ) affected by different lipid packing modes in (A) cav1_45_6 (θ = 69°), (B) cav1_65_3 (θ = 57°), (C) cav1_65_3 (θ = 43°), and (D) cav1_45_4 (θ = 33°). Caveolin-1 and surrounding lipids are illustrated in cartoon and licorice representations, respectively. The phosphate atoms on these lipids are shown as orange spheres to emphasize the position of the head groups and the CSD domain is in white. The structures in (A), (B), and (C) have the canonical ρ_1 (343°) and ρ_2 (111°) values with W98, L102, A105, and L106 on H1 and I114 and Y118 on H2 at the interface. These residues are highlighted in stick presentation. Most contacts between the two helices are found in regions close to the linker residues, namely between A105, L106, and I114. Linker residue I109 (cyan) can form the interface in these structures as well. The structure in (D) has its ρ_1 and ρ_2 values at 35° and 329°, respectively. This rotation angle pair is less predominant in the simulations. It has a compact structure and poses aromatic residues such as F92 and F124 (spheres) at the H1-H2 interface.

Even with similar θ values, there is no guarantee that caveolin-1 is structurally equivalent. In fact, they differ by the H1 (ρ_1) and H2 (ρ_2) rotation angles. Figure 3.6 presents the (ρ_1, ρ_2) distributions in the five sets of simulations. The simulations start with entirely different (ρ_1, ρ_2) values but exhibit preferred rotation angles. For ρ_1, they are 343° and 213°, and for ρ_2, they are 111° and 327° with (343°, 111°) being the most

frequently visited values. The value of ρ_1 at 343° translates to a structure where W98, L102, A105, and L106 on H1 are present at the H1-H2 interface. With the slightly less populated ρ_1 value at 213°, the interface on H1 consists of T93, Y97, and S104. Structures with these two ρ_1 angles can also be distinguished by the relative position of R101 sidechain to the caveolin-1 molecular plane. At either of the ρ_1 values, the caveolin-1 structures have R101 sidechains just outside of the H1-H2 interface. Following the definition of the helical axis in Figure 3.2A, the plane normal is $\boldsymbol{n} = \boldsymbol{a}_1 \times \boldsymbol{a}_2$, and the location of R101 sidechain is always on the same side as $\boldsymbol{n}$ when $\rho_1 = 343°$ and the opposite side when $\rho_1 = 213°$ (Figure 3.7A and B). The difference between structures with ρ_2 at 111° and 327° is manifested by the location of the Y118 sidechain. $\rho_2 = 111°$ puts the Y118 sidechain in between the helices, whose sidechain is able to form π-π interactions with either W98 ($\rho_1 = 343°$) or Y97 ($\rho_1 = 213°$) from H1. When present at the interface, the Y118 sidechain can also interact (via H-bonds) with the lipid head groups located in between the two helices. These interactions attract lipid molecules to the space between the two helices, which would in turn help stabilize the caveolin-1 U-shaped configuration. Compared to $\rho_2 \approx 111°$, the population of caveolin-1 with ρ_2 at 327° is less and more spread out around this value (Figure 3.7). The structures with $\rho_2 \approx 327°$ have I113 and F124 sidechains on H2 at the H1-H2 interface. These hydrophobic residues either interact with H1 residues (Figure 3.5D) or with the lipid acyl chains partitioned in between the two helices if θ is large enough.

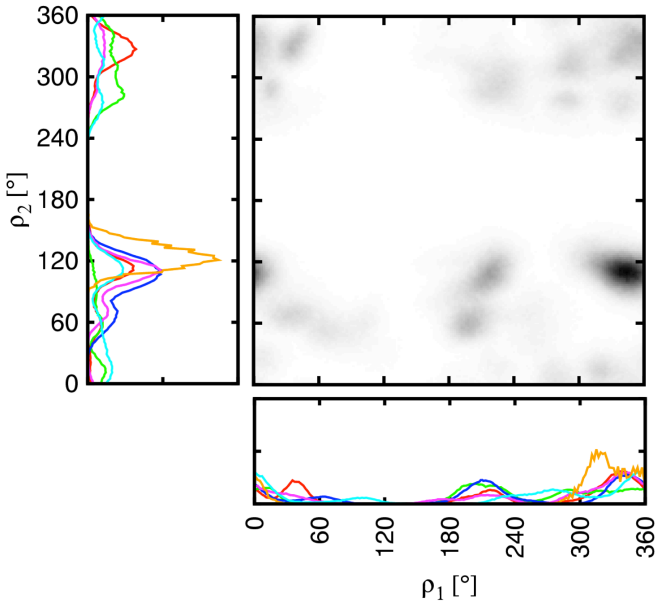


Figure 3.6. The average two-dimensional ρ_1 and ρ_2 distribution in all the multiple simulations. The bottom and left panels show the distributions of ρ_1 and ρ_2 respectively. The distributions from different systems are distinguished by colors with system cav1_45 in red, cav1_55 in green, cav1_65 in blue, cav1_75 in magenta, cav1_85 in cyan, and the 1-μs simulation in orange.

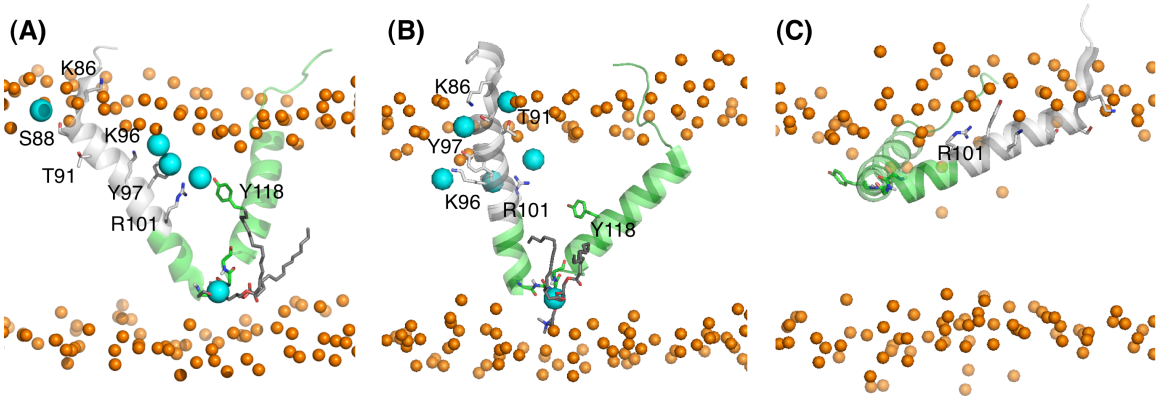


Figure 3.7. (A) and (B) molecular snapshots showing interactions between caveolin-1 and the lipid molecules. H1 is anchored to the membrane by interactions with lipid head groups at both ends of the helix. In the N-terminus of H1, polar and charged residues from the CSD domain are responsible for making interactions with the lipids of the top leaflet. These residues are K86, S88, T91, K96, Y97, and R101, and are highlighted in licorice presentation. Lipid phosphate groups that interact with these residues are shown in cyan spheres. Among these residues, K96, Y97 and R101 as well as Y118 on H2 are responsible for the membrane thinning effect at the top leaflet and the rest are involved in

causing the increase in thickness H1 N-terminus. H-bonds can be formed between the linker backbone atoms and the lipid head groups at the bottom leaflet and this acts to secure the position of H1 C-terminus. (C) A caveolin-1 structure lacking the protein-lipid interactions at the bottom leaflet.

Although all the above rotation angles are accessible in the simulations, the most frequented structures have (343°, 111°). A closer look at these structures (Figure 3.5A-C) reveals that interactions between the ends of the helices adjacent to the linker residues (G108, I109, and P110) are always maintained. These interactions are mediated by residues A105 and L106 on H1 and I114 on H2. Because the linker is short (only three residues), I109 can interact with A105 and I114 via its two β-carbon branches. This interaction is particularly pronounced in caveolin-1 structures having the most populated θ angles of 45°-64°. Such interactions are lost in the tightly packed or extended caveolin-1 structures (see Figure 3.5D). This finding also supports recent NMR studies showing the necessity of having a β-branched hydrophobic residue at this position [124].

Another type of H1-H2 contact observed in the simulations were π-π interactions mediated by aromatic residues. For example, in structures with (343°, 111°), the residues W98 (H1) and Y118 (H2) found in the middle of the helices can form a favorable contact especially if θ is towards the lower end of the preferred θ range from 45° to 64°. Other aromatic residue pairs seen in the simulations include F92-F124, Y97-Y118, F99-F124, W98-F124, and Y100-Y118. They are mostly associated with compact caveolin-1 structures that have θ less than 35° (Figure 3.5D). Interestingly, all of the H1 aromatic residues involved in forming π-π interactions are from the CSD domain. Four of them (Y97 to Y100) span one turn of the helix and only one of the four can be at H1-H2

interface at any given time. Because of the tendency of these aromatic residues to form inter-helical interactions and the fact that there is always more than one of them available outside of the H1-H2 interface, they may participate in the oligomerization process of caveolin-1.

The structure preference observed in the multiple 80-ns simulations is largely maintained in the 1-μs simulation. Both simulations also exhibit degrees and time scales of structural fluctuations that are alike. The 1-μs simulation starts with the snapshot at 35 ns in system cav1_65_3, which has the following structural parameters θ = 42°, ρ_1 = 4°, and ρ_2 = 124°. Such a configuration is chosen because all of its structural parameters are within the preferred range seen in the multiple cav1_ simulations. After the simulation starts, there is an increase of θ within the first 20 ns towards 60° (Figure 3.4C), followed by oscillations between 45° and 69° with an average at 53°±5° (Figure 3.3A and 3.4C). The amplitude and cycle length of the oscillations are quite similar to those observed in the shorter MD simulations, and so does the coupling between θ fluctuation and $N_{tail/head}$. Some degree of coupling between θ and ρ_1 is also seen (Figure 3.4C), because H1 can rotate to accommodate the inserted lipid chains. Even with these fluctuations, the ρ_1 and ρ_2 values still remain close to those observed in in the shorter simulations (ρ_1 = 343° and ρ_2 = 111°). Therefore, the interfacial residues discussed in the shorter simulations (i.e. Y118) are still relevant in the longer simulations.

3.3.3. Overall orientation of caveolin-1 revealed by multiple MD simulations

With the representative structure of caveolin-1 described above, the next step is to characterize its orientation in a membrane bilayer. A majority of the simulations show that the molecular planes of caveolin-1 are nearly perpendicular to the membrane. As illustrated in Figure 3.3B, most caveolin-1 molecules exhibit ϕ values at ~70° and only a few have ϕ values under 40°; a zero ϕ value indicates a caveolin structure lying parallel to the membrane. The average ϕ and its standard deviation for each system are shown in Table 3.2. The only system showing an average ϕ below 40° is cav1_75_6. In this system, caveolin-1 exhibits decreasing ϕ for the first 20 ns and then remains constant for the rest of the simulation at 22° ± 7°. Such a decrease in ϕ is linked to the initial Z_{COM} of the linker residues, which was in the top leaflet and stayed that way during the entire 80-ns simulation without having the chance to interact with bottom leaflet lipid headgroups. In fact, structures in other systems (e.g., cav1_45_3, cav1_85_4, cav1_85_5, and cav1_85_8) that access orientations with $\phi \leq 40°$ often have their linker-lipid interactions in the bottom lipid leaflet lost. It is interesting that caveolin-1 structures with the predominant ϕ values always have their R101 sidechains pointing towards the positive Z-direction even when they have different ρ_1 angles (Figure 3.7A and B). Such an orientation places the R101 sidechain underneath the top leaflet lipid headgroups, following the electrostatic interactions between R101 and lipids as discussed below. In addition to a similar ϕ average, the systems also share considerable ϕ fluctuations in the range between ϕ = 50° and 80°. An estimated time scale of the fluctuations is ~50 ns, meaning that it takes ~50 ns to reach from the lowest ϕ to the highest ϕ in one trajectory. ϕ fluctuations with comparable range and time scale are also seen in the 1-μs simulation, and this excludes the possibility that the observations of ϕ fluctuations are made by

chance in the multiple cav1_ systems. Altogether, our analysis suggests that caveolin-1 orientation is highly dynamic in the membrane and most of the time it exhibits an insertion angle at 70°.

Table 3.2. Average ϕ in each of the 50 $Cav1_{82\text{-}136}$ MD simulation systems including the averages and standard errors of the ten replicas in each simulation set.

	45	55	65	75	85
1	62°±12°	66°±9°	76°±9°	70°±7°	74°±8°
2	80°±6°	69°±6°	74°±6°	81°±6°	81°±6°
3	60°±9°	63°±8°	80°±9°	72°±10°	81°±5°
4	59°±5°	68°±12°	80°±6°	61°±5°	49°±12°
5	57°±14°	71°±9°	85°±4°	82°±6°	68°±13°
6	80°±7°	83°±5°	73°±5°	22°±7°	62°±8°
7	77°±8°	85°±4°	80°±8°	81°±5°	81°±6°
8	82°±5°	68°±5°	71°±7 °	77°±6°	51°±10°
9	77°±6°	68°±7°	75°±6°	71°±11°	72°±7°
10	75°±5°	75°±10°	84°±4°	67°±7°	72°±6°
Avg.± S.E.	71°±10°	71°±7°	78°±4°	68°±17°	69°±11°

With ϕ at ~70°, most caveolin-1 structures are fully embedded in the bilayer with the termini in the headgroup region of the top leaflet and the linker residues buried in the bottom leaflet. The insertion is illustrated by the *Z*-positions of several key residues' COM (Z_{COM}) including those at the linker (G108, I109, and P110) and the four tryptophan residues (W85, W98, W115, and W128). Figure 3.8A plots the Z_{COM} distributions of these residues. It is remarkable that the linker residues show similar *Z*-distributions, even with different initial Z_{COM} values. The peaks of these distributions are at -5 Å, and lie between the bilayer center (dashed line) and the phosphocholine (PC) groups (orange box) of the bottom leaflet. W85 is located outside of the PC headgroup with a Z_{COM} of 23.8 Å and thus is solvent exposed. W128 has a Z_{COM} of 10.3 Å and straddles the interface between the PC headgroup and the beginning of the acyl chain

region. In contrast, both W98 and W115 are located completely within the acyl chain hydrophobic core. The Z_{COM} distribution of W115 peaks at -0.25 Å, which is 9.5 Å below that of W98. This is because, first of all, sequence-wise, W115 is closer to the linker residues. Secondly, H2 is tilted further away from the *Z*-axis than H1 (see below), and this buries W115 deeper in the hydrophobic core of the membrane. Together with the predominant ϕ value at ~70°, the Z_{COM} distributions indicate that a caveolin-1 molecule is inserted in but does not span across the entire membrane bilayer. This result is also in line with the data from previous fluorescence microscopy immunodetection experiments [148].

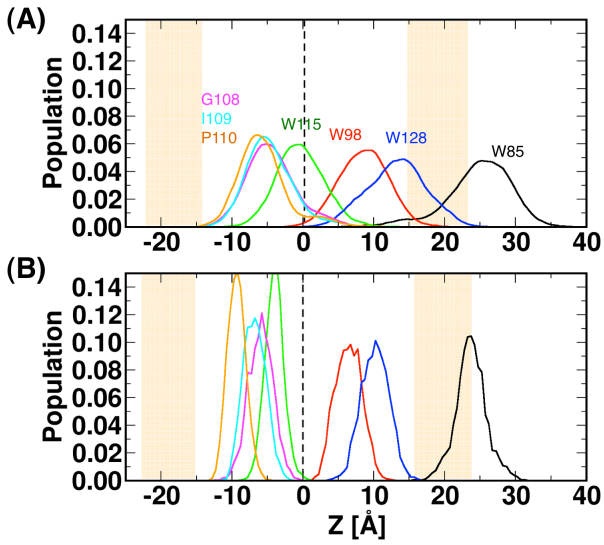


Figure 3.8. The Z_{COM} distributions for W85 (black), W98 (red), W115 (green), and W128 (blue) and the linker residues G108 (magenta), I109 (cyan), and P110 (orange) in (A) the multiple Cav1$_{82\text{-}136}$ systems and (B) the 1-μs simulation. The histograms in (A) are averaged over all 50 multiple simulations. The dashed black line shows the bilayer center. The orange bars indicate the regions of lipid phosphocholine groups. An illustration of the approximate relative positions of these residues on Cav1$_{82\text{-}136}$ can be found in Figure 3.2B.

The Z_{COM} distributions of the tryptophan residues are also well correlated with their fluorescence emission measurements. As shown in Figure 3.9, the λ_{max} values from the fluorescence emission spectra of each tryptophan mutant are 341±0.25 (W85), 337±0.10 (W98), 334±0.33 (W115), and 340±0.75 nm (W128), which reflect the degree of solvent exposure of tryptophan residues. Based on the values, they are usually divided into three classes: (I) 330-333 nm, (II) 340-343 nm, and (III) 347-350 nm [149]. Tryptophan residues with λ_{max} in class I are generally regarded as solvent inaccessible and buried in the hydrophobic core of a membrane, and a class III λ_{max} value is indicative of a tryptophan residue fully exposed to the aqueous solution. In between these two are the class II tryptophan residues, which are mostly at the lipid/water interface (headgroup region). The λ_{max} value of W98 lies between classes I and II, indicative of a position slightly below the interfacial headgroup region. The λ_{max} value of W115 is consistent with class I revealing that it has a deep location in the bilayer. In addition, W85 and W128 reside most likely in the lipid head groups according to their λ_{max} values (class II). Importantly, the progressively greater blue shifts for W85, W128, W98, and W115 are in very good agreement with the Z_{COM} distributions observed in the simulations (Figure 3.8). Therefore it is clear that the simulations are conveying relevant information about caveolin-1's topology and orientation in the membrane.

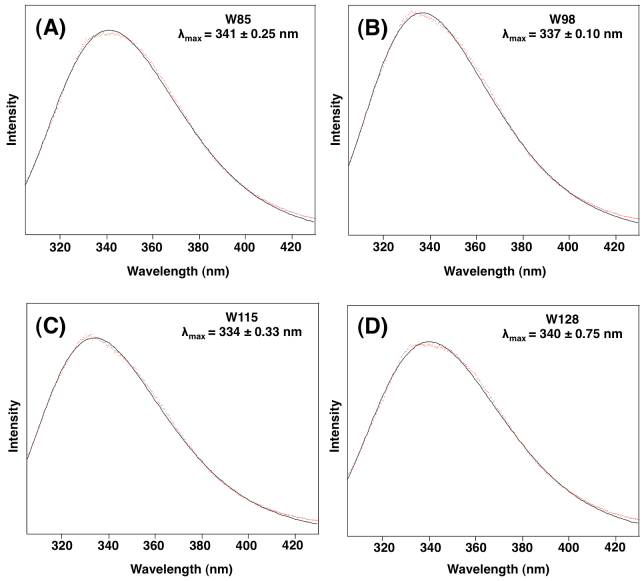


Figure 3.9. Fluorescence emission spectra of (A) W85, (B) W98, (C) W115, and (D) W128 in wild-type $Cav1_{62\text{-}178}$ in 2.0% (w/w) 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine/1,2-dihexanoyl-*sn*-glycero-3-phosphocholine bicelles $\boldsymbol{q}$ = 0.5; 20 mM phosphate pH 7.0.

Besides the insertion angle and insertion depth, characterizing the orientation of caveolin-1 requires two more parameters, α and β, demonstrating the tilting of caveolin-1 in its molecular plane (Figure 3.2B). The distributions of α and β are shown in Figure 3.10. The preferred α in almost all the systems lies between 90° to 120°. The one-dimension α distributions (Figure 3.10) reveal that the primary peaks are at 105°, 104°, 101°, 105°, and 103° for the cav1_45, 55, 65, 75, and 85 systems, respectively. With caveolin-1 inserted nearly perpendicular to the membrane, this corresponds to a caveolin-1 orientation with its H1 helix roughly parallel to the *Z*-axis. Such a structural preference is stabilized by interactions formed at both ends of H1. The CSD domain at the N-terminal end of H1 contains multiple polar and charged residues such as K86, S88, T91, K96, Y97, and R101. These residues interact with the lipid head groups of the top leaflet

and anchor the N-terminus of H1 at the top leaflet lipid-water interface (Figure 3.7A and B). On the other hand, the C-terminus of H1 is directly connected to the linker residues, which have the ability to form H-bonds with lipid headgroups of the bottom leaflet. Although the average Z_{COM} of the linker residues is -4.9±3.3 Å and is above the *Z*-position of the PC groups in the bottom leaflet as indicated in Figure 3.8, the interactions are possible as the membrane embedded caveolin-1 introduces thinning of the surrounding bottom lipid leaflet, which is discussed further in the membrane perturbation section below. These interactions between H1 and lipid molecules in opposite leaflets help secure the positioning of the H1 helix in the membrane. When either of the anchor interactions is compromised, H1 can access a larger spectrum of α values as exemplified in system cav1_45_3, in which the anchorage at the bottom leaflet is lost and consequently the range of accessible α values becomes 28° to 109° (Figure 3.7C), much larger than the 90° to 120° range observed for systems with the linker anchored in the bottom leaflet. Most systems show similar β distributions as well. The primary values of β in the five simulation sets were 53°, 59°, 54°, 58°, and 44°, respectively, and the difference between α and β (Δαβ = 52°, 45°, 47°, 47°, and 59° for systems cav1_45, 55, 65, 75, and 85) reflects the accessible crossing angle θ range (from 45° to 64°) in all the simulations. Clearly there is more variability in β than was observed in α. In addition, β was strongly dependent on θ. In general, systems with large θ had small β and vice versa, meaning that there may be multiple minima along β when α is held fixed. Altogether, our analysis indicates that caveolin-1 tilts in its molecular plane with a majority of the molecules having α ≈ 103° and β ≈ 55°.

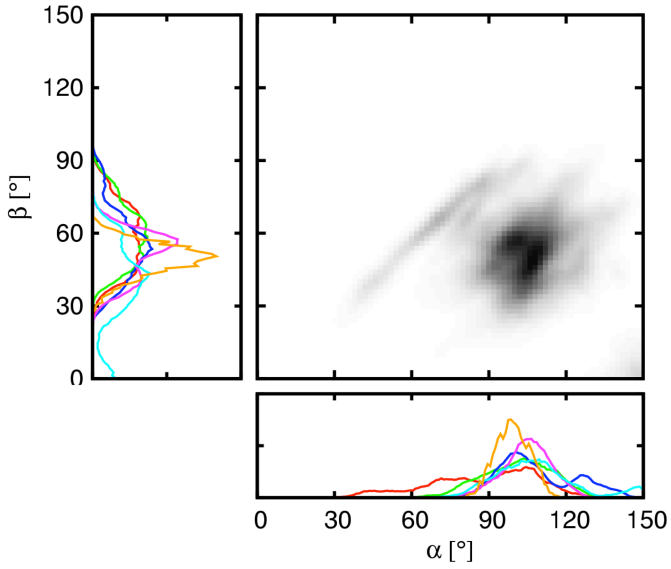


Figure 3.10. The average two-dimensional α and β distributions in all the multiple simulations. The bottom and left panels show the distributions of α and β respectively. The distributions from different systems are distinguished by colors with system cav1_45 in red, cav1_55 in green, cav1_65 in blue, cav1_75 in magenta, cav1_85 in cyan, and the 1-μs simulation in orange.

The preferred orientation obtained from the multiple cav1_ simulations is also stable at least on the microsecond time scale, as no major change is observed during the 1-μs simulation. In the entire simulation, caveolin-1 remains inserted, as manifested by the average ϕ and the Z_{COM} of the linker residues of 63°±6°and -7±1 Å (Figure 3.3B and 3.8B). There is a decline in ϕ over the first 60 ns from 78° to 53°, but afterwards, it varies only between 50° and 70°. Again this agrees well with the shorter 80-ns simulations. The primary values for the linker residue Z_{COM} are -5.8 Å, -6.8 Å, and -9.3 Å for G108, I109, and P110, respectively. Although the values are slightly bigger in magnitude than those obtained by averaging over all the trajectories in the cav1_ simulations including the ones with caveolin-1 afloat at the lipid-water interface, they are accessible in many of the cav1_ systems. Indeed, systems that exhibit the canonical structure and orientation show

similar Z_{COM} values of the linker residues as the ones seen in the 1-μs simulation. The Z_{COM} distributions for the tryptophan residues imply that W89 likely locates at the hydrophobic-hydrophilic interface and W98 and W115 are buried in the hydrophobic core. From the distributions, most of the W128 population has a Z_{COM} value below 17Å, the Z_{COM} of the top leaflet headgroups; a portion of the population is actually solvent accessible due to the thinning in the top leaflet invoked by the interactions between caveolin-1 residues such as R101 and Y118 and the lipid head groups (Figure 3.7A and B). These observations also correlate well with the tryptophan fluorescence emission measurements. Besides the comparable insertion angle and depth, the wobbling of caveolin-1 in its molecular plane is seen as well, with α and β values (α ≈ 99° and β ≈ 49°) very close to those in the multiple 80-ns simulations (α ≈ 103° and β ≈ 55°) (Figure 3.10). Overall, the configurations including their fluctuations from the multiple 80-ns simulations are correctly reflected in the 1-μs simulation. The similarity shared between the results indicates that the configurational preference obtained from the multiple simulations is indeed physically relevant and the impact from the initial caveolin-1 configuration is minimal.

3.3.4. Membrane perturbation induced by caveolin-1 insertion

The caveolin-1 insertion has an impact on lipid packing at the protein-lipid interface, and deformation of the membrane around the caveolin-1 molecule is seen. This is illustrated by the bilayer thickness profiles taken from system cav1_65_3, in which the caveolin-1 molecule displays the preferred configurations discussed above (Figure 3.11A). As a matter of fact, all systems with fully inserted caveolin-1 molecules show a considerable

amount of membrane thinning, especially in regions near the *XY*-projection of the basal portion of the U-shaped caveolin-1. In this region, lipid head groups from both the top and bottom leaflets bend towards the hydrophobic core of the membrane as shown by the individual leaflet thickness profiles (Figure 3.7A and B and 3.11B and C). At the top leaflet, the thinning mostly occurs in the region between H1 and H2 caused by electrostatic and H-bond interactions between residues, R101 (H1) and Y118 (H2) in particular, and the lipid head groups (Figure 3.7A and B and 3.11B). At the bottom leaflet, H-bonds between the linker residue *backbone* atoms and the lipid head groups are mostly responsible for pulling the polar lipid head groups inwards to the hydrophobic core, resulting in the thinning effect (Figure 3.7A and B and 3.11C).

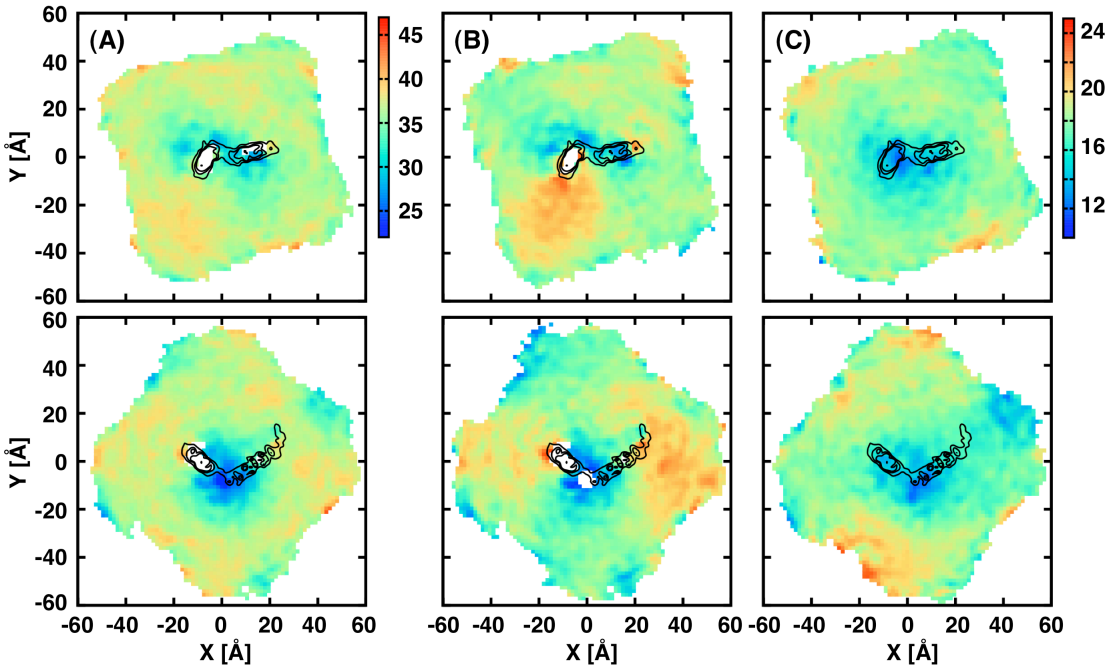


Figure 3.11. The thickness profiles of (A) the bilayer, (B) the top leaflet, and (C) the bottom leaflet in systems cav1_65_3 (top) and cav1_55_8 (bottom). Prior to making the profiles, the caveolin-1 molecules are aligned so that the vector connecting the centers of H1 and H2 coincides with the *X*-axis. To calculate the profile, a two-dimensional grid is placed on the *XY* plane on top of the protein. For each grid point, the averaged *Z*-positions of the phosphate atoms at the top and bottom leaflets are computed. These *Z*-

values represent the thickness of the top and bottom leaflets, respectively. The difference indicates the entire bilayer thickness. The color scale in the plots goes from blue to red with increasing thickness. The density of the protein is also plotted on the *XY* plane as black contour lines to show caveolin-1 responsible for membrane perturbation.

In contrast, regions near the outer edges of caveolin-1 projection on the *XY*-plane exhibit moderate hydrophobic thickness increase, mostly at the top leaflet near the N- and C-termini of caveolin-1 (Figure 3.11B). Residues accountable for the thickness increase include K86, S88, and T91 in H1 and S133 and K135 close to the C-terminus. In a rigid membrane without the ability to make any adaptations, these residues would be exposed to the aqueous solution with the preferred caveolin-1 configuration. However, lipid bilayers are fluidic and can either compress or expand along the thickness dimension depending on local interactions, which in this case are the salt bridges and H-bonds between the aforementioned terminal residues and the lipid head groups, therefore resulting in the membrane thickness expansion seen in the simulations. The location of expansion is closely associated with the position of these residues. For example, in system cav1_55_8, a prevailing H1 rotation at 181° places K86, S88, and T91 at the opposite side of the caveolin-1 molecular plane as in system cav1_65_3 (Figure 3.7A and B), and the location of membrane thickness increase follow that of the polar residues on H1. The impact of H2 rotation on the membrane thickness increase is not as prominent as S133 and K135 are located in the loop region of the C-terminus and are more flexible, but increased membrane thickness is always observed near the C-terminal region of caveolin-1 in nearly all simulation systems.

Thickness deformation is not the only perturbation caused by caveolin-1 insertion into the membrane. The insertion also stretches the top and bottom lipid leaflets in an

asymmetric fashion. On average, there are always 3 to 4 less lipid molecules at the top leaflet due to the U-shape conformation adopted by caveolin-1. Because of this unique shape, the cytoplasmic side of caveolin-1 is wide open and has a much larger cross-sectional area (A_{prot} = 104 ± 29 Å^2, 15 ≤ Z ≤ 19 Å) than the extracellular side (A_{prot} = 0 Å^2, -19 ≤ Z ≤ -15 Å). It is then intuitive that when incorporated in a planar bilayer, the caveolin-1 is likely to create lateral tension imbalance in the opposite leaflets and perhaps this drives the curvature formation in the bilayer. This is, however, not observed in the multiple 80-ns or the 1-μs simulations as the $P2_1$ image transformation is employed in the former to allow the exchange of lipid molecules between the opposing leaflets and the size of the simulation cell in both simulations is not large enough for a global curvature of the membrane to be seen. Nevertheless, the difference in lipid numbers and A_{prot} in the opposite leaflets illustrate the ability of caveolin-1 to cause a positive intrinsic curvature.

3.4. Summary and Conclusions

The caveolin protein family is an indispensable component of caveolae. Their misfunction can lead to severe disease conditions [112-115]. Because of its role as the inducer and principal component of caveolae, the structure and the action mechanism of caveolin molecules have been under the scrutiny of researchers for years. Despite the efforts, however, no atomic structure of caveolin molecule is available, possibly due to the fact that it is highly dynamic and also its association with the membrane. Therefore, the primary goal of this study was to reveal the structure and dynamics of caveolin-1 in a membrane bilayer environment. The study also investigates the impact of caveolin-1 insertion on the surrounding membrane bilayers in the hope of gaining insights into

caveolin-induced membrane curvature formation mechanism. The strategy adopted in this study is to perform multiple MD simulations to generate favorable caveolin-1 configurations and a 1-μs long simulation of a cavelin-1 system with a preferred configuration to further examine the stability and the dynamics of such a configuration. The simulation results are well correlated with the tryptophan fluorescence emission measurements.

First, the simulations show that the caveolin-1 prefers to adopt a U-shape conformation in membranes with a wide range of preferred θ angle from 45° to 64° (Figure 3.3A and 3.4). This is a very important finding as there has been considerable doubt as to whether a conformation such as this could persist in the membrane. Our studies show that stable yet dynamic U-shape conformations are possible and that the linker residue backbone atoms (G108-P110) make H-bonds to the lipid head groups of a locally thinned bilayer.

Second, the simulations also confirm that caveolin-1 is inserted in but does not span the membrane with the insertion angle at ~70° (Figure 3.3B). Both the N- and C-termini of caveolin-1 face the cytoplasm and are located at the hydrophobic-hydrophilic interface of the top bilayer leaflet, while the linker residues are found beneath the lipid head groups of the opposite leaflet (Figure 3.7 and 3.8). The mode of insertion observed in the current simulations is also supported by the tryptophan fluorescence emission experimental data (Figure 3.9). Such a finding is significant as it shows that caveolin does not completely span the bilayer.

Third, the structural analyses results show that there are four possible H1-H2 interfaces manifested by four pairs of H1 rotation angle (ρ_1) of 343° and H2 ρ_2 of 111°

being the predominant. The structures with such a rotation angle preference have polar residues such as R101 and Y118 very close to or inside the space between the H1 and H2 helices (Figure 3.7A and B). These residues are involved in causing local membrane thinning around caveolin-1.

Lastly, caveolin-1 insertion causes local membrane disruption. Part of the disruption is reflected in the local membrane deformation. It includes thinning of the membrane, which occurs in the area in between the H1 and H2 helix as a result of specific residue-lipid interactions, and membrane thickness expansion near both the N- and C-termini of caveolin-1. This deformation pattern may play a role in caveolin-1 oligomerization in a sense that the favorable local environments of the monomers are similar and it could help them recognize each other with ease. The membrane disruption is also reflected in the lipid number difference in the opposite leaflets due to the different cross-sectional areas of caveolin-1 in them. This illustrates the potential of the caveolin-1 molecules to asymmetrically stretch the bilayer and induce membrane curvature.

Chapter IV

Molecular Dynamics Studies of Ion Permeation in VDAC [2]

[2] Reused from *Biophysical Journal*, Huan Rui, Kyu Il Lee, Richard W. Pastor, and Wonpil Im, **100**. 2011. pp 602-610. Copyright (2011). With permission from Elsevier Science.

Summary

The voltage dependent anion channel (VDAC) in the outer membrane of mitochondria serves an essential role in transport of metabolites and electrolytes between the cell matrix and mitochondria. To examine its structure, dynamics, and mechanisms underlying its electrophysiological properties, we have performed a total of 1.77 μs molecular dynamics simulations of human VDAC isoform 1 in DOPE/DOPC mixed bilayers in 1M KCl solution with transmembrane potentials of 0, ±25, ±50, ±75, and ±100 mV. The calculated conductance and ion selectivity are in good agreement with the experimental measurements. In addition, ion density distributions inside the channel reveal possible pathways for different ion species. Based on these observations, a mechanism underlying the anion selectivity is proposed; both ion species are transported across the channel, but the rate for K^+ is smaller than that of Cl^- because of the attractive interactions between K^+ and residues on the channel wall. This difference leads to the anion selectivity of VDAC.

4.1. Introduction

The voltage dependent anion channels (VDACs) are 30-kDa transmembrane (TM) proteins found in the outer membrane of mitochondria. They provide important conducting pathways for simple ions such as K^+, Cl^- and Ca^{2+} [150], and small molecules like pyruvate and ATP [150] between the cytosol and the intermembrane space in mitochondria. There is also evidence that VDACs are involved in cancer cell metabolism [151] and release of apoptotic proteins from mitochondria [152]. Isoforms of VDACs exist in many organisms. In human mitochondria, three VDAC isoforms (hVDAC1, hVDAC2 and hVDAC3) have been identified [153]. Although they all provide exchange pathways for ions and metabolites between the mitochondria and cell matrix, their expression profiles and physiological properties are different [154]. VDAC1 shows characteristic electrophysiological properties (e.g., single channel conductance, anion selectivity, and voltage dependence) that are shared in VDACs across different eukaryote species from yeast to human [155]. Therefore, it is presently the most extensively studied VDAC isoform. Studies of VDACs from rat liver mitochondria reveal moderate anion selectivity. The selectivity ratio between Cl^- and K^+ is approximately 2:1 in 1.0:0.1 M KCl asymmetric solution [156], and the single-channel conductance ranges from 3.9 nS to 4.5 nS in 1.0 M KCl solution [157]. When inserted into planar phospholipid membranes with low TM potential (~±10 mV), VDACs are open and preferentially transport anions. Increasing the TM potential to ~±30 mV, however, shifts the channels into a slightly cation selective "closed" state with reduced single channel conductance of 1-2 nS [156]. Given the gating threshold of VDAC and the mitochondria resting

transmembrane potential of about -180 mV [158], VDACs are presumably closed at the resting state in the mitochondrial outer membrane.

hVDAC1 was first cloned and expressed in human B-lymphocyte cells [159]. It consists of 283 amino acids, with alternating hydrophobic and hydrophilic residues in many regions, as consistent with β-strands in a TM β-barrel [160]. Recent NMR and X-ray structures of human and mouse VDAC1 [161-163] confirm the β-barrel architecture of the channel. In both structures of hVDAC1 [161, 163], there are 19 β-strands with the first and last strands parallel to each other, forming a cylindrical barrel of ~30 Å in diameter. A part of the N-terminus forms a short α-helix that extends into the channel lumen (Figure 4.1). Of the 179 residues residing in the pore lining, 17 are acidic and 20 are basic, as might be anticipated for a channel that is mildly anion selective but also transports cations. VDAC1 orientation in the mitochondria membrane, i.e., whether both the N- and C-termini face the cytosol [164, 165] or the intermembrane side [166], is still under debate.

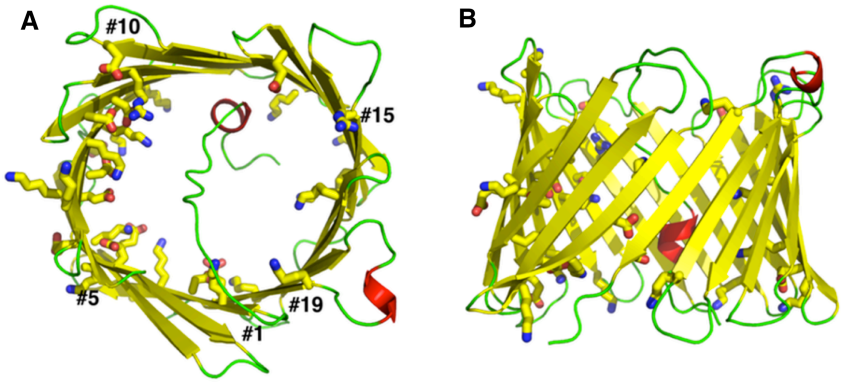


Figure 4.1. Molecular representation of hVDAC1 from the first NMR model (PDB:2K4T): (*A*) top view and (*B*) side view, with α-helices (*red*), β-strands (*yellow*) and loops (*green*). Charged residues that line the channel wall are highlighted in stick representations. The figures were produced using the molecular visualization program

PyMOL [167].

Before the atomic structures of VDACs were determined, models for ion selectivity and voltage dependence were primarily based on functional studies [168, 169]. It is now possible to examine the ion permeation at the atomic level. Based on the crystal structure of mouse VDAC1 [162] and continuum electrostatic calculations, Choudhary et al. [170] rationalized the anion selectivity of the wild type channel by computing the free energy required to transfer different ion species from bulk solution into the channel. Since the calculations were performed on a static structure, structural dynamics, which may alter the electrophysiological properties of the channel, was not considered in computing the free energies. In addition, the free energies were calculated by placing a single ion along the pore axis, neglecting its lateral freedom as well as the effects from other ions. This work presents results from a total of 1.77 μs molecular dynamics (MD) simulations of hVDAC1 in DOPE/DOPC mixed bilayers bathed in 1M KCl solution at TM potentials of 0, ±25, ±50, ±75, and ±100 mV. The results are discussed in terms of the channel stability and dynamics in membranes, its electrophysiological properties, ion diffusion constants inside the pore, and the molecular mechanisms underlying its anion selectivity.

4.2. Computational Methods

4.2.1. Simulation of hVDAC1 in explicit membrane bilayers without transmembrane potential

The structure of hVDAC1 was taken from the first model of the NMR structures (PDB:2K4T) [161] using the PDB Reader module in CHARMM-GUI (http://www.charmm-gui.org) [171] with a positively charged N-terminus and a negatively charged C-terminus. The channel was centered at $Z = 0$ and reoriented to make its pore axis parallel to the membrane normal (i.e., the Z-axis). Both the N- and C-termini were placed in the lower half of the simulation cell with $Z < 0$. To best resemble the lipid compositions used in the structure determination [161], a mixed lipid bilayer containing 1,2-dioleoyl-glycero-3-phosphatidylethanolamine (DOPE), 1,2-oleoylphosphatidylcholine (DOPC), and cholesterol [172] was generated with a ratio of 40:40:1 (PE:PC:Chol) by CHARMM-GUI Membrane Builder [137]. Both the top and bottom leaflets contain 81 lipid molecules, with the same lipid composition ratios. 1.0 M KCl was also added to hydrate and neutralize the system (217 K^+, 219 Cl^-, and approximately 12,500 waters), resulting in an 88.9 × 88.9 × 85.7 $Å^3$ simulation system. To enhance sampling of hVDAC1 conformation and ion motions, three independent hVDAC1 systems, denoted S1_0, S2_0, and S3_0, were built, with total atom numbers of 63,749, 63,710, and 63,884 respectively; the suffix indicates that the applied voltage is 0. The number of water molecules is different in each system because the systems were generated independently with lipid molecules of different conformations and positions by CHARMM-GUI Membrane Builder. After 700-ps equilibration, each system was subjected to an initial 5-ns production. All the simulations were carried out in NPγT

ensemble (i.e., constant pressure P_z = 1.0 atm long *Z*-axis, surface tension γ = 20 dyn/cm [173], and temperature T = 303.15 K) using biomolecular simulation program CHARMM [174] with all-atom parameter set PARAM22 [84] including the dihedral cross-term corrections (CMAP) [175], C27r lipid parameters [86], and the modified TIP3P water model [176]. Trajectories were generated with a 2-fs time step and bonds with hydrogen were constrained with the SHAKE algorithm. We used the same options for non-bonded interactions in the input scripts provided by the CHARMM-GUI Membrane Builder [137]; the van der Waals interactions were smoothly switched off at 11−12 Å by a force switching function [141] and the electrostatic interactions were calculated using the particle-mesh Ewald (PME) method with a mesh size of about 1 Å for fast Fourier transformation, κ = 0.34 $Å^{-1}$, and a sixth-order B-spline interpolation [177].

4.2.2. Simulation of hVDAC1 in explicit membrane bilayers with transmembrane

After 5-ns equilibration, nine different TM potentials (V_{mp}) ranging from -100 mV to 100 mV in 25 mV increments were applied to the three systems, resulting in a total of 27 systems; e.g., S2_n100 is originated from S2_0 with V_{mp} = -100 mV. For each system, a 65-ns MD trajectory was then generated in the NPγT ensemble with the same conditions described above. Applying TM potentials greatly enhances the sampling of ion crossing events and therefore increases the statistical significance of the calculated electrophysiological properties of the channel. Nevertheless, it should be stressed that because of the gap between the current simulation times (tens of nanoseconds) and the time required for the “open” to “closed” transition of the channel (seconds), the simulations with high TM potentials are unlikely to provide any gating information or to

alter the channel's conductance and selectivity significantly. Nine different TM potential are employed in order to improve the conformational sampling of hVDAC1 as well as to evaluate the channel's *I-V* characteristics at various TM potentials. Figure 4.2 shows the last snapshots of system S1_n100, S2_0, and S3_p100.

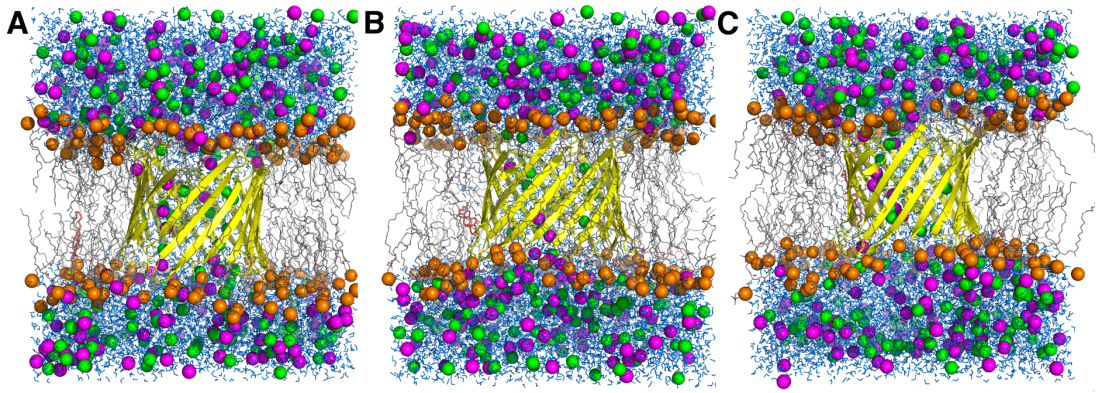


Figure 4.2. Molecular representation of systems (*A*) S1_n100, (*B*) S2_0, and (*C*) S3_p100 at the end of the simulations: hVDAC1, *yellow ribbon*; phosphates of DOPE and DOPC molecules, *orange spheres*; remainder of the lipids, *grey lines*; cholesterols, *red lines*; water, *marine*; K^+, *magenta*; and Cl^-, *green*. The figures were produced with PyMOL [167].

V_{mp} was established by applying an external constant electric field across the membrane system [178]; V_{mp} = V (*Z* = -Lz/2) - V (*Z* = Lz/2), where Lz is the system size in the *Z*-axis. The direction of the electric field was parallel to the membrane normal. To verify if V_{mp} was implemented correctly during the simulations, we calculated the electrostatic potential along the *Z*-axis, $\mathrm{V}(z)$, by solving Poisson's equation,

$$\frac{d^2\mathrm{V}(z)}{dz^2} = -4\pi\rho_{sys}(z) \qquad (4\text{-}1)$$

where $\rho_{sys}(z)$ denotes the system charge distribution along the *Z*-axis. The $V(z)$ profiles and the potential differences across the bilayers in Figures 4.3 and 4.4 validate the applied V_{mp} during the simulations.

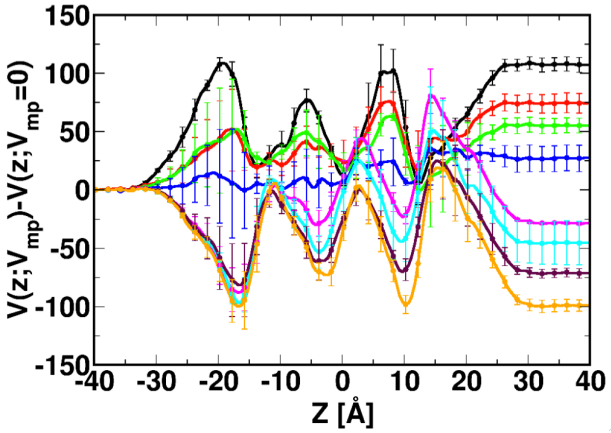


Figure 4.3. Average electrostatic potential profiles of the systems with V_{mp} = -100 (*black*), -75 (*red*), -50 (*green*), -25 (*blue*), 25 (*magenta*), 50 (*cyan*), 75 (*maroon*) and 100 mV (*orange*) along the *Z*-axis. The average electrostatic potential profile from the systems with V_{mp} = 0 is subtracted from the original V(z; V_{mp}) profiles (Figure 4.4) to illustrate the influence of the applied electric field on the electrostatic potential throughout the simulation system. The established potential, i.e., V(-40; V_{mp}) - V(40; V_{mp}), during the simulations are -104±4, -72±5, -53±4, -25±7, 31±1, 48±11, 74±2, and 102±3mV for the applied V_{mp} = -100, -75, -50, -25, 25, 50, 75, and 100 mV.

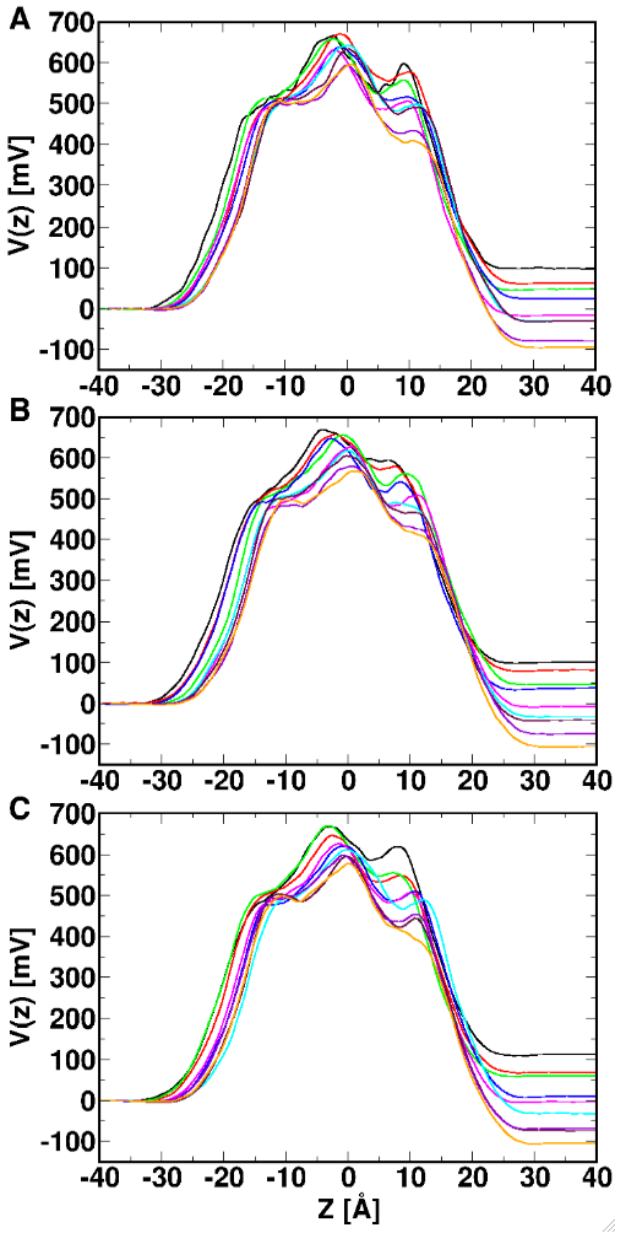


Figure 4.4. Electrostatic potential V(z) profiles along the *Z*-axis in the simulation set S1 (*top*), S2 (*middle*) and S3 (*bottom*). Systems with different TM potentials are presented using different color schemes: V_{mp} = -100 mV (*black*), -75 mV (*red*), -50 mV (*green*), -25 mV (*blue*), 0 mV (*magenta*), 25 mV (*cyan*), 50 mV (*maroon*), 75 mV (*violet*), and 100 mV (*orange*). Because the resulting V(z) is sensitive on the choice of the integration region along the Z-axis, the calculated profiles cover different *Z*-ranges starting from *Z* = -40.0 Å to values near 40.0 Å. To clearly indicate the potential difference across the simulation cell, all the *Z*-ranges are extended to 40.0 Å with the V(z) value at their ending value between *Z* = 30.0 Å and *Z* = 35.0 Å.

4.3. Results and Discussion

4.3.1. Structural stability and dynamics of the hVDAC1-membrane systems

The overall stability of hVDAC1 during the simulations with different V_{mp} was examined by the root mean-square deviations (RMSD) of the β-barrel backbone atoms. The RMSD are averaged for the last 60 ns and plotted with standard deviations for all systems in Figure 4.5A. The averaged RMSD are similar for different V_{mp}, ranging from 1.6 ± 0.1 Å in system S2_p50 to 2.5 ± 0.3 Å in system S1_n25. Although the overall RMSD is small, the β-barrel is flexible in some regions, especially the channel wall near the N-terminal α-helix (-4.0 Å ≤ Z ≤ -2.0 Å). In particular, the changes of the X-Y cross-section of the channel in this region suggest a large breathing motion; i.e., the ratios of the short and long axis (b/a) under different TM voltage conditions vary from 0.65 to 1.0 with an average of 0.88 (Figure 4.6). The breathing motion is likely related to the dynamics of the N-terminal helix (residue 6 to 10), which is believed to play a key role in channel gating as a voltage sensor [155]. In all the simulations, its center of mass position varied from -10.0 Å to 6.4 Å along the Z-axis (Figure 4.7A and B), while the hydrophobic interactions between Tyr7 and Leu10 in the α-helix and the pore lining residues Val143, Leu150, and Ala170 are well maintained (Figure 4.7C). Interestingly, the helix movement is biphasic; since the overall charge of the N-terminal region (residue 1 to 23) is positive, the helix moves toward more negative Z direction when $V_{mp} < 0$, and vice versa.

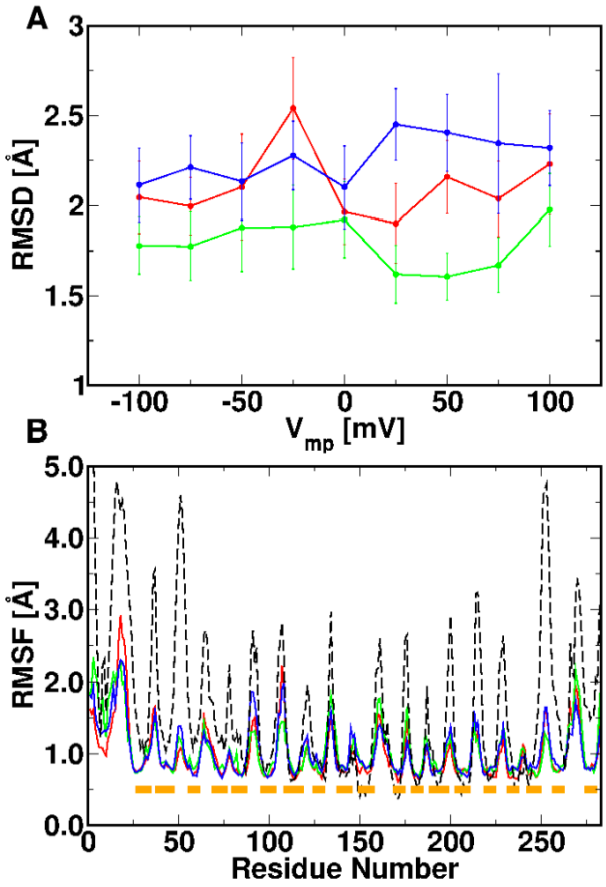


Figure 4.5. (*A*) Average backbone RMSD in all systems for the last 60 ns, with S1 (*red*), S2 (*green*), and S3 (*blue*). (*B*) Backbone RMSF for systems S1_n100 (*red*), S1_0 (*green*) and S1_p100 (*blue*), and from 20 NMR models (PDB ID: 2K4T) (*black dash*). The NMR models' RMSF was calculated with respect to the average NMR structure. The residue span of the β-strands is shown in boxes (*orange*).

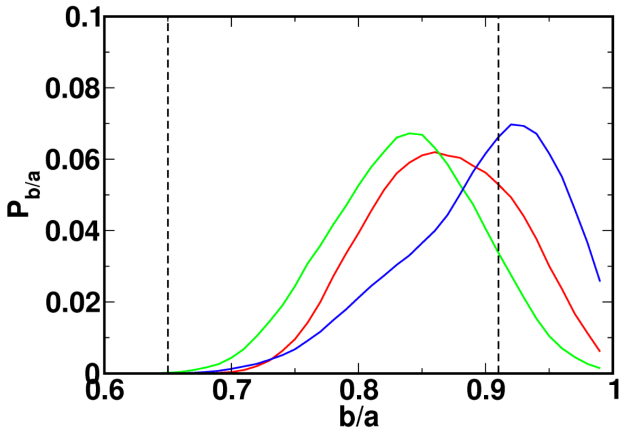


Figure 4.6. The normalized population histogram of the ratio between the short axis and the long axis of the ellipse formed by the β-barrel. The ratios are collected for systems under all TM potentials in each simulations set: S1 (*red*), S2 (*green*), and S3 (*blue*). The dashed lines show the span of the ratios calculated from the NMR structure ensemble.

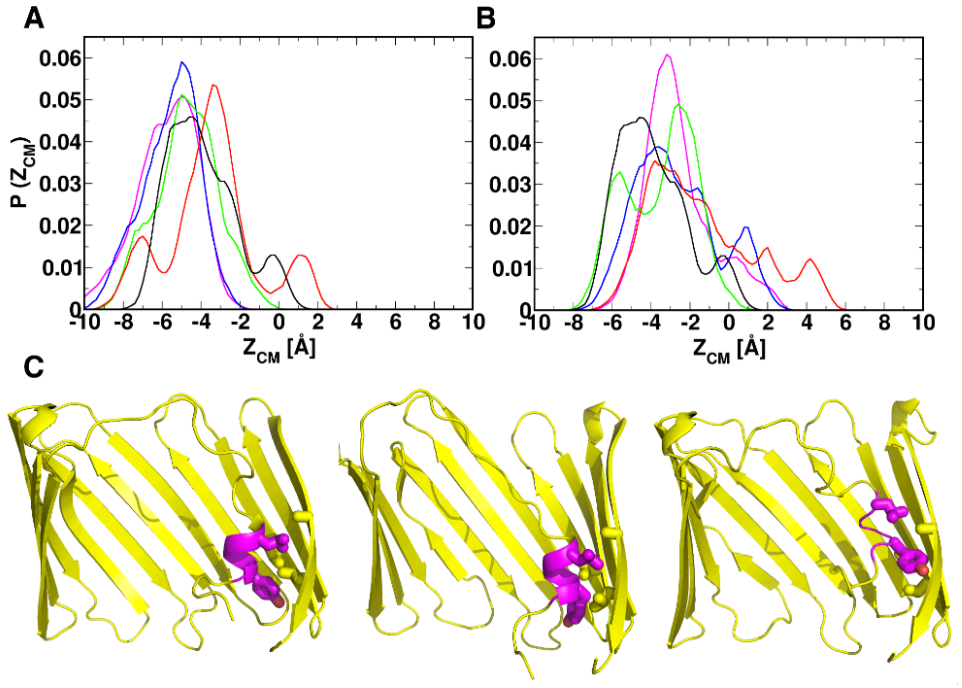


Figure 4.7. (A and B) Population of the Z-position (Z_{CM}) of the N-terminal α-helix. Z_{CM} are collected for systems with (*A*) V_{mp} = 0 mV (*black*), -25 mV (*red*), -50 mV (*green*), -75 mV (*blue*), and -100 mV (*magenta*), and (*B*) V_{mp} = 0 mV (*black*), 25 mV (*red*), 50 mV (*green*), 75 mV (*blue*), and 100 mV (*magenta*). (*C*) Interaction between the N-terminal α-helix residues (magenta sticks with Tyr7 (bottom) and Leu10 (top)) and the pore lining residues (yellow sticks with Val143 (middle), Leu150 (bottom), and Ala170 (top)): (*left*) Z_{CM} = -8.0 Å, (*middle*) Z_{CM} = -4.0 Å, and (*right*) Z_{CM} = 4.0 Å.

Figure 4.5B plots the backbone root mean-square fluctuations (RMSF) of S1_n100, S1_0 and S1_p100 for the last 60 ns. Similar RMSF are observed for systems with different V_{mp}, indicating minimal structural change upon the applied TM potentials on the present simulation timescale. In each system, the RMSF differ for residues

residing in different secondary structures; as expected, they are smaller for residues in the channel wall (~0.5 Å) than those in loops or both termini (~2.0 Å). The RMSF calculated from the 20 NMR structure models [161] exhibit similar trends in the conformational flexibilities of different channel regions (dashed lines in Figure 4.5B). The loop/termini RMSF are larger in the NMR structure ensemble than in the MD simulations because the paucity of NMR restraints in those regions leads to the larger conformational uncertainties. Nonetheless, together with the similar trends in the RMSF from MD and NMR, we show that the simulations, with and without TM potentials, maintain the channel stability with the integrity of its secondary structures with reasonable dynamics during the simulations. The generated trajectories, therefore, can be used to further analyze the electrophysiological properties and ion permeation mechanism of the channel.

4.3.2. Electrophysiological properties of hVDAC1 calculated from MD simulations

The electrophysiological properties of hVDAC1 are examined in terms of its conductance level and ion current ratio. The conductance (G) of hVDAC1 was calculated based on the accumulated ion crossing number ($N_{K/Cl}$) for each ion type (Figure 4.8) and Ohm's Law,

$$G = \frac{I_{\text{total}}}{V_{\text{mp}}} = \frac{(N_K - N_{Cl})q}{V_{\text{mp}}\tau} \tag{4-2}$$

where I_{total} is the total current and q is the charge of a monovalent ion. $N_{K/Cl}$ is either increased or decreased every 2 ps by the net number of ions that cross Z = 0 in the positive direction of the Z-axis. With $N_{K/Cl}$ in a given time interval τ, one can estimate the average current carried by each ion using Equation 4-2.

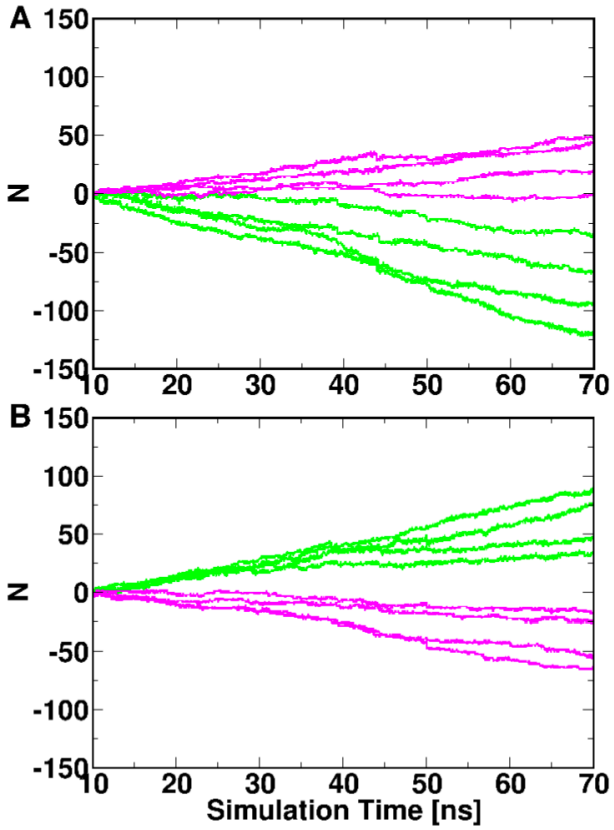


Figure 4.8. Accumulated ion crossing number for K^+ (*magenta*) and Cl^- (*green*) in the last 60 ns of the set S1 under (*A*) positive and (*B*) negative TM potentials. From top to bottom in (*A*) are V_{mp} = 100, 75, 50, 25 mV for N_K and V_{mp} = 25, 50, 75, and 100 mV for N_{Cl}. From top to bottom in (*B*) are V_{mp} = -100, -75, -50, -25 mV for N_{Cl} and V_{mp} = -25, -50, -75, and -100 mV for N_K.

As shown in Figure 4.9, the conductance values are similar in all systems with an average at 4.8±0.2 nS except in S3_n75 (6.5 nS), S3_n50 (7.0 nS), and S3_p75 (2.8 nS). The noticeable deviation from the mean in these three systems arises from stochastic ion crossing events with limited simulation timescales. Nonetheless, the calculated conductance values agree well with experimental values of about 3.9 nS to 4.5 nS in the same KCl concentration [155]. In addition, the fact that the MD conductance is very similar to the conductance (4.9±0.6 nS) from longer-time Brownian dynamics (BD) [179]

supports that the averaged MD conductance is statistically reliable even with the stochastic nature of ion crossing events with small V_{mp} and limited simulation timescales. The slightly higher value of the computed *G* than the experimental values is partly associated with the low viscosity of the TIP3 water model (see below for more details) [180].

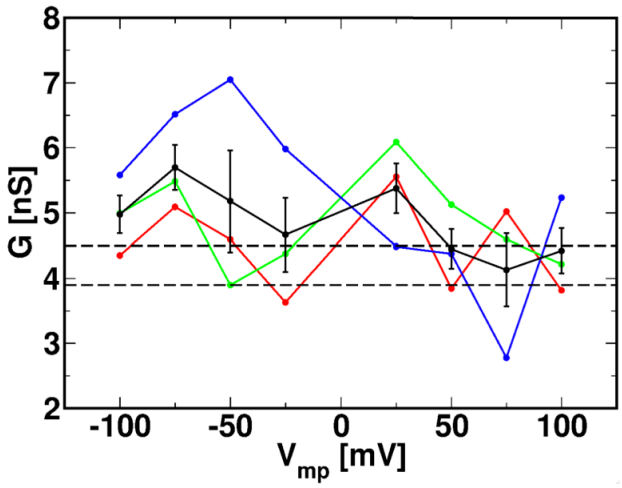


Figure 4.9. Conductance calculated from the last 60 ns in simulation sets S1 (*red*), S2 (*green*), and S3 (*blue*) as a function of applied voltage V_{mp}, with average conductance and standard error (*black*). The dashed lines indicate the range of experimental single channel conductance.

Figure 4.10 shows the total current (I_{total}) and the currents carried by K^+ (I_K) and Cl^- (I_{Cl}) as a function of applied voltage. Although the difference in current magnitudes is within the standard errors at most voltages, there is a small asymmetry of the currents (and thus conductance) at positive and negative TM potentials (Tables 4.1 and 4.2); a similar level of asymmetry was obtained in the BD simulations of hVDAC1 under the same conditions [179]. Such conductance asymmetry was also observed in OmpF in *E. coli* [181, 182]. This suggests that hVDAC1 passes K^+ ions more easily in the negative

direction of the *Z*-axis, while the situation is reversed in the case of Cl^-. The ion selectivity (ρ) can be estimated from the ratio between I_{Cl} and I_K: $\rho = I_{Cl}/I_K$. Assuming the linearity of the *I-V* curve, ρ can be evaluated by the ratio of the slopes of individual ion currents. Using least square linear regression, the calculated slopes of I_{Cl} and I_K are 3.2 and 1.6 respectively, giving a ratio of 2.0 ± 0.1. This ratio is in excellent agreement with the 2:1 anion selective ratio estimated experimentally [155].

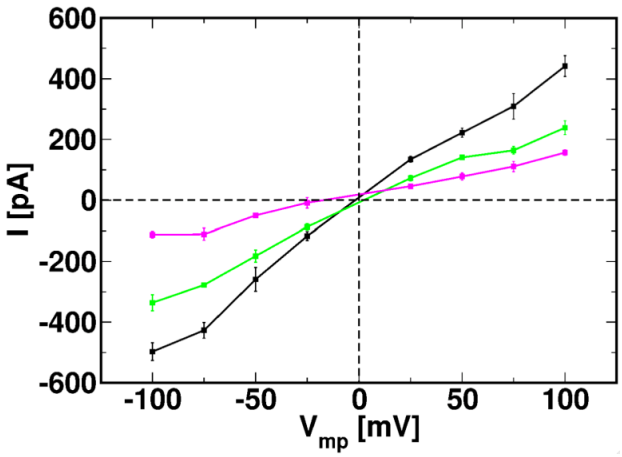


Figure 4.10. Current-voltage (*I-V*) relationship for the total (*black*), K^+ (*magenta*) and Cl^- (*green*) currents. The average current and standard error at each voltage point are calculated from three independent simulations.

Table 4.1. Average currents [in pA] and standard errors among three replicates at various TM potentials (V_{mp}).

V_{mp}	I_{total}		I_K		I_{Cl}	
	$-V_{mp}$	$+V_{mp}$	$-V_{mp}$	$+V_{mp}$	$-V_{mp}$	$+V_{mp}$
100	-498±35	442±42	-132±17	166±10	-366±26	275±33
75	-427±32	310±52	-138±14	126±25	-289±20	183±32
50	-259±48	223±19	-56±17	88±14	-203±31	134±5
25	-117±17	134±12	13±17	53±4	-104±5	82±12

Table 4.2. Individual ion currents [in pA] at various TM potentials.

	S1		S2		S3	
	I_K	I_{Cl}	I_K	I_{Cl}	I_K	I_{Cl}
n100	-118	-318	-112	-387	-166	-393
n75	-128	-254	-120	-291	-166	-323
n50	-51	-179	-29	-166	-88	-264
n25	5	-96	3	-112	-45	-104
p25	45	94	59	94	53	59
p50	67	126	115	142	83	136
p75	174	203	118	227	88	120
p100	147	382	171	251	182	342

Given the structural variability in the NMR model structures (Figures 4.5B and 4.6), it is reasonable to ask whether the electrophysiological properties calculated from MD simulations using only one model are biologically relevant and can be compared with experimental data. Clearly, the current MD simulations do not sample the barrel conformations with low *b*/*a* ratios (~0.65) that exist in the NMR models. It would be computationally prohibitive to perform MD simulations with all NMR models and examine the influence of structural variance on channel transport properties. However, our coarse-grained BD simulation study with all NMR models [179] indicates that the transport properties of the model used in the present MD simulation study closely represent those calculated from the ensemble of all the NMR models. Therefore, the ion transport properties from the current MD simulations are meaningful and can be directly compared with experimental measurements.

4.3.3. Diffusion constants inside the hVDAC1 pore

Another important determinant of electrophysiological property of the channel is the position dependent ion diffusion constant, $D(z)$, inside the pore. Note that $D(z)$ is an important input for coarse-grained BD or PNP (Poisson-Nernst-Planck) approaches to calculate electrophysiological properties of ion channels [181, 182]. The evaluation of $D(z)$, however, is not straightforward because the effects of systematic forces (e.g., ion-protein interactions) must be removed [183]. Following reference [182], $D(z)$ is obtained from the variance of a shifted Gaussian distribution along the *Z*-axis,

$$D(z) = \frac{\left\langle \left[\Delta z(\tau) - \left\langle \Delta z(\tau) \right\rangle \right]^2 \right\rangle}{2\tau} \tag{4-3}$$

where $\Delta z(\tau) = z(t+\tau) - z(t)$, τ is a suitably small time interval (several ps), and $\left\langle \Delta z(\tau) \right\rangle$ is the average *Z*-displacement of an ion per unit time τ arising from systematic forces. In free solution, $\left\langle \Delta z(\tau) \right\rangle$ = 0, and Equation 4-3 reduces to the familiar Einstein relation [184].

Figure 4.11 shows the averaged diffusion constant profile along the *Z*-axis for both K^+ and Cl^- at V_{mp} = 0. The diffusion constants of K^+ and Cl^- at ± 40 Å (bulk solution) are 0.30 $Å^2$/ps and 0.29 $Å^2$/ps, respectively. These values are independent of τ and approximately 1.5 times larger than the experimental values in bulk solution, D_K = 0.19 $Å^2$/ps and D_{Cl} = 0.2 $Å^2$/ps, at 298.15 K [185]. The discrepancy is the result of the different temperature (303.15 K) used in the simulations, the low viscosity of TIP3P water [180], and other errors in the potential energy function. As expected, the choice of τ modulates the calculated diffusion constants of both ion species inside the channel pore. With a small τ (e.g., τ = 1 ps), both ion types experience a 30% reduction of diffusion

constant (D_K = 0.21 Å^2/ps, D_{Cl} = 0.21 Å^2/ps), while the reduction is 50% (D_K = 0.15 Å^2/ps, D_{Cl} = 0.15 Å^2/ps) when a larger time interval is used (e.g., τ = 4 ps). The diffusion constant calculated from the larger τ represents the mobility of ions in the pore region more precisely compared to the one calculated using a smaller *τ*. This is because the autocorrelation of an ion's velocity is non-zero in a small time interval. In other words, the velocity of an ion is dependent on its previous velocity (i.e., the system memory), and the evaluation of *D* from Equation 4-3 is biased [186]. Following the previous molecular dynamics study of OmpF in *E. coli* [182], we determined the diffusion constants for K^+ and Cl^- in the VDAC pore with a time interval τ = 4 ps. The resulting diffusion profiles along the *Z*-axis can be used as inputs for BD or PNP calculations.

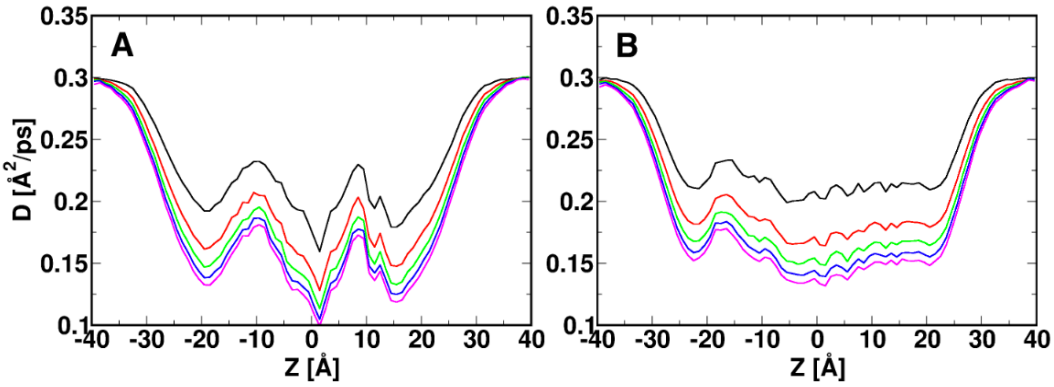


Figure 4.11. Diffusion constant profiles for (*A*) K^+ and (*B*) Cl^- along the *Z*-axis. The diffusion constant *D* is calculated using different time intervals: τ = 1 ps (*black*), τ = 2 ps (*red*), τ = 3 ps (*green*), τ = 4 ps (*blue*) and τ = 5 ps (*magenta*). Using time intervals τ = 4 ps and τ = 5 ps, a 50% decrease of *D* is observed when ions move from the bulk region to the VDAC pore region.

4.3.4. Mechanisms of hVDAC1's anion selectivity

The good agreement between the calculated and experimental electrophysiological properties of hVDAC1 supports an investigation into the mechanisms underlying the anion selectivity of the channel. We approach the problem by examining the ion number density profile, ion hydration number profile, and the cross-sectional ion charge distribution along the *Z*-axis.

The number density profiles of both K^+ and Cl^- ions along the *Z*-axis averaged for three independent simulations with $V_{mp} = 0$ (S1_0, S2_0 and S3_0) are shown in Figure 4.12A. Not surprisingly, the average density of Cl^- inside the pore is higher than that of K^+, consistent with the anion selectivity of the channel. We also calculated the average number of ions inside the pore by integrating the density over the pore region (-15.0 Å ≤ Z ≤ 15.0 Å). Table 4.3 lists the resulting numbers of K^+ and Cl^- inside the channel pore and the ratios between them. The average ratio of 1.3 is 35% smaller than the average ratio of 2.0 calculated from the ion currents (Table 4.1). This implies that K^+ ions stay in the channel longer than Cl^-. These differences in crossing times of the two ion species appear to arise from the attractive interactions between K^+ and the inner wall of the channel as seen in the K^+ density. In the middle of the channel, there is a region (-2.0 Å ≤ Z ≤ 2.0 Å) with high K^+ density, whereas the density sharply decreases when ions move from the middle to either ends of the channel (-12.0 Å ≤ Z ≤ -8.0 Å or 8.0 Å ≤ Z ≤ 12.0 Å). This indicates that there are free energy barriers for K^+ ions near both openings of the channel, which prevent K^+ from either entering or leaving the channel. In contrast, the density (and corresponding free energy profile) for Cl^- is much smoother, and thus, transport of Cl^- ions across the channel is more rapid.

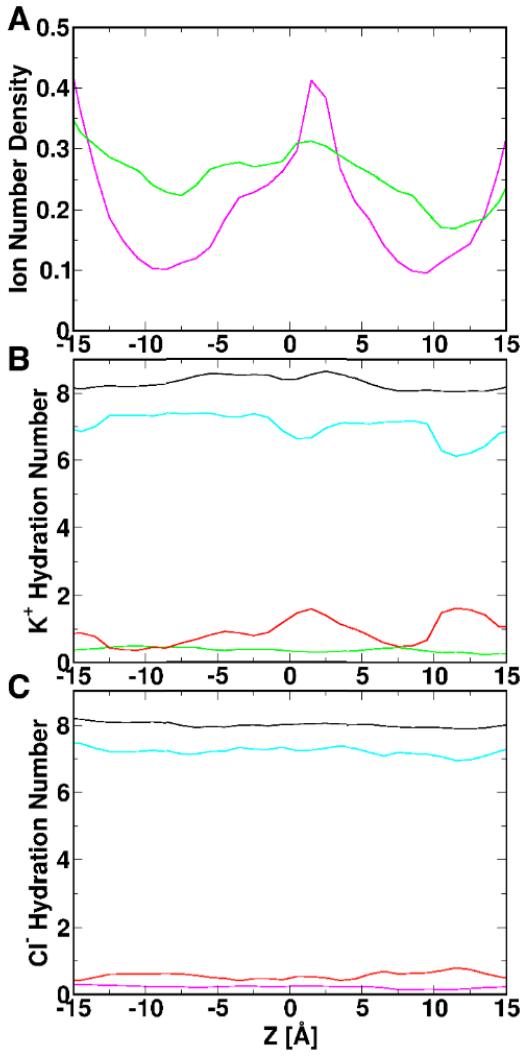


Figure 4.12. (*A*) Ion density profiles along the *Z*-axis in the pore region (-15.0 Å ≤ Z ≤ 15.0 Å): K^+ (*magenta*) and Cl^- (*green*). The densities are averaged for three independent simulations with $V_{mp} = 0$. (*B*) K^+ ion hydration number profiles and (*C*) Cl^- ion hydration number profiles along the *Z*-axis in systems with $V_{mp} = 0$. The hydration partners for K^+ include water oxygen (*cyan*), protein oxygen (*red*) and Cl^- ions (*green*), while those for Cl^- include water oxygen (*cyan*), protein nitrogen (*red*) and K^+ ions (*magenta*). The total hydration number (*black*) is also shown. A cutoff of 4.0 Å is used to count the hydration partners.

Table 4.3. Average number of K^+ and Cl^- ions inside the pore and their ratios (n_{Cl}/n_K). Standard errors are also shown with the averages.

	n_K	n_{Cl}	n_{Cl}/n_K
n100	8.1±0.8	9.2±0.7	1.1±0.2
n75	6.9±0.6	9.5±0.4	1.4±0.2
n50	6.1±0.7	8.5±0.4	1.4±0.2
n25	7.0±0.3	10.0±0.3	1.4±0.1
0	5.4±0.5	7.5±0.6	1.4±0.3
p25	5.6±0.1	6.8±0.2	1.2±0.1
p50	5.8±0.2	7.6±0.9	1.3±0.2
p75	6.0±0.0	8.2±1.0	1.4±0.2
p100	6.5±1.1	7.8±0.3	1.2±0.3

The variation of K^+ and Cl^- density profiles correlates with the different hydration conditions for the ions along the pore axis. In channels with narrow passageways like gramicidin A [187] and KscA [188], ions are desolvated upon entering to the channel or the selectivity filter, and then form favorable interactions with the channel, i.e., the channel provides the solvation partners. To examine such hydration effects in hVDAC1 with a much larger aqueous pore, we calculated the hydration number as a function of the ion's *Z*-position. The total hydration number is further decomposed to the contributions from water molecules, counterions, and protein atoms (i.e., oxygen atoms for K^+ and nitrogen atoms for Cl^-). The hydration profiles for systems with $V_{mp} = 0$ are shown in Figures 4.12B and C. The total hydration numbers for both K^+ ($N_{K,total} = 8.3 \pm 0.2$) and Cl^- ($N_{Cl,total} = 8.0 \pm 0.1$) at any position along the *Z*-axis inside the pore are similar. These numbers are also similar to the total hydration numbers of both ion species in the bulk KCl solution ($N_{K,total} = 8.3 \pm 0.0$ and $N_{Cl,total} = 8.4 \pm 0.1$). Note that the protein hydration for K^+ reaches its peak in the middle of the channel where the highest K^+ density is found. This indicates that the K^+ ions are close to the channel wall, possibly as a result of

specific attractive interactions. In contrast, protein hydration of Cl^- is constant throughout the channel.

To determine the specific residues involved in ion hydration, the cross-sectional ion charge distributions were evaluated for three slabs along the pore axis: (A) -8.0 Å $\leq Z \leq$ -4.0 Å; (B) -2.0 Å $\leq Z \leq$ 2.0 Å; and (C) 4.0 Å $\leq Z \leq$ 8.0 Å (Figure 4.13). Note that the charge distributions result from the summation of both K^+ and Cl^-, so that zero density at a certain region reflects the fact that there are equal amounts of K^+ and Cl^- in the region. Near the openings of the channel (A and C), the overall positive charge density (ρ_+) is relatively small, with the maxima at 6.0×10^{-3} e/Å^3 in both slab A and C. The major ρ_+ are distributed in two zones on the *X*-*Y* plane. One is in the vicinity of the first five β-strands of hVDAC1, and the other is close to β-strands #11 to #14, which are on the opposite side of the channel wall (Figure 4.13). While ρ_+ has a clear position preference, the negative charge density (ρ_-) is distributed throughout the pore, with slightly higher values near the channel wall due to the protein-ion interactions. The scenario shifts in the middle of the channel (slab B), where concentrated positive charge is found. The maximum value of ρ_+ reaches 1.7×10^{-2} e/Å^3, 4.3 times larger than the maximum ρ_- in the slab and 2.8 times larger than the maximum ρ_+ observed in slabs A and C. Further investigation indicates that this accumulation arises from attractive electrostatic interactions between K^+ and negatively charged residues on the first five β-strands, especially Asp30 and Glu84. Interestingly, mutations of the corresponding residues, Asp30 to Lys and Lys84 to Glu, in VDAC1 from *Saccharomyces cerevisiae* increase and decrease the anion selectivity of the channel respectively [189]. The present simulation studies of hVDAC1 show that these two residues form strong attractive interactions with

K^+, which make K^+ permeation less effective. In contrast, Cl^- ions pass the channel relatively freely without strong interactions with residues on the channel wall. Consequently, hVDAC1 conducts more Cl^- ions than K^+ ions in a given time interval, resulting in the observed anion selectivity.

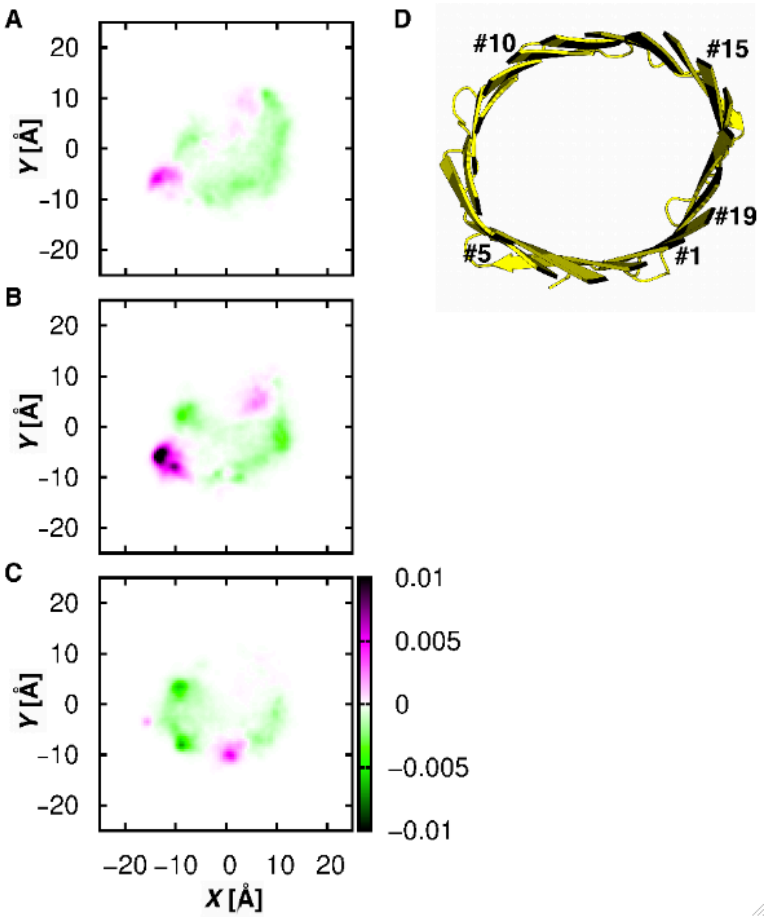


Figure 4.13. Cross-sectional ion distributions on the *X*-*Y* plane in three regions along the *Z*-axis: (*A*) 4.0 Å ≤ *Z* ≤ 8.0 Å, (*B*) -2.0 Å ≤ *Z* ≤ 2.0 Å, and (*C*) -8.0 Å ≤ *Z* ≤ -4.0 Å. The charge distributions are calculated from S2_0. (*D*) A molecular representation of hVDAC1 channel with its β-strands numbered. The first β-strand from the N-terminus of the channel is numbered as #1.

An alternative way of examining the selectivity mechanism is to characterize the pathway that each ion type takes to travel across the channel. In a wide channel pore

containing multi-ions, an overlap of ion positions during the simulation offers information on the average pathway of ions. For this reason, we superimposed ion trajectories from the last 10 ns in system S1_n100 (Figure 4.14). Although it is not quantitative, the superposition of ions demonstrates approximately the three-dimensional ion density inside the pore. From Figure 4.14, there are three clusters of the most populated K^+ ions: the first is near Asp30 on β-strand #1 toward the channel opening with $Z < 0$; the second is near Glu84 on β-strand #5 in the middle of the channel; and the third is between Asn207 on β-strand #14 and Asp16. These clusters are consistent with the observed positive charge density in Figure 4.13. There are no clearly defined paths for K^+ and Cl^- in hVDAC1, compared to the ones observed in the OmpF channel from *E. coli* [182]. The hVDAC1 channel has an aqueous conducting pathway and transports both ion species. However, the pathway for K^+ is more restricted than that of Cl^-. K^+ ions enter the channel from either side, but they cannot proceed directly as do Cl^- ions, although the two have similar averaged diffusion constants inside the pore. Once the K^+ ions are in the pore, they start to form attractive interactions with the aforementioned four residues. As a result, their effective diffusion is reduced. Alternatively, the passageway of Cl^- ions is not as restricted. The majority of Cl^- ions pass through the channel with minimal interactions with the protein, and therefore their effective diffusion is faster than that of the K^+ ions. The difference in the effective diffusion rates of K^+ and Cl^- determines the anion selective nature of the channel.

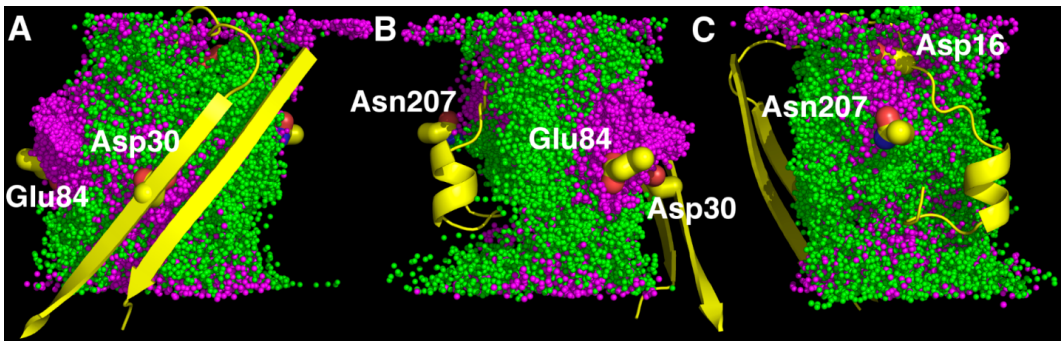


Figure 4.14. Overlay of ion trajectories for the last 10-ns simulations in S1_n100. The trajectories are sampled every 10 ps, resulting in a superposition of 1,000 snapshots. (*A*), (*B*) and (*C*) are three different views by 120° rotation. K^+ (*magenta*) and Cl^- (*green*) are shown in sphere presentation. To indicate the positions of involved residues, a reference segment (*yellow*) of the channel is also shown, including residues numbering from 1 to 36 and from 273 to 285. Important residues (*yellow*) such as Asp16, Asp30, Glu84 and Asn207 are presented in stick representation.

Although it is suggested that the channel's anion selectivity mainly arises from the net positive charges on the pore lining [155], the proposed mechanism in the present study supplements rather than contradicts this prevailing view. Figure 4.15 shows the one-dimensional multi-ion PMF [190], $\mathcal{W}_{\alpha,1D}(Z)$, for both K^+ and Cl^- from their equilibrium distributions along the *Z*-axis, calculated using

$$\mathcal{W}_{\alpha,1D}(Z) = -k_B T \ln\left[C_\alpha(Z)/C_{ref}\right], \tag{4-4}$$

where α is the ion type and C_{ref} is set to the bulk concentration so that $\mathcal{W}_{\alpha,1D}(Z) = 0$ in the bulk solution. $\mathcal{W}_{\alpha,1D}(Z)$ represents the mean potential acting on an ion along the channel axis, including contributions from protein-ion, ion-ion, and ion-solvent interactions and the volume exclusion effect (i.e., variation in ion accessible cross-sectional area). As shown in Figure 4.15, the free energy barriers for Cl^- at both entrances of VDAC are slightly lower than that for K^+, probably due to the net positive charges of the pore. This

limits K^+ from entering the channel. In addition, the K^+ PMF shows a free energy well with a depth of 0.7 kcal/mol in the middle of the channel, resulted from the attractive interactions between K^+ and negatively charged residues including Asp30 and Glu84. Compared to the Cl^- PMF, which is smooth inside the pore, such a free energy well for K^+ makes the K^+ permeation slow. These two features of the K^+ PMF are likely to be the origin of hVDAC1's mild anion selectivity. The spatial preference of both ion species inside the pore can also be illustrated by the two-dimensional multi-ion PMF maps along the pore axis (Figure 4.16). Both the observation of homogeneous Cl^- PMF throughout the pore and the existence of potential wells for K^+ in the middle of channel further corroborate the proposed selectivity mechanism, which is largely dictated by electrostatic interactions between ions and the charged residues on the channel wall. In fact, the importance of electrostatics underlying VDAC's anion selectivity is clearly illustrated by the electrostatic potential maps calculated from the Poisson-Boltzmann equation [191] (Figure 4.17).

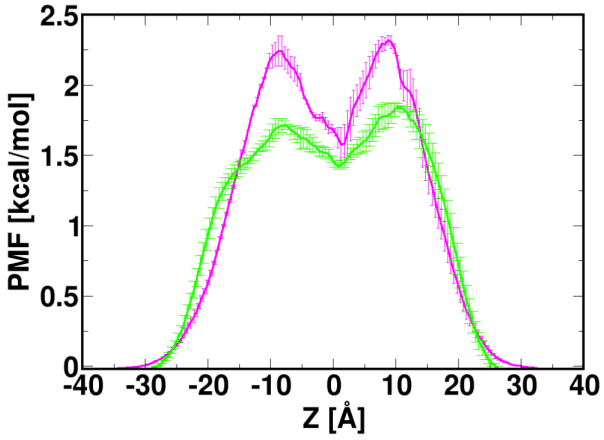


Figure 4.15. The averaged one-dimensional multi-ion PMF for K^+ (*magenta*) and Cl^- (*green*) calculated from simulations at zero transmembrane potential.

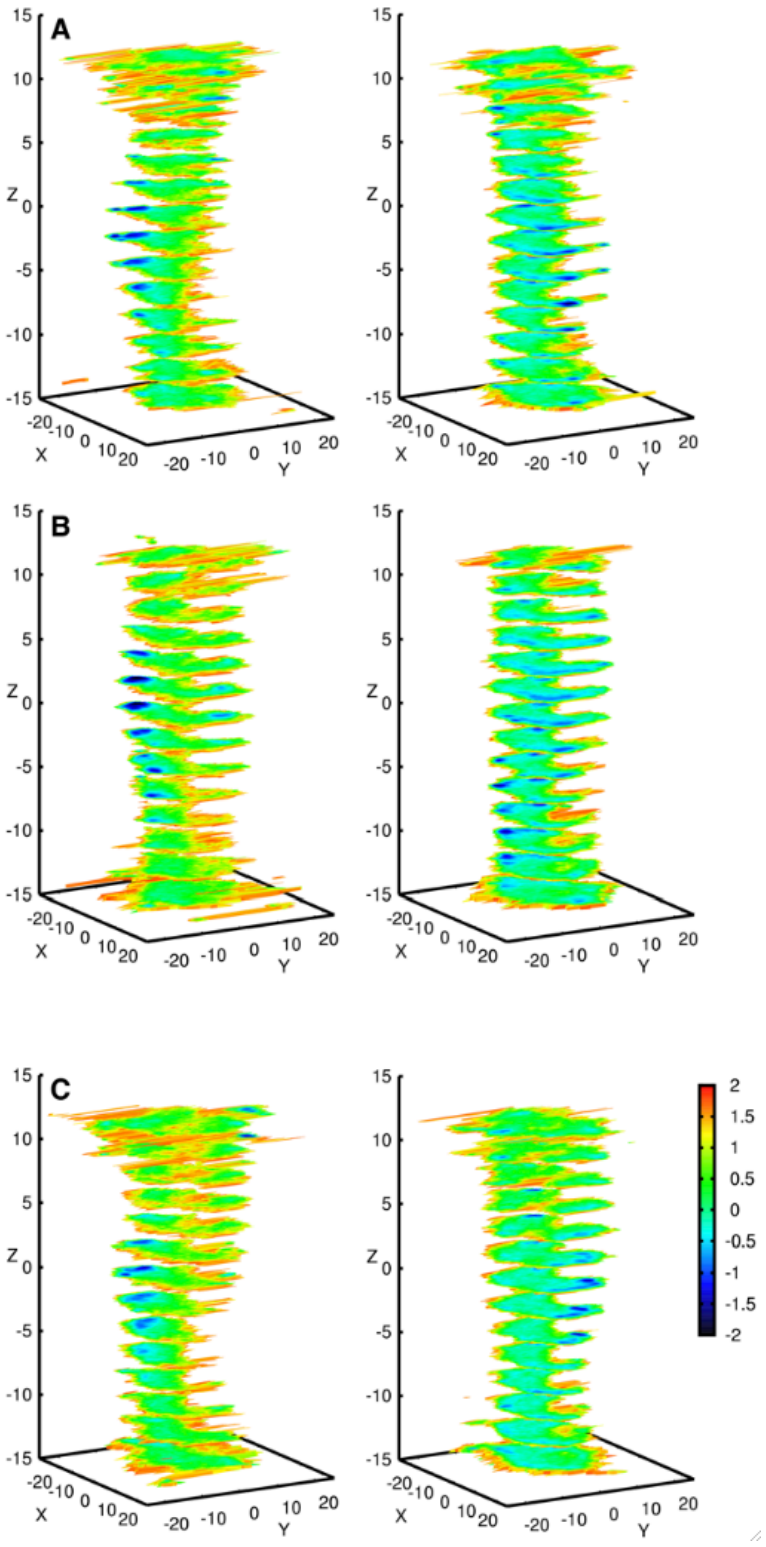


Figure 4.16. Two-dimensional multi-ion PMF profiles [in kcal/mol] for both K^+ (*left*) and Cl^- (*right*) in systems (*A*) S1_0, (*B*) S2_0, and (*C*) S3_0. The PMF profiles are calculated using $\mathcal{W}_{\alpha,\mathbf{2D}}\left(X,Y\right) = -k_{\mathbf{B}}T\,\mathbf{ln}\left[C_{\alpha}\left(X,Y\right)\middle/C_{\mathbf{ref}}\right]$ for 2 Å slabs along the Z-axis,

ranging from -14 Å to 14 Å, in which $k_{\mathbf{B}}$, $C_{\alpha}\left(X,Y\right)$, and C_{ref} are the Boltzmann constant, local ion concentration at (X, Y), and the bulk ion concentration. The PMF plots from different simulations are very similar, indicating a good convergence of ion distribution inside the pore.

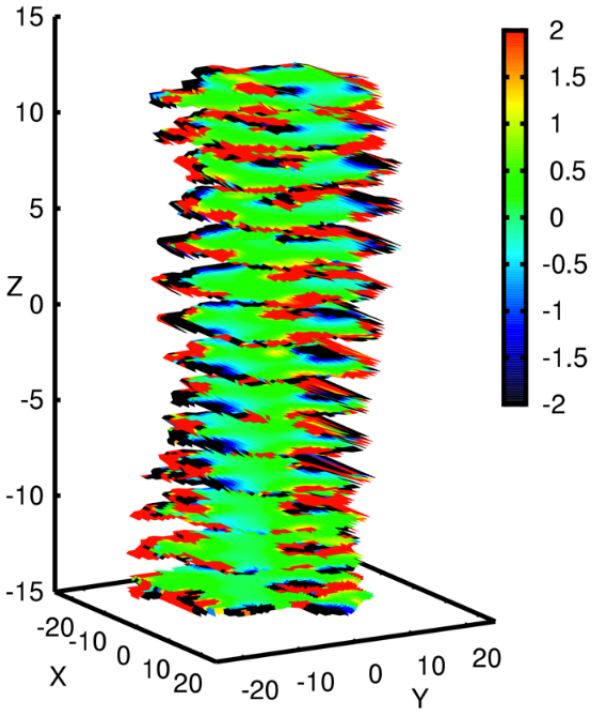


Figure 4.17. Two-dimensional electrostatic potential profile probed by a unit charge inside the channel pore calculated based on hVDAC1 NMR model #1. The electrostatic potential is computed using the PBEQ solver module in CHARMM-GUI. Regions with low electrostatic potential (*blue*) are favored by K^+ ions while zero electrostatic potential (*green*) indicates there is no specific preference for different ion species. The electrostatic potential profiles are well correlated with the calculated two-dimensional multi-ion PMF.

4.4. Summary and Conclusions

To investigate the ion permeation and selectivity of hVDAC1 and its underlying mechanisms, we have performed sets of three independent all-atom MD simulations of hVDAC1 embedded in DOPE/DOPC/cholesterol mixed membrane bilayers at TM potentials of 0, ±25, ±50, ±75, and ±100 mV. We generated a total of 1.77-μs trajectories and analyzed them to elucidate the channel's stability, electrophysiological properties, ion diffusion constants inside the channel pore, and a possible ion selectivity mechanism.

The overall hVDAC1 structure remained stable in each simulation, as corroborated by its backbone RMSD and RMSF (Figures 4.5A and B). The qualitative agreement between the RMSF calculated from different structural information (i.e., the NMR models and the MD trajectories) illustrates adequate sampling of hVDAC1 conformations.

The good stability and sampling of hVDAC1 allows for the evaluation of important properties of the channel with acceptable accuracy. The channel conductance, ion current ratio, and ion diffusion constant can be readily calculated. The conductance and ion current ratio of hVDAC1 are in good agreement with experimental measurements, revealing a functional channel with a conductance level around 4.8 nS and an approximate 2:1 selectivity between Cl^- and K^+ (Figures 4.9 and 4.10). Position dependent ion diffusion constants (Figure 4.11) were obtained from a shifted Gaussian distribution (Equation 4-3) in order to separate stochastic and systematic forces. These diffusion constants can be used in BD or PNP calculations for rapid evaluations of the channel's electrophysiological properties.

Based on the calculated ion density profile, ion hydration profile, cross-sectional ion charge density distributions, and position dependent diffusion constants, we propose a possible mechanism for ion permeation and selectivity of hVDAC1. In this mechanism, both ion species can travel across the channel. The determinant of the channel's anion selectivity is the transport efficiency of different ions. Cl^- is transported more efficiently than its counterions, which are slowed down by attractive ion-protein interactions in the middle of the pore. Residues involved in K^+-protein interactions are Asp16, Asp30, Glu84 and Asn207 (Figures 4.12, 4.13, and 4.14). Mutation studies also suggest the importance of Asp16, Asp30 and Glu84 in yeast VDAC anion selectivity [189]. These residues appear to serve as traps for K^+ ions as they pass through the middle of the channel. It is such specific K^+-protein interactions that increase the average time K^+ ions spend in the pore. Together with the higher free energy barrier for K^+ to enter the channel (Figure 4.15), the K^+-protein attractive interactions results in the difference in ion transport rate for K^+ and Cl^- and determines the anion selectivity of hVDAC1.

Chapter V

Protein Dynamics and Ion Traffic in Bacterioferritin[3]

[3] Reused from *Biochemistry*, Huan Rui, Mario Rivera, and Wonpil Im, **51**. 2012. pp 9900-9910. Copyright (2012). With permission from American Chemical Society.

Summary

Bacterioferritin (Bfr) is a spherical protein composed of 24 subunits and 12 heme molecules. Bfrs contribute to regulate iron homeostasis in bacteria by capturing soluble but potentially toxic Fe^{2+} and by compartmentalizing it in the form of a bioavailable ferric mineral inside the protein's hollow cavity. When iron is needed, Fe^{3+} is reduced and mobilized into the cytosol as Fe^{2+}. Hence, key to the function of Bfr is its ability to permeate iron ions in and out of its interior cavity, which is likely imparted by a flexible protein shell. To examine the conformational flexibility of Bfrs in a native-like environment and the way in which the protein shell interacts with monovalent cations, we have performed molecular dynamics (MD) simulations of BfrB from *Pseudomonas aeruginosa* (Pa BfrB) in K_2HPO_4 solutions at different ionic strengths. The results indicate the presence of coupled thermal fluctuations (dynamics) in the 4-fold and B-pores of the protein, which is key to enable passage of monovalent cations through the protein shell using B-pores as conduits. The MD simulations also show that Pa BfrB ferroxidase centers are highly dynamic and permanently populated by transient cations exchanging with other cations in the interior cavity, as well as the solution bathing the protein. Taken together, the findings suggest that Fe^{2+} traffic across the Pa BfrB shell via B-pores and that the ferroxidase pores enable capture and oxidation of Fe^{2+}, followed by translocation of Fe^{3+} to the interior cavity, aided by the conformationally active H130.

5.1. Introduction

Iron is the fourth most abundant metal on Earth and an indispensable cofactor in the active sites of proteins and enzymes, where it functions as an integral component of physiological processes such as respiration, DNA synthesis, gene regulation, degradation of peroxides and superoxides, oxygen transport and oxygen activation, and degradation of xenobiotics [192]. Fe^{2+} is soluble (~ 0.1 M at pH 7.0) but readily oxidized by molecular oxygen to the Fe^{3+} form, which is extremely insoluble (~10^{-18} M at pH 7.0) [192, 193]. The oxidation of Fe^{2+} by O_2 generates hydrogen peroxide, which is also a byproduct of aerobic respiration; Fe^{3+} can be reduced to Fe^{2+} by reducing agents in the cell and rapidly react with hydrogen peroxide to produce Fe^{3+} and the hydroxyl radical, thus creating a cycle conducive to the formation of the highly reactive $^{\cdot}OH$ which indiscriminately attacks biological molecules [193, 194]. Thus, the potential toxicity of Fe^{2+} and the insolubility of Fe^{3+} pose enormous challenges to living cells. These challenges have been largely ameliorated by evolution of the ferritins and ferritin-like molecules, which specialize in the handling and storage of iron. These molecules function by oxidizing the soluble but potentially toxic Fe^{2+} using O_2 or H_2O_2 as electron acceptors and by storing the otherwise insoluble Fe^{3+} in the form of a mineral in a compartment (internal cavity) that is isolated from cellular processes. When necessary, iron stored in the internal cavities of ferritin-like molecules can be reduced, mobilized from the interior cavity and incorporated into cellular processes [192].

The crucial contribution of ferritin-like molecules to iron metabolism is evident in their presence in all three domains of life with remarkable conservation of structure in spite of very low conservation in sequence [195]. Three types of ferritin-like molecules are found

in bacteria, bacterioferritin (Bfr), ferritin (Ftn), and DNA binding proteins from starved cells (Dps) [192]. Bfrs and Ftns are spherical proteins composed of 24 subunits that encase a hollow cavity of approximately 80 Å in diameter where the iron mineral is stored. Dps are also spherical proteins but are composed of 12 subunits encasing a hollow cavity approximately 45 Å in diameter. Structures of Bfr from several organisms [196-202] reveal that the canonical fold of a Bfr subunit (Figure 5.1A) is composed of a four-helix bundle (helices A to D) and a short C-terminal helix E nearly perpendicular to the bundle. Each subunit harbors a ferroxidase center in which a di-Fe^{2+} moiety is oxidized to a di-Fe^{3+} species, prior to their storage as an iron mineral in the Bfr interior cavity. A unique property of the Bfrs, which only occur in archaea and in bacteria [195], is that they bind heme between two subunits (subunit-dimer) (Figure 5.1B) such that the 24-mer protein consists of 12 subunit dimers and 12 heme molecules.

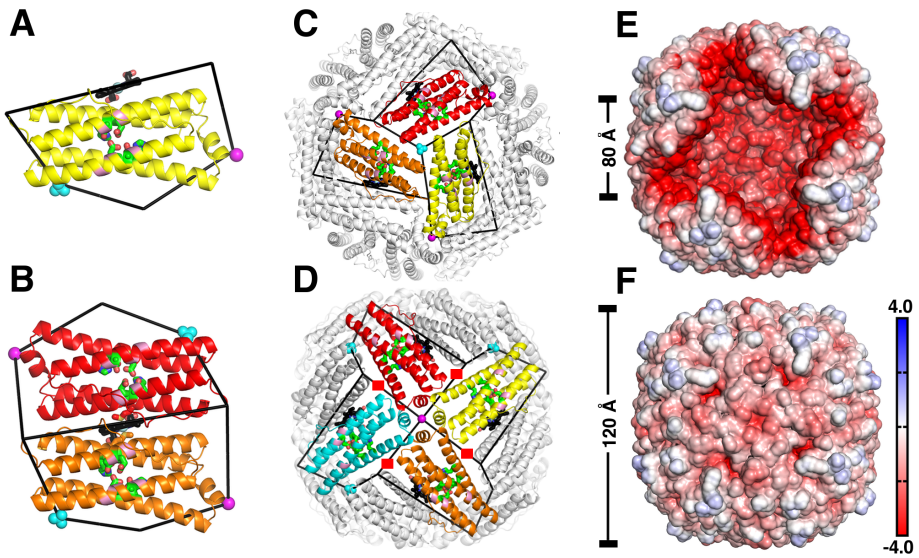


Figure 5.1. The bacterioferritin architecture. (A) View of a Pa BfrB subunit illustrating the ferroxidase center (*green*) and the ferroxidase pore (*pink*), which enables access of iron from the protein exterior to the ferroxidase center; heme is in *black*, K^+ in the 4-fold pore and PO_4^{3-} in the 3-fold pore are shown as magenta and cyan spheres, respectively. (B) Subunit dimer showing heme coordinated axially by a M52 from each subunit. (C)

View along a 3-fold symmetry axis depicting a 3-fold pore and a phosphate ion. (D) View along a 4-fold symmetry axis illustrating a 4-fold pore and the associated K^+ ion. Each 4-fold pore is surrounded by 4 B-pores, highlighted by red squares. (E) The electrostatic potential mapped on the interior surface of Pa BfrB. Four front subunits were removed for a clear view of the internal cavity. (F) The electrostatic potential mapped on the exterior surface of a Pa BfrB with the same orientation as in (E). Negative electrostatic potential is in red, positive electrostatic potential is in blue, and the unit is kcal/(mol·*e*). The figures were produced using the molecular visualization program PyMOL [167].

The protein shells of ferritin-like molecules function similar to cell membranes in that they separate the iron mineral from the cytosol. Four types of pores are thought to communicate the interior cavity of a 24-mer assembly with the exterior milieu: eight 3-fold pores (Figure 5.1C), six 4-fold pores (Figure 5.1D), a ferroxidase pore in each subunit (Figure 5.1A) and four B-pores located near each of the 4-fold pores (Figure 5.1D) [202]. Notably, the surface lining the interior cavity (interior surface) of Bfrs exhibits a very large negative potential (Figure 5.1E) [202], whereas the surface of the protein exposed to the protein exterior (exterior surface) shows less pronounced positive and negative electrostatic potentials (Figure 5.1F).

Structural studies showed that the 3-fold pores of eukaryotic ferritins are lined with conserved negatively charged residues. Substitution of these conserved residues caused decreased rates of iron uptake, leading to the suggestions that Fe^{2+} ions may traffic in and out of eukaryotic ferritins via 3-fold pores [203-206]. In comparison, the function of the 3-fold pores in the Bfrs, which are lined by layers of alternating positive and negative charge, is much less clear. Sulfate and phosphate ions have been observed in 3-fold pores in the crystal structures of *Pseudomonas aeruginosa* BfrB (Pa BfrB) [207], Pa FtnA [208] and *Desulfovibrio desulfuricans* Bfr (Dd Bfr) [209], suggesting that in the

Bfrs and bacterial Ftns the 3-fold pores may be conduits for anion traffic across the 24-mer shell. The 4-fold pores of Bfrs and bacterial Ftns are also structurally different from those of eukaryotic ferritins. The 4-fold pores in eukaryotic ferritins are lined by hydrophobic residues and are narrower than the 3-fold pores. In agreement, substantial barriers for the entry of Fe^{2+} and Na^{+} ions into eukaryotic ferritin 4-fold pores have been calculated [210]. In contrast, the 4-fold pores of Bfrs typically contain hydrophilic residues and in some instances are observed to bind monovalent and divalent cations in X-ray crystal structures. The presence of these positively charged ions in the 4-fold pores of Bfrs has led to the hypothesis that these channels may serve as conduits for the traffic of Fe^{2+} in and out of the Bfr shell [207, 211, 212]. The protein shells of bacterial ferritins (Bfr and Ftn) also have channels at the intersection of three subunits, not aligned with any axis of symmetry in the structure. These channels have been termed B-pores [209]. In all known Bfr structures the B-pores are lined by hydrophilic residues and have a relatively large number of negatively charged residues. Mg^{2+} ions coordinated by water molecules have been observed within the B-pores of *Azotobacter vinelandii* (Av) [212] and *Mycobacterium smegmatis* (Ms) Bfrs [196]. A more recent structure of Pa BfrB in complex with its cognate electron donor partner, Pa ferredoxin (Bfd), shows Na^{+} ions in the B-pores, where they are coordinated by three negatively charged residues lining the narrowest section of the pores [213]. Although the function of the B-pores is unknown, it has been noted that B-pores are sufficiently large to accommodate a Fe^{2+} ion, thus suggesting that B-pores may function to facilitate trafficking of iron in and out of Bfrs and bacterial Ftns.

Clearly, the different pores in the ferritin structure are likely important to the dual function of iron uptake and release exhibited by ferritins and ferritin-like molecules. However, current understanding of how the pores enable these functions is limited. Although the perception of ferritin-like molecules has evolved from that of a rigid cage to that of a dynamic supramolecular assembly, virtually nothing is known about their dynamic properties, and how dynamics enable their function. In an attempt to bridge these gaps, to the best of our knowledge, we report the first molecular dynamics (MD) simulation study of the 24-mer Pa BfrB. Results from these MD simulations reveal that significant cooperative dynamics at 4-fold pores and B pores are linked to the traffic of K^+ ions across the protein shell, using B-pores as conduits. Moreover, the simulations show highly dynamic ferroxidase centers permanently populated by transient K^+ ions, which can access the centers from the interior cavity as well as from the protein exterior. The traffic and distribution of K^+ observed during the simulations is discussed in the context of iron uptake and release from the Bfrs.

5.2. Computational Methods

5.2.1. Constructing the Pa BfrB system

The Pa BfrB structure used in the simulations (PDB:3ISF) [197] was retrieved using the *PDB Reader* module in CHARMM-GUI (www.charmm-gui.org) [171]. The terminal residues including M1, D157, and D158 were missing in some of the Pa BfrB subunits. These residues were added using the IC BUILD command in CHARMM [45] and were patched with acetylated N-termini and N-methylamide C-termini, respectively. Since the MD simulations of Pa BfrB were carried out in K_2HPO_4 solution (see below), PO_4^{3-} ions were added in the 3-fold pores in place of the SO_4^{2-} ions found in the crystal structure of the iron-soaked Pa BfrB structure (PDB:3IS8) [197]; this structure was not used in the simulations because it was obtained from Pa BfrB containing ~600 iron atoms inside the central cavity [197]. Hence, the structure used in the simulations consists of a protein shell made of 24 Pa BfrB subunits, 12 heme molecules, K^+ in each of the six 4-fold pores, and PO_4^{3-} in each of the eight 3-fold pores (Figure 5.1A-D).

5.2.2. Estimating initial ion positions from GCMC/BD simulations

Having a proper starting system configuration for a MD simulation is extremely important. In the case of Pa BfrB, how to place the ions to initialize the MD simulations is the key issue, because Pa BfrB has a net negative charge of +378 and an uncorrelated system could take a long time to equilibrate. Under such a scenario, randomly placing the ions through out a Pa BfrB may not be an optimal option. Therefore, GCMC/BD

simulations were employed to generate the initial ion positions that follow the electrochemical potential in the system for the subsequent MD simulations.

The structure of Pa BfrB was centered in the origin of a cubic simulation box with 160 Å on each side. The vectors connecting two 4-fold pores on the opposite sides of the protein shell were aligned with the *X*-, *Y*-, and *Z*-axis. The CHARMM-GUI *GCMC/BD ion simulator* module [214] was employed to calculate the grid-based potential maps for the grand canonical Monte Carlo / Brownian dynamics (GCMC/BD) simulations. The grid spacing was set to 0.5 Å. To account for the influence of different protein dielectric constants (ε_p) on the electrostatic potential, a series of ε_p (ε_p = 2, 4, 10, 20, and 40) were used in the calculations, resulting in five sets of potential maps. They were denoted by E2, E4, E10, E20, and E40. Each set contains the electrostatic potential ϕ_{sf}, the reaction field potential ϕ_{rf}, and the core-repulsive steric potential U_{core}. From these, a space dependent multi-ion potential of mean force, $\mathcal{W}(\{\mathbf{r}_i\})$, can be evaluated by

$$\mathcal{W}(\{\mathbf{r}_i\}) = \sum_i q_i\left[\phi_{sf}(\mathbf{r}_i) + \phi_{rf}(\{\mathbf{r}_i\})\right] + \sum_i U_{core}(\mathbf{r}_i) + \sum_i \sum_{j\neq i} U_{ij}(|\mathbf{r}_i - \mathbf{r}_j|), \tag{5-1}$$

where r_i is the position of the i^{th} ion and U_{ij} denotes the pair-wise ion-ion interactions [181]. $\mathcal{W}(\{\mathbf{r}_i\})$ was used to guide the GCMC/BD simulations following the BD equation of motion [215, 216],

$$\frac{d\mathbf{r}_i}{dt} = -\frac{D_i}{k_B T}\cdot\frac{\partial \mathcal{W}(\{\mathbf{r}_i\})}{\partial \mathbf{r}_i} + \zeta(t) \tag{5-2}$$

with D_i and $\zeta(t)$ representing the diffusion constant of the i^{th} ion and the stochastic water bombardment on ions.

Five GCMC/BD simulations (GCMC/BD_E2-E40) with 30,000 steps each were performed using the last step input file from the *GCMC/BD ion simulator* module [214]. The positions of the protein atoms, heme molecules, and crystal ions in the pores were fixed. In each simulation K^+ and Cl^- were added to the system from the GCMC buffer region equivalent in size to the simulation box and distributed by the subsequent BD simulations. K^+ instead of Fe^{2+} were used as the cation species in the GCMC/BD and the following MD simulations because of (i) the accessibility of accurate force field parameters for K^+ but not for Fe^{2+} and (ii) the similarity of K^+ and Fe^{2+} in that they are both soluble cations. The anion in the GCMC/BD simulations is Cl^-, which is different from the HPO_4^{2-} present in the MD simulations. This is because there are no accurate Lennard-Jones and short-range ion-ion interaction parameters for K_2HPO_4 available in GCMC/BD, but those parameters for KCl are well established and frequently used [181, 214]. The total number of ions after each cycle was plotted for each system (Figure 5.2). A snapshot of ion positions was selected after the ion number in each system reached a plateau value. To generate the initial ion configuration for the MD simulation systems, half of the Cl^- were randomly chosen and discarded, and the other half was replaced with HPO_4^{2-} to keep the total negative charge constant. Table 5.1 shows the number of ions in each system. Clearly, the resulting number of ions is strongly dependent on ε_p and increases with decreasing ε_p. Instead of choosing a particular ε_p, all systems with different ε_p are considered in this study to reflect the different ionic strengths.

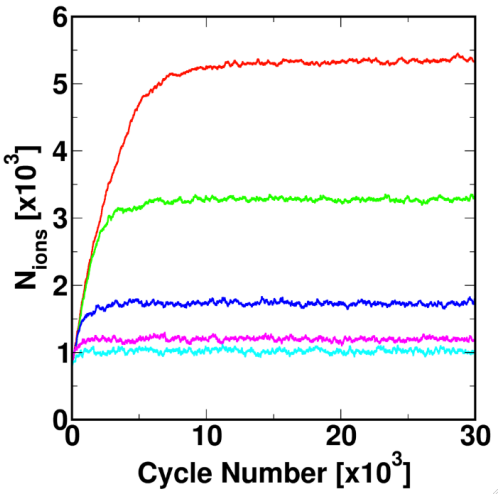


Figure 5.2. Total number of ions (N_{ions}) during GCMC/BD simulations for systems GCMC/BD_E2 (*red*), _E4 (*green*), _E10 (*blue*), _E20 (*magenta*), and _E40 (*cyan*).

5.2.3. MD simulations of Pa BfrB in K_2HPO_4 Solution

In the preparation of the all-atom MD simulations, each of the five resulting systems from the GCMC/BD simulations (MD_E2-E40) was immersed in a pre-equilibrated water box of the same size as in the GCMC/BD systems (160 × 160 × 160 $Å^3$). Water molecules within 2.4 Å from the ions and the Pa BfrB heavy atoms were removed. The systems were then subjected to a 210-ps equilibration cycle with decreased positional harmonic restraints on heavy atoms excluding water oxygen. During the equilibration process some water molecules infiltrated the Pa BfrB shell, creating vacuum pockets in the K_2HPO_4 solution inside and outside Pa BfrB. Therefore, the water molecules in the systems were removed after the first cycle of equilibration followed by addition of water in the manner described above except that 2.0 Å was used as a distance cutoff instead of 2.4 Å, and a second cycle of 780-ps equilibration was conducted. At the end of this cycle, small vacuum regions were still observed, but only in the central cavity. Therefore, a

sphere of pre-equilibrated water with a radius of 40 Å was appended, and newly added water molecules within 2.4 Å of any heavy atoms in the system were deleted. The whole system was then equilibrated without restraints for 20 ps, and was followed by a 40-ns production. The system information is summarized in Table 5.1.

Table 5.1. Numbers of K^+ ($N_{K\text{-}total}$), HPO_4^{2-} ($N_{phos\text{-}total}$), cavity K^+ ($N_{K\text{-}cavity}$), cavity HPO_4^{2-} ($N_{phos\text{-}cavity}$), water molecules (N_{water}), and total atoms ($N_{atom\text{-}total}$) in the MD simulations systems.

	$N_{K\text{-}total}$	$N_{phos\text{-}total}$	$N_{K\text{-}cavity}$	$N_{phos\text{-}cavity}$	N_{water}	$N_{atom\text{-}total}$
E2	2,840	1,231	1,430	0	110,172	404,112
E4	1,800	711	918	1	112,343	406,465
E10	1,042	332	512	1	113,992	408,380
E20	788	205	334	2	114,487	408,849
E40	714	168	246	3	114,478	408,526

All the simulations were carried out in NPT ensembles using NAMD2.8 [217] with the CHARMM all-atom parameter set PARAM22 [84] including the dihedral cross-term corrections (CMAP) [85] and a modified TIP3P water model [218]. The van der Waals interactions were smoothly switched off at 10-12 Å by a force switching function [141] and the electrostatic interactions were calculated using the particle-mesh Ewald method with a mesh size of ~1 Å. The temperature (300 K) and pressure (1 atm) were kept constant during all the simulations by Langevin dynamics and the hybrid Nose-Hoover Langevin piston method, respectively. The Langevin damping coefficient was set to 1 ps^{-1}; the decay period and damping timescale were 50 fs and 25 fs, respectively.

5.2.4. Principal Component Analysis

Principal component analysis (PCA) was performed to investigate the intrinsic dynamics of the Pa BfrB monomers and the assembly. Prior to calculating the covariance matrix for PCA, we superimposed all the Pa BfrB structures to the starting structure to remove the protein's translation and rotation during the simulations. For each of the Pa BfrB monomers, the coordinates of all the C_α atoms (**R**) were recorded at each trajectory snapshot,

$$\mathrm{R} = \begin{bmatrix} r_{1,1} & r_{2,1} & r_{3,1} & \cdots & r_{3N-2,1} & r_{3N-1,1} & r_{3N,1} \\ r_{1,2} & r_{2,2} & r_{3,2} & \cdots & r_{3N-2,2} & r_{3N-1,2} & r_{3N,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ r_{1,T} & r_{2,T} & r_{3,T} & \cdots & r_{3N-2,T} & r_{3N-1,T} & r_{3N,T} \end{bmatrix} \tag{5-3}$$

where $r_{3n-2,t}$, $r_{3n-1,t}$, and $r_{3n,t}$ represent the X, Y, Z-coordinates of the n^{th} C_α atom at time t, and N (N = 158) is the total number of C_α atoms in a subunit. Using **R**, a covariance matrix (Σ) can be constructed,

$$\Sigma = \begin{bmatrix} \mathrm{cov}(r_1,r_1) & \mathrm{cov}(r_2,r_1) & \cdots & \mathrm{cov}(r_{3N},r_1) \\ \mathrm{cov}(r_1,r_2) & \mathrm{cov}(r_2,r_2) & \cdots & \mathrm{cov}(r_{3N},r_2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}(r_1,r_{3N}) & \mathrm{cov}(r_2,r_{3N}) & \cdots & \mathrm{cov}(r_{3N},r_{3N}) \end{bmatrix} \tag{5-4}$$

where the covariance of r_i and r_j, $\mathrm{cov}(r_i,r_j)$, over a total of T time points is defined by

$$\mathrm{cov}(r_i,r_j) = \frac{\sum_{t=1}^{T}(r_{i,t}-\bar{r}_i)(r_{j,t}-\bar{r}_j)}{T-1} \tag{5-5}$$

and $\overline{r_i}$ and $\overline{r_i}$ are the time averages of all the $r_{i,t}$ and $r_{j,t}$ respectively. The eigenvalues and eigenvectors of the matrix Σ were computed. The eigenvector corresponding to the largest eigenvalue was then recorded and grouped into three groups for each C_α atom. The vector formed by these three values gives the principal direction of motion of a specific C_α.

The calculation of the principal motional modes of the entire Pa BfrB assembly is rooted on the same concept. Rather than using the coordinates of all the C_α atoms in the assembly, the center of mass coordinates of helical fragments were used. To distinguish motions occurring in different portions on helices A to D, each of them was divided into three fragments with each one containing approximately 10 residues. Helix E contains only six residues and thus the entire helix was used. The center of mass coordinates of these helical fragments were recorded and used to construct the **R** matrix with N = 13. The rest of the calculation follows the description given above.

5.3. Results and Discussion

5.3.1. Structure stability and dynamics of Pa BfrB

The calculated backbone root-mean-squared deviation (RMSD) of Pa BfrB bathed in K_2HPO_4 solution indicates that the protein is stable during the simulations (Table 5.2). Even the systems with most flexibility (E2 and E4) show RMSD values well under 4.0 Å. The greater RMSD in these systems likely stems from enhanced flexibility in specific regions of a Pa BfrB subunit, as can be seen in the per-residue RMS fluctuations (RMSF) in Figure 5.3A-C. Nonetheless, the characteristic 432-point symmetry of the Bfr assembly is preserved during the simulations, which is a strong indicator that the integrity of the Pa BfrB structure is well maintained.

Table 5.2. BfrB Backbone RMSD (Å) in the MD simulation systems.

	E2	E4	E10	E20	E40
MD	3.8±0.4	2.3±0.1	2.3±0.2	1.7±0.1	1.6±0.0

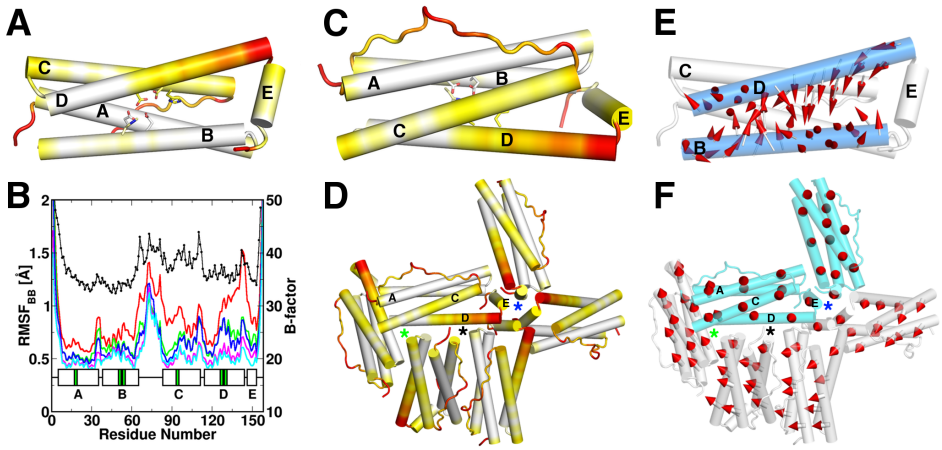


Figure 5.3. Intrinsic flexibility in Pa BfrB. The per-residue RMSF observed in system MD_E2 mapped onto a subunit viewed from the (A) interior and (C) exterior surface of

Pa BfrB. Flexibility increases in the color scale from white to red. (B) Per-residue RMSF in a subunit of Pa BfrB plotted for systems MD_E2 (*red*), E4 (*green*), E10 (*blue*), E20 (*magenta*), and E40 (*cyan*). Per-residue B factors from the X-ray structure (PDB:3ISF) are shown in the black trace; helices A to E are indicated as boxes and ferroxidase center residues are highlighted in green. (D) Per-residue RMSF mapped on six subunits of the 24-mer assembly show relatively large dynamic behavior near 4-fold and B-pores, and considerable less dynamic activity near 3-fold pores. To facilitate visualization, the pores are highlighted by black (B-), green (3-fold) and blue (4-fold) stars. (E) Principal motional mode of C_α atoms on helices B and D (*blue*) in subunit H in MD_E40 shown by arrows. (F) The principal motional mode of six subunits of the 24-mer assembly shown in the same orientation as in (D). The helices were segmented into 10-residue pieces and the motions were calculated based on the centers of mass of these structural segments. The arrows (*red*) indicate the directions of motion. Subunits exhibiting overall motion along and perpendicular to the assembly plane are colored *white* and *cyan*, respectively.

A per-residue RMSF plot (Figure 5.3B) illustrates the relative fluctuations in a Pa BfrB subunit for each of the MD systems (E2, E4, E10, E20, and E40). In addition to the N- and C-termini, it is apparent that specific regions of a subunit are significantly more dynamic than average. The relative extent of per-residue RMSF mapped onto a BfrB subunit is shown in Figure 5.3A and C with white representing the lowest propensity and red the highest propensity to fluctuate. In addition to the N- and C-termini, the long loop connecting helices B and C (BC loop) and the C-terminus of helix D exhibit the highest tendency to fluctuate, followed by the C-terminal half of helix D, and by the N-terminal half of helix C. The sections with higher than average flexibility also exhibit higher B-factors than the rest of the subunit in the Pa BfrB crystal structure (black trace in Figure 5.3B). It is also interesting that the mobility of these regions increases with the increase of K_2HPO_4 concentration (Figure 5.3B). Nevertheless, analyses of the principal motional modes of C_α atoms reveal absence of correlated motions within each of the helices and among different helices in a subunit (Figure 5.3E).

Although no clear functional implications can be derived from fluctuations within an individual subunit, their significance emerges in the context of the 24-mer assembly. This is illustrated in Figure 5.3D by a minimal assembly of six Pa BfrB subunits containing a 3-fold pore (green star), a 4-fold pore (blue star), and a B-pore (black star). Clearly, the regions that exhibit increased mobility in each of the subunits constitute the pores in the Bfr structure. This is particularly evident in the area surrounding the 4-fold pores, including the B-pores. In comparison, the regions comprising the 3-fold pores have much less dynamical activity. These observations suggest that while the less mobile sections of a Pa BfrB subunit may be important to maintain the structural integrity of the four-helix bundle and the 24-mer assembly, the fluctuations of high mobility residues impart the dynamic behavior to the pores in the Bfr shell. This aspect is further elaborated below after some salient points regarding each of the pores have been addressed. In addition to the fluctuation of residues near the pores, neighboring subunits can also adjust their orientations relative to one another, effectively creating breathing motions in the 24-mer protein. Figure 5.3F illustrates the principal motional modes of subunits within the minimal Pa BfrB assembly that includes each of the 3-fold, 4-fold and B-pores. Motions of subunits that are in immediate vicinity are usually coupled (share similar directionality), although in some cases adjacent subunits may move in orthogonal directions.

The 3-fold, 4-fold, and B-pore conformations are mostly maintained during the simulations. As shown in Figure 5.4A, the 3-fold pore is lined by three subunits; specifically, the C-terminus of helix C forms the outermost layer of the pore, while the N-terminus of helix D forms the innermost layer. Near its narrowest section, there is a PO_4^{3-}

replacing the SO_4^{2-} observed in the crystal structure. During the course of the simulations, the PO_4^{3-} is always coordinated by the side chains of R117 and K121 (Figure 5.4A) and does not exit the 3-fold pores. As shown in Figure 5.4B, a 4-fold pore is formed by four E-helices adopting an orientation that is more or less parallel to the pore axis. The BfrB crystal structure contains a K^+ ion in each 4-fold pore where it is coordinated by the side chains of N148 and Q151. During the simulations many of these K^+ stay inside the 4-fold pores as in the crystal structure (Figure 5.4B). In some cases, however, K^+ ions exit the pore. When this happens, the side chains of N148 and Q151 in E-helices located on opposite sides of the pore interact, leading to temporary obstruction of the 4-fold pore. Unlike 3-fold or 4-fold pores, B-pores are not at a crystallographic axis of symmetry. As shown in Figure 5.4C, the side chains of D34, E60, E66, D132, and E129 in BfrB protrude the pore interior thus creating the narrowest section of the B-pores. In the course of all simulations K^+ ions traverse B-pores aided by transient coordinative interactions with these negatively charged residues (Figure 5.4C). The fashion in which K^+ ions traffic through B-pores in our simulations is in agreement with recent experimental observations showing that Na^+ ions are coordinated by D34 from one subunit and by D132 and T136 from another subunit in the B-pores of Pa BfrB in complex with its cognate electron transfer Pa Bfd [213]. Along the same vein, Mg^{2+} ions have also been observed within the B-pores of Av BfrB [212] and Ms BfrB [196], where they are coordinated by water molecules. These observations taken together strongly suggest that B-pores are conduits for ion traffic in and out of the Pa BfrB central cavity.

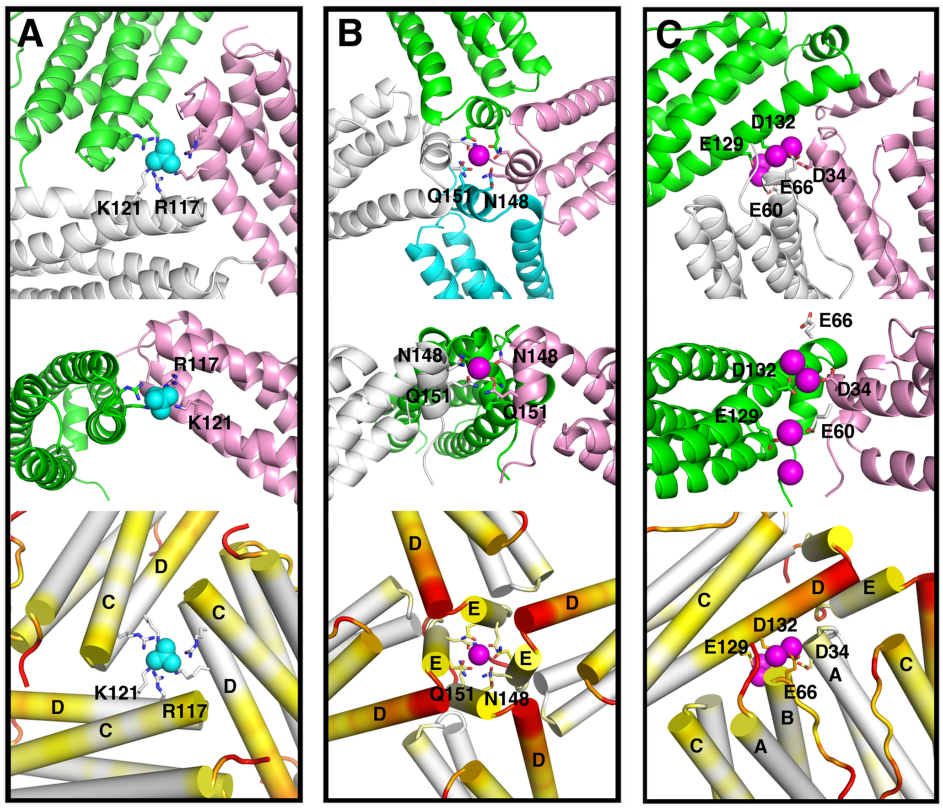


Figure 5.4. Architecture and dynamic properties of the pores in BfrB taken from a snapshot in system MD_E40. (A) 3-fold, (B) 4-fold, and (C) B-pores. Each pore type is shown viewed from the protein exterior (top) and from a transverse perspective (middle) with the internal cavity at the bottom and with the subunit closer to the viewer removed for clarity. (Bottom) The intrinsic flexibility of each type of pore is highlighted by mapping the RMSF fluctuations (see Figure 5.3B) from lowest (white) to highest (red). The phosphate ions in the 3-fold pores (cyan spheres) are coordinated by the side chains of R117 and K121 from each of the three subunits forming the pore. The K^+ ions (magenta spheres) in the 4-fold pores are coordinated by N148 and N151 from each of the four E-helices forming the pore. B-pores harbor negatively charged residues including D34, E60, E66, D132, and E129. These residues attract K^+ ions and translocate them across the protein shell during the simulations.

The bottom row in Figure 5.4 clearly illustrates that 3-fold pores show modest dynamic behavior. In contrast, a significant concentration of fluctuating domains is observed in the external perimeter of 4-fold pores, extending toward B-pores. Key residues contributing to the enhanced dynamics observed in these regions are located in

the C-terminus of helix D and the loop connecting helices D and E; specifically, residues Q137 to G145. Moreover, in high ion concentration systems (e.g., MD_E2) the C-termini of helices D undergo frequent transitions between folded and unfolded states; Figure 5.5 documents this transition and its impact on a 4-fold and one of its surrounding B-pores. The unfolding events in helices D enhance the mobility of the entire E helices: one pair of diagonally opposed E helices comes close to one another while the other pair is separated further apart. When the C-termini of helices D refold, the corresponding E helices return to their original positions. In extreme cases, the folding and unfolding events yield a ratio of up to 0.58 for the distance between two diagonally opposed E helices in close proximity relative to the distance separating two E-helices moved further apart. (Figure 5.5B and C). Despite their dynamic nature, the 4-fold pores are mostly closed during the simulations as the deformation caused by their dynamics always places two E-helices in close contact. The folding-unfolding equilibrium of the C-terminal portions of helices D affects the conformation of B-pores as well. Unfolding of the C-termini in helices D expands the passageway across the B-pores (Figure 5.5B), while refolding either returns the pore to its initial closed state or keeps the pore open by deforming the nearby 4-fold pore (Figure 5.5C). The dynamic opening and narrowing of B-pores supports the notion that ions traffic across the Pa BfrB shell using B-pores as conduits. Further evidence will be provided in the following section where the analysis of ion distribution and motions is presented.

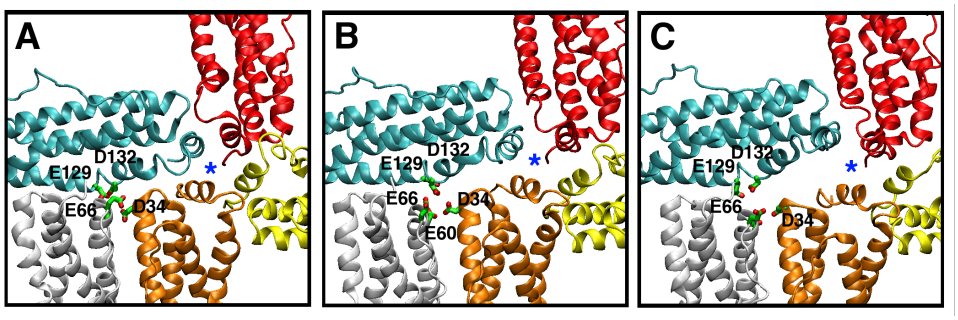


Figure 5.5. Snapshots illustrating the dynamic behavior of the 4-fold and B-pores. (A) A 4-fold pore (*blue star*) and one of its surrounding B-pores highlighted by its negatively charged residues in *green*, as seen in the crystal structure. (B) Melting of the C-terminal sections of helices D (Figure 5.3C) opens up the B-pore and increases the mobility of the entire E helix in the cyan subunit. This marks the start of the 4-fold pore deformation. (C) The E helix in the cyan subunit moves further away from its neighboring E helix (orange) toward the D-helix in the same subunit, which is now refolded. The displacement of the E helix results in deformation of the 4-fold pore and the further expansion of the B-pore. All the structure snapshots were taken along the MD_E2 trajectory.

5.3.2. Ion distributions in the Pa BfrB systems

Although the interior surface of the protein shell exhibits a relatively large negative electrostatic potential (Figure 5.1E), the largest concentrations of negatively charged residues occur in the middle of each subunit near the ferroxidase center and in the perimeters surrounding 3-fold, 4-fold, and B-pores. Because of the significant electrostatic potential gradients created by these residues, during the simulations, K^+ ions migrate toward the interior surface of BfrB and penetrate into the protein shell (black dashed lines in Figure 5.6) even though the majority of them stayed away from the protein shell in the GCMC/BD simulations with rigid BfrB (black line in Figure 5.6A). This is evident in Figure 5.6A, which plots the potassium ion concentration (C_K) along the radial distance (R) from the center of the interior cavity for systems GCMC/BD_E10 (black) and MD_E2-E40. Clearly, systems with different ionic strengths share similar

profiles of C_K, strongly suggesting that ion distribution has reached a steady state in all the systems and that the concentration profiles represent the real likelihood of having ions in different regions in a given system.

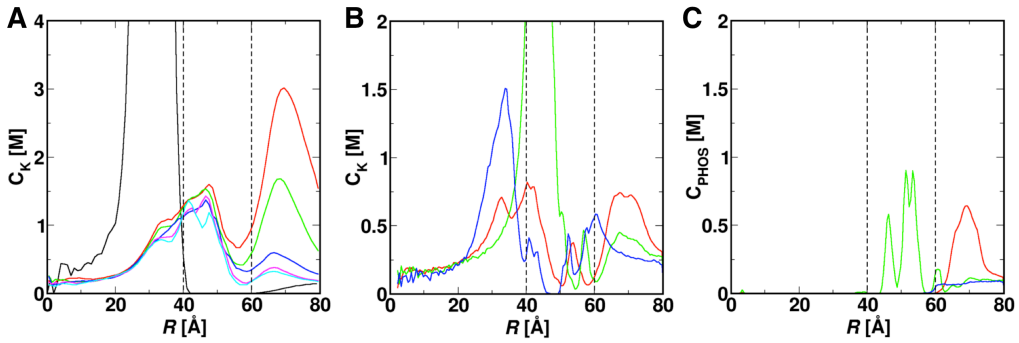


Figure 5.6. Potassium and phosphate ion concentration profiles (C_K and C_{PHOS}). C_K in (A) was calculated along the radial distance (R) from the center of the interior cavity with a bin size of 0.5 Å. The results for systems MD_E2 (*red*), E4 (*green*), E10 (*blue*), E20 (*magenta*), E40 (*cyan*), and GCMC/BD_E10 (*black*) are shown. To calculate C_K and C_{PHOS} in (B) and (C), a cone is constructed with apex at the center of BfrB interior cavity and axis collinear with the pore axis. The volume enclosed by the cone is then divided into 0.5-Å slabs along its axis and concentrations of K^+ and phosphate ions are computed inside each slab. Results are shown for the 3-fold (*green*), 4-fold (*blue*), and B-pores (*red*) in MD_E10. The estimated positions of the inner and outer protein surfaces along R are indicated by the dashed black lines.

K^+ also concentrate near the 3-fold, 4-fold, and B-pores. This is clearly shown in the C_K profiles along the axes of the 3-fold, 4-fold, and B-pores in Figure 5.6B. The peaks observed in the profiles at R = 32, 43 and 48 Å denote accumulation of K^+ near 4-fold, B- and 3-fold pores, respectively. The accumulation of K^+ near the three types of pores can be viewed from the protein interior (Figure 5.7A) and exterior (Figure 5.7C), where the K^+ density is rendered in mesh (magenta) and the HPO_4^{2-} density in cyan. Figure 5.7A also shows that the regions with high C_K, especially those near the 3-fold pores, B-pores, and ferroxidase centers are connected by bridges with lower C_K. Importantly, in all

simulations there is a dynamic flow that exchanges ions between these "reservoirs". There is also occasional exchange of K^+ between the 4-fold pore and the rest of the reservoirs.

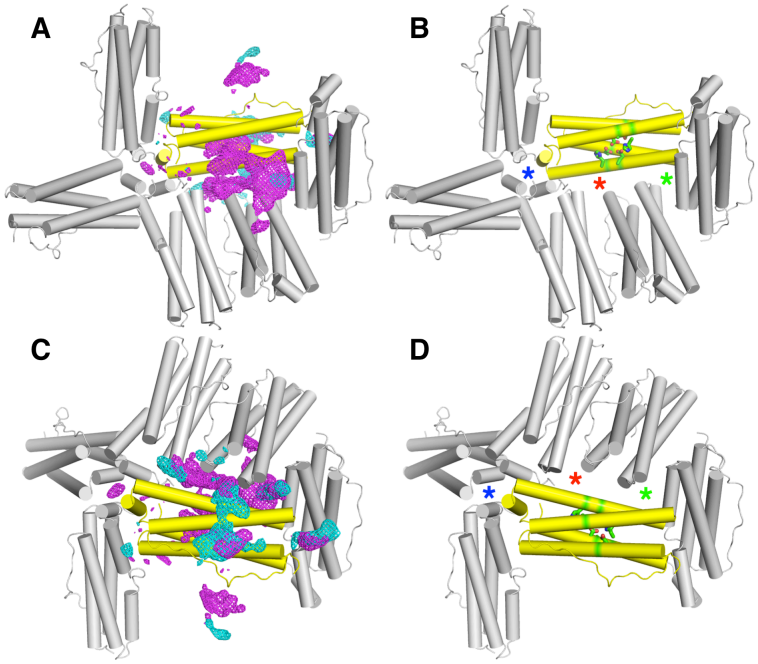


Figure 5.7. Potassium ions concentrated near pores in BfrB are in dynamic exchange. (A) Representative density of K^+ (magenta mesh) and HPO_4^{2-} (cyan mesh) in a minimal portion of the Pa BfrB assembly bearing a 3-fold, 4-fold and a B-pore viewed from the interior cavity. An identical view devoid of ion density is shown in (B) to facilitate identification of the 3-fold, 4-fold, and B-pores, which are highlighted by green, blue and red stars, respectively. This view also shows the ferroxidase center residues (green) of subunit A (yellow). The views in (C) and (D) are from the protein exterior.

In the context of the above-described observations, the significance of a network of negative potential on the inner surface of Pa BfrB is that it provides routes for cations to move along the interior surface. It is therefore plausible that Fe^{3+} ions exiting the ferroxidase center into the interior cavity [197, 201, 202] move along the interior surface in a similar fashion to nucleation sites where the ferric mineral grows. Similarly, Fe^{2+} that

enters the cavity via B-pores may be efficiently transported to a ferroxidase center where it is oxidized to Fe^{3+} and subsequently transported to a nucleation site.

Unlike K^+, which localize mainly within the subunit four-helix bundles or at the interior surface of BfrB, HPO_4^{2-} usually remain associated with the protein exterior surface, near the ferroxidase and B-pores (Figure 5.7C). As indicated above, the phosphate ions present within the 3-fold pores remain in their positions throughout the simulations (Figure 5.6C). The fact that not many HPO_4^{2-} are inside the central cavity is probably the result of the initial placement of these anions. Since the electrostatic potential inside the cavity is very negative (Figure 5.1E), HPO_4^{2-} are located predominantly in the solution bathing the protein exterior at the start of the simulations (Table 5.1). A few phosphate ions get fairly close to the protein during the simulations, especially in regions near the outermost layer of the B-pores. In systems with the highest HPO_4^{2-} concentration (MD_E2), the anion can co-appear with K^+ in the outer layer of the B-pores, but unlike K^+ the anions do not penetrate the B-pores. Given the predominantly negative potential within the B-pores, it is unlikely that phosphate ions use this conduit to traffic across Pa BfrB. It is tempting to speculate that phosphate ions utilize 3-fold pores and there are two plausible reasons why this traffic is not observed in the simulations: First, the timescale of the simulation is short. Second, the possibility exists that phosphate access to the interior cavity is promoted by the presence of iron concentrations within the cavity. As iron accumulates in the cavity, the positive charge carried by iron ions compensates and eventually exceeds the negative potential created by the negatively charged Pa BfrB interior surface, thus allowing the entrance of phosphate, which is incorporated into the ferric mineral.

5.3.3. Ion trafficking and its relation to Pa BfrB dynamics

The simulations suggest that K^+ moves in and out of Pa BfrB exclusively through the B-pores (Table 5.3). The path traveled by K^+ across a B-pore consists of a series of negatively charged residues: E66 at the outermost layer, D34 and D132 in the middle, and E60 and E129 in the interior opening (Figure 5.4C). Among these residues, D34 and D132 are also found to coordinate Na^+ in the crystal structure of Pa BfrB in complex with Pa Bfd [213]. In the crystal structure of Av BfrB, three of these residues D34, E66 and D132, and two additional ones, E135 and D139, are probably responsible for the presence of the Mg^{2+} ions inside the B-pores [212]. Considering the ability of these negatively charged residues to attract cations into the B-pores and further transport them across the protein shell, it is reasonable to suggest that Fe^{2+} ions enter and/or exit the BfrB central cavity via B-pores.

Table 5.3. Number of K^+ ions escaped/entered ($N_{escaped}/N_{entered}$) via the B-pores during the 40-ns trajectories in the MD systems E2 to E40.

	E2	E4	E10	E20	E40
$N_{escaped}$	63	56	63	23	0
$N_{entered}$	5	2	4	3	4

The net number of K^+ moving out of the protein cage displays a slight dependence on the initial C_K inside the Pa BfrB cage. In general, higher interior C_K results in more K^+ exiting the interior cavity, which is in agreement with the intuitive notion that K^+ flux occurs in response to concentration gradients across the protein shell. Also in agreement with the idea of flux controlled by a chemical potential is the observation that the total

numbers of K^+ released in systems MD_E2, E4, and E10 in the 40-ns duration are not significantly different because the C_K in the external solution is high and comparable to the concentration of K^+ in the interior.

5.3.4. Communication between the ferroxidase center and the K_2HPO_4 solution

The ferroxidase centers communicate with the K_2HPO_4 solution by allowing K^+ to enter from either the exterior or the interior K_2HPO_4 solution. At the beginning of the simulations, all the ferroxidase centers were devoid of cations. As the simulations proceed, there is one incidence of K^+ penetrating into the ferroxidase centers from the external K_2HPO_4 solution via the previously identified ferroxidase pore [197]. The low frequency of this occurrence is probably linked to the fact that K^+ entering the protein from the exterior solution has to work against the concentration gradient. K^+ enters the pore steered by one of the ferroxidase center residues, E94, and reaches a pocket within the four-helix bundle, which is adjacent to the ferroxidase center and harbors two negatively charged residues, E44 and D90. These observations are in agreement with the notion derived from X-ray crystallographic and kinetic studies in solution that the ferroxidase pores in Pa BfrB serve as entry ports of iron into the ferroxidase center [197].

In the simulations K^+ ions enter the ferroxidase centers from the interior K_2HPO_4 solution with much higher frequency. The events follow a well-defined path composed of several negatively charged residues (E47, D50, and E129) on helices B and D near the inner opening of the ferroxidase centers. Permeation of K^+ into the ferroxidase center from the interior cavity creates a unique organization of cations in and around the ferroxidase centers as depicted in Figure 5.8. Four sites are favored by K^+; two of them

(① and ②) are close to the previously identified ferroxidase iron binding sites, the third (③) is near the exterior surface, and the fourth (④) is located adjacent to the ferroxidase center in the pocket created by E44 and D90, close to H130. The observation that cations such as K^+ are stabilized within the apo-ferroxidase centers is in line with the crystal structure of Pa FtnA, where Na^+ ions were found in the ferroxidase centers [208]. In this context, the presence of K^+ within the ferroxidase centers is not only expected, but it provides an important correlation to experimental observations which largely validate our approach and additional insights derived from the MD simulations.

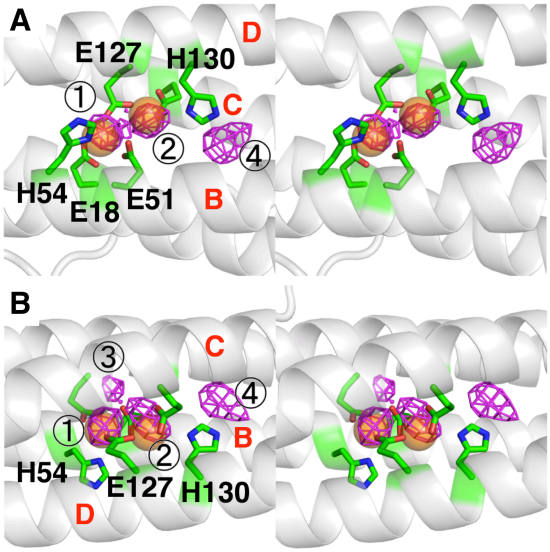


Figure 5.8. Potassium ions exchange in and out of ferroxidase centers. Stereo views showing the distribution of potassium ions (magenta mesh) close to the ferroxidase ligands (green) viewed from (A) the interior cavity of BfrB with helices B and D in the front and (B) the side of the subunit with helices C and D in the front. In (B), the interior cavity is at the bottom and the protein exterior at the top. (A) and (B) are related by a 90° rotation along the horizontal axis. Iron ions in the ferroxidase center of Pa BfrB (PDB:3IS8) are shown as orange spheres. Different helices are indicated by red letters.

Although sites ①-④ are always observed with K^+ bound, these K^+ do not necessarily remain at these sites during the entire course of the simulations. Rather, they exchange sites, predominantly between sites ② and ④. K^+ at site ② can also be released and replaced by another K^+ coming in from the interior cavity. The dynamic reorganization of K^+ in and around the ferroxidase centers is likely facilitated by H130 side chains. Two dominant side chain conformations of H130 are detected in all the simulations, with χ_1 angles at -180° and -60° (Figure 5.9) that correspond to the "gate closed" and "gate open" conformations in the crystal structures of Pa BfrB obtained with and without iron in the ferroxidase center [197]. The transition between these two conformations swings the imidazole ring of H130 and disrupts interactions between K^+ and the ferroxidase center residues at the K^+ binding sites, especially site ②, which harbors the K^+ that exchanges frequently with the K_2HPO_4 solution in the interior cavity.

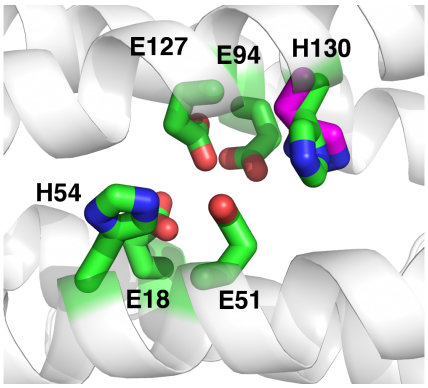


Figure 5.9. H130 in its "gate-closed" (*magenta*) and "gate-open" (*green*) conformations. For a clear view, K^+ ions are not included.

The above observations taken together provide unique insights into the influence exerted by dynamics on the function of the ferroxidase centers in Pa BfrB. Cations such as K^+ and Na^+ are likely ubiquitously present in the ferroxidase centers, to where they

gain access from either side of the protein shell. The monovalent ions flow in and out of the ferroxidase centers in large part aided by the dynamic pedaling of the H130 side chain. In this context, the translocation of K^+ from the ferroxidase center to the interior cavity observed during the simulations, which is gated by the conformational exchange of the H130 side chain, is reminiscent of the mechanism proposed for gating Fe^{3+} from the ferroxidase center to the interior cavity [197]. Hence, the MD simulations lend support to the gating role played by the side chain of H130 and provide unprecedented insight into the protein motions that enable cations to traffic through ferroxidase centers.

5.4. Summary and Conclusions

Bacterioferritin (Bfr) regulates cytosolic iron concentrations by oxidizing Fe^{2+} ions, compartmentalizing the Fe^{3+} as a mineral in its interior, and when necessary releasing Fe^{2+} to the cytosol for their incorporation in metabolism. Hence, the Bfr shell separates the iron mineral from the cytosolic environment and protects it from indiscriminate reduction by reducing agents in the cell. In order to function as effective regulators of cytosolic iron concentrations, Bfrs must allow the efficient and directional flow of iron, either into their cavity (iron uptake) or out to the cytosol (iron mobilization). In the process of iron uptake, Fe^{2+} is thought to access the ferroxidase centers via ferroxidase pores and bind to ferroxidase center ligands to form di-Fe^{2+} complexes. These are subsequently oxidized to di-Fe^{3+} moieties, which are then translocated from the ferroxidase centers to the interior cavity, where the iron mineral is formed. Structural and kinetic studies carried out with Pa BfrB [197] have shown that H130, a ferroxidase center ligand, coordinates iron when its side chain is in a "closed gate" conformation; a change to its "open gate" conformation allows Fe^{3+} to move from the ferroxidase center to the interior cavity. Hence, the ferroxidase center of Pa BfrB is thought to function in the dual capacity of both a pore and a catalytic site [197]. In comparison, kinetic and structural studies with *E. coli* Bfr (Ec Bfr) showed that iron bound to the ferroxidase centers of Ec Bfr is not translocated into the interior cavity, but functions as a cofactor that catalyzes the oxidation of Fe^{2+} that gained access to the interior cavity via a yet unknown type of pore (3-fold, 4-fold or B-) in its structure [199, 219].

The current MD simulations are performed in K_2HPO_4 solutions. Although K^+ and Fe^{2+} are different in many ways, they share a certain degree of similarity. Both ions are

soluble and carry positive charges. We suggest that the results obtained for K^+ in our simulations may be indicative of the behavior of Fe^{2+} as it traffics in and out of the Bfr molecule. Clearly, the rich redox chemistry of iron (Fe^{2+}/Fe^{3+}) cannot be replicated using K^+. Nevertheless, the traffic of K^+ ions across the Bfr structure observed in our simulations is either in agreement with prior experimental observations, or adds new insights, such as the traffic of K^+ through B-pores.

The simulation results reveal dynamically active ferroxidase centers (devoid of iron) in Pa BfrB, which are readily accessed by K^+ ions from the solution bathing the 24-mer, or from the interior cavity. In the simulations K^+ in the external solution is seen to enter the ferroxidase center through the ferroxidase pore, an observation that is in agreement with prior structural reports suggesting that as part of the iron uptake process Fe^{2+} ions access the ferroxidase center via ferroxidase pores [197, 202]. The MD simulations also reveal that K^+ in the ferroxidase centers is in dynamic exchange with K^+ in the interior cavity. Access and exit of K^+ to and from ferroxidase centers is facilitated by the dynamic behavior of the H130 side chain (Figure 5.9), which is seen to periodically alternate between a conformation similar to the iron-bound (closed gate) state in the iron-bound BfrB crystal structure and a conformation similar to the iron-free (open gate) state in the crystal structure of iron-free BfrB [197]. Consequently, the dynamic nature of the ferroxidase centers in Pa BfrB is reflected correctly in the MD simulations. The differences in the iron uptake process observed for Pa and Ec Bfr, despite their nearly identical structures, cannot be explained from the available data. Nevertheless, future efforts aiming at understanding the source of the differences is likely to provide

important additional insights that will illuminate the details of how bacteria regulate cytosolic iron concentrations.

The recovery of Fe^{2+} from the Bfr cavity (iron mobilization) requires specific protein-protein interactions that enable electron transfer from the [2Fe, 2S] cluster of Bfr-associated Bfd to the iron mineral in Bfr [213, 220]. Although Fe^{2+} is thought to exit the protein via pores in the structure, the type of pore used as conduit is unknown. Crystal structures of Bfrs typically exhibit monovalent or divalent cations in the 4-fold pores, which has led to the suggestion that Fe^{2+} may utilize 4-fold pores to traffic across the Bfr protein shell. The MD simulations show that K^+ ions do not cross the Pa BfrB shell via 4-fold pores. Instead K^+ ions traverse the BfrB shell using B-pores, assisted by conserved negatively charged residues. The simulations also suggest that K^+ traffic through B-pores is enabled by folding-unfolding fluctuations in the C-termini of helices D, which cause the B-pores to expand and contract (Figure 5.5). The folding-unfolding fluctuations of helices D also cause lateral displacement of the E-helices, placing two opposed E-helices closer to each other while the other two are separated further apart (Figure 5.5). Such displacement of the E-helices has the effect of narrowing the 4-fold pores. This not only causes some of the K^+ ions to exit the pores early in the simulations, but also prevents ion traffic through them. It is important to note that the folding-unfolding fluctuations in the C-termini of helices E are facilitated by the architecture of the 4-fold pores, which absorb the lateral translation of E-helices with minimal perturbation to the 24-mer assembly. Hence, the coordinated fluctuations in and around 4-fold pores allow the B-pores to undergo "breathing" motions that enable ion traffic through the Bfr shell. In this context,

it is noteworthy that B-pores are not present in the structures of eukaryotic ferritins, where ion traffic is thought to occur through 3-fold pores.

References

1. Krogh A, Larsson B, von Heijne G, Sonnhammer ELL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. J Mol Biol. 2001;305(3):567-80.

2. Overington JP, Al-Lazikani B, Hopkins AL. How many drug targets are there? Nat Rev Drug Discov. 2006;5(12):993-6.

3. Lenard J, Singer SJ. Protein conformation in cell membrane preparations as studied by optical rotatory dispersion and circular dichroism. P Natl Acad Sci USA. 1966;56(6):1828-35.

4. Maddy AH, Malcolm BR. Protein conformations in the plasma membrane. Science. 1965;150(3703):1616-8.

5. Deisenhofer J, Epp O, Miki K, Huber R, Michel H. Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3A resolution. Nature. 1985;318(6047):618-24.

6. White SH. Biophysical dissection of membrane proteins. Nature. 2009;459(7245):344-6.

7. White SH. Membrane proteins of known 3D structure 1998 [updated 02/24/2013]. Available from: http://blanco.biomol.uci.edu/mpstruc/listAll/list.

8. Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI. OPM: orientations of proteins in membranes database. Bioinformatics. 2006;22(5):623-5.

9. Reyes N, Ginter C, Boudker O. Transport mechanism of a bacterial homologue of glutamate transporters. Nature. 2009;462(7275):880-5.

10. Hilf RJ, Dutzler R. X-ray structure of a prokaryotic pentameric ligand-gated ion channel. Nature. 2008;452(7185):375-9.

11. Bocquet N, Nury H, Baaden M, Le Poupon C, Changeux JP, Delarue M, et al. X-ray structure of a pentameric ligand-gated ion channel in an apparently open conformation. Nature. 2009;457(7225):111-4.

12. Andersen OS, Koeppe RE, 2nd. Bilayer thickness and membrane protein function: an energetic perspective. Annu Rev Biophys Biomol Struct. 2007;36:107-30.

13. Spector AA, Yorek MA. Membrane lipid composition and cellular function. J Lipid Res. 1985;26(9):1015-35.

14. Simons K, van Meer G. Lipid sorting in epithelial cells. Biochemistry. 1988;27(17):6197-202.

15. Montaville P, Jamin, N. Determination of Membrane Protein Structures Using Solution and Solid-State NMR. In: Lacapère J-J, editor. Membrane protein structure determination: methods and protocols. New York: New York: Humana Press, c2010; 2010. p. 261-82.

16. Kang C, Li Q. Solution NMR study of integral membrane proteins. Curr Opin Chem Biol. 2011;15(4):560-9.

17. Hong M. Structure, topology, and dynamics of membrane peptides and proteins from solid-state NMR Spectroscopy. J Phys Chem B. 2007;111(35):10340-51.

18. Ader C, Pongs O, Becker S, Baldus M. Protein dynamics detected in a membrane-embedded potassium channel using two-dimensional solid-state NMR spectroscopy. Biochim Biophys Acta. 2010;1798(2):286-90.

19. Cong Y, Zhang Q, Woolford D, Schweikardt T, Khant H, Dougherty M, et al. Structural mechanism of SDS-induced enzyme activity of scorpion hemocyanin revealed by electron cryomicroscopy. Structure. 2009;17(5):749-58.

20. Serysheva, II, Ludtke SJ, Baker ML, Cong Y, Topf M, Eramian D, et al. Subnanometer-resolution electron cryomicroscopy-based domain models for the cytoplasmic region of skeletal muscle RyR channel. P Natl Acad Sci USA. 2008;105(28):9610-5.

21. Yu X, Jin L, Zhou ZH. 3.88 A structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. Nature. 2008;453(7193):415-9.

22. Zhang X, Settembre E, Xu C, Dormitzer PR, Bellamy R, Harrison SC, et al. Near-atomic resolution using electron cryomicroscopy and single-particle reconstruction. P Natl Acad Sci USA. 2008;105(6):1867-72.

23. Beck M, Lucic V, Forster F, Baumeister W, Medalia O. Snapshots of nuclear pore complexes in action captured by cryo-electron tomography. Nature. 2007;449(7162):611-5.

24. Murphy GE, Leadbetter JR, Jensen GJ. In situ structure of the complete Treponema primitia flagellar motor. Nature. 2006;442(7106):1062-4.

25. Han X, Bushweller JH, Cafiso DS, Tamm LK. Membrane structure and fusion-triggering conformational change of the fusion domain from influenza hemagglutinin. Nature Structural Biology. 2001;8(8):715-20.

26. Arkhipov A, Shan Y, Das R, Endres NF, Eastwood MP, Wemmer DE, et al. Architecture and membrane interactions of the EGF receptor. Cell. 2013;152(3):557-69.

27. Endres NF, Das R, Smith AW, Arkhipov A, Kovacs E, Huang Y, et al. Conformational Coupling across the Plasma Membrane in Activation of the EGF Receptor. Cell. 2013;152(3):543-56.

28. Alberts B. The cell as a collection of protein machines: preparing the next generation of molecular biologists. Cell. 1998;92(3):291-4.

29. Janin J, Bahadur RP, Chakrabarti P. Protein-protein interaction and quaternary structure. Q Rev Biophys. 2008;41(2):133-80.

30. Goodsell DS, Olson AJ. Structural symmetry and protein function. Annu Rev Bioph Biom. 2000;29:105-53.

31. Levy ED, Pereira-Leal JB, Chothia C, Teichmann SA. 3D complex: a structural classification of protein complexes. PLoS Comput Biol. 2006;2(11):e155.

32. Browne WR, Feringa BL. Making molecular machines work. Nat Nano. 2006;1(1):25-35.

33. Sali A, Glaeser R, Earnest T, Baumeister W. From words to literature in structural proteomics. Nature. 2003;422(6928):216-25.

34. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A resolution. Cell. 1999;98(6):811-24.

35. Ben-Shem A, Garreau de Loubresse N, Melnikov S, Jenner L, Yusupova G, Yusupov M. The structure of the eukaryotic ribosome at 3.0 A resolution. Science. 2011;334(6062):1524-9.

36. Forkey JN, Quinlan ME, Shaw MA, Corrie JE, Goldman YE. Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization. Nature. 2003;422(6930):399-404.

37. Byeon I-JL, Hou G, Han Y, Suiter CL, Ahn J, Jung J, et al. Motions on the Millisecond Time Scale and Multiple Conformations of HIV-1 Capsid Protein: Implications for Structural Polymorphism of CA Assemblies. J Am Chem Soc. 2012;134(14):6455-66.

38. Alber F, Dokudovskaya S, Veenhoff LM, Zhang W, Kipper J, Devos D, et al. Determining the architectures of macromolecular assemblies. Nature. 2007;450(7170):683-94.

39. Tama F, Gadea FX, Marques O, Sanejouand YH. Building-block approach for determining low-frequency normal modes of macromolecules. Proteins. 2000;41(1):1-7.

40. Bahar I, Rader AJ. Coarse-grained normal mode analysis in structural biology. Curr Opin Struct Biol. 2005;15(5):586-92.

41. Miao L, Schulten K. Transport-related structures and processes of the nuclear pore complex studied through molecular dynamics. Structure. 2009;17(3):449-59.

42. Rui H, Rivera M, Im W. Protein dynamics and ion traffic in bacterioferritin. Biochemistry. 2012;51(49):9900-10.

43. MacKerell AD, Roux B. Atomistic models and force fields. In: Becker OM, editor. Computational biochemistry and biophysics. New York: M. Dekker; 2001.

44. Verlet L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. Phys Rev. 1967;159(1):98-103.

45. Brooks BR, Brooks III CL, MacKerell JAD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: The Biomolecular Simulation Program. J Comput Chem. 2009;30:1545-614.

46. Kale L, Skeel R, Bhandarkar M, Brunner R, Gursoy A, Krawetz N, et al. NAMD2: Greater scalability for parallel molecular dynamics. J Comput Phys. 1999;151(1):283-312.

47. Case DA, Cheatham TE, 3rd, Darden T, Gohlke H, Luo R, Merz KM, Jr., et al. The Amber biomolecular simulation programs. J Comput Chem. 2005;26(16):1668-88.

48. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics. 2013.

49. Barth E, Kuczera K, Leimkuhler B, Skeel RD. Algorithms for Constrained Molecular-Dynamics. J Comput Chem. 1995;16(10):1192-209.

50. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. J Comput Phys. 1977;23:327-41.

51. Alder BJ, Wainwright TE. Phase Transition for a Hard Sphere System. J Chem Phys. 1957;27(5):1208-9.

52. McCammon JA, Gelin BR, Karplus M. Dynamics of folded proteins. Nature. 1977;267(5612):585-90.

53. Klepeis JL, Lindorff-Larsen K, Dror RO, Shaw DE. Long-timescale molecular dynamics simulations of protein structure and function. Current Opinion in Structural Biology. 2009;19(2):120-7.

54. Dror RO, Pan AC, Arlow DH, Borhani DW, Maragakis P, Shan Y, et al. Pathway and mechanism of drug binding to G-protein-coupled receptors. P Natl Acad Sci USA. 2011;108(32):13118-23.

55. Schaeffer RD, Fersht A, Daggett V. Combining experiment and simulation in protein folding: closing the gap for small model systems. Curr Opin Struct Biol. 2008;18(1):4-9.

56. Henzler-Wildman KA, Thai V, Lei M, Ott M, Wolf-Watz M, Fenn T, et al. Intrinsic motions along an enzymatic reaction trajectory. Nature. 2007;450(7171):838-44.

57. Zasloff M. Antimicrobial peptides of multicellular organisms. Nature. 2002;415(6870):389-95.

58. Gordon YJ, Romanowski EG, McDermott AM. A review of antimicrobial peptides and their therapeutic potential as anti-infective drugs. Curr Eye Res. 2005;30(7):505-15.

59. Wang Z, Wang GS. APD: the Antimicrobial Peptide Database. Nucleic Acids Res. 2004;32:D590-D2.

60. Iwanaga S, Muta T, Shigenaga T, Miura Y, Seki N, Saito T, et al. Role of Hemocyte-Derived Granular Components in Invertebrate Defense. Ann Ny Acad Sci. 1994;712:102-16.

61. Kokryakov VN, Harwig SSL, Panyutich EA, Shevchenko AA, Aleshina GM, Shamova OV, et al. Protegrins - Leukocyte Antimicrobial Peptides That Combine Features of Corticostatic Defensins and Tachyplesins. FEBS Letters. 1993;327(2):231-6.

62. Fahrner RL, Dieckmann T, Harwig SS, Lehrer RI, Eisenberg D, Feigon J. Solution structure of protegrin-1, a broad-spectrum antimicrobial peptide from porcine leukocytes. Chem Biol. 1996;3(7):543-50.

63. Harwig SS, Waring A, Yang HJ, Cho Y, Tan L, Lehrer RI. Intramolecular disulfide bonds enhance the antimicrobial and lytic activities of protegrins at physiological sodium chloride concentrations. Eur J Biochem. 1996;240(2):352-7.

64. Mani R, Cady SD, Tang M, Waring AJ, Lehrert RI, Hong M. Membrane-dependent oligomeric structure and pore formation of beta-hairpin antimicrobial peptide in lipid bilayers from solid-state NMR. Proc Natl Acad Sci. 2006;103(44):16242-7.

65. Mani R, Tang M, Wu X, Buffy JJ, Waring AJ, Sherman MA, et al. Membrane-bound dimer structure of a beta-hairpin antimicrobial peptide from rotational-echo double-resonance solid-state NMR. Biochemistry. 2006;45(27):8341-9.

66. Yamaguchi S, Hong T, Waring A, Lehrer RI, Hong M. Solid-state NMR investigations of peptide-lipid interaction and orientation of a ss-sheet antimicrobial peptide, protegrin. Biochemistry. 2002;41(31):9852-62.

67. Jang H, Ma B, Woolf TB, Nussinov R. Interaction of protegrin-1 with lipid bilayers: membrane thinning effect. Biophys J. 2006;91(8):2848-59.

68. Khandelia H, Kaznessis YN. Structure of the antimicrobial beta-hairpin peptide protegrin-1 in a DLPC lipid bilayer investigated by molecular dynamics simulation. Biochim Biophys Acta. 2007;1768(3):509-20.

69. Rui H, Lee J, Im W. Comparative molecular dynamics simulation studies of protegrin-1 monomer and dimer in two different lipid bilayers. Biophys J. 2009;97(3):787-95.

70. Gidalevitz D, Ishitsuka YJ, Muresan AS, Konovalov O, Waring AJ, Lehrer RI, et al. Interaction of antimicrobial peptide protegrin with biomembranes. P Natl Acad Sci USA. 2003;100(11):6302-7.

71. Kirkwood JG. Statistical mechanics of fluid mixtures. J Chem Phys. 1935;3(5):300-13.

72. Torrie GM, Valleau JP. Non-Physical Sampling Distributions in Monte-Carlo Free-Energy Estimation - Umbrella Sampling. J Comput Phys. 1977;23(2):187-99.

73. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules .1. The Method. J Comput Chem. 1992;13(8):1011-21.

74. Lee J, Ham S, Im W. Beta-hairpin restraint potentials for calculations of potentials of mean force as a function of beta-hairpin tilt, rotation, and distance. J Comput Chem. 2009;30(8):1334-43.

75. Lazaridis T. Effective energy function for proteins in lipid membranes. Proteins. 2003;52(2):176-92.

76. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins. 1999;35(2):133-52.

77. Lee J, Im W. Restraint potential and free energy decomposition formalism for helical tilting. Chem Phys Lett. 2007;441:132-5.

78. Jo S, Kim T, Iyer VG, Im W. Software news and updates - CHARNIM-GUI: A web-based grraphical user interface for CHARMM. J Comput Chem. 2008;29(11):1859-65.

79. Jo S, Kim T, Im W. Automated builder and database of protein/membrane complexies for molecular dynamics simulations. PLoS ONE. 2007;2(9):e880.

80. Jo S, Lim JB, Klauda JB, Im W. CHARMM-GUI Membrane Builder for Mixed Bilayers and Its Application to Yeast Membranes. Biophys J. 2009;97:50-8.

81. Dolan EA, Venable RM, Pastor RW, Brooks BR. Simulations of membranes and other interfacial systems using P2(1) and pc periodic boundary conditions. Biophys J. 2002;82(5):2317-25.

82. Klauda JB, Kucerka N, Brooks BR, Pastor RW, Nagle JF. Simulation-based methods for interpreting x-ray data from lipid bilayers. Biophys J. 2006;90(8):2796-807.

83. Skibinsky A, Venable RM, Pastor RW. A molecular dynamics study of the response of lipid bilayers and monolayers to trehalose. Biophys J. 2005;89(6):4111-21.

84. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B. 1998;102(18):3586-616.

85. Mackerell AD, Feig M, Brooks CL. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J Comput Chem. 2004;25(11):1400-15.

86. Klauda JB, Brooks BR, MacKerell AD, Venable RM, Pastor RW. An ab initio study on the torsional surface of alkanes and its effect on molecular simulations of alkanes and a DPPC bilayer. J Phys Chem B. 2005;109(11):5300-11.

87. DeLano WL. The PyMOL Molecular Graphics System Palo Alto, CA, USA.2002. Available from: http://www.pymol.org.

88. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of *n*-Alkanes. J Comp Phys. 1977;23:327-41.

89. Tang M, Waring AJ, Hong M. Intermolecular packing and alignment in an ordered beta-hairpin antimicrobial peptide aggregate from 2D solid-state NMR. J Am Chem Soc. 2005;127(40):13919-27.

90. Gawrisch K, Gaede HC, Mihailescu M, White SH. Hydration of POPC bilayers studied by H-1-PFG-MAS-NOESY and neutron diffraction. Eur Biophys J Biophy. 2007;36(4-5):281-91.

91. Kucerka N, Liu Y, Chu N, Petrache HI, Tristram-Nagle S, Nagle JF. Structure of fully hydrated fluid phase DMPC and DLPC lipid bilayers using X-ray scattering from oriented multilamellar arrays and from unilamellar vesicles. Biophys J. 2005;88(4):2626-37.

92. Lee J, Im W. Transmembrane helix tilting: Insights from calculating the potential of mean force. Phys Rev Lett. 2008;100(1):018103.

93. Janin J. Surface and inside Volumes in Globular Proteins. Nature. 1979;277(5696):491-2.

94. Lee J, Chen J, Brooks CL, 3rd, Im W. Application of solid-state NMR restraint potentials in membrane protein modeling. J Magn Reson. 2008;193(1):68-76.

95. Andersen OS, Koeppe RE. Bilayer thickness and membrane protein function: An energetic perspective. Annu Rev Bioph Biom. 2007;36:107-30.

96. Langham AA, Ahmad AS, Kaznessis YN. On the nature of antimicrobial activity: A model for protegrin-1 pores. J Am Chem Soc. 2008;130(13):4338-46.

97. Jang H, Ma B, Lal R, Nussinov R. Models of toxic beta-sheet channels of protegrin-1 suggest a common subunit organization motif shared with toxic alzheimer beta-amyloid ion channels. Biophys J. 2008;95(10):4631-42.

98. McQuarrie DA. Statistical Mechanics. New York: Harper and Row, New York; 1976.

99. Oldfield E. Chemical shifts in amino acids, peptides, and proteins: from quantum chemistry to drug design. Annu Rev Phys Chem. 2002;53:349-78.

100. Wu CH, Ramamoorthy A, Gierasch LM, Opella SJ. Simultaneous characterization of the amide 1H chemical shift, 1H-15N dipolar, and 15N chemical shift interaction tensors in a peptide bond by three-dimensional solid-state NMR spectroscopy. J Am Chem Soc. 1995;117:6148-9.

101. Lee DK, Wei Y, Ramamoorthy A. A tow-dimensional magic-angle decoupling and magic-angle turning solid-state NMR method: an application to study chemical shift tensors fro peptides that are nonselectively labeled with 15N isotope. J Phys Chem B. 2001;105:4752-62.

102. Asakura T, Yamazaki Y, Seng KW, Demura M. Determination of the mutual orientation of the 15N and 13C NMR chemical shift tensors of 13C-15N double labeled model peptide for silk fibroin from the dipolar-coupled powder patterns. J Mol Struct. 1998;39:179-90.

103. Parton RG, Simons K. The multiple faces of caveolae. Nat Rev Mol Cell Biol. 2007;8(3):185-94.

104. Anderson RG. The caveolae membrane system. Annu Rev Biochem. 1998;67:199-225.

105. Patron RG, de Pozo, M. A. Caveolae as plasma membrane sensors, protectors, and organizers. Nat Rev Mol Cell Biol. 2013;14:98-112.

106. Krajewska WM, Maslowska I. Caveolins: structure and function in signal transduction. Cell Mol Biol Lett. 2004;9(2):195-220.

107. Williams TM, Lisanti MP. The Caveolin genes: from cell biology to medicine. Ann Med. 2004;36(8):584-95.

108. Fielding CJ, Fielding PE. Caveolae and intracellular trafficking of cholesterol. Adv Drug Deliv Rev. 2001;49(3):251-64.

109. Fujimoto T. Calcium pump of the plasma membrane is localized in caveolae. J Cell Biol. 1993;120(5):1147-57.

110. Le PU, Guay G, Altschuler Y, Nabi IR. Caveolin-1 is a negative regulator of caveolae-mediated endocytosis to the endoplasmic reticulum. J Biol Chem. 2002;277(5):3371-9.

111. Liu P, Rudick M, Anderson RG. Multiple functions of caveolin-1. J Biol Chem. 2002;277(44):41295-8.

112. Ahn M, Kim H, Matsumoto Y, Shin T. Increased expression of caveolin-1 and-2 in the hearts of Lewis rats with experimental autoimmune myocarditis. Autoimmunity. 2006;39(6):489-95.

113. Bonuccelli G, Casimiro MC, Sotgia F, Wang C, Liu M, Katiyar S, et al. Caveolin-1 (P132L), a common breast cancer mutation, confers mammary cell invasiveness and defines a novel stem cell/metastasis-associated gene signature. Am J Pathol. 2009;174(5):1650-62.

114. Gaudreault SB, Dea D, Poirier J. Increased caveolin-1 expression in Alzheimer's disease brain. Neurobiol Aging. 2004;25(6):753-9.

115. Shatz M, Liscovitch M. Caveolin-1: a tumor-promoting role in human cancer. Int J Radiat Biol. 2008;84(3):177-89.

116. Drab M, Verkade P, Elger M, Kasper M, Lohn M, Lauterbach B, et al. Loss of caveolae, vascular dysfunction, and pulmonary defects in caveolin-1 gene-disrupted mice. Science. 2001;293(5539):2449-52.

117. Fra AM, Williamson E, Simons K, Parton RG. De novo formation of caveolae in lymphocytes by expression of VIP21-caveolin. P Natl Acad Sci USA. 1995;92(19):8655-9.

118. Dupree P, Parton RG, Raposo G, Kurzchalia TV, Simons K. Caveolae and sorting in the trans-Golgi network of epithelial cells. Embo J. 1993;12(4):1597-605.

119. Monier S, Parton RG, Vogel F, Behlke J, Henske A, Kurzchalia TV. VIP21-caveolin, a membrane protein constituent of the caveolar coat, oligomerizes in vivo and in vitro. Mol Biol Cell. 1995;6(7):911-27.

120. Couet J, Li S, Okamoto T, Ikezu T, Lisanti MP. Identification of peptide and protein ligands for the caveolin-scaffolding domain. Implications for the interaction of caveolin with caveolae-associated proteins. J Biol Chem. 1997;272(10):6525-33.

121. Dietzen DJ, Hastings WR, Lublin DM. Caveolin is palmitoylated on multiple cysteine residues. Palmitoylation is not necessary for localization of caveolin to caveolae. J Biol Chem. 1995;270(12):6838-42.

122. Parton RG, Hanzal-Bayer M, Hancock JF. Biogenesis of caveolae: a structural model for caveolin-induced domain formation. J Cell Sci. 2006;119(Pt 5):787-96.

123. Sargiacomo M, Scherer PE, Tang Z, Kubler E, Song KS, Sanders MC, et al. Oligomeric structure of caveolin: implications for caveolae membrane organization. P Natl Acad Sci USA. 1995;92(20):9407-11.

124. Lee J, Glover KJ. The transmembrane domain of caveolin-1 exhibits a helix-break-helix structure. Biochim Biophys Acta. 2012;1818(5):1158-64.

125. Hoop CL, Sivanandam VN, Kodali R, Srnec MN, van der Wel PC. Structural characterization of the caveolin scaffolding domain in association with cholesterol-rich membranes. Biochemistry. 2012;51(1):90-9.

126. Spisni E, Tomasi V, Cestaro A, Tosatto SC. Structural insights into the function of human caveolin 1. Biochem Biophys Res Commun. 2005;338(3):1383-90.

127. Le Lan C, Neumann JM, Jamin N. Role of the membrane interface on the conformation of the caveolin scaffolding domain: a CD and NMR study. FEBS Lett. 2006;580(22):5301-5.

128. Okamoto T, Schlegel A, Scherer PE, Lisanti MP. Caveolins, a family of scaffolding proteins for organizing "preassembled signaling complexes" at the plasma membrane. J Biol Chem. 1998;273(10):5419-22.

129. Diefenderfer C, Lee J, Mlyanarski S, Guo Y, Glover KJ. Reliable expression and purification of highly insoluble transmembrane domains. Analytical biochemistry. 2009;384(2):274-8.

130. Marley J, Lu M, Bracken C. A method for efficient isotopic labeling of recombinant proteins. J Biomol NMR. 2001;20(1):71-5.

131. Truhlar SM, Cervantes CF, Torpey JW, Kjaergaard M, Komives EA. Rapid mass spectrometric analysis of 15N-Leu incorporation fidelity during preparation of specifically labeled NMR samples. Protein Sci. 2008;17(9):1636-9.

132. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995;6(3):277-93.

133. Goddard TD, Kneller DG. SPARKY 3. San Francisco: University of California.

134. Wishart DS, Sykes BD. The 13C chemical-shift index: a simple method for the identification of protein secondary structure using 13C chemical-shift data. J Biomol NMR. 1994;4(2):171-80.

135. Lakowicz JR. Principles of Fluorescence Spectroscopy. 3rd ed. New York, NY: Springer; 2006.

136. Ladokhin AS, Jayasinghe S, White SH. How to measure and analyze tryptophan fluorescence in membranes properly, and why bother? Analytical biochemistry. 2000;285(2):235-45.

137. Jo S, Lim JB, Klauda JB, Im W. CHARMM-GUI Membrane Builder for mixed bilayers and its application to yeast membranes. Biophys J. 2009;97(1):50-8.

138. Feller SE, Zhang YH, Pastor RW, Brooks BR. Constant-Pressure Molecular-Dynamics Simulation - the Langevin Piston Method. J Chem Phys. 1995;103(11):4613-21.

139. Klauda JB, Venable RM, Freites JA, O'Connor JW, Tobias DJ, Mondragon-Ramirez C, et al. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. J Phys Chem B. 2010;114(23):7830-43.

140. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of Simple Potential Functions for Simulating Liquid Water. J Chem Phys. 1983;79(2):926-35.

141. Steinbach PJ, Brooks BR. New Spherical-Cutoff Methods for Long-Range Forces in Macromolecular Simulation. J Comput Chem. 1994;15(7):667-83.

142. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A Smooth Particle Mesh Ewald Method. J Chem Phys. 1995;103(19):8577-93.

143. Shaw DE, Dror RO, Salmon JK, Grossman J, Mackenzie KM, Bank JA, et al., editors. Millisecond-scale molecular dynamics simulations on Anton. Proceedings of the Conference on High Performance Computing Networking, Storage and Analysis; 2009: ACM.

144. Martyna GJ, Klein ML, Tuckerman M. Nose-Hoover Chains - the Canonical Ensemble Via Continuous Dynamics. J Chem Phys. 1992;97(4):2635-43.

145. Krautler V, Van Gunsteren WF, Hunenberger PH. A fast SHAKE: Algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations. J Comput Chem. 2001;22(5):501-8.

146. Shan YB, Klepeis JL, Eastwood MP, Dror RO, Shaw DE. Gaussian split Ewald: A fast Ewald mesh method for molecular simulation. J Chem Phys. 2005;122(5).

147. Tuckerman M, Berne BJ, Martyna GJ. Reversible Multiple Time Scale Molecular-Dynamics. J Chem Phys. 1992;97(3):1990-2001.

148. Aoki S, Thomas A, Decaffmeyer M, Brasseur R, Epand RM. The role of proline in the membrane re-entrant helix of caveolin-1. J Biol Chem. 2010;285(43):33371-80.

149. Burstein EA, Vedenkina NS, Ivkova MN. FLUORESCENCE AND THE LOCATION OF TRYPTOPHAN RESIDUES IN PROTEIN MOLECULES. Photochemistry and Photobiology. 1973;18(4):263-79.

150. Hodge T, Colombini M. Regulation of metabolite flux through voltage-gating of VDAC channels. J Membr Biol. 1997;157(3):271-9.

151. Linden M, Gellerfors P, Nelson BD. Pore protein and the hexokinase-binding protein from the outer membrane of rat liver mitochondria are identical. FEBS Lett. 1982;141(2):189-92.

152. Shimizu S, Narita M, Tsujimoto Y. Bcl-2 family proteins regulate the release of apoptogenic cytochrome c by the mitochondrial channel VDAC. Nature. 1999;399(6735):483-7.

153. Decker WK, Bowles KR, Schatte EC, Towbin JA, Craigen WJ. Revised fine mapping of the human voltage-dependent anion channel loci by radiation hybrid analysis. Mamm Genome. 1999;10(10):1041-2.

154. Sampson MJ, Decker WK, Beaudet AL, Ruitenbeek W, Armstrong D, Hicks MJ, et al. Immotile sperm and infertility in mice lacking mitochondrial voltage-dependent anion channel type 3. J Biol Chem. 2001;276(42):39206-12.

155. Colombini M. Voltage gating in the mitochondrial channel, VDAC. J Membr Biol. 1989;111(2):103-11.

156. Colombini M. Purification of VDAC (voltage-dependent anion-selective channel) from rat liver mitochondria. J Membr Biol. 1983;74(2):115-21.

157. Roos N, Benz R, Brdiczka D. Identification and characterization of the pore-forming protein in the outer membrane of rat liver mitochondria. Biochim Biophys Acta. 1982;686(2):204-14.

158. Petit PX, O'Connor JE, Grunwald D, Brown SC. Analysis of the membrane potential of rat- and mouse-liver mitochondria by flow cytometry and possible applications. Eur J Biochem. 1990;194(2):389-97.

159. Blachly-Dyson E, Zambronicz EB, Yu WH, Adams V, McCabe ER, Adelman J, et al. Cloning and functional expression in yeast of two human isoforms of the outer mitochondrial membrane channel, the voltage-dependent anion channel. J Biol Chem. 1993;268(3):1835-41.

160. Mannella CA, Forte M, Colombini M. Toward the molecular structure of the mitochondrial channel, VDAC. J Bioenerg Biomembr. 1992;24(1):7-19.

161. Hiller S, Garces RG, Malia TJ, Orekhov VY, Colombini M, Wagner G. Solution structure of the integral human membrane protein VDAC-1 in detergent micelles. Science. 2008;321(5893):1206-10.

162. Ujwal R, Cascio D, Colletier JP, Faham S, Zhang J, Toro L, et al. The crystal structure of mouse VDAC1 at 2.3 A resolution reveals mechanistic insights into metabolite gating. P Natl Acad Sci USA. 2008;105(46):17742-7.

163. Bayrhuber M, Meins T, Habeck M, Becker S, Giller K, Villinger S, et al. Structure of the human voltage-dependent anion channel. P Natl Acad Sci USA. 2008;105(40):15370-5.

164. Depinto V, Prezioso G, Thinnes F, Link TA, Palmieri F. Peptide-Specific Antibodies and Proteases as Probes of the Transmembrane Topology of the Bovine Heart Mitochondrial Porin. Biochemistry. 1991;30(42):10191-200.

165. McDonald BM, Wydro MM, Lightowlers RN, Lakey JH. Probing the orientation of yeast VDAC1 in vivo. FEBS Lett. 2009;583(4):739-42.

166. Stanley S, Dias JA, Darcangelis D, Mannella CA. Peptide-Specific Antibodies as Probes of the Topography of the Voltage-Gated Channel in the Mitochondrial Outer-Membrane of Neurospora-Crassa. Journal of Biological Chemistry. 1995;270(28):16694-700.

167. DeLano WL. The PyMOL Molecular Graphics System. Palo Alto, CA, USA.2002.

168. Mannella CA. Structural analysis of mitochondrial pores. Experientia. 1990;46(2):137-45.

169. Zizi M, Thomas L, Blachly-Dyson E, Forte M, Colombini M. Oriented channel insertion reveals the motion of a transmembrane beta strand during voltage gating of VDAC. J Membr Biol. 1995;144(2):121-9.

170. Choudhary OP, Ujwal R, Kowallis W, Coalson R, Abramson J, Grabe M. The electrostatics of VDAC: implications for selectivity and gating. J Mol Biol. 2010;396(3):580-92.

171. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. J Comput Chem. 2008;29(11):1859-65.

172. Rostovtseva TK, Kazemi N, Weinrich M, Bezrukov SM. Voltage gating of VDAC is regulated by nonlamellar lipids of mitochondrial membranes. J Biol Chem. 2006;281(49):37496-506.

173. Rui H, Im W. Protegrin-1 orientation and physicochemical properties in membrane bilayers studied by potential of mean force calculations. J Comput Chem. 2010;(in press).

174. Brooks BR, Brooks CL, 3rd, Mackerell AD, Jr., Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. J Comput Chem. 2009;30(10):1545-614.

175. MacKerell AD, Feig M, Brooks CL. Improved treatment of the protein backbone in empirical force fields. J Am Chem Soc. 2004;126(3):698-9.

176. Durell SR, Brooks BR, Bennaim A. Solvent-Induced Forces between Two Hydrophilic Groups. J Phys Chem. 1994;98(8):2198-202.

177. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald potential. J Chem Phys. 1995;103:8577-92.

178. Roux B. The Membrane Potential and its Representation by a Constant Electric Field in Computer Simulations. Biophys J. 2008;95(9):4205-16.

179. Lee KI, Rui H, Pastor RW, Im W. Brownian dynamics simulations of ion transport through the VDAC. Biophys J. 2011;100(3):611-9.

180. Feller SE, Pastor RW, Rojnuckarin A, Bogusz S, Brooks BR. Effect of electrostatic force truncation on interfacial and transport properties of water. J Phys Chem. 1996;100(42):17011-20.

181. Im W, Roux B. Ion permeation and selectivity of OmpF porin: a theoretical study based on molecular dynamics, Brownian dynamics, and continuum electrodiffusion theory. J Mol Biol. 2002;322(4):851-69.

182. Im W, Roux B. Ions and counterions in a biological channel: A molecular dynamics simulation of OmpF porin from Escherichia coli in an explicit membrane with 1 M KCl aqueous salt solution. J Mol Biol. 2002;319(5):1177-97.

183. Roux B, Karplus M. Ion-Transport in a Gramicidin-Like Channel - Dynamics and Mobility. J Phys Chem. 1991;95(12):4856-68.

184. Einstein A. Investigation on the theory of Brownian movement. New York: Dover Publications; 1926.

185. Lobo VMM, Ribeiro ACF, Verissimo LMP. Diffusion coefficients in aqueous solutions of potassium chloride at high and low concentrations. J Mol Liq. 1998;78(1-2):139-49.

186. Lyubartsev AP, Laaksonen A. Concentration effects in aqueous NaCl solutions. A molecular dynamics simulation. J Phys Chem. 1996;100(40):16410-8.

187. Andersen OS, Koeppe RE. Molecular Determinants of Channel Function. Physiol Rev. 1992;72(4):S89-S158.

188. Roux B, MacKinnon R. The cavity and pore helices the KcsA K+ channel: Electrostatic stabilization of monovalent cations. Science. 1999;285(5424):100-2.

189. Blachly-Dyson E, Peng SZ, Colombini M, Forte M. Selectivity Changes in Site-Directed Mutants of the Vdac Ion Channel - Structural Implications. Science. 1990;247(4947):1233-6.

190. Egwolf B, Luo Y, Walters DE, Roux B. Ion Selectivity of alpha-Hemolysin with beta-Cyclodextrin Adapter. II. Multi-Ion Effects Studied with Grand Canonical Monte Carlo/Brownian Dynamics Simulations. J Phys Chem B. 2010;114(8):2901-9.

191. Jo S, Vargyas M, Vasko-Szedlar J, Roux B, Im W. PBEQ-Solver for online visualization of electrostatic potential of biomolecules. Nucleic Acids Res. 2008;36(Web Server issue):W270-5.

192. Andrews SC, Robinson AK, Rodríguez-Quiñones F. Bacterial Iron Homeostasis. FEMS Microbiol Rev. 2003;27:215-37.

193. Touati D. Iron and oxidative stress in bacteria. Arch Biochem Biophys. 2000;373(1):1-6.

194. Lewin A, Moore GR, Le Brun NE. Formation of protein-coated iron minerals. Dalton Trans. 2005(22):3597-610.

195. Andrews SC. The Ferritin-like Superfamily: Evolution of the Biological Iron Storeman from a Ruberythrin-Like Ancestor. Biochim et Biophys Acta. 2010;1800:691-705.

196. Janowski R, Auerbach-Nevo T, Weiss MS. Bacterioferritin from Mycobacterium smegmatis contains zinc in its di-nuclear site. Protein Science. 2008;17(7):1138-50.

197. Weeratunga SK, Lovell S, Yao HL, Battaile KP, Fischer CJ, Gee CE, et al. Structural Studies of Bacterioferritin B from Pseudomonas aeruginosa Suggest a Gating Mechanism for Iron Uptake via the Ferroxidase Center. Biochemistry. 2010;49(6):1160-75.

198. Cobessi D, Huang LS, Ban M, Pon NG, Daldal F, Berry EA. The 2.6 A resolution structure of Rhodobacter capsulatus bacterioferritin with metal-free dinuclear site and heme iron in a crystallographic 'special position'. Acta Crystallogr D Biol Crystallogr. 2002;58(Pt 1):29-38.

199. Crow A, Lawson TL, Lewin A, Moore GR, Le Brun NE. Structural basis for iron mineralization by bacterioferritin. J Am Chem Soc. 2009;131(19):6808-13.

200. Frolow F, Kalb AJ, Yariv J. Structure of a unique twofold symmetric haem-binding site. Nat Struct Biol. 1994;1(7):453-60.

201. Liu HL, Zhou HN, Xing WM, Zhao JF, Li SX, Huang JF, et al. 2.6 A resolution crystal structure of the bacterioferritin from Azotobacter vinelandii. FEBS Lett. 2004;573(1-3):93-8.

202. Macedo S, Romao CV, Mitchell E, Matias PM, Liu MY, Xavier AV, et al. The nature of the di-iron site in the bacterioferritin from Desulfovibrio desulfuricans. Nat Struct Biol. 2003;10(4):285-90.

203. Levi S, Santambrogio P, Corsi B, Cozzi A, Arosio P. Evidence that Residues Exposed on the Three-Fold Channels Have Active Roles in the Mechanism of Ferritin Iron Incorporation. Biochem J. 1996;317:467-73.

204. Treffry A, Bauminger ER, Hechel D, Hodson NW, Nowik I, Yewdall SJ, et al. Defining the Roles of the Threefold channels in Iron Uptake, Iron Oxidation and Iron-Core Formation in Ferritin: A Study Aided by Site-Directed Mutagenesis. Biochem J. 1993;296:721-8.

205. Tosha T, Behera RK, Ng HL, Bhattasali O, Alber T, Theil EC. Ferritin Protein Nanocage Ion Channels: Gating by N-Terminal Extensions. J Biol Chem. 2012;287(16):13016-25.

206. Hempstead PD, Yewdall SJ, Fernie AR, Lawson DM, Artymiuk PJ, Rice DW, et al. Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. J Mol Biol. 1997;268(2):424-48.

207. Weeratunga S, Lovell S, Yao H, Battaile KP, Fischer CJ, Gee CE, et al. Structural Studies of Bacterioferritin B (BfrB) from *Pseudomonas aeruginosa* Suggest a Gating Mechanism for Iron Uptake via the Ferroxidase Center. Biochemistry. 2010;49:1160-75.

208. Yao H, Jepkorir G, Lovell S, Nama PV, Weeratunga S, Battaile KP, et al. Two distinct ferritin-like molecules in Pseudomonas aeruginosa: the product of the bfrA gene is a bacterial ferritin (FtnA) and not a bacterioferritin (Bfr). Biochemistry. 2011;50(23):5236-48.

209. Macedo S, Romão CV, Mitchell E, Matias PM, Liu MY, Xavier AV, et al. The Nature of the Di-Iron Site in the Bacterioferritin from *Desulfovibrio desulfuricans*. Nat Struct Biol. 2003;10:285-90.

210. Takahashi T, Kuyucak S. Functional properties of threefold and fourfold channels in ferritin deduced from electrostatic calculations. Biophys J. 2003;84(4):2256-63.

211. Liu H-L, Zhou H-N, Xing W-M, Zhao J-F, Li S-X, Huang J-F, et al. 2.6 Å Resolution Crystal Structure of the Bacterioferritin from *Azotobacter vinelandii*. FEBS Lett. 2004;573:93-8.

212. Swartz L, Kuchinskas M, Li H, Poulos TL, Lanzilotta WN. Redox-dependent structural changes in the Azotobacter vinelandii bacterioferritin: new insights into the ferroxidase and iron transport mechanism. Biochemistry. 2006;45(14):4421-8.

213. Yao H, Wang Y, Lovell S, Kumar R, Ruvinsky AM, Battaile KP, et al. The Structure of the BfrB-Bfd Complex Reveals Protein-Protein Interactions Enabling Iron Release from Bacterioferritin. J Am Chem Soc. 2012.

214. Lee KI, Jo S, Rui H, Egwolf B, Roux B, Pastor RW, et al. Web interface for brownian dynamics simulation of ion transport and its applications to beta-barrel pores. J Comput Chem. 2012;33(3):331-9.

215. Ermak DL, Mccammon JA. Brownian Dynamics with Hydrodynamic Interactions. J Chem Phys. 1978;69(4):1352-60.

216. Chandrasekhar S. Stochastic problems in physics and astronomy. Rev Mod Phys. 1943;15(1):0001-89.

217. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005;26(16):1781-802.

218. Durell SR, Brooks BR, Benniam A. Solvent-induced forces between two hydrophillic groups. J Phys Chem. 1994;98:2198-202.

219. Le Brun NE, Wilson MT, Andrews SC, Guest JR, Harrison PM, Thomson AJ, et al. Kinetic and structural characterization of an intermediate in the biomineralization of bacterioferritin. FEBS Lett. 1993;333(1-2):197-202.

220. Weeratunga SK, Gee CE, Lovell S, Zeng Y, Woodin CL, Rivera M. Binding of Pseudomonas aeruginosa apobacterioferritin-associated ferredoxin to bacterioferritin B promotes heme mediation of electron delivery and mobilization of core mineral iron. Biochemistry. 2009;48(31):7420-31.

Appendices

Appendix A

A List of Publications

1. "Probing Caveolin-1 Structures in Membrane Bilayers Using Molecular Dynamics Simulations" **Rui, H.**, Root, K. T., Glover, K. J., and Im, W. (To be submitted)
2. "Protein Dynamics and Ion Traffic in Bacterioferritin" **Rui, H.**, Rivera, M., and Im, W. *Biochemistry.* (2012). 51:9900-9910.
3. "Web Interface for Brownian Dynamics Simulation of Ion Transport and Its Applications to Beta-Barrel Pores" Lee, K. I., Jo, S., **Rui, H.**, Egwolf, B., Roux, B., Pastor, R. W., and Im, W. *J. Comput. Chem*. (2012). 33:331-339.
4. "Membrane tension, lipid adaptation, conformational changes and energetics in MscL gating" **Rui, H.**, Kumar, R., and Im, W. *Biophys J.* (2011). 101:671-679.
5. "Brownian dynamics simulations of ion transport through the VDAC" Lee, K. I., **Rui, H.**, Pastor, R. W., and Im, W. *Biophys J.* (2011). 100:611-619.
6. "Molecular dynamics studies of ion permeation in VDAC" **Rui, H.**, Lee, K. I., Pastor, R. W., and Im, W. *Biophys J.* (2011). 100:602-610.
7. "Cholesterol flip-flop: insights from free energy simulation studies." Jo, S.*, **Rui, H.***, Lim, J. B., Klauda, J. B., and Im, W. *J Phys Chem B*. (2010). 114:13342-13348.

8. "Structural, NMR Spectropic and Computational Investigation of Hemin Loading in the Hemophore HasAp from Pseudomonas aeruginosa" Jepkorir, G., Rodriguez, J. C., **Rui, H.**, Im, W., Lovell, S., Battaile, K. P., Alontaga, A. Y., Yukl, E. T., Moenne-Loccoz, P., and Rivera, M. *J. Am. Chem. Soc.* (2010). 132:9857-9872.

9. "Protegrin-1 Orientation and Physicochemical Properties in Membrane Bilayers Studied by Potential of Mean Force Calculations" **Rui, H.** and Im, W. *J. Comput. Chem.* (2010). 31:2859-2867.

10. "Novel Free Energy Calculations to Explore Mechanisms and Energetics of Membrane Protein Structure and Function" Im, W., Lee, J., Kim, T., and **Rui, H.** *J. Comput. Chem.* (2009). 30:1622-1633.

11. "Comparative Molecular Dynamics Simulation Studies of Protegrin-1 Monomer and Dimer in Different Lipid Bilayers" **Rui, H.**, Lee, J. and Im, W. *Biophys. J.* (2009). 97:787-795.

* Both authors contribute equally to the work.

Appendix B

License Terms and Conditions

JOHN WILEY AND SONS LICENSE TERMS AND CONDITIONS

Mar 26, 2013

This is a License Agreement between Huan Rui ("You") and John Wiley and Sons ("John Wiley and Sons") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by John Wiley and Sons, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3112110747360
License date	Mar 18, 2013
Licensed content publisher	John Wiley and Sons
Licensed content publication	Journal of Computational Chemistry
Licensed content title	Protegrin-1 orientation and physicochemical properties in membrane bilayers studied by potential of mean force calculations
Licensed copyright line	Copyright © 2010 Wiley Periodicals, Inc.
Licensed content author	Huan Rui,Wonpil Im
Licensed content date	Jun 29, 2010
Start page	2859
End page	2867
Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No
Total	0.00 USD
Terms and Conditions	

TERMS AND CONDITIONS

This copyrighted material is owned by or exclusively licensed to John Wiley & Sons, Inc. or one of its group companies (each a "Wiley Company") or a society for whom a Wiley Company has exclusive publishing rights in relation to a particular journal (collectively WILEY"). By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the billing and payment terms and conditions established by the Copyright Clearance Center Inc., ("CCC's Billing and Payment terms and conditions"), at the time that you opened your Rightslink account (these are available at any time at http://myaccount.copyright.com)

Terms and Conditions

1. The materials you have requested permission to reproduce (the "Materials") are protected by copyright.

2. You are hereby granted a personal, non-exclusive, non-sublicensable, non-transferable, worldwide, limited license to reproduce the Materials for the purpose specified in the licensing process. This license is for a one-time use only with a maximum distribution equal to the number that you identified in the licensing process. Any form of republication granted by this licence must be completed within two years of the date of the grant of this licence (although copies prepared before may be distributed thereafter). The Materials shall not be used in any other manner or for any other purpose. Permission is granted subject to an appropriate acknowledgement given to the author, title of the material/book/journal and the publisher. You shall also duplicate the copyright notice that appears in the Wiley publication in your use of the Material. Permission is also granted on the understanding that nowhere in the text is a previously published source acknowledged for all or part of this Material. Any third party material is expressly excluded from this permission.

3. With respect to the Materials, all rights are reserved. Except as expressly granted by the terms of the license, no part of the Materials may be copied, modified, adapted (except for minor reformatting required by the new Publication), translated, reproduced, transferred or distributed, in any form or by any means, and no derivative works may be made based on the Materials without the prior permission of the respective copyright owner. You may not alter, remove or suppress in any manner any copyright, trademark or other notices displayed by the Materials. You may not license, rent, sell, loan, lease, pledge, offer as security, transfer or assign the Materials, or any of the rights granted to you hereunder to any other person.

4. The Materials and all of the intellectual property rights therein shall at all times remain the exclusive property of John Wiley & Sons Inc or one of its related companies (WILEY) or their respective licensors, and your interest therein is only that of having possession of and the right to reproduce the Materials pursuant to Section 2 herein during the continuance of this Agreement. You agree that you own no right, title or interest in or to the Materials or any of the intellectual property rights therein. You shall have no rights hereunder other than the license as provided for above in Section 2. No right, license or interest to any trademark, trade name, service mark or other branding ("Marks") of WILEY or its licensors is granted hereunder, and you agree that you shall not assert any such right, license or interest with respect thereto.

5. NEITHER WILEY NOR ITS LICENSORS MAKES ANY WARRANTY OR REPRESENTATION OF ANY KIND TO YOU OR ANY THIRD PARTY, EXPRESS, IMPLIED OR STATUTORY, WITH RESPECT TO THE MATERIALS OR THE ACCURACY OF ANY INFORMATION CONTAINED IN THE MATERIALS, INCLUDING, WITHOUT LIMITATION, ANY IMPLIED WARRANTY OF MERCHANTABILITY, ACCURACY, SATISFACTORY QUALITY, FITNESS FOR A PARTICULAR PURPOSE, USABILITY, INTEGRATION OR NON-INFRINGEMENT AND ALL SUCH WARRANTIES ARE HEREBY EXCLUDED BY WILEY AND ITS LICENSORS AND WAIVED BY YOU.

6. WILEY shall have the right to terminate this Agreement immediately upon breach of this Agreement by you.

7. You shall indemnify, defend and hold harmless WILEY, its Licensors and their respective directors, officers, agents and employees, from and against any actual or threatened claims, demands, causes of action or proceedings arising from any breach of this Agreement by you.

8. IN NO EVENT SHALL WILEY OR ITS LICENSORS BE LIABLE TO YOU OR ANY OTHER PARTY OR ANY OTHER PERSON OR ENTITY FOR ANY SPECIAL, CONSEQUENTIAL, INCIDENTAL, INDIRECT, EXEMPLARY OR PUNITIVE DAMAGES, HOWEVER CAUSED, ARISING OUT OF OR IN CONNECTION WITH THE DOWNLOADING, PROVISIONING, VIEWING OR USE OF THE MATERIALS REGARDLESS OF THE FORM OF ACTION, WHETHER FOR BREACH OF CONTRACT, BREACH OF WARRANTY, TORT, NEGLIGENCE, INFRINGEMENT OR OTHERWISE (INCLUDING, WITHOUT LIMITATION, DAMAGES BASED ON LOSS OF PROFITS, DATA, FILES, USE, BUSINESS OPPORTUNITY OR CLAIMS OF THIRD PARTIES), AND WHETHER OR NOT THE PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. THIS LIMITATION SHALL APPLY NOTWITHSTANDING ANY FAILURE OF ESSENTIAL PURPOSE OF ANY LIMITED REMEDY PROVIDED HEREIN.

9. Should any provision of this Agreement be held by a court of competent jurisdiction to be illegal, invalid, or unenforceable, that provision shall be deemed amended to achieve as nearly as possible the same economic effect as the original provision, and the legality, validity and

enforceability of the remaining provisions of this Agreement shall not be affected or impaired thereby.

10. The failure of either party to enforce any term or condition of this Agreement shall not constitute a waiver of either party's right to enforce each and every term and condition of this Agreement. No breach under this agreement shall be deemed waived or excused by either party unless such waiver or consent is in writing signed by the party granting such waiver or consent. The waiver by or consent of a party to a breach of any provision of this Agreement shall not operate or be construed as a waiver of or consent to any other or subsequent breach by such other party.

11. This Agreement may not be assigned (including by operation of law or otherwise) by you without WILEY's prior written consent.

12. Any fee required for this permission shall be non-refundable after thirty (30) days from receipt.

13. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) form the entire agreement between you and WILEY concerning this licensing transaction and (in the absence of fraud) supersedes all prior agreements and representations of the parties, oral or written. This Agreement may not be amended except in writing signed by both parties. This Agreement shall be binding upon and inure to the benefit of the parties' successors, legal representatives, and authorized assigns.

14. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall prevail.

15. WILEY expressly reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

16. This Agreement will be void if the Type of Use, Format, Circulation, or Requestor Type was misrepresented during the licensing process.

17. This Agreement shall be governed by and construed in accordance with the laws of the State of New York, USA, without regards to such state's conflict of law rules. Any legal action, suit or proceeding arising out of or relating to these Terms and Conditions or the breach thereof shall be instituted in a court of competent jurisdiction in New York County in the State of New York in the United States of America and each party hereby consents and submits to the personal jurisdiction of such court, waives any objection to venue in such court and consents to service of process by registered or certified mail, return receipt requested, at the last known address of such party.

Wiley Open Access Terms and Conditions

All research articles published in Wiley Open Access journals are fully open access: immediately freely available to read, download and share. Articles are published under the terms of the Creative Commons Attribution Non Commercial License. which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes. The license is subject to the Wiley Open Access terms and conditions: Wiley Open Access articles are protected by copyright and are posted to repositories and websites in accordance with the terms of the Creative Commons Attribution Non Commercial License. At the time of deposit, Wiley Open Access articles include all changes made during peer review, copyediting, and publishing. Repositories and websites that host the article are responsible for incorporating any publisher-supplied amendments or retractions issued subsequently.
Wiley Open Access articles are also available without charge on Wiley's publishing platform, **Wiley Online Library** or any successor sites.

Use by non-commercial users

For non-commercial and non-promotional purposes individual users may access, download, copy,

v1.7

If you would like to pay for this license now, please remit this license along with your payment made payable to "COPYRIGHT CLEARANCE CENTER" otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or money order referencing your account number and this invoice number RLNK500980061.
Once you receive your invoice for this order, you may pay your invoice by credit card. Please follow instructions provided at that time.

Make Payment To:
Copyright Clearance Center
Dept 001
P.O. Box 843006
Boston, MA 02284-3006

For suggestions or comments regarding this order, contact RightsLink Customer Support: customercare@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Gratis licenses (referencing $0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

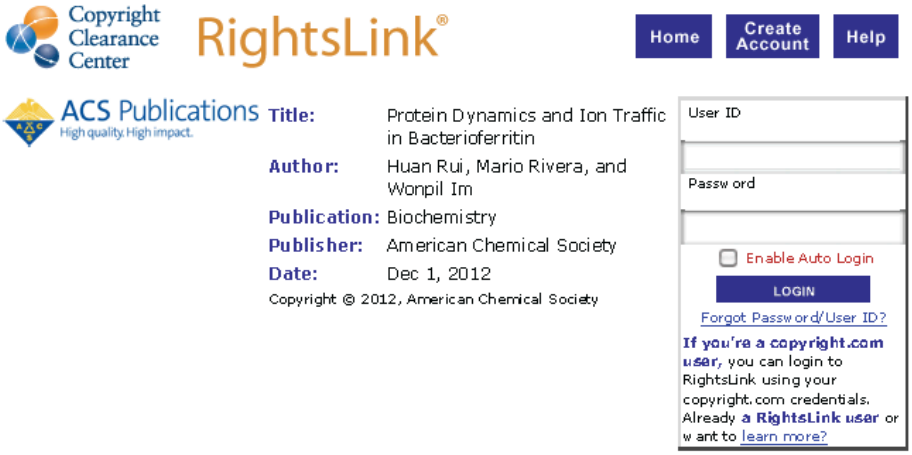
Title:	Molecular Dynamics Studies of Ion Permeation in VDAC
Author:	Huan Rui,Kyu Il Lee,Richard W. Pastor,Wonpil Im
Publication:	Biophysical Journal
Publisher:	Elsevier
Date:	Feb 2, 2011

Order Completed

Thank you very much for your order.

This is a License Agreement between Huan Rui ("You") and Elsevier ("Elsevier") The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

License number	Reference confirmation email for license number
License date	Mar 18, 2013
Licensed content publisher	Elsevier
Licensed content publication	Biophysical Journal
Licensed content title	Molecular Dynamics Studies of Ion Permeation in VDAC
Licensed content author	Huan Rui,Kyu Il Lee,Richard W. Pastor,Wonpil Im
Licensed content date	2 February 2011
Licensed content volume number	100
Licensed content issue number	3
Number of pages	9
Type of Use	reuse in a thesis/dissertation
Portion	full article
Format	both print and electronic
Are you the author of this Elsevier article?	Yes
Will you be translating?	No
Order reference number	
Title of your thesis/dissertation	PROTEIN STRUCTURE AND DYNAMICS IN NATIVE-LIKE ENVIRONMENTS AND THE INDICATION ON PROTEIN FUNCTIONS STUDIED BY COMPUTATIONAL SIMULATIONS
Expected completion date	Apr 2013
Estimated size (number of pages)	186
Elsevier VAT number	GB 494 6272 12
Billing Type	Invoice
Billing address	2030 Becker Drive
	Lawrence, KS 66047
	United States
Permissions price	0.00 USD
VAT/Local Sales Tax	0.00 USD
Total	0.00 USD



PERMISSION/LICENSE IS GRANTED FOR YOUR ORDER AT NO CHARGE

This type of permission/license, instead of the standard Terms & Conditions, is sent to you because no fee is being charged for your order. Please note the following:

- Permission is granted for your request in both print and electronic formats, and translations.
- If figures and/or tables were requested, they may be adapted or used in part.
- Please print this page for your records and send a copy of it to your publisher/graduate school.
- Appropriate credit for the requested material should be given as follows: "Reprinted (adapted) with permission from (COMPLETE REFERENCE CITATION). Copyright (YEAR) American Chemical Society." Insert appropriate information in place of the capitalized words.
- One-time permission is granted only for the use specified in your request. No additional uses are granted (such as derivative works or other editions). For any other uses, please submit a new request.

BACK CLOSE WINDOW