PROBING THE BIOPHYSICAL PROPERTIES OF THE MURINE VOLTAGE DEPENDENT ANION CHANNEL (MVDAC1) STRUCTURE: INSIGHT FROM CONTINUUM ELECTROSTATICS AND MARKOV STATE MODELS.

by

Om P. Choudhary

B.Tech., Indian Institute of Technology Guwahati, Guwahati, 2007

Submitted to the Graduate Faculty of

the Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2013

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Om P. Choudhary

It was defended on

November 15, 2013

and approved by

Dr. Michael Grabe, Associate Professor, Department of Biological Sciences

Dr. Ivet Bahar, Professor, Department of Computational Biology

Dr. Lillian Chong, Associate Professor, Department of Chemistry

Dr. Christopher J. Langmead, Associate Professor, Department of Computer Science, CMU

Dissertation Director: Dr. Michael Grabe, Associate Professor, Department of Biological

Sciences

PROBING THE BIOPHYSICAL PROPERTIES OF THE MURINE VOLTAGE DEPENDENT ANION CHANNEL (MVDAC1) STRUCTURE: INSIGHT FROM CONTINUUM ELECTROSTATICS AND MARKOV STATE MODELS.

Om P. Choudhary, PhD

University of Pittsburgh, 2013

The voltage-dependent anion channel (VDAC) is the major pathway mediating the transfer of metabolites and ions across the mitochondrial outer membrane. The open state of the channel passes millions of ATP molecules per second and is anion selective, while the closed state exhibits no detectable ATP flux and is cation selective.

The high-resolution structure of VDAC1 revealed a 19 stranded β-barrel with an α-helix occupying the pore. To probe VDAC1's biophysical properties, we first carried out continuum electrostatics calculations on the murine VDAC1 (mVDAC1) structure. Poisson-Boltzmann (PB) calculations show that the ion transfer free energy through the channel is favorable for anions, suggesting that mVDAC1 represents the open state. This claim is buttressed by Poisson–Nernst–Planck calculations that predict a high single-channel conductance indicative of the open state and an anion selectivity of 1.75, nearly a two-fold selectivity for anions over cations. These calculations were then repeated on mutant channels and gave selectivity changes in accord with experimental observations.

We next investigated two proposals for how the channel gates between the open and the closed state. Both models involve the movement of the N-terminal helix, but neither motion produced the observed voltage sensitivity, nor did either motion result in a cation-selective channel, which is observed experimentally. Thus, we were able to rule out certain models for channel gating, but the true motion is yet to be determined.

Finally, to understand ATP permeation through VDAC, our collaborators solved the structure of mVDAC1 in the presence of ATP revealing a low-occupancy binding site. We first carried out long, unbiased, multi-microsecond simulations of mVDAC1 in the presence of ATP on the Anton Supercomputer. Guided by the mVDAC1-ATP co-crystal coordinates, we then initiated hundreds of molecular dynamics (MD) simulations to construct a Markov state model (MSM) of ATP permeation. These MSM results show a high ATP flux generated from multiple pathways through the channel, consistent with our structural data and previously reported physiological rates. Continuum calculations and fully atomistic MSM both demonstrate that the solved structures are indicative of an open state of the channel.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to express my deepest gratitude to my advisor Dr. Michael Grabe for his excellent guidance and providing me with an intellectually invigorating research environment in the laboratory. His enthusiasm and motivation has helped me tremendously in my research, and I could not have imagined having a better advisor and mentor for my doctoral studies. His advice on my research as well as on my career has been invaluable to me.

My sincere thanks goes to Dr. Joshua Adelman for playing a key role in my scientific progress and development. This research project would not have been possible without his invaluable assistance and support. I thank him for all the detailed discussions these past years, and for his expert advice on MSMs and simulation methods. In many ways I have learnt much from him.

I thank my lab-mates Dr. Keith Callenberg, Dr. Seungho Choe, Naomi Latorraca, Ying Sheng, and Jianmin Sun, who made my stay in the lab more enjoyable and helped me along the way.

I thank our collaborators Dr. Jeff Abramson, Dr. Rachna Ujwal, Dr. Aviv Paz, and Dr. Jacques-Philippe Colletier for their experiments, insightful comments, and suggestions that helped shape my research work.

I would also like to thank Dr. Rob Coalson and William Kowallis for their generous help with the continuum electrostatic calculations.

I am very thankful for my thesis committee members, Dr. Ivet Bahar, Dr. Lillian Chong, and Dr. Christopher J. Langmead, for their encouragement, guidance, and hard questions.

Finally, I dedicate this research to my family, particularly to my parents. Words cannot express how grateful I am for all the sacrifices they've made for me. They are, and will always be, my source of inspiration.

1.0   INTRODUCTION

Mitochondria are present in all eukaryotic cells and serve two major functions: producing energy and programmed cell death. The interior of the mitochondria is separated from the cytosol by the mitochondrial outer membrane (MOM), and accordingly the mitochondrial functions are governed by flux of metabolites across the MOM. The dominant pathway for metabolite permeation across the MOM is the voltage-dependent anion channel (VDAC). Voltage-dependent anion channels (VDACs) are the most highly expressed protein in the MOM, where they facilitate the flow of adenosine triphosphate (ATP), adenosine diphosphate (ADP), and other metabolites between the cytosol and the intermembrane space [1,2].

In addition to its energetic role as a channel for metabolite permeation, VDAC appears to have a more complex purpose, serving as a scaffold for proteins and molecules that modulate mitochondria permeability, and thereby its function [3,4]. Mutations in VDAC have been implicated in cancers and cardiovascular disease as well as apoptotic cell death [5]. This cell death/survival role has implicated VDAC in the metabolic stresses of cancer and cardiovascular disease specifically as well as mitochondrial-dependent apoptotic cell death in general [5]. Thus, the ability to manipulate VDAC's function in a rational manner will have important implications for novel therapeutics to modulate cell survival in different diseases.

1

1.1 ROLE OF VDAC IN CELLULAR FUNCTION

1.1.1 Cellular respiration

Cellular activity requires a high rate of ATP flux through the outer mitochondrial membrane. Thus, it is logical that VDACs are optimized for ATP transport across the mitochondrial outer membrane.

Several enzymes such as hexokinase and glycerol kinase use ATP produced by the mitochondria for their functions [6,7]. These enzymes play a crucial role in regulation of gycolysis, which is the first stage of cellular respiration. Hexokinase and glycerol kinase associate with the outer mitochondrial membrane by binding VDAC [8]. The hydrophobic N-terminal domain of hexokinase is known to bind residue E72 on VDAC [8], and both glycerol kinase and hexokinase share a common binding site on VDAC [9]. The covalent modification of VDAC with N, N'-dicyclohexylcarbodiimide (DCCD) inhibits hexokinase binding [10], and E72 has been identified as the site of DCCD modification [11]. The binding of hexokinase and glycerol kinase to VDAC increases VDAC's permeability and thus, accessibility of mitochondrially generated ATP to these enzymes [12]. Accordingly, VDAC knockouts give rise to cause respiratory defects, sterility, and embryonic lethality [13].

1.1.2 Apoptosis

1.1.2.1 VDAC regulation by the Bcl-2 family of proteins VDAC is thought to play a role in mitochondrial-dependant apoptotic cell death making VDAC a potential therapeutic target for inducing cell death in cancerous cells. The hallmark of the apoptosis is the release of cytochrome c from the mitochondria to the cytosol, which results in the activation of proteolytic caspases. Mitochondrial permeability of the cytochrome c, and ultimately cell death, is controlled by members of the Bcl-2 family of proteins [3]. Pro-apoptotic members of the Bcl-2 family, Bax and Bak, open VDAC and the anti-apoptotic member Bcl-x_L closes VDAC through direct binding [4]. Additionally, Bax and Bak appear to permit cytochrome c to pass through VDAC out of artificial liposomes.

There are two major models that propose the interaction between the Bcl-2 family of

proteins and VDAC.

Studies carried out on yeast cells proposed a direct interaction between Bcl-2 family of proteins and VDAC [4, 14]. The first model (Fig 1.1 A) associate a direct interaction between the pro-apoptotic members of the Bcl-2 family and VDAC. Shimizu *et.al.,* proposed that Bax induces a novel VDAC-containing channel [4, 14] that allows cytochrome c release from the mitochondria into the cytosol. Though, in complete disagreement with this finding, mammalian VDAC reconstituted in planer phospholipid membrane does not interact with Bax under a variety of conditions [15], and there is no elecrophysiological or structural data indicating that VDAC might form a pore larger than in its open configuration.

A second model (Fig 1.1 B) proposed that VDAC closure is induced by Bcl-2 family protein, tBid [7, 16]. This results in swelling, and subsequent rupture of the outer mitochondrial membrane due to the accumulation of the products from mitochondrial activity such as phosphocreatine [7, 16]. Experiments on VDAC embedded in planer phospholipid membrane have shown that the tBid protein induces closure of VDAC [15]. This leads to permeabilization of outer membrane and release of cytochrome c into the cytosol. Adding anti-apototic proteins such as Bcl-x_L promote the channel opening by restoring the ATP-ADP exchange, and thus rescuing the cells [16].

Thus, the first model accounts the ability of Bcl-2 family proteins (Bax and Bak) to be pro-apoptotic, whereas the second model proposes Bcl-x$_L$ to be anti-apoptotic. However, a physical interaction of VDAC with Bcl-2, Bcl-x_L, or tBid has not been demonstrated so far, and the mechanism by which these protein alter VDAC permeability is yet to be understood.

1.2   STRUCTURE OF THE VDAC CHANNEL

1.2.1   VDAC isoforms

In mammals, there are three VDAC isoforms (VDAC1, VDAC2, and VDAC3) with VDAC1 being the prototypical channel common to all eukaryotes. These three isoforms share greater than 80% sequence homology, yet they have distinct physiological roles, electrophysiological

properties, and relative abundance and distribution between cell types [17–19]. VDAC2 is the dominant isoform expressed in the brain, displaying normal gating patterns relative to VDAC1, but there is an additional population with lower conductance [20].

VDAC3 is less well characterized than VDAC1 and VDAC2, but appears to have distinct gating properties and reduced metabolite permeability. VDAC1 is by far the most well-studied isoform in the family, and many of the fundamental electrophysiological properties including single channel conductance, selectivity, and voltage dependence are remarkably conserved across different organisms ranging from yeast to mammal [21, 22]. In spite of this basic conservation, there are species related differences resulting in some variations when reconstituted in planar phospholipid membranes [21, 22].

1.2.2 VDAC structure

Prior to 2008, direct structural insight into the operation of VDAC has come from low-resolution electron microscopy (EM) [23, 24] and atomic force microscopy [25]; however, three mammalian structures were reported [26–28], making it possible to ask unprecedented questions concerning VDAC1's molecular workings. All three structures show that VDAC1 is a β-barrel formed by 19 β-strands with an ~27 Å pore diameter in excellent agreement with previous EM [23, 24] and atomic force microscopy [25]. Mouse VDAC1 (mVDAC1) is the highest-resolution structure (2.3 Å) revealing an N-terminal helix situated in the pore domain where it adheres to the wall of the barrel [26]. In mVDAC1 structure, the helix has a distinct kink in the middle at Leu10 and Gly11, and Arg15 reaches across this kink to form hydrogen bonds that most likely stabilize the entire conformation [26]. The helix is highly polar with 4 positive residues and 2 negative residues making it an excellent candidate for the voltage sensor as suggested by numerous studies [24, 29, 30]. At both mouths, the pore is ~27 Å wide, but it narrows to 14 Å near the center of the channel at the helix (Fig. 4.1). The dimension of the fully extended ATP is also about 14 Å, which suggests that the ATP can pass through mVDAC1 sterically unhindered.

This observation suggests that mVDAC1 may represent the open state of the channel. However, it is possible that permeation is not controlled through steric interactions, but

rather by the electrostatic field in the channel created by the 30 charged residues that line the pore. Structural changes could alter this field, thereby drastically affecting the permeation properties of ATP and ADP, which carry a formal charge of –4 and –3, respectively.

1.3 ELECTROPHYSIOLOGICAL PROPERTIES OF VDAC

1.3.1 VDAC conductance

The fundamental electrophysiological properties of the channel are remarkably conserved across different isoforms. All VDAC channels exhibit a large conductance 0.45–0.58 nS in 0.1 M KCl and 4–5 nS in 1.0 M KCl [21]. Experiments in reconstituted bilayer have shown that the channel is maximally open at 0 mV, and it enters a lower-conductance state above + 30 mV and below –30 mV resulting in a bell shaped open probability curve. Here, we refer to the high-conductance state as the open state and the lower-conductance state, or possibly states, as the closed state. While the open state facilitates the flux of $\sim 2 \times 10^6$ ATP molecules per second under high ATP concentration, the conductance drops to near zero in the closed state despite the persistent flow of small ions [1].

1.3.2 VDAC selectivity

The selectivity of the open state of the channel differs from the closed state. VDAC is weakly anion selective in the open state (prefers Cl^- over K^+ by a factor of 2 at 1.0 M KCl and by a factor of 5 at 0.1 M KCl), but the closed state of the channel is weakly cation selective [21]. An extensive study of yeast VDAC revealed that residues responsible for selectivity are spread throughout the primary sequence and that the high-impact mutants involved changes to the residue charge [32].

1.3.2.1 Biophysical basis of selectivity The mVDAC1 structure is ~ 14 Å at the center of the pore [26], and is lined with 17 acidic and 20 basic residues. It is likely that the selectivity of the ions is governed by the strong electrostatic nature of the channel, but the
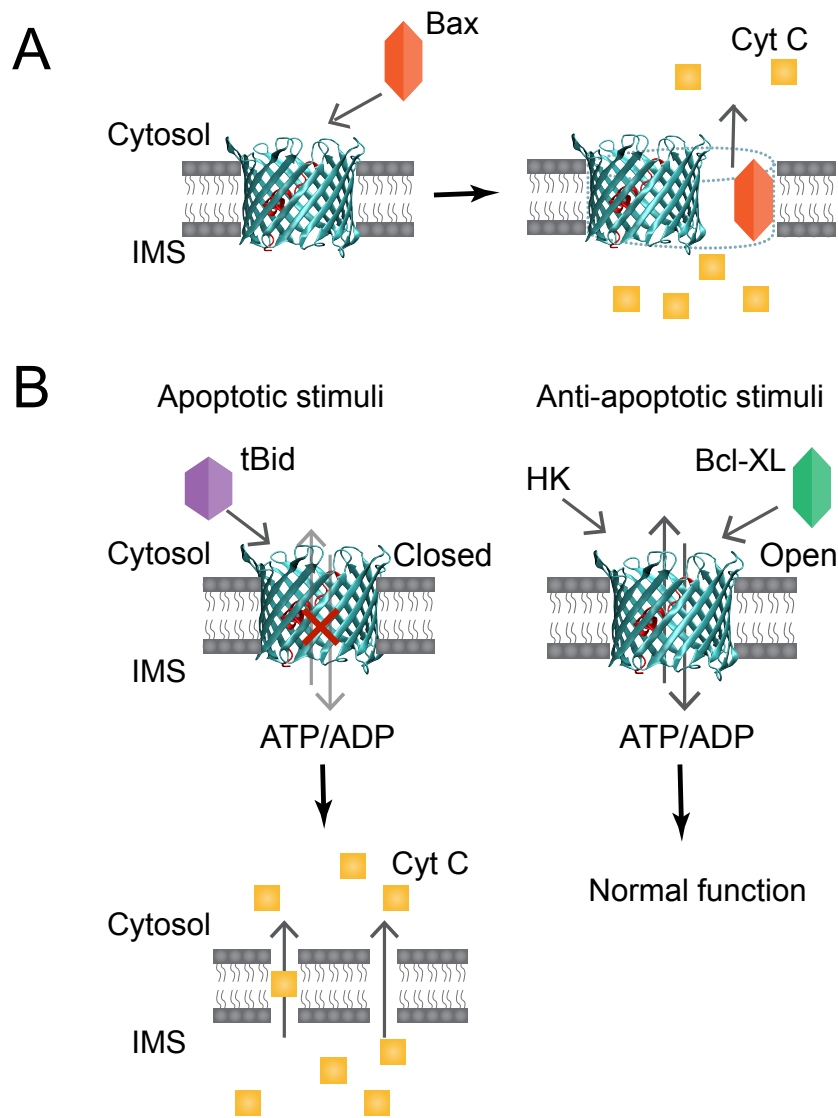
Figure 1.1: Summary of VDAC's role in apoptosis. Figure is modified from reference [31]. (A) Model 1: Pro-apoptotic protein Bax interacts with the VDAC forming bigger channels, resulting in cytochrome c permeation across the outer membrane. (B) Model 2: Apoptotic stimuli like tBid induces VDAC closure and prevent the efficient exchange of ATP and ADP. This results in swelling and rupture of the outer membrane, and release of cytochrome c into the cytosol. Anti-apoptotic proteins like Bcl-x_L promote the open configuration of VDAC, and restore the metabolite exchange and the normal functioning of the cells.

large pore-size would allow the transport of both anions and cations. The uneven distribution of electrostatic charge on the walls of the channel may result in preferred conduits for anions in the open state [32]. The selectivity change upon channel closure could be a result of change in the charge distribution on the wall of the channel, as well as a change in the diameter of the pore.

Rui and coworkers used all-atom MD simulations to investigate the mechanism underlying the anion selectivity of the open state of the channel [33]. Based on ion density profile and cross-sectional ion charge distribution, they proposed that Cl^- ions move through the channel more efficiently than K^+ ions, which are slowed down by the favorable ion-protein interactions at the center of the pore [33]. Acidic residues such as D16, D30, and E84 traps K^+ ions as they pass through the channel and increase their transit time. Mutation of these residues is shown to alter the anion selectivity of the channel [32].

We further investigated the residues implicated in channel selectivity in sec. 3.1.3.

1.3.3   Voltage gating

Voltage biases VDAC between an open state that allows ATP passage and a closed state that does not [2]. Moreover, the channel becomes slightly cation selective in the closed state [21]. Several studies imply that VDAC1's voltage sensor is the N-terminal helix [24, 29, 30], while other studies indicate that the residues involved in gating are distributed throughout the primary sequence [34]. To our knowledge, the most specific gating models involve movement of the N-terminal helix, but even within this subclass of models, the degree of motion and exact molecular position of the α-helix is highly debated [24, 26, 35]

A signature of the closing motion is the gating charge or sensor valence associated with the movement. The sensor valence is equal to the fraction of the membrane electric field that the charges on the voltage sensor pass through during gating. One charge moving through the entire membrane electric field would contribute a valence of 1, but since the voltage-sensor charges need not pass through the entire field this value often takes on noninteger values. The voltage dependence of the channel is directly proportional to its valence, with a high sensor valence corresponding to a steeply voltage-dependent protein. The valence of

the VDAC voltage sensor varies slightly with the subtype, but it is estimated to be between 2.5 and 4.5 charge units based on single channel recordings [21, 28] compared to 12–14 for Shaker voltage-gated potassium channels [36].

With regard to gating, the helix is flanked by two proline residues on the N-terminal side and three glycine residues at the C-terminal end, making it ideally suited to be a mobile element. The helix position was not resolved in the NMR structure, which is consistent with a high degree of mobility [24, 28, 37]. Therefore, it is quite possible that the N-terminal helix, which has four basic and two acidic residues, is the voltage sensor and metabolite gate as suggested by past work [29, 30]. localized It is also known that VDAC is open under high pH conditions and it closes at acidic pH values less than 5 [38]. However, the extent to which pH regulate VDAC under cellular conditions at the MOM is uncertain, and little is known about the site of action for proton binding.

In sec. 3.2.1, we use computational methods to ask whether N-terminal helix movement could account for the channel's voltage sensitivity and changes in ion selectivity.

1.3.3.1 Factors affecting gating There are a environmental factors that affect the open state probability of VDAC, namely polyanions, nucleotides, and other VDAC modulators such as soluble proteins. When added to VDAC constituted in planer bilayer at a very low concentration, VDAC modulators induces VDAC closures. These modulators such as dextran sulfate bind very tightly to VDAC and increase the voltage dependence of gating while biasing the channel towards the closed state [39]. They thus reduce the permeability of VDAC, and inhibit the metabolite flux across the outer mitochondrial membrane. Nucleotides such as NADH, NADPH, ADP, and NADH interacts with VDAC. NADH, and NADPH increase the voltage-dependence of the channel by two-fold [40], and also induce VDAC closure resulting in $\sim$ 6-fold reduction in the permeability of the mitochondrial outer membrane [41].

1.4 VDAC NUCLEOTIDE BINDING SITES (NBS)

1.4.1 Location of NBS

Current-noise spectral analysis showed that ATP and ADP decrease the ionic conductance through the open state of the channel indicating the presence of one or more nucleotide-binding sites (NBS) [40,42]. The localization and the role of nucleotide binding sites in VDAC was further investigated using MALDI-TOF analysis [43]. Yehezkel and coworkers showed that photoreactive ATP analog, benzoyl-benzoyl ATP (BzATP) binds to three proteolytically cleaved VDAC fragments, suggesting the presence of atleast three NBS [43]. These three putative NBS were locallized to the N-terminal region (residues 19-25, site1), the β-barrel wall (residues 110-120, site2), and the C-terminal region (residues 271-283, site3) [43]. The C-terminal site is suggested to bind NADH, whereas the other two sites are suggested to bind ATP. Further, the N-terminal site in VDAC has the same sequence as the Walker A nucleotide-binding motif, G(X/G)XGXGKT, and covalent modification of lysine on this site (K20) with Fluorescein $5^{'}$-isothiocyanate (FITC) inhibited BzATP labeling suggesting the involvement of K20 in ATP binding [43]. This lysine is highly conserved across various species, and mutating this residue in yeast VDAC reportedly decreased the channel's anion selectivity [13, 32, 44].

1.4.2 Functions of NBS

To understand the role of the residues implicated in the N-terminal NBS, Yehezkel and coworkers carried out site-directed mutagenesis (G21A, G23A, and K20S) of mVDAC1 expressed in yeast cells [45]. While ATP and NADH permeated through all mutant channels, the K20S-VDAC mutant showed a decreased affinity to ATP binding compared to native VDAC [45]. ATP synthesis by cells expressing K20S-VDAC mutant was also significantly lower (43%) than cells expressing native VDAC [45]. Not surprisingly, the mutant strain experienced growth difficulties when ATP production depended on the mitochondrial activity [45]. The poor growth of the cells was not due to reduced production of VDAC, but due to dysfunctional nature of mutant channels [45]. Thus, the N-terminal NBS regulates VDAC,

9

thereby mitochondrial activity and metabolite permeation across the outer mitochondrial membrane.

The C-terminal NBS of VDAC is proposed to bind NADH. NADH is known to promote VDAC closure, and this could explain VDAC's sensitivity to cytosolic NADH levels [41, 46]. Other functions of VDAC NBS could be interactions with key metabolic enzymes, such as hexokinases, or pro- and anti-apoptotic Bcl-2 family of proteins [6, 7, 47, 48]. Further, it is suggested that cytosolic and mitochondrial ATP and ADP modulate VDAC activity through interacting with the NBS [2, 41, 49]. The NBS thus provides a regulatory mechanism for metabolite transport across the mitochondrial membrane.

2.0 THEORY AND METHODS

2.1 POISSON-BOLTZMANN (PB) CALCULATIONS

We used the program APBS [50] to carry out continuum electrostatic calculations on the mVDAC1 structure [Protein Data Bank (PDB) ID 3EMN]. Calculations were performed using the PB equation in the linearized limit in the presence of 100 mM salt concentration. The equation was solved using a finite difference method on a $161 \times 161 \times 161$ grid with two levels of focusing. The grid spacing at the finest level of focusing was 0.56 Å. The PARSE parameter set was used to assign partial charges to the protein [51] using the PDB2PQR Web server with standard protonation states for all residues [52]. The influence of the membrane was included as a low-dielectric slab of dielectric value $\epsilon_m = 2$ [53]. Water was assigned a dielectric value of $\epsilon_w = 80$, and the protein dielectric constant was assigned values of $\epsilon_p = 2, 5,$ and 10 in separate calculations. Based on the structure, a suitable path for the ion through the pore was selected (see green path in 3.1 C and D).

We have adopted the convention that the N- and C-termini face the intermembrane space; this is represented by negative z values in Fig. 3.1. The center of the channel is always positioned near z = 0 Å. Ion transfer free energies are calculated by first computing the total energy of the ion plus protein system and then subtracting off the energy of the solo ion in bulk solution and the energy of the channel without the ion [54, 55]. The voltage dependence of the hypothetical gating motions in Fig. 3.6 was calculated using the same setup described above; however, the far-field boundary conditions were manipulated to impose an applied membrane potential in the program APBS as described previously in reference [53]. The interaction energy of the protein charges with the field due to the membrane potential is calculated as follows. First, the electrostatic potential profile across the membraneprotein

system is determined in the absence of any protein charges by calculating the PB equation with all of the protein charges set to zero. In this case, a uniform charge density is added to the intermembrane space to give rise to the desired membrane potential. Second, the protein charges are reintroduced and overlaid on the calculated electrostatic map. The interaction energy is given by the sum over all of the protein partial charges multiplied by the local membrane potential value calculated from APBS.

2.2 POISSON-NERNST-PLANCK (PNP) CALCULATIONS

We used PNP calculations to compute the steady-state ionic flux through mVDAC1 using code written in the Coalson laboratory [56,57]. Current-voltage curves were computed under bi-ionic conditions with the cytosol and intermembrane space maintained at 0.1 and 1.0 M KCl, respectively, and curves were also computed under symmetric 1.0 M KCl solutions. The PNP equations were solved on a $135 \times 135 \times 185$ grid with a grid spacing of 0.83 Å. The channel structures and spatial orientations were identical to those used in the PB calculations; however, the channels were flipped 180 to simulate the insertion with the opposite topology. As with the PB calculations, a low dielectric slab representing the membrane extended from -14 to +14 Å along the z-axis, and it was assigned a dielectric value of 2. The protein dielectric value was set to 5. A membrane potential was imposed across the membrane, and the value was varied from -5 to +15 mV to construct the currentvoltage relation. The bulk diffusion coefficient for potassium and chloride ions was set to 19.5×10^{-6} cm2/s, and the value in the channel was reduced by 40% based on our MD results. The zone of reduced diffusion was defined by a cylinder of radius 18.5 Å centered on the channel and extending 36 Å along the axis of the channel from z = -19 to +17 Å. For each imposed membrane potential value, the PNP equations were solved via relaxation methods using anywhere from 36,000-47,000 iterations. Solutions under both symmetric and asymmetric conditions exhibited excellent convergence.

2.3 MD SIMULATIONS

2.3.1 Simulation systems for computing ion diffusion coefficient

We used the CHARMM-GUI to align mVDAC1 along the z-axis, embed the protein in a dimyristoylphosphatidylcholine (DMPC) lipid bilayer, and then solvate the entire system in a hexagonal cell [58]. We enforced electroneutrality by adding 0.15 M KCl (14 positive ions and 17 negative ions). The approximate system size is 75 Å × 75 Å × 66 Å, and it consists of 34,228 atoms. Simulations were carried out using NAMD with the CHARMM27 parameter set and TIP3P water molecules [59]. Initially, a conjugate gradient minimization was carried out for 3000 steps on the full system. During equilibration, using Langevin dynamics, we increased the temperature every 50 steps by 1 K until it reached 303 K. The time step was 2 fs, and a 10 Å cutoff was used for van der Waals interactions. All electrostatic energies were calculated using the particle mesh Ewald summation. After reaching 303 K, an additional 1 ns of simulation was carried out using the NPT ensemble with a Langevin piston set to 1 bar. We equilibrated a chloride ion in the channel pore at two positions: z = 0 and z = +2. From each setup we ran a 2–ns simulation and recorded positions every 100 fs. For both simulations, the ion remained in the channel. For each ion trajectory, we calculated $R^2(t) = (r(t) - r(t = 0))^2$ for t = 0 to 1 ns, where r is the chloride ion's vector position in time. We then stepped through the time series data and recalculated $R^2(t)$ to arrive at an average squared displacement $\langle R^2(t) \rangle$. The slope of this curve was set equal to 6D to estimate the chloride diffusion coefficient in the channel, D. We repeated this process for a potassium-chloride ion pair in a cubic water box of length 20 Å to estimate D in bulk solution. Our calculations in the channel were ∼40% smaller than our estimate of the bulk value, and this 40% reduction was used to model the diffusion coefficient of the ions in the channel for the PNP calculations. This amount of reduction is very close to the value used by Im and Roux in their study of porins [60].

2.3.2 Simulation systems for Markov state model (MSM) construction

Models of the VDAC were generated using the crystal structure of mouse VDAC (mVDAC1) (Protein Data Bank ID 3EMN) [26]. The channel was centered at the origin and oriented such that the pore axis was aligned with the z axis of the coordinate system with both N- and C-termini laying in the lower half of the simulation cell ($z < 0$). All ionizable groups were left in the dominant protonation state at pH 7.0, except residue E73, which was protonated since it faces lipid. A single ATP molecule with a coordinating Mg^{2+} ion was then introduced at one of several sites. Models of VDAC with ATP at the center of the channel AS1 (pre-refined VDAC-ATP co-crystal structure), and above the N-terminal α-helix (near mass spectrometry site2), were generated from an an initial refinement of the co-crystal data. As described in Table 2.2, the pre-refined ATP structure was placed in different regions of the channel to initiate additional MD simulations for Markov state model construction. The Mg^{2+} was positioned to coordinate the β- and γ-phosphate oxygens of the ATP in accordance with experiments on ATP in solution [61] and previous simulations [62].

Each model of VDAC with ATP-Mg^{2+} was then embedded in a DMPC lipid bilayer and solvated with a neutralizing quantity of 0.15 M KCl using the CHARMM-GUI Membrane Builder [63]. The box size and geometry along with additional information for each model is given in Table 2.1.

Models were parameterized using the CHARMM additive force field, with the C22 parameter set [64] with CMAP corrections [65] for the protein and the C36 parameter set for the lipids [66]. Parameters for the ATP were taken from the nucleic acids section of the C27 parameter set [67, 68], and water molecules were represented by the TIP3P model [69].

2.3.3 Molecular dynamics simulation details

From the initial set of models (Table 2.1), a number of MD simulations were performed on two different computational platforms. Approximately 453 simulations, ranging in length from 40-130 ns, were used to construct a Markov state model (MSM) of ATP permeation through VDAC (Sec. 2.4.3 and Table 2.2). Additionally, a number of long trajectories ranging from 500-5000 ns were generated using the Anton supercomputer [57] at the Pittsburgh

Supercomputing Center (Table 2.1). For models of VDAC with ATP positioned in the upper bulk above the channel (UB), or the lower bulk (LB), below the channel, the ATP coordinates from model AS1 were translated 30 Å along the z axis in the positive or negative directions, respectively. To do this, the ATP-Mg^{2+} configuration was rigidly translated and/or rotated in the simulation box. All of the simulations reported here share a common equilibration protocol, which was carried out using NAMD 2.7 [70]. First, a 2000 step energy minimization was performed using the conjugate gradient algorithm during which all non-hydrogen atoms of the protein and ATP-Mg^{2+} were harmonically restrained to the pre-equilibration model coordinates ($k = 5$ kcal/mol/$Å^2$). The system was then gradually heated to 303 K over 0.3 ns with the same set of external restraints used during the minimization. After thermalizing the system, the harmonic restraints were gradually removed over the subsequent 2 ns. During each of these steps, an additional external force was applied to all water molecules outside of the hydrated channel to prevent them from entering into the protein-membrane interface or penetrating the hydrophobic core of the membrane.

Model ID	Time μs	Initial ATP locations	Applied Voltage (mV)	# of atoms	unit cell
Seeds$_{MSM}$	4.84	Table 2.2	0	36,000	hexagonal
AS1$_{MSM}$	4.84	pore	0	57,000	rectangular
UB1$_{Anton}$	0.61	UB	0	57,000	rectangular
UB2$_{Anton}$	1.99	UB	50	57,000	rectangular
UB3$_{Anton}$	1.99	UB	100	57,000	rectangular
LB1$_{Anton}$	4.17	LB	50	57,000	rectangular
LB2$_{Anton}$	3.11	LB	-50	57,000	rectangular

Table 2.1: VDAC setups for MD simulations

During this phase, the temperature was thermostated by Langevin dynamics with a damping coefficient of 5 ps^{-1}. A constant pressure of 1 atm was maintained with a Nosé-Hoover Langevin piston barostat with a piston period and decay timescale of 200 fs and 100

fs, respectively. The three orthogonal dimensions of the periodic cell were allowed to fluctuate independently during the simulation. The dynamics were propagated with a multiple time-stepping algorithm in which bonded and short-range, non-bonded interactions were evaluated every 2 fs, and long-range electrostatics interactions every 4 fs. Short-range non-bonded interactions were truncated smoothly with a spherical cutoff radius of 10 Å and a switching distance of 8 Å. Long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method, with a grid density of approximately 1 Å^3. All hydrogen atom-heavy atom bond lengths were constrained using the SHAKE algorithm [71]. Subsequent production simulations used to construct the MSM employed the same parameters, but without biasing forces.

Before moving systems to Anton to carry out long MD simulations, the equilibration protocol above was extended 20 ns in NAMD followed by an additional 40 ns of simulation using the molecular dynamics package Desmond version 2.2.9.1 [72]. In Desmond, the dynamics were propagated with constant temperature (312 K) and pressure (1 atm) using the Martyna-Tobias-Klein (MTK) method [73], with a coupling constant of 2.0 ps for the thermostat and barostat, and a 2 fs time step. The barostat was isotropic in x and y, but independent in z. Hydrogen positions were constrained using the M-SHAKE algorithm with a tolerance of 1.0×10^{-8}. Long-range electrostatics were computed using the PME method with 90 Fourier mesh points along each cell axis. Nonbonded and electrostatic interactions were cutoff at a radius of 10 Å.

Production simulations on Anton were initiated from the final snapshot of the corresponding equilibration runs on Desmond. A multi-timestep integration scheme was used to propagate the dynamics based on a Nosé-Hoover thermostat and the MTK method to maintain constant pressure. Long-range electrostatic interactions were calculated using the k-space Gaussian split Ewald (GSE) method [74].

2.3.4 Temperature-accelerated molecular dynamics

Our initial simulations revealed that ATP remained tightly associated with the N-terminal helix when simulations were initiated from a pre-refined model, AS1. Therefore, we used

temperature-accelerated molecular dynamics (TAMD) to rapidly sample alternative conformations of ATP throughout the channel, starting from the AS1 state. Previously, TAMD has been successfully used to generate a broad coverage of conformational space for other biomolecular systems [75, 76]. Briefly, TAMD involves simulating two physical systems simultaneously: (1) the full molecular system (channel, ATP, etc.) represented by its cartesian coordinates $\vec{x}$, and (2) a fictitious particle described by coordinates, $\vec{z}$, that exist in a subspace of the full system described by the coordinate transformation $\vec{\theta}(\vec{x})$. The full molecular system, including the ATP molecule, is simulated at a temperature T. Meanwhile, the fictitious particle, here related to the ATP as described below, is simulated at an elevated temperature $\bar{T} > T$. The two systems are coupled via a harmonic restraint, and the dynamics of the full system evolve on the following extended energy potential:

$$U_\kappa(\vec{x}, \vec{\theta}) = V(\vec{x}) + \frac{\kappa}{2} \sum_{i=1}^{N} [\theta_i(\vec{x}) - z_i]^2,$$ (2.1)

where κ is the harmonic coupling strength, $V(\vec{x})$ is the standard molecular force field described in the previous section, and N is the dimension of $\vec{z}$ and $\vec{\theta}$. The fictitious particle evolves according to the following Langevin equation:

$$\gamma m \dot{\vec{z}} = \kappa \left[\vec{\theta}(\vec{x}) - \vec{z} \right] + \vec{\eta}(t, \bar{T}),$$ (2.2)

where γ is a friction coefficient, η is a random thermal noise term at temperature $\bar{T}$, and m is the effective mass of the fictitious particle.

For this study, we chose the center-of-mass (COM) of the phosphate tail of the ATP and the COM of its purine ring as our collective coordinates comprising the space defined by $\vec{\theta}$. Thus, as the fictitious particle, described by $\vec{z}$, explores phase space, it will apply a force on the ATP molecule in the true system through the coupling term in Eq. 2.1. If κ is chosen so that $\vec{z}$ moves slower than $\vec{x}$ and $\vec{\theta}(\vec{x}) \approx \vec{z}$, TAMD generates a trajectory in $\vec{\theta}(\vec{x})$ that explores the collective variable space at the elevated temperature $\bar{T}$ on the correct free energy landscape at T. For a deeper theoretical description of this method please refer to Ref. [77].

We carried out two TAMD simulations starting from the AS1 state, and each ran for

25 ns. The extended system in the collective coordinates space were simulated at $\bar{T} = 5T$ with $\kappa = 50$ kcal/mol·$\mathring{A}^2$ and $\gamma = 250$ ps^{-1}. The ATP rapidly moved from its starting conformation and sampled a large volume of the channel, but the RMSD of VDAC was indistinguishable from standard MD simulations run without the fictitious particle. Conformations of the system were taken from these runs to initiate local sampling of the channel with conventional MD to build the Markov state model in Sec. 2.4.2; however, no dynamics from the TAMD simulations were actually used in the construction of the MSM.

2.4 MARKOV STATE MODELS

2.4.1 Introduction to Markov state models

MD simulations are a useful tool for modeling thermodynamic and kinetic properties of biomolecules. For a system of interacting particles, their trajectories are determined by numerically solving Newton's equation of motion. Unfortunately, MD simulation takes very small timesteps (fs) to sample the system at an atomistic resolution leading to a sampling problem, where the achieved timescales are restricted to a few nanoseconds a day on a typical computer. Another problem associated with MD simulations is convergence. MD simulations are stochastic, so a single long trajectory is inadequate to determine a biological process.

Markov state models (MSMs) are one of the potential solution to these problems [78,79]. MSMs are a probability distribution of all the conformation a molecule can take, with rate of transitions between them. The MSMs are typically constructed from multiple short simulations that can be generated independently. The conformations that interconvert rapidly are then grouped into states, and the connectivity of these states is determined by number of transitions from one state to another [78,79]. MSMs are also a rigorous way for approaching problems that exhibit metastability. During the process of decomposing the conformational space into metastable states, statistics are generates at each step. Adaptive sampling algorithm then calculates the uncertainties in each step of the process using Bayesian statistics

and runs additional simulations that leads to a reduction in this uncertainty [80–82].

Once a converged MSM is obtained, once can determine most probable conformation at the equilibrium, and use network theory to analyze MSMs. Such an approach has been extensively used to study microsecond to millisecond folding of small proteins.

Here we describe the first application of this methodology to study a nucleotide (ATP) permeation through a membrane channel (mVDAC1). To model ATP permeation through VDAC, we use unbiased MD simulations to build a MSM, in which ATP permeation is represented by transitions between a set of distinct, well-populated conformational states. Such a model can be built from many unbiased equilibrium simulations, dividing ATP conformations into sets of states, and enumerating transitions between states. This simulation method is ideal for running hundreds of simulations simultaneously in parallel, achieving perfect scaling due to the independent nature of the simulations.

2.4.2 Constructing a Markov state model:MSMBuilder2 protocol

we use the software MSMBuilder2 [83,84] (available at https://simtk.org/home/msmbuilder/) with local modifications to construct a MSM of the permeation process. The major steps in a construction of any MSM are as follows:

1. Cluster molecular dynamics simulation data into small sets of states using a hybrid k-centers/k-medoid algorithm.

   The first step in constructing a MSM is identifying conformational states using a hybrid k-centers/k-medoid algorithm. It is important to achieve kinetic clustering to ensure that the clusters are free from any internal energy barriers. The k-centers algorithm minimizes the worst-case clustering error as described by the following objective function:

$$f_{max}(\sigma) = max_i d(x_i, \sigma(x_i)) \tag{2.3}$$

   where d(x,y) is the RMSD distance between two conformation x and y, and $\sigma(x)$ is a mapping function that maps conformations to the clusters. K-medoid algorithm minimizes ensemble-average deviation from the cluster centers using the following objective

19

function:

$$f_{med}(\sigma) = (1/N)(\sum_{i=1}[d(x_i, \sigma(x_i)^2]) \tag{2.4}$$

Simultaneously optimizing both K-centers/K-medoid algorithm is shown to mitigate erroneously fast kinetics.

2. Restrict the MSM to a maximal ergodic subgraph.

   After clustering the data into a set of states, MSMBuilder2 next identifies a maximal ergodic subgraph. It is the maximal set of states, such that there exists a pathway from state i to state j and vice versa. An ergodic MSM has no artifacts in equilibrium estimate of the states. MSMBuilder2 uses Tarjan's algorithm to identify maximal ergodic subgraph, and restricts data to it.

3. Check model consistency by computing relaxation timescales for a series of lagtimes.

   The implied timescales are the timescales for transitions between different set of states, and are an indicator if a model is Markovian. The standard approach is to compute implied timescales plot for a series of lagtimes.

   The relaxation times, or implied timescales, $\hat{t}$, are computed from the eigenvalues as follows:

$$\hat{t}_i = -\frac{\tau}{\ln \lambda_i(\tau)}, \tag{2.5}$$

   where λ_i is an eigenvalue, and τ is the lagtime.

   If a model is Markovian at a given lagtime (Markov time), implied timescales remain constant for greater lagtime. The implied timescales for a non-Markovian model tends to increase, and not level with lagtime.

4. Estimate transition probability and count matrices using a maximum likelihood reversible estimator.

   The entries of a transition probability matrix, $P_{ij}(\tau)$ at a given lagtime τ, are the probabilities of transition from state i to state j in time τ. The standard method for estimating

20

transition probability matrix is $P_{i,j}(\tau) = C_{i,j}/C_i$, where C_i are the transitions originating in state i. However, it is possible that $P(\tau)$ might not satisfy detailed balance.

$$\pi_i P_{i,j} = \pi_j P_{j,i} \tag{2.6}$$

where π_i is the equilibrium population of state i. To ensure reversibility, the corresponding transition count matrix should be symmetric. MSMBuilder2 ensures reversibility by computing a symmetric count matrix X from observed count matrix C, using a maximum likelihood estimator. X is computed by maximizing the likelihood of X given C, L(X/C) as follows:

$$L(X/C) = \prod_{ij} \left(\frac{X_{i,j}}{X_i} \right)^{C_{ij}} \tag{2.7}$$

Maximizing this likelihood gives,

$$X_{ij} = \frac{C_{ij} + C_{ji}}{\frac{C_i}{X_i} + \frac{C_j}{X_j}} \tag{2.8}$$

2.4.3   Data set for Markov state model construction

A kinetic model of the permeation of ATP through VDAC was constructed using a Markov state model built from 453 independent simulations each lasting 40-130 ns (40 μs aggregate simulation time). These independent simulations (Table 2.2) were initiated from a non-equilibrium distribution of starting conformations, selected from either:

1. An equilibrated model of the pre-refined crystallographic ATP-bound VDAC structure.
2. Conformations selected from the temperature-accelerated MD simulations.
3. Initial analysis of earlier simulations using adaptive sampling methods [80–82].
4. Heuristic measures used to identify poorly sampled regions of conformational space.

   More specifically, the seeds were selected as follows:

1. **Seeds $1-5$**

   We initiated 5 independent MD simulations from AS1, and the production run were carried out for 100 ns each.

2. **Seeds 6 − 45**

We carried out two independent temperature accelerated molecular dynamics simulations (TAMD) to aid ATP explore the channel and the bulk regions outside the channel. The ATP positions from all the TAMD trajectories were first geometrically and then, kinetically clustered into 40 states using the software MSMbuilder. We then initiated 40 independent MD simulations from a representative confirmation of each state. Each simulation was carried out between 40 ns and 100 ns, resulting in a total aggregate time of 2.9 μs.

3. **Seeds 46 − 99**

Each seed was initialized from a random conformation within the channel (z coordinate of the center-of-mass of the ATP > -20 Å and < 20 Å after centering VDAC at $z = 0$ Å) taken from the trajectories resulting from seeds 6-45. The simulations were carried out for 130 ns each.

4. **Seeds 100 − 129**

From the above 99 simulations, we visually inspected the ATP density inside the channel and started new simulations in the regions with poor sampling. Independent MD simulations, each in two ATP orientations were started, majority of which lie over the N-terminal α-helix. These seeds were each run for 92-94 ns each.

5. **Seeds 130 − 153**

The N-terminal helix is hypothesized to play an important role in channels permeation process, so we made it high priority to sample the region around the N-terminal helix. From the 5 μs long VDAC production run, we computed the distribution of the distance between the residue K20 on the N-terminal helix, and the beta phosphate atom of the ATP. The peak of this distribution lies at d ~ 4 Å. Next, we constructed a MSM from an aggregate 13 μs of data collected from 129 seeds using a cutoff radius of 4.5 Å, and selected generators with K20-PB distance less than 4.5 Å. 24 such generators were identified, each of which were run independently for 100 ns.

6. **Seeds 154 − 264**

After an initial 14.2 μs of production run from 129 seeds, we wanted to use a systematic method for determining additional sampling. To do this we constructed a MSM

and identified states responsible for most uncertainty in transition count matrix using adaptive sampling method, states with high internal energy barrier, and states with low incoming/outgoing/self transition counts. 110 such states were identified, and simulated for 100 ns each.

7. **Seeds 265 − 283**

   we reconstructed a MSM and computed the relaxation timescales as the function of lag time. We identified the state involved in the slowest transition, and initiated 5 independent MD simulations from it. We also identified 15 additional states using adaptive sampling approach and initiated additional MD simulations from each. All 20 seeds were run for 100 ns each.

8. **Seeds 284 − 303**

   We initiated 20 independent MD simulations from conformations near the ATP binding site2 identified from mass spectrometry, and the production run were carried out for 100 ns each.

9. **Seeds 304 − 415**

   We computed the highest flux pathways for ATP permeation from IMS to cytoplasm, and vice-versa using transition path theory. we analyzed top 100 pathways in each case, and initiated additional simulations from the rate-limiting step (lowest flux along a pathway) of each pathway.

10. **Seeds 416 − 453**

    We visually inspected the channel resulting from seeds 1-415 to identify regions of low ATP density. Simulations were initiated from representative conformations near the lower boundary of the channel.

    A summary of all the starting conformation is shown in the Table 2.2.

Seed ID	length per seed	Notes
1-5	100 ns	Initiated from equilibrated AS1 model.
6-45	40-100 ns	Starting conformation from TAMD trajectories.
46-99	130 ns	From random conformations within the channel.

100-129	92-94 ns	In regions with low visual ATP density.
130-153	100 ns	Near the N-terminal α-helix.
154-264	100 ns	Using adaptive sampling method, and states with low transition counts.
265-283	100 ns	Using adaptive sampling method, and states involved in slowest transition.
284-303	100 ns	Initiated from conformations near ATP binding site2 identified from mass spectrometry.
304-415	130 ns	Initiated from bottlenecks of top 100 flux pathways identified using Transition path theory.
416-453	130 ns	In regions with low visual ATP density.

Table 2.2: MSM dataset

Each starting conformation was assigned a seed number, and its velocities were drawn from the Boltzmann distribution at $303K$ to ensure divergent trajectories, even when starting from the same initial coordinates.

Coordinates were saved every 4 ps, yielding over 10 million conformations. The length of the channel traversed by individual simulations along the z-axis and the total coverage of the channel are shown in Fig. 4.2.

2.4.4 Markov state model coupled to a continuum bath

The MSMBuilder2 software [83, 84] was used to construct a Markov state model of ATP permeation through mVDAC1 using the following protocol:

1. Cluster molecular dynamics trajectories using a hybrid k-centers/k-medoids algorithm.
2. Restrict clusters to those inside of the channel that form the maximal ergodic subset of states.

3. Estimate transition and count matrices using a maximum likelihood reversible estimator.

4. Couple atomistic MSM to a continuum bath of ATP.

Snapshots from all trajectories were aligned to the β-barrel of the mVDAC1 apo structure, and then clustering was carried out on the ATP coordinates using the RMSD between ATP molecules as a metric. The channel, including the N-terminal tail, remained quite stable during all simulations, and based on this observation, we chose not to include the protein degrees of freedom in the state space. We first subsampled the data set by a factor of 500 (2 ns separation between trajectory snapshots), and clustering with a cut-off radius of 6.5 Å resulted in 836 states, each defined by a representative 'generator' ATP conformation. The remaining snapshots were then assigned to these initial states.

In order to efficiently focus our computational efforts on ATP movement in the channel, we chose to initiate more ATP simulations in the pore than in aqueous solution. This leads to a poor estimate of the influx of ATP into the channel from solution due to incomplete sampling. To overcome this issue and accurately capture ATP movement into and out of the channel, we developed a method for coupling an atomistic MSM to a continuum description of free ATP diffusion in solution. First, we defined the channel pore as a cylinder running from z = -18 Å to +20 Å. All states whose COM fell outside the channel were lumped into a single state corresponding to the empty channel. The empty channel is labeled state N, and states in the pore are labeled 1 through $N - 1$.

Next, the dynamical trajectories were projected onto the state space, and we constructed the count matrix $C_{ij}(\tau)$, whose elements correspond to the number of observed transitions of ATP from state i to state j after a lag time, τ. Once an initial count matrix at a given τ was constructed, states were trimmed from the MSM to produce a maximal ergodic subset of states [84, 85]. These ergodically trimmed MSMs typically had ~200 states. From the count matrix, we used a maximum likelihood reversible estimator (MLE) to obtained the

transition probability matrix, $P_{ij}(\tau)$ [84]:

$$\boldsymbol{P} = \begin{pmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,N-1} & P_{1,N} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,N-1} & P_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ P_{N-1,1} & P_{N-1,2} & \cdots & P_{N-1,N-1} & P_{N-1,N} \\ P_{N,1} & P_{N,2} & \cdots & P_{N-1,N} & P_{N,N} \end{pmatrix}, \tag{2.9}$$

The element $P_{ij}(\tau)$ is the conditional probability that the system is in state j at time $t + \tau$ given that it was in state i at time t. More generally, the transition probability matrix determines the time evolution of the system according to:

$$\boldsymbol{p}(t + \tau) = \boldsymbol{p}(t)\boldsymbol{P}, \tag{2.10}$$

where $\mathbf{p}(t)$ is a vector representing the probability of each state in the MSM at time t and $\mathbf{p}(t+\tau)$ is the distribution after a time τ. The MLE algorithm ensures that equilibrium dynamics generated from $\boldsymbol{P}(\tau)$ obey detailed balance. The MSM is coupled to a continuum bath through elements $P_{i,N}$ (ATP transitions to bath), $P_{N,i}$ (ATP transitions into the channel), and $P_{N,N}$ (self transitions for empty channel), which must all be considered carefully.

$\underline{P_{N,i}}$: ATP transitions into the empty channel. First, we consider the elements of $P_{N,i}$ that involve ATP *entering* the empty channel. Let the total flux of ATP into the channel from the upper bath (A) be J_A, and the total flux into the empty channel from the lower bath (B) is J_B. We use Fick's law [86] to estimate this flux assuming that the channel presents a flat surface of radius R to solution:

$$J = 4\pi R c D_{ATP}, \tag{2.11}$$

where c is the concentration of ATP in bath A or B and D_{ATP} is the diffusion coefficient of ATP in water, 20 Å^2/ns [87]. There are N_A upper boundary states that can exchange with bath A, and N_B lower boundary states that can exchange with bath B. Assuming that the incoming flux is equally distributed among all boundary states, the incoming rate into any state adjacent to bath A is J_A/N_A, and the incoming rate for states adjacent to bath B

is J_B/N_B. For our analysis, we assume that all states within 4 Å of the upper boundary of the MSM can exchange with bath A and all states within 4 Å of the lower MSM boundary can exchange with bath B. The rate (number per unit time) of entry for ATP into state i from bath A is then $\rho_N \frac{J_A}{N_A}$, where ρ_N is the probability that the channel is empty so that an incoming ATP can be received in state i. Finally, the one step probability value is $P_{N,i}$:

$$P_{N,i} = \frac{4\pi R c_A D_{ATP}}{N_A}\tau, \tag{2.12}$$

where τ must be sufficiently small such that $P_{N,i}$ is less than 1. In practice, $P_{N,i}$ is on the order of 1%, since our MSM is constructed at a lag time that is much smaller than the mean time for ATP entry into the channel. We see here that the incoming flux has an explicit dependence on the concentration of ATP in bath A, and the flux into states adjacent to bath B have a similar concentration dependence.

$P_{N,N}$: self transitions of the empty channel. The requirement that the row elements normalize to 1 determines the value of the self transition probability for the empty channel:

$$P_{N,N} = 1 - \left(\sum_{\{i\}_A} \frac{J_A}{N_A} + \sum_{\{i\}_B} \frac{J_B}{N_B} \right)\tau = 1 - (J_A + J_B)\tau. \tag{2.13}$$

$P_{i,N}$: ATP transitions out of the channel. Next, we consider the elements $P_{i,N}$ that involve ATP *exiting* the channel, leaving it empty. Since we initiated many simulations from within the pore domain, we observed tens of thousands of exit events out of the channel. Therefore, we believe that $P_{i,N}$ elements in Eq. 2.9 determined directly from the MLE step in MSMBuilder2 are a good estimate of this state dependent jump probability. To test this assumption, we plotted the average jump probability out of the channel, where we averaged $P_{i,N}$ over all states in 2 Å windows along the z axis (Fig. 2.1).

There is a smooth decrease in the jump probability as the states penetrate deeper into the channel from either the upper or lower face of the channel, which is expected since it is harder to escape the channel the deeper the penetration. ATP in states near the upper bath never escape to the lower bath, and states near the lower bath never exhibit transitions to the upper bath. Also, states at the upper face of the channel have a nearly 50% probability of jumping out of the channel, which is intuitive if the ATP is only weakly interacting with the
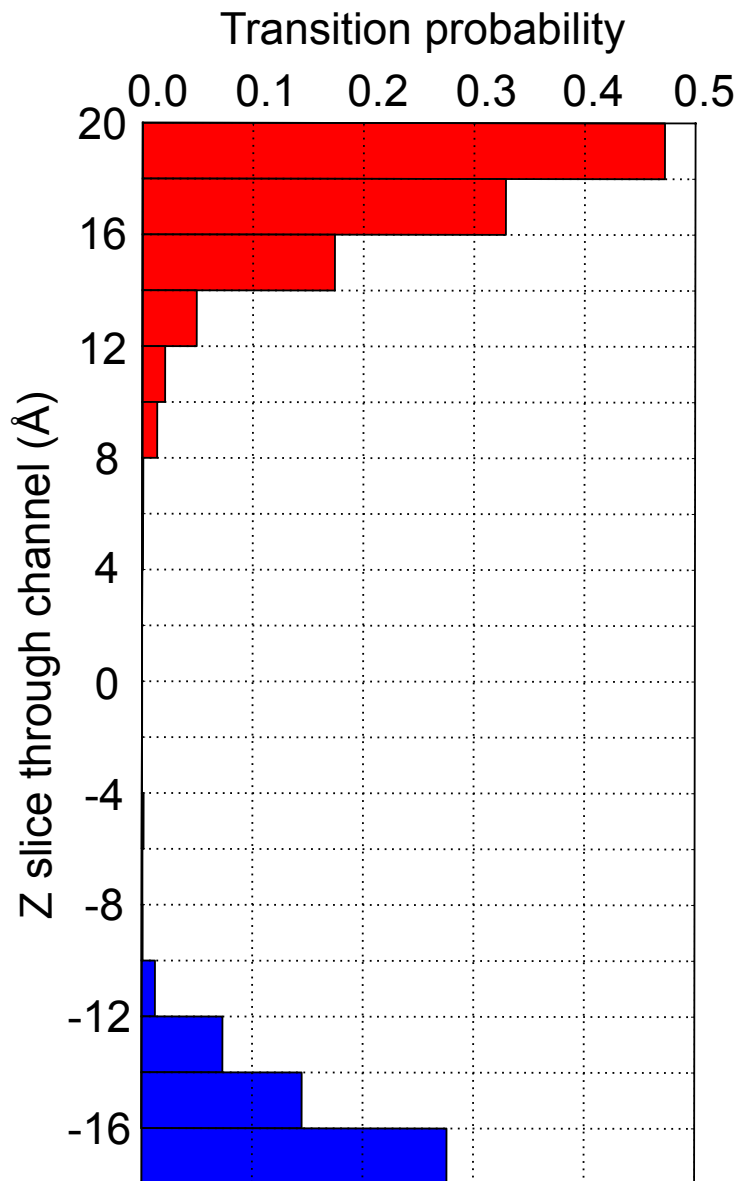
Figure 2.1: Transition probability for exit from the channel. The transition probability for ATP to exit state i and enter the bath, leaving channel empty, was computed using the MLE estimator as stated in Sec. 2.4.2 using a 5 ns lag time. We then averaged these values in 2 Å slices along the z axis from the top to the bottom of the channel. Red bars indicate transitions to the upper bath (cytosol), and blue bars indicate transitions to the lower bath (intermembrane space). Between -10 and +8 Å no exit transitions were observed. The highest probability for exiting the pore occurs at the channel boundaries.

pore at the edge of the channel. At the lower mouth, which is closer to the N-terminal helix, the probability of exiting the channel is smaller, which reflects that these states are still under some electrostatic influence of the pore, which we described in an earlier manuscript [88]. Based on these results, we believe that we have sampled these exit transitions well.

2.4.5 Calculation of the mean first passage time (MFPT)

The mean first passage time (MFPT) from state i to state j, $m_{i,j}$, is defined as the average time taken to reach state j for the first time given that the system was initially in state i. The MFPT between two states can be determined by solving the following linear system of equations [78]:

$$m_{i,j} = P_{i,j}\tau + \sum_{k \neq j} P_{i,k}\left(m_{k,j} + \tau\right). \tag{2.14}$$

We are interested in the MFPT for ATP movement from the upper bath to the lower bath, and vice versa. Unfortunately, Eq. 4.4 cannot be applied directly to our MSM, since transitions to/from the empty channel, N, do not distinguish whether ATP exited/entered to the upper or lower bath. However, we can modify our MSM to take the separate baths into account since we know the exchange probabilities with each bath.

We split state N into two separate states: A where ATP is in the upper bath, and B where ATP in the lower bath. We then self-consistently compute all MFPTs including $m_{A,B}$ and $m_{B,A}$. However, there are several transition probability elements that must be reconsidered: $P_{A,A}$, $P_{B,B}$, $P_{i,A}$, $P_{i,B}$, $P_{A,i}$, and $P_{B,i}$. Both $P_{A,B}$ and $P_{B,A}$ are zero, as we never observe movement from one bath to the other in one step. $P_{A,i}$ and $P_{B,i}$ are given by Eq. 2.12 with the corresponding ATP concentration values, c_A or c_B, respectively. $P_{i,A}$ and $P_{i,B}$ are determined directly from the simulation data as described in Sec. 2.4.4. Based on the data in Fig. 2.1, we know ATP exiting states at the top of the channel only enter bath A and ATP exiting states at the bottom of the channel only enter bath B. This uniquely determines $P_{i,A/B}$ from $P_{i,N}$ based on the location of state i. The last terms to be determined are the self transition probabilities for ATP in states A and B, which are given

by the normalization condition:

$$P_{A,A} = 1 - \sum_{i \neq A} P_{A,i}, \tag{2.15}$$

With these terms identified, the MFPT for arriving in bath B starting from all other states can be determined from the following set of self-consistent equations:

$$\begin{pmatrix} P_{A,A} - 1 & P_{A,1} & \cdots & P_{A,N-1} \\ P_{1,A} & P_{1,1} - 1 & \cdots & P_{1,N-1} \\ \vdots & \vdots & \ddots & \vdots \\ P_{N-1,A} & P_{N-1,1} & \cdots & P_{N-1,N-1} - 1 \end{pmatrix} \begin{pmatrix} m_{A,B} \\ m_{1,B} \\ \vdots \\ m_{N-1,B} \end{pmatrix} = \begin{pmatrix} -\tau \\ -\tau \\ \vdots \\ -\tau \end{pmatrix}, \tag{2.16}$$

where the element $m_{A,B}$ on the right hand side is the quantity of interest, and the MFPT from B to A, $m_{B,A}$, can be determined from a similar set of equations.

2.4.6   Transition path theory (TPT) analysis

We computed the highest flux pathways taken by ATP as it passes through the channel using transition path theory [89, 90] as implemented in MSMBuilder2 [83, 84, 91]. We started with the original MSM defined on the states that fall between -18 and +20 Å along the z axis, but rather than coupling to a continuum bath, we extended the atomistic MSM 4 Å in the positive, Ω^+ ($+20\,\text{Å} < z \leq +24\,\text{Å}$), and negative directions, Ω^- ($-22\,\text{Å} \geq z < -18\,\text{Å}$). We then computed all the pathways from all states in Ω^- to all states in Ω^+ (Fig. 4.8), and vice versa.

Briefly, the algorithm computes the forward-committor probability q^+ and the backward-committor probability q^- of the states. The forward-committor probability of the state i during Ω^+ to Ω^- transition, q_i^+, is defined as the probability if the system is in state i, will reach state Ω^- next rather than Ω^+. Similarly the backward-committor probability of the state i, q_i^-, is defined as probability if the system is in state i, it was in state Ω^+ previously and not Ω^-. Using the committor probabilities we can calculate the effective flux along an

edge i, j contributing to the Ω^+ to Ω^- transitions as:

$$f_{ij} = \pi_i q_i^- P_{ij} q_j^+ \tag{2.17}$$

where π_i is the stationary state population of state i, and P_{ij} is the probability of transition from state i to state j. The net flux from q_i^+ to q_i^- across a trajectory is sum total of net flux along the edges contributing to that trajectory. The total flux from q_i^+ to q_i^- is then decomposed into a set of pathways, along with their fluxes using a greedy backtracking decomposition algorithm [92, 93].

During our analysis we removed pathways from the analysis which contained transitions with zero counts in the count matrix, $C_{ij}(\tau)$. This occasionally occurs because the MLE method for computing the transition probability matrix will assign a finite probability value to a transition with zero counts if there are counts in the reverse direction. Finally, we analyzed the top 143 pathways, which account for 70% of the total flux. In Fig. 4.8C, we grouped the pathways into four categories (paths 1-4) based on which residue the ATP interacts with just prior to escape to the cytoplasm. We included a 5th path, which does not interact with the central portion of the N-terminal helix during permeation.

3.0   THE ELECTROSTATICS OF VDAC: IMPLICATIONS IN SELECTIVITY AND GATING

Despite the windfall of VDAC1 structural information, there are still numerous outstanding questions concerning the mechanism by which VDAC1 differentiates between ions, how it conducts ATP at high rates, and how the membrane voltage regulates its conduction state. The most pressing question is whether the latest high-resolution structure represents the open state or the closed state of the channel. To better understand VDAC1's biophysical properties, we have carried out a number of continuum electrostatics calculations on mVDAC1. As we describe below, the channel is anion selective, and it has a large single-channel conductance that is most consistent with the open state. A key aspect of any computational model is the ability to address a wide range of experimental data. Continuum electrostatic calculations provide a rapid means to compute the currentvoltage characteristics for a number of mutant channels, and our results are in good agreement with experimental trends. The molecular rearrangements that accompany voltage gating are not known. We examined two hypothetical models that involve motions of the N-terminal helix, and our calculations indicate that neither motion fully accounts for the channel's observed voltage sensitivity. Thus, we were able to rule out some of the proposed molecular motions, but to gain an understanding of the true mechanism will require extensive probing by a number of different techniques.

3.1 CONTINUUM ELECTROSTATICS SUGGEST THAT THE MVDAC1 STRUCTURE IS OPEN

3.1.1 Cation and anion energy profiles

The mVDAC1 structure has been proposed to be in an open conformation [26], which is known to be anion selective. We wanted to quantitatively probe this claim by determining if ion passage through the channel was more conducive to anions or cations. Poisson-Boltzmann electrostatic calculations are a fast, effective way of calculating such ion transfer free energies and for understanding the role of the protein architecture in the process [54, 55]. We computed the energy required to transfer a chloride-sized monovalent anion from bulk water through the central pore of mVDAC1. The total free energy of transfer consists of a Born solvation term, which corresponds to stripping waters away from the ion as it passes through the channel, and an electrostatic term, which corresponds to the interaction of charges on the channel with the charge on the ion. The proximity of the ion to the low dielectric of the membrane could drastically affect the permeation energetics, so we embedded the channel in a water-impermeable, uniform slab of dielectric 2, which closely mimics the properties of the membrane.

The geometry of the system is depicted in Fig. 3.1. The ion (green sphere) was translated from -40 Å to +40 Å, and the total electrostatic energy was calculated every 1 Å. The channel was centered at the origin, and z = 0 and 20 Å are indicated by arrows in panel C. While included in all of the calculations, the membrane is not pictured but its extent is indicated by grey bars in Fig. 3.1. Our choice of orientation along the z-axis is arbitrary since the orientation of VDAC in the membrane is still under debate (see below). For consistency, we define the negative direction to be the side of the membrane that contains the N and C-termini. This convention is held throughout. We see from the set of curves in Fig. 3.1 that the transfer free energy is energetically favorable resulting in a 2.5 k_BT stabilization of the anion when it is at z = -7.5 Å (1 k_BT = 0.6 kcal/mol at room temperature).

Interestingly, the profile is not symmetric with respect to the center of the channel. The N-terminal helix occupies the lower half of the channel very close to the energy minima. It
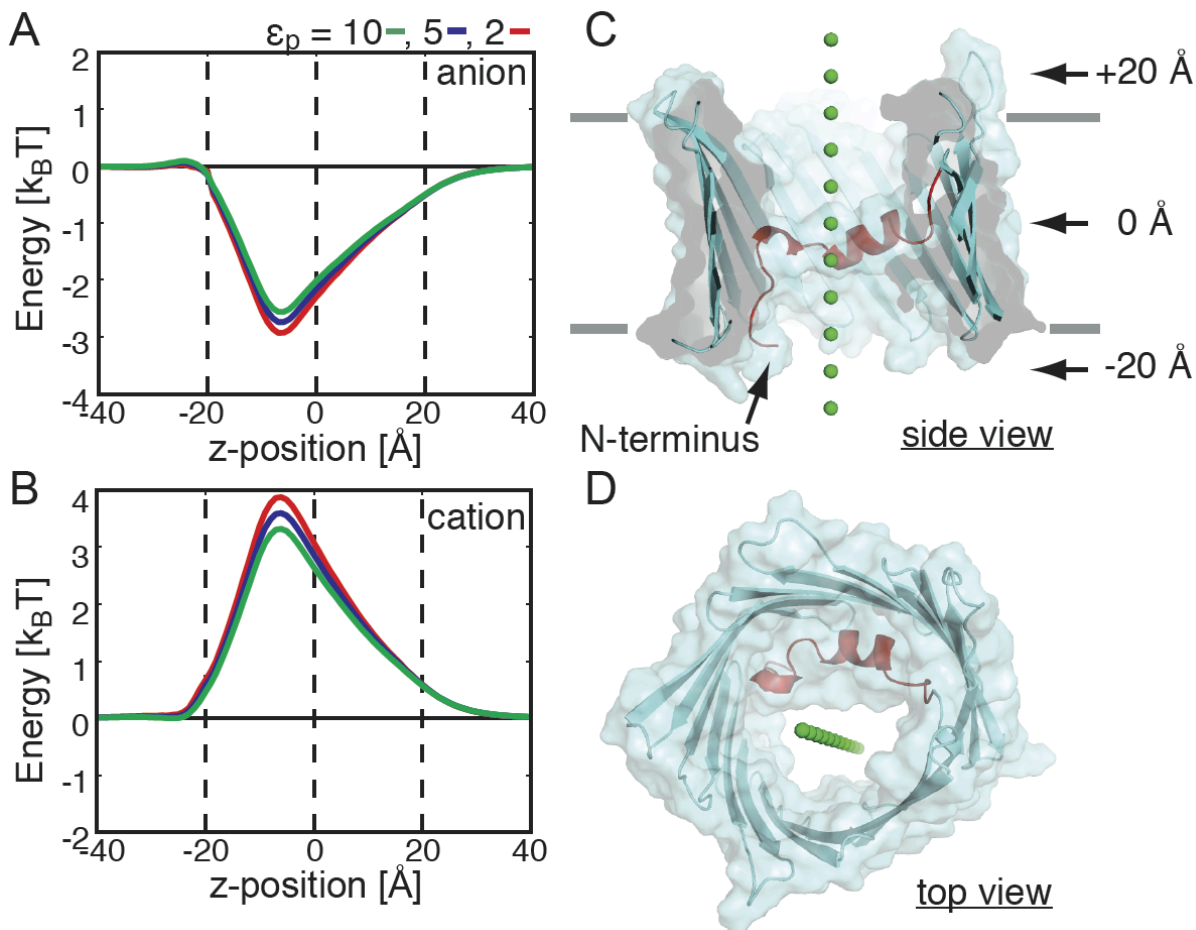
Figure 3.1: mVDAC1 is selective for anions. Ion transfer free energies calculated through mVDAC1 for a chloride-sized anion (A) and a potassium-sized cation (B). Energies were calculated using the Poisson-Boltzmann equation after embedding mVDAC1 in a low dielectric, $\epsilon_m = 2$, slab corresponding to the membrane. The path for both ions through the channel is pictured in panels C (side view) and D (top view). Varying the protein dielectric constant, ϵ_p, had little effect on the free energy profiles. For all calculations, z = 0 corresponds to the center of the channel, and the channel extends from -20 to +20 Å. These positions are indicated by dashed lines in panels A and B.

is generally thought that the dielectric constant of protein ranges from 2 to 20 [94], but it is also understood that proteins are heterogeneous [95]. We explored the affect of the choice of protein dielectric constant, p, on our results by varying it from 2 to 10. Paradoxically, increasing the protein dielectric value destabilizes the negative ion by 0.5 k_BT (red curve compared to green curve in panel A). While this is only a small change, it results from the decreased electric field in the center of the channel that accompanies increasing the protein dielectric value. In Fig. 3.2, we show that the membrane has very little effect on the permeation energetics since the ion remains surrounded by a significant amount of water during penetration due to the large pore size.

Next, we carried out all of the free energy calculations using a potassium-sized monovalent cation. Interestingly, the channel presents a free energy barrier to cation movement (Fig. 3.1 B). The maximum energy difference between the curves in panel A and B is 5.5-7 k_BT, which is energetically significant and shows that the channel is selective for anions. This result corroborates the claim that the x-ray structure is indeed the open state [26], and that it is not the cation-selective closed state. The anion and cation energy profiles in panels A and B are nearly mirror opposites of each other suggesting that the Born solvation energy is very small, which we verified in separate calculations in Fig. 3.2. Therefore, the free energy profile is dominated by the electrostatic interaction with charged groups on the channel. For a narrow pore like the KcsA potassium channel, the Born solvation term is very important since all the water molecules are essentially stripped away during translocation [54]. In our case, the pore is large enough that ions can retain their hydration shell as they move through the channel resulting in a negligible Born energy.

The PB electrostatic calculations show that the mVDAC1 x-ray structure is anion selective, suggesting that the structure is open. While the transfer free-energy profile in Fig. 3.1 has a clear minimum in the electrostatic component of the free energy, its depth is very weak, 2.5 k_BT. An ATP molecule, with a charge of -4, will experience a well depth four times deeper than that of a chloride anion, implying that a shallow minimum for chloride would be a much deeper for ATP. Close coordination by magnesium or other cations could reduce this well, but if too deep, it could lead to increased ATP dwell times and decrease the channel conductance. The true test of the channel's conduction state will come from a detailed study
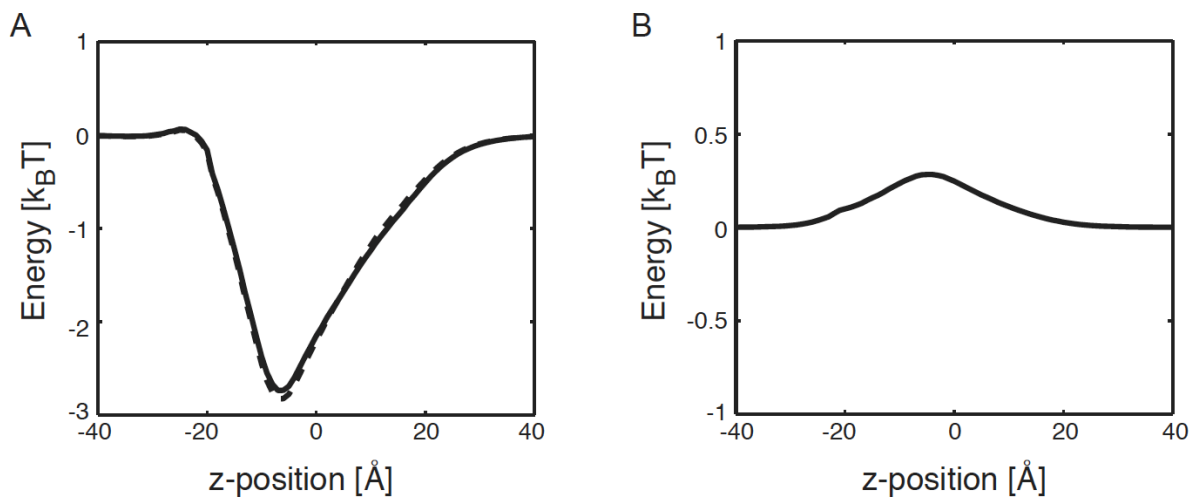
Figure 3.2: Protein charges dominate the ion transfer free energy. (a) Ion transfer free-energy profiles through mVDAC1 for a chloride-sized anion in the absence (dashed curve) and presence (continuous curve) of the membrane. The membrane has very little effect on the permeation energetics. (b) Born solvation energy for anion permeation. The energy peaks near the N-terminal helix (-7.5 Å) at 0.25 k_BT. Such a small value indicates that ions remain largely solvated during the transfer from one side of the pore to the other. a and b indicate that the electrostatic interactions of the permeant ions with the permanent charges on mVDAC1 dominate the ion transfer free energy. In both panels that protein dielectric, ϵ_p, was set to 5.

of ATP and ADP permeation energetics. Unfortunately, due to the increased complexity and conformational flexibility, continuum approaches are not appropriate for studying ATP's interactions with the channel. Fully atomistic simulations will be required to properly address this question.

3.1.2 Cation and anion pathways through the channel

The ion paths shown in Fig. 3.1 were chosen as straight lines that passed approximately through the middle of the channel. However, the pore is wide, and it is likely that the local electric fields guide anions and cations along different paths that are not straight. PB calculations were used to determine the electric field in the pore, and in Fig. 3.3, the potential contours were plotted at different heights along the channel: $z = +10$ Å (panel A), 0 Å (panel B), and -10 Å (panel C). Contours were drawn in 35 mV intervals between -35 mV and 70 mV. At all levels, the pore is dominated by positive potentials showing that the respective anion and cation transfer free energies along single paths in Fig. 3.1 are representative of the true energetics of ion permeation. For instance, at $z = -10$ Å the potential is 70 mV or greater throughout 90% of the pore indicating that the anion minimum and cation maximum in Fig. 3.1 are present for nearly all paths through the channel. For reference, a 60-75 mV potential corresponds to a 2-3 k_BT energy for a monovalent anion at room temperature, in excellent accord with the values in Fig. 3.1 A. Interestingly, panels A and B indicate that cations will flow along the side of the channel opposite the helix (negative y values), while anions will occupy a larger space closer to the helix (positive y values). The 0 mV contour forms the dividing line between favorable and non-favorable positions for both types of ions.

3.1.3 The permeation energetics of mutant channels

VDAC has an open-state chloride-to-potassium ion selectivity ratio of 1.7-1.9 in a 1.0 to 0.1 M asymmetric salt gradient. [21] These selectivity values are modest compared to ion channels such as the Shaker voltage-gated potassium channel (1000:1) [96], which has a selectivity filtera narrow stretch made up of a handful of residues responsible for selecting one ion over another via interactions of the carbonyl backbone with the permeating ion [97]. Blachly-
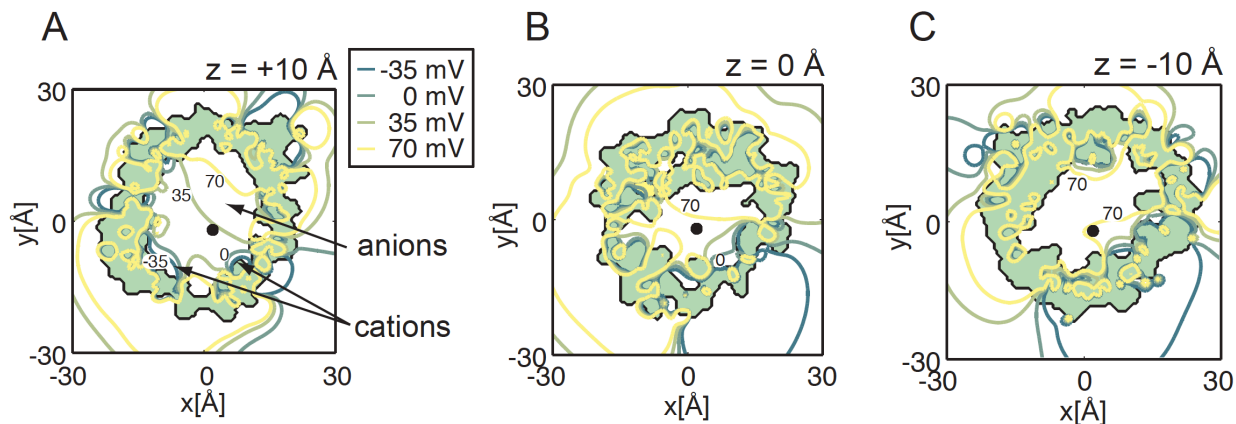
Figure 3.3: The electrostatic potential at three different heights in the pore. The potential contours due to the protein charges were calculated as in Fig. 3.1 and then plotted at $z = +10$ Å (a), $z = 0$ Å (b) and $z = -10$ Å (c). Each view is along the pore's long axis, and the protein interior is shaded light green. The ion path used for all calculations in Figs. 3.1, 3.2, and 3.4 is indicated by a black dot in each panel. Four equally spaced isocontours were chosen with the green representing negative potential values and the yellow values representing more positive values. The contour at 0 mV is indicated in (a) and (b), while the +70 -mV contour is shown in each. The 0 -mV contour divides the region favorable to anions (positive values at positive y values) and cations (negative values at negative y values). The pore is dominated by positive potential values at all heights, especially at $z = -10$ Å. The ion path falls along contour values that are representative of the average value at each height. At $z = +10$ Å the value is +30 mV, near the middle of the full range of values. In (b), the positive potentials occupy more of the cross-sectional area, and the electrostatic potential at the ion pathway increases to +50 mV. In (c), close to the N-terminal helix, the potential is entirely positive. This is in accord with the anion free-energy minimum and cation barrier observed in Fig. 3.1 A and B.

Dyson and coworkers showed that mutations throughout the primary sequence of VDAC from Saccharomyces cerevisiae (scVDAC1) could affect channel selectivity, and the largest changes involved charge mutations (such as lysine to glutamic acid) [32]. Additionally, the effects from multiple mutations were additive, which is suggestive of an electrostatic mechanism in which the electric fields from multiple residues add linearly at the site of the permeating ion. This mechanism is in stark contrast to those of potassium channels, which are believed to depend on the details of packing geometry and fluctuations of the channel around the ion [98,99]. We investigated those residues that were previously shown to contribute to anion selectivity.

Blachly-Dyson et al. carried out site-directed mutagenesis and identified 14 sites that affect selectivity and 12 that, when mutated, showed little or no change in selectivity [32]. We aligned mVDAC1 to scVDAC1 and found that 11 of the 14 residues implicated in selectivity are on the bottom half of the channel (the overall sequence identity is 25% and the sequence similarity is 51%). Interestingly, the bottom half of the channel is the most favorable location for anions in the pore as shown by the energy minimum in Fig. 3.1 C. As can be seen in Fig. 3.4 C, we made in silico mutations to some of the homologous residues in mouse VDAC1 using VMD [100], and then we recalculated the free energy of ion transfer to compare against calculations on the wild-type structure. Experimentally, mutating scVDAC1 residues K19E or K61E elicits changes in the reversal potential, which is directly related to a decrease in the channel's selectivity (see Table 3.1).

We mutated the corresponding residues in mVDAC1, K20E, and K61E and plotted the change in the free-energy profile compared to that of wild type (Fig. 3.4 A). The changes in the free-energy profile are striking. The free-energy minimum near -7.5 Å is decreased by 1.5 and 2.5 k_BT relative to that of the wild-type curve (red) for K20E (blue) and K61E (green), respectively. Moreover, both mutations present a slight barrier to permeation along the chosen path at z=-20 Å. Blachly-Dyson and coworkers also reported 12 nonselective sites where mutating basic residues to acidic ones, or vice versa, had little affect on the channel's selectivity. Our electrostatic calculations on a few of the corresponding mouse mutants (K110E, A134E, and N207E) resulted in very small changes to the free-energy profile, less than 0.4 k_BT in each case (see Table 3.1). All of the ion transfer free energies in Fig. 3.4
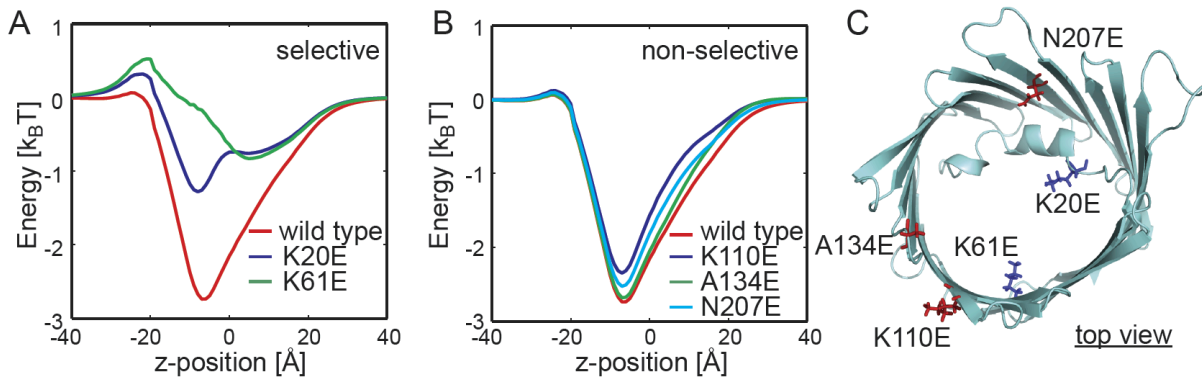
Figure 3.4: Altered ion transfer free energies through mVDAC1 match measured changes to selectivity. (a) Free-energy profile for a chloride anion moving through the wild-type (red) and two mutant channels: K20E (blue) and K61E (green). K20E and K61E significantly destabilize the anion in the channel, and this is consistent with experimental measurements. (b) Mutations K110E, A134E, and N207E have little effect on the free-energy profile, which is in excellent agreement with experiment. (c) All residues are pictured. The selective positions (blue) are much closer to the permeation pathway, while the nonselective positions (red) are farther away.

B are very similar to the wild-type curve, indicating that these mutations do not alter the biophysical properties of the channel. Thus, our calculations, like the experiments, show that these residues play a minor role in VDAC1 selectivity.

S.cerevisiae mutation	mVDAC1 mutation	ΔV_{rev} Ref. [32]	ΔV_{rev} fig3.5A	ΔV_{rev} fig3.5B	$\Delta\Delta G(k_B T)$ fig3.4
K19E	K20E	-12.3	-6.4	-7.0	1.5
K61E	K61E	-8.4	-4.1	-5.5	2.6
K19E/ K61E/ K95E	K20E/ K61E/ K96E	NA	-13.4	-19.0	NA
K108E	K110E	-0.8	-2.5	-0.8	0.4
K132E	K134E	-0.1	-1.4	~0.0	0.2
K205E	N207E	-1.0	-1.8	-0.3	0.05

Table 3.1: Summary of mutant channel properties. We considered three mutations that affected selectivity (upper three rows) and three that did not (lower three rows). The original mutagenesis was carried out on scVDAC1, and we present the homologous mouse VDAC1 residues mutated for the present study. The change in reversal potential, Vrev, under 1.0 to 0.1 M KCl conditions are based on experiments from Ref. [32], PNP calculations with the channel orientation in Fig. 3.5A, and PNP calculations with the orientation in Fig. 3.5B. While Ref. [32] studied K19E, K61E, and K95E separately, the triple mutant was not studied; hence, this value is listed as not available (NA). The final column is the change in the minimum energy for the mutant channels compared to the wild-type channel based on the transfer free energies in Fig. 3.4.

3.1.4 PNP theory suggests an open state

Relating our calculated changes in ion free energy profiles to changes in selectivity is not straightforward; however, a change of more than 1 k_BT should bring about a macroscopic change in measured currents. To probe aspects of conductance and selectivity more deeply, we turned to a related continuum approach known as Poisson-Nernst-Planck theory. PNP

computes the steady state concentration profiles of each ionic species based on the net electrical potential distribution, which depends in a self-consistent manner on the ion concentration profiles. From the computed concentration profiles, the ionic flux through the channel in different physiological conditions can be calculated [56]. Since the ionic concentration profiles are 3-dimensional densities, these calculations do not suffer from picking a single arbitrary pathway, but rather the ions occupy the most favorable positions in the pore during permeation. We first calculated the ionic flux through wild-type mVDAC1 for a wide range of voltages in symmetric salt conditions corresponding to a 0.1 M KCl solution. Throughout this work, all membrane voltages refer to the value in the intermembrane space with respect to the cytosol. The current-voltage relation predicted a single-channel conductance of 1.06 nS, and the experimental open-state value is 0.45-0.58 nS in 0.1 M KCl3. It has been found from previous studies that PNP theory yields conductance values that are typically 1.5-2 times higher than experimental values for large pores [60, 101]. Therefore, given the limitations of this theory, our calculations again suggest that mVDAC1 resides in the open state.

The theoretical conductance values presented here are larger than those predicted by experiment for the open channel. PNP is a mean field theory that neglects dynamical ion-ion correlations, and it is generally thought that this can lead to a larger current flux through the pore [60,102]. Despite being inflated with respect to experiment, our calculations are in good agreement with other researchers' findings that PNP theory overestimates the experimental values. For instance, Im and Roux calculated a conductance value for OmpF that was 1.8 times higher than that of experiments carried out in 0.2 M KCl [60], while our value is 1.82.3 times higher than values reported for VDAC in 0.1 M KCl [21]. Nevertheless, such a large conductance value is even more incompatible with the closed state conductance found to be 0.22 nS in 0.1 M KCl [103]. If the mVDAC1 structure truly represents the closed state, this would imply that our PNP calculations are off by a factor of 5, and to our knowledge, such a large discrepancy between theory and experiment has never been reported. Thus, we feel that our conductance calculations suggest that mVDAC1 represents the open structure. Interestingly, while PNP overestimates the flux, it appears to do a better job of determining the relative ratio of the anion to cation flux. In our case, we predict reversal potentials

within 2 to 4 mV of the value reported for scVDAC1, and Im and Roux determined a value within 2 mV of the recorded value for OmpF [60].

3.1.5   A closer look at VDAC selectivity

Next, we solved the PNP equations under asymmetric conditions with the 0.1 M KCl in the cytosol and 1.0 M KCl in the intermembrane space. In this situation, a net current will result if the channel selects one ion type over the other. The applied voltage required to oppose this current is termed the reversal potential, and it is a direct measure of the channels selectivity [96]. The solid black curve in Fig. 3.5 A shows that mVDAC1 is indeed anion selective as suggested by the curves in Fig. 3.1. The reversal potential is 11.9 mV, and the ratio of the anion to cation conductance is about 1.75. This is in excellent agreement with the experimental value of 10.2 mV determined for scVDAC1 [32]. Since the charges on the channel are asymmetric, it is possible that the channels selectivity could depend on its orientation in the membrane. This concern is particularly relevant since the orientation of the channel in the MOM is debated [37, 104, 105]. We flipped the channel in the membrane and recalculated the current-voltage curve (solid black curve in Fig. 3.5B). In this configuration, the reversal potential increases to 14.1 mV indicating a slight gain in anion selectivity. Thus, mVDAC1 is more selective when the energy minimum in Fig. 3.1A is facing the low concentration bath, but what this means for channel function in the MOM is difficult to say at this time.

We also calculated the current-voltage curves for the point mutants investigated in Fig. 3.4 using both channel orientations. These curves are plotted on the respective panels in Fig. 3.5. Both K20E and K61E produced a leftward shift in the reversal potential indicating that they reduce the channels anion selectivity (red curves in both panels); however, these changes are more pronounced for the orientation in panel B. Our calculations predict that K20E evokes a shift 1.5 times larger than K61E, and this is exactly what is observed experimentally [32]. In general, K110E and A134E produced small changes in the reversal potential, as observed experimentally, but there are orientation dependent differences (see Table 3.1). One should note that the homologous sites for A134 and N207 in scVDAC1 are K132 and
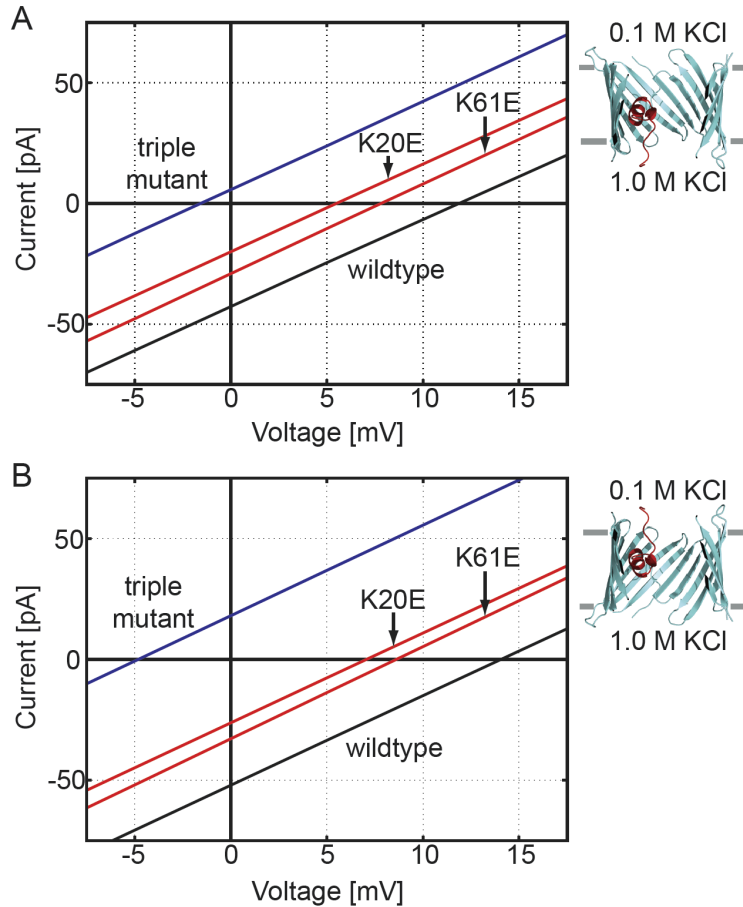
Figure 3.5: Wild-type and mutant current-voltage curves calculated using PNP theory. The cytosol was held at 0.1 M KCl and the intermembrane space at 1.0 KCl. The applied voltage is the value in the intermembrane space relative to the cytosol. Under these conditions, positive reversal potentials indicate that the channel is anion selective. (A) The channel was embedded in a dielectric slab with the N and C-termini facing the intermembrane space. The black line is the wild-type curve, and the red curves correspond to K20E and K61E. The leftward shift of both red curves indicates that both mutants reduce the anion selectivity of mVDAC1. The blue curve corresponds to the K20E/K61E/K96E triple mutant. This curve has a negative reversal potential indicating that the channel has been engineered to be cation selective. (B) The orientation of the channels in the membrane were reversed compared to panel A. In this case, the N and C-termini face the upper bath. Interestingly, the reversal potentials are all shifted compared to those in panel A demonstrating that the channels selectivity depends on its orientation in the membrane.

K205, so acidic substitution results in an overall change in charge of -2, but only -1 for the present calculations. This difference will have very little affect on the reversal potential shifts reported using the orientation in Fig. 3.5 B, but it would make the shifts calculated using the configuration in Fig. 3.5 A even larger. Finally, we wanted to convert mVDAC1 into a cation-selective channel. We created a triple mutant channel, K20E/K61E/K96E, based on three of the highest impact mutations identified by Blachly-Dyson and coworkers [32]. As can be seen from the blue curves in Fig. 3.5 A and B, this hypothetical channel has a negative reversal potential showing that we were able to convert the anion channel into a cation channel through electrostatic manipulation of the pore residues.

Our analysis illustrates that not all pore-facing residues contribute equally to channel selectivity. This was first apparent from our observation that most residues that affect selectivity in the Blachley-Dyson et al. study are on the lower half of the channel [32]. From Fig. 3.4 C one can see that A134 and N207 are both pore-facing residues, but they are at the outer mouth of the channel near the upper solution. Acidic substitution of these residues has a modest effect on the reversal potential when the channel is in the configuration in Fig. 3.5 A, but it has almost no effect on selectivity when the channel is in the opposite orientation. Meanwhile, K20E and K61E impact selectivity regardless of the orientation of the channel. We believe that these results can be explained by the proximity of the mutated residue to the channel's natural energy minima at -7.5 Å (Fig. 3.1 A). Residues that increase or decrease the depth of this minimum appear to have the largest effect on selectivity. Interestingly, the calculated reversal potential shifts elicited by point mutations found to affect selectivity, while significant, are still smaller than the experimental shifts (see Table 3.1).

We feel that there are two primary reasons for this inconsistency. First, the charged , pore-facing residues on scVDAC1 and mVDAC1 are not strictly conserved, which could lead to a different electrostatic profile in the pore of mVDAC1. Second, we do not know the true rotamer conformations of the point mutants. Incorrectly modeled side chains could heavily influence the electrostatics along the permeation pathway. Additionally, nonmutated residues may adopt new rotamer conformations in response to a point mutant, and this could not be accounted for. All of these effects could explain the differences between theory and experiment. Nonetheless, there is a clear delineation between the selective mutants and the

nonselective ones.

3.1.6 mVDAC1 selectivity and channel orientation

Our results show that VDAC1 selectivity depends on the orientation of the channel in the membrane. The channel is slightly less selective when the N and C-termini face the high-concentration bath (equivalent to 2 mV shift in reversal potential compared to both termini facing the low concentration bath). Since the orientation of VDAC1 in the mitochondria is currently debated [37,104–106], the physiological relevance of this result is unclear. However, it is obvious that our calculated reversal potential shifts correspond most closely to the experimental results when the N- and C-termini face the low-concentration bath (see Table 3.1). We believe that this finding could have consequences for our understanding of VDAC's topology in the MOM. Theory could be used to predict mutant channels that produce large shifts in the reversal potential when the N- and C-termini are presented to the cytoplasm and very little shift when presented to the inner-membrane space. Conversely, a set of mutant channels could be engineered that have the opposite property. Expressing and carefully recording these mutants directly from the MOM may provide a means to test the channel's orientation in vivo. Experimentally, we would expect large reversal potential shifts for one orientation and negligible shifts for the opposite arrangement. The control experiments would come from the point mutations that predict the opposite result.

3.2 WHAT IS VDAC1′S VOLTAGE SENSOR AND HOW DOES IT MOVE?

Next, we hoped to gain deeper insight into the operation of mVDAC1 by considering the gating motion and its dependence on membrane voltage. A signature of the closing motion is the gating charge or sensor valence associated with the movement. The sensor valence is equal to the fraction of the membrane electric field that the charges on the voltage sensor pass through during gating. One charge moving through the entire membrane electric field would contribute a valence of 1, but since the voltage-sensor charges need not pass through

the entire field this value often takes on noninteger values. The voltage dependence of the channel is directly proportional to its valence, with a high sensor valence corresponding to a steeply voltage-dependent protein. The valence of the VDAC voltage sensor varies slightly with the subtype, but it is estimated to be between 2.5 and 4.5 charge units based on single channel.

We considered two hypothetical gating motions that have been suggested in the literature. First, looking at the surface representation of mVDAC1 along the channel axis, as in Fig. 3.1 D, it is clear that an ATP molecule can pass through the channel sterically unhindered [26, 107]. However, moving the N-terminal helix into the center of the pore would obstruct this pathway. Ujwal and coworkers proposed a hypothetical closed state model in which the N-terminal helix was rigidly rotated by pivoting about the C-terminus of the helix just before the beginning of the first β-strand [26]. While this movement is speculative, electrostatic interactions between the N-terminal helix and charged residues on the wall of the β-barrel opposite the helix might account for such a rearrangement. These two states can be seen in Fig. 3.6 A and B.

The theory for using continuum electrostatic calculations to determine the voltage dependence of such movements was developed previously [108]. Briefly, a modified PB equation is solved to account for the transmembrane potential, and a series of calculations are carried out to isolate the interaction energy of the charges on the protein with this field [108]. Several studies have applied this theory to particular channels [53, 109, 110], and here we implement the procedure, using the electrostatics package APBS [50], to determine the sensor valance of this motion (gating motion 1). Figure 3.6C shows the energy difference between the hypothetical state and the starting x-ray structure, $\Delta E = E_{hypo.state}$ - E_{mVDAC1}. This is only the energy difference due to the interaction of the protein charges with the membrane electric field [108]. The switch between the states in panel a shows almost no voltage dependence (blue curve in Fig. 3.6C). The slope of this line determines the corresponding voltage sensor valence, which is 0 in this case. This strongly supports the notion that gating motion 1 is not correct. Second, N-terminal antibody studies on native MOM indicate that gating involves the helix exiting the β-barrel [24, 35], so we considered a second motion (gating motion 2) in which the helix moves into the upper bath as pictured on the right side of Fig. 3.6B. This
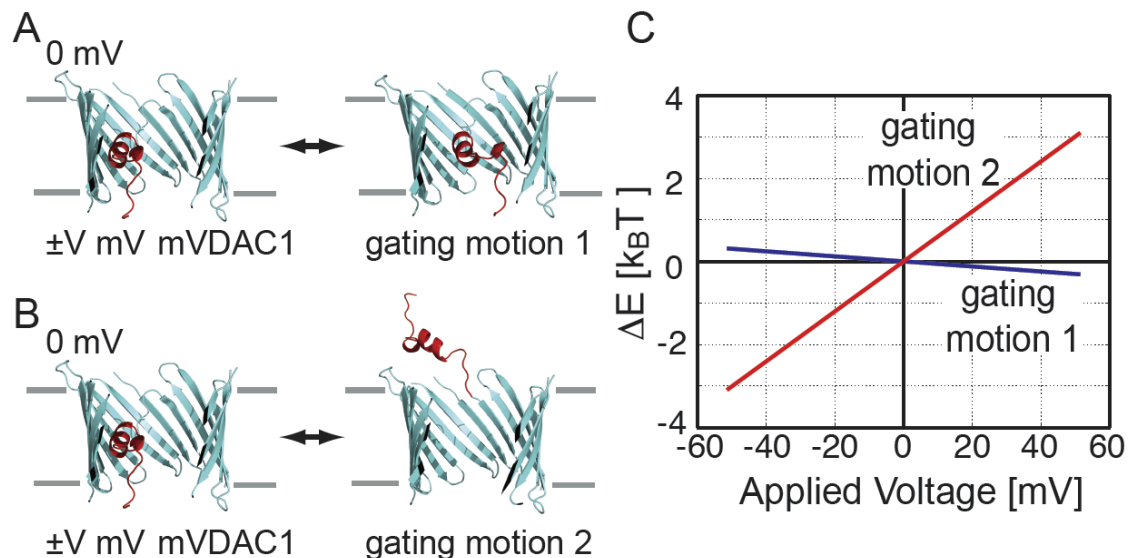
Figure 3.6: Voltage dependence of the mVDAC1 x-ray structure compared to two hypothetical gating motions. (A) mVDAC1 is pictured on the left and a hypothetical closed state suggested by Ujwal et al. [26] is on the right (gating motion 1). (B) mVDAC1 is pictured on the left and the N-terminal helix (red) has been removed from the pore on the right (gating motion 2). (C) Poisson-Boltzmann calculations were carried out to determine the membrane potentials contribution to the energy difference between these sets of structures ($\Delta E = E_{hypo.state}$ - E_{mVDAC1}). The outer bath was held at 0 mV and the inner bath was varied from -50 to +50 mV. To represent the hypothetical state in panel B, the helix was deleted from mVDAC1 as discussed in the text. The energy difference between the states in A is represented by the blue curve, and the difference between the states in B is represented by the red curve. Gating motion 1 shows no voltage dependence, while gating motion 2 has a voltage-sensor valence of 1.5.

situation is equivalent to removing the helix from mVDAC1, since the electrostatic potential of the outer bath is nearly zero everywhere. Figure 3.6C shows that positive potentials stabilize the helix in the outer bath, whereas decreasing the membrane potential stabilizes the x-ray crystal structure (red curve). This happens because the helix has a net charge of + 2. Decreasing the membrane potential from +50 to -50 mV results in a 6 k_BT stabilization of the x-ray structure compared to that of the hypothetical state. The sensor valence of gating motion 2 is only 1.5, which is again too weak to be the true motion, but it could form part of the motion as discussed below [21, 28].

With regard to motion 1, it is attractive to imagine that the N-terminal helix moves into the middle of the channel to block metabolite passage; however, this motion has no voltage dependence. Nonetheless, it is possible that a similar final state does result in channel closure. The N-terminal helix is amphipathic due to a distribution of both basic and acidic residues along its length, and therefore, simply rotating the helix in the membrane electric field could produce a sizable gating charge. Thus, the true motion could involve swinging the helix out into the pore and rotating it. Meanwhile, gating motion 2 also has some very attractive features despite producing only about half of the observed sensor valence. The reactivity of antibodies to the N-terminus increases with the lysis of the outer membrane, indicating that the N-terminus is exposed to or accessible from the inner surface of the MOM [104]. Therefore, a model of gating has been suggested that involves the helix leaving the pore prior to entering the closed state [35]. The helix may then associate with the lipid bilayer as suggested by immuno-EM [24]. This second step would require the helix to move through the membrane electric field, and it may produce enough charge movement to account for the missing one to two fundamental charge units in gating motion 2. As further confirmation that both of these motions are not correct, we calculated the electrostatic properties of the end configurations in Fig. 3.7. Both models are still anion selective, while the true closed state is cation selective. This suggests that neither of the proposed structural models is correct.
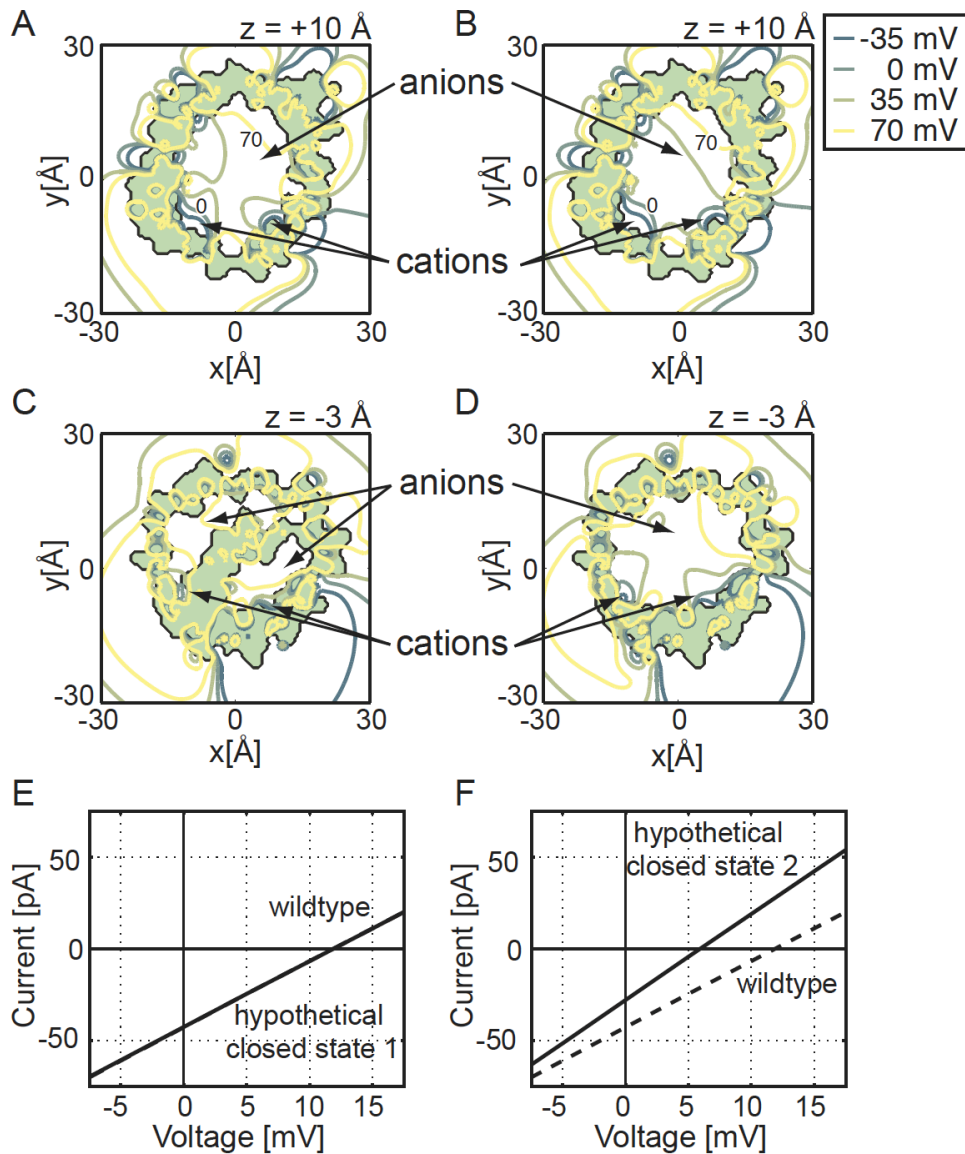
Figure 3.7: Biophysical properties of the two hypothetical closed states. The electrostatic potential of hypothetical closed state 1 was calculated and plotted at z = +10 Å (A) and z = -3 Å (C), and the isocontours for hypothetical closed state 2 were plotted at z = +10 Å (B) and z = -3 Å (D). Each channel is viewed along the pore's long axis, and the protein interior is shaded light green. We see in (C) that the N-terminal helix occludes the middle of the pore of hypothetical closed state 1, but (B) and (D) show that hypothetical closed state 2 has an unobstructed pore at all levels.

Figure 3.7: As in Fig. 4, four equally spaced isocontours are depicted with dark green curves representing negative potentials and yellow curves representing more positive potentials. From (C) we see that closed model 1 has only a very small portion of the pore favorable to cation passage. Currentvoltage curves for closed state 1 (E) and closed state 2 (F) were also carried out using PNP calculations. The conductance of hypothetical closed state 1 (continuous line) is so similar to that of wild type (dashed line) that it is hard to distinguish both curves. Due to the larger pore diameter, hypothetical closed state 2 has an increased conductance, indicated by a steeper slope (continuous curve) compared to that of the wild type (dashed). Since the currentvoltage curve in (F) crosses the x-axis at a positive potential, the channel is anion selective as is the wild-type channel.

3.2.1 Continuum electrostatics rule out hypothetical closed states

We showed in Fig. 3.6 that the hypothetical gating motions do not generate the observed gating charge; however, a second property of the closed state is that it is cation selective. To determine this, we computed the electrostatic potential in the pores at two different z values (-3 and +10 Å). Figure 3.7 A and B shows that the protein interior (light green) is identical for both models at the outer slice; however, the pore of hypothetical closed state 1 is noticeably more positive (left panel) suggesting that it would be more anion selective. At the inner slices, we can see that the helix occupies the pore in the Ujwal model (Fig. 3.7 C), but that the pore is less obstructed in hypothetical closed state 2 (Fig. 3.7 D). As with the x-ray structure, positive potentials dominate at the inner slices for both models. These positive values are due in part to the preponderance of basic residues lining the inner half of the β-barrel.

PNP calculations carried out on the hypothetical closed states show that they are both anion selective (continuous lines require positive voltages to stop ion flow in Fig. 3.7 E and F). In accord with the electrostatic potentials, hypothetical closed state 2 is less anion selective than state 1 and the mVDAC1 x-ray structure. Moreover, the conductance of state 2 is greater than those of hypothetical closed state 1 and the mVDAC1 structure as indicated

by the steeper currentvoltage curve. This is to be expected, since model 2 no longer has the helix occupying the pore domain. Thus, these hypothetical models fail to reproduce the selectivity, voltage dependence, and conductance changes required for the closed state.

4.0 THE MECHANISM OF ATP PERMEATION THROUGH VDAC

Each day, humans turnover an amount of ATP equivalent to their body weight [107], which requires a large flux of metabolites to cross the mitochondrial outer membrane. VDAC is expressed in all eukaryotes [111] and is the principal conduit for this flux from the mitochondrial intermembrane space (IMS) to the cytosol. In addition to its bioenergetic role, VDAC acts as a scaffold, influencing mitochondrial and cellular physiology through its interactions with numerous proteins and small molecules [4, 112, 113]. In accordance with these crucial functions, VDAC knockouts in mice give rise to respiratory defects, sterility, and embryonic lethality [13].

In the absence of a membrane potential, VDAC1 adopts a high conductance state (450-580 pS in 100 mM KCl) that is anion selective (1.7-1.9 anion over cation) and capable of passing 10^5-10^6 ATP molecules per second [2, 21]. Membrane voltages above or below $\pm$ 30 mV, as well as several other environmental cues, induce a conformation(s) with reduced conductance (220 pS in 100 mM KCl) that is slightly cation selective and lacks detectable ATP flux [2, 21, 103]. The structures of VDAC1 have provided additional insight into the architecture and function of the channel [26–28] revealing a 19 stranded β-barrel with an N-terminal α-helix in the central pore adjacent to the wall (Fig. 4.1). The channel has a 27 Å diameter facing both the cytosol and IMS but tapers to 14 Å in the center of the channel owing to the presence of the helix (Fig. 4.1). Continuum calculations [88] and atomistic MD simulations [33] both indicate that the solved structures are anion selective with high single-channel conductance, suggesting that the reported VDAC1 structures represent the open state of the channel. Nonetheless, it is not known if the VDAC1 structures allow ATP passage and, if so, how they achieve such a high rate of ATP flux.
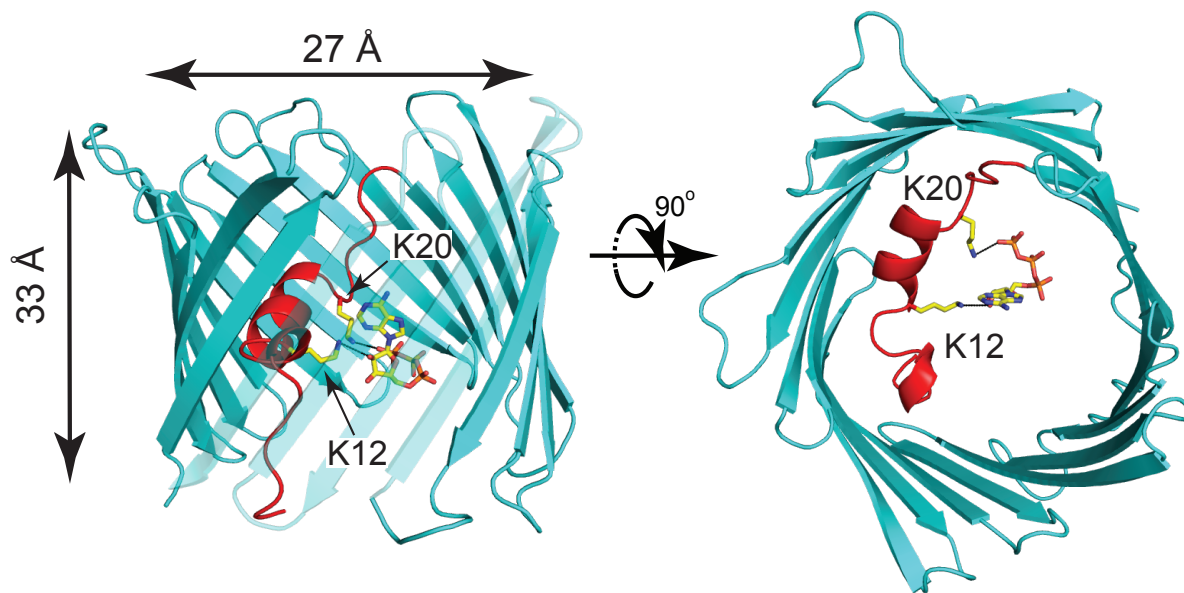
Figure 4.1: Cartoon representations of the refined mVDAC1 in complex with ATP. The left image shows a view parallel to the membrane plane, while the right is a view from the cytoplasm. ATP is bound to mVDAC1 at the center of the pore, forming hydrogen bonds with K12 and K20. The N-terminal helix of mVDAC1 is red, side chains are shown in stick representation, and the rest of the polypeptide chain is in cyan. To facilitate visualization of ATP strands 7-10 are partially transparent. The pore height and width are indicated on the left image.

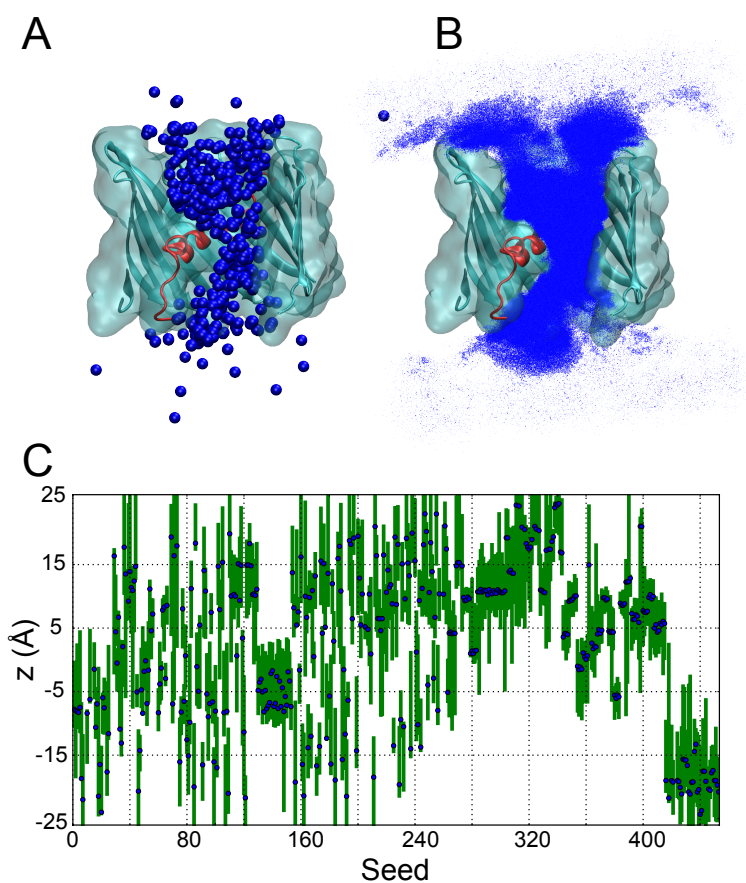Figure 4.2: Initial placement and ATP coverage of the channel for the MSM construction. (A) Blue spheres represent the β-phosphate atom of all 453 ATP configurations used to initiate MD simulations. The mVDAC1 channel is represented with the β-barrel cyan and the N-terminal helix red. The majority of simulations start with ATP in the pore domain. Membrane boundaries are indicated, but lipid and water are not shown for clarity. (B) Blue dots represent the β-phosphate of ATP plotted every 20 ps from all MD simulations initiated in A. The aggregate simulation time was 40 μs, and there are 10,000,000 ATP configurations represented. The densest regions are in the channel around the helix, indicating that configurations in the pore domain were highly sampled. (C) Extent of ATP motion during each individual simulation. The starting z position (blue dot) and range of motion (green bar) of the COM of the ATP for each of the 453 MD simulations. The ATP undergoes 10-20 Å movements in many simulations, but never crosses the entire channel.

4.1 ATP BINDING SITE IN VDAC

Experiments using current fluctuations analysis [42], structure specific noise generation [40], and mass spectrometry [43], suggested that ATP weakly interacts with the N-terminal α-helix and a patch of basic residues on the β-barrel wall. To identify potential binding sites, our collaborators from Abramson lab undertook the structural characterization of mVDAC1 crystals soaked for one hour in 50 mM ATP, revealing a low-affinity ATP binding site on the N-terminal α-helix. Both the adenosine and ribose rings are coordinated by K12, while the phosphate oxygen interacts with K20 (Fig. 4.1). This ATP binding site displays weak electron density with a low real-space map correlation (0.67) inferring that it is a low-affinity binding site and that ATP is highly mobile in the pore. Other than the addition of the lone ATP binding site, the channel remains virtually unaltered with respect to the ligand-free structure [26] having a 0.36 Å root-mean-squared deviation (RMSD).

4.2 SINGLE-CHANNEL CONDUCTANCE AND SELECTIVITY

While the open state of the channel possesses a slight preference for anions over cation, there is some evidence that the electrophysiological properties (conductance, selectivity) of VDAC depends on the bulk salt concentration, and the presence of nucleotide in the channel [21, 114, 115]. Rostovtseva and Colombini measured the single channel conductance of open VDAC reconstituted into planar phospholipid membranes [2]. In presence of 0.1 M KCl, the addition of ATP on the cis side increased the single-channel conductance even though the current contributed by ATP flux was minor [2]. Upon further increasing ATP concentration, conductance decreases again to the same value as in absence of ATP [2]. This biphasic nature of the conductance curve indicates two ATP-binding sites in VDAC - one that increases the conductance, and other that diminishes it. The presence of ATP inside the channel could further induce the channel opening, or facilitate the ion diffusion through electrostatic interactions, causing an increase in ionic conductance. Addition of NADPH has been shown to decrease the mean current through VDAC [40]. This could be due to the fact

that NADPH induces channel closure resulting in $\sim$ 6-fold reduction in the permeability [41].

VDAC has an open-state chloride-to-potassium ion selectivity ratio of 1.7-1.9 in a 1.0 to 0.1 M asymmetric salt gradient [21].

To understand the effects of ATP on the conductance and selectivity of VDAC, we carried out two independent series of MD simulations at varying voltage (-100 to 100 mV in steps of 25 mV). The simulations were run for 60 ns each in presence of 1.0 M KCl. In the first setup, ATP was present in the channel in the most probable conformation near the N-terminal α-helix (see Sec. 4.4.2) and the second set of simulations were carried out in absence of ATP.

For both sets of simulations, the conductance was calculated as a slope of I/V plot. The conductance in the absence of ATP was 2.83 nS, and this value dropped to 1.62 nS in presence of ATP. The reduction was most likely due to the steric effect of ATP interfering with the flow of K^+ and Cl^- ions. The calculated conductance value of 2.83 differs slightly from the experimental value of $\sim$3.8 in 1.0 M KCl. The ion selectivity in absence of ATP was $\sim$2.2 for anions over cations, and $\sim$1.3 in presence of ATP. The reduction in the selectivity in presence of ATP could be due to electrostatic repulsion between negatively charged ATP and incoming anions. The selectivity ration of 2.2 is in excellent agreement with the 1.7-1.9 anion selectivity ratio estimated experimentally [21], and by Wompil Im and coworkers using MD simulations on hVDAC1 [33].

4.3 LONG MULTI-MICROSECOND MD SIMULATIONS

4.3.1 ATP permeation

In an effort to monitor ATP flux through the pore, long, multi-microsecond MD simulations were carried out using the Anton special-purpose supercomputer [116]. The apo mVDAC1 structure was embedded in a homogenous DMPC bilayer with both termini facing the lower bath, which corresponds to the IMS [117]. A single Mg^{2+} coordinated ATP molecule (5 mM effective concentration) was placed in the upper or lower bath. Several independent

simulations lasting 0.6-4.8 μs were initiated from these configurations (Table 2.1). In each case, ATP entered the pore but failed to permeate the entire channel, even under the application of a membrane potential, as is commonly used in translocation studies (Fig. 4.3). For the majority of each simulation, the phosphate tail of the ATP remained in contact with basic residues K12, R15 and K20 on the N-terminal α-helix, in agreement with the crystallographic data. Meanwhile, the adenine ring processively explored different portions of the channel including residues on the β-barrel wall opposite the helix (Fig. 4.3, Movie 1). These simulations confirm the importance of the N-terminal α-helix in forming favorable interactions with ATP.

4.4   MARKOV STATE MODELS

As an alternative approach to determine if, and how, ATP permeates the mVDAC1 structure, we employed a computational method known as Markov state modeling [78,79,118]. A MSM of ATP movement through the channel was constructed by combining the results of 453 short simulations, each lasting 40-130 ns. Multiple simulations were initiated from an ensemble of ATP configurations located throughout the pore and bathing solutions, including the position observed in the co-crystal structure. We generated 40 μs of aggregate simulation time, resulting in 10 million snapshots saved at 4 ps intervals. All snapshots were aligned to a common reference structure by superposing the non-hydrogen atoms of the β-barrel.

Next, a hybrid k-means/k-medoid algorithm within the software MSMBuilder [119] was used to cluster the positions of all ATP molecules within the pore domain of the channel with a 6.5 Å cluster radius, resulting in 210 conformational states. The transition probability matrix, $P_{ij}(\tau)$, for ATP moving to state j in time given that it was initially in state i was constructed from the simulation data using a maximum likelihood reversible estimator [119]. To focus the computational effort on the slow transitions, we further developed a framework for coupling an atomistic MSM with a continuum treatment of ATP diffusion in bulk solution, similar to a previous study of proton transport in rotary ATPases [120]. Additional details concerning MSM construction can be found in the Theory and Methods section.
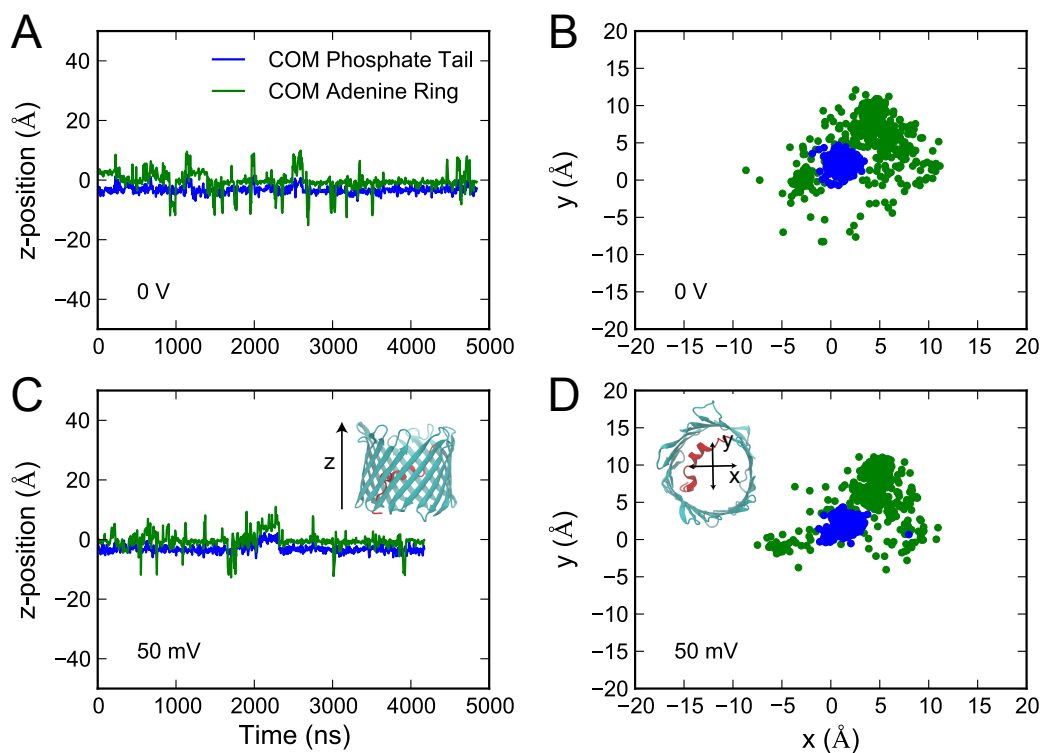
Figure 4.3: ATP fails to permeate mVDAC1 on the microsecond timescale. (A,C) The z-position of the COM of the triphosphate tail and adenine ring are plotted versus time for two simulations run with either a 0 or 50 mV electric field. The channel was centered at z = 0 and aligned along the z-axis. (B,D) The same trajectories, projected onto the x-y plane. For the simulation performed with a 50 mV field, the ATP was initiated in the upper bath; the simulation at 0 mV was initiated in the pore. Despite their different initial positions within the simulation box and applied voltage, the ATP in both simulations bind to the N-terminal helix and sample a similar set of conformations. The phosphate tail remain largely immobile over the full length of the simulation, whereas the adenine ring interacts with a number of moieties within the channel.

4.4.1 ATP mobility

ATP shows significant mobility in the pore. In 20% of the simulations, the ATP moved greater than 15 Å along the z-axis (Fig. 4.2), resulting in either entry into the pore from bulk solvent or exit to the upper or lower baths (see Movie 2).

4.4.2 Stationary state distribution and implied timescales

The left eigenvalues and eigenvectors of $\boldsymbol{P}(\tau)$ are related to the natural relaxation timescales and transitions between states in the MSM, respectively. For an ergodic and irreducible Markov state model, there is a single eigenvector with an eigenvalue of 1:

$$\boldsymbol{\pi}\boldsymbol{P}(\tau) = \boldsymbol{\pi}. \tag{4.1}$$

where $\boldsymbol{\pi}$ is the stationary distribution vector plotted in Fig. 4.4A. This vector determines the equilibrium probability of finding ATP in a given state. The equilibrium distribution of ATP in the channel was determined from the transition probability matrix, and the probability of each of the 210 conformational states in the MSM is shown in Fig. 4.4A.

The dominant state has a modest occupancy of 16%, while that of the next 9 most populated states ranges from 2-6%. The longest Anton simulation (4.8 μs) visits many, but not all, of these states at a similar frequency indicating that they can interconvert on the microsecond timescale (Fig. 4.4A), while also validating our MSM approach. The two most stable configurations show ATP interacting with N-terminal α-helix residues K12, R15, and K20 (Fig. 4.4B), consistent with our structural observations. The 8 following top states, however, are structurally distinct, which likely explains the weakness of ATP binding in the co-crystal structure. We note, in particular, that the 9th most populated state corresponds to a configuration where the phosphate tail interacts with β-barrel residues K113 and K115 (Fig. 4.4B), which were previously identified using mass spectroscopy [43].

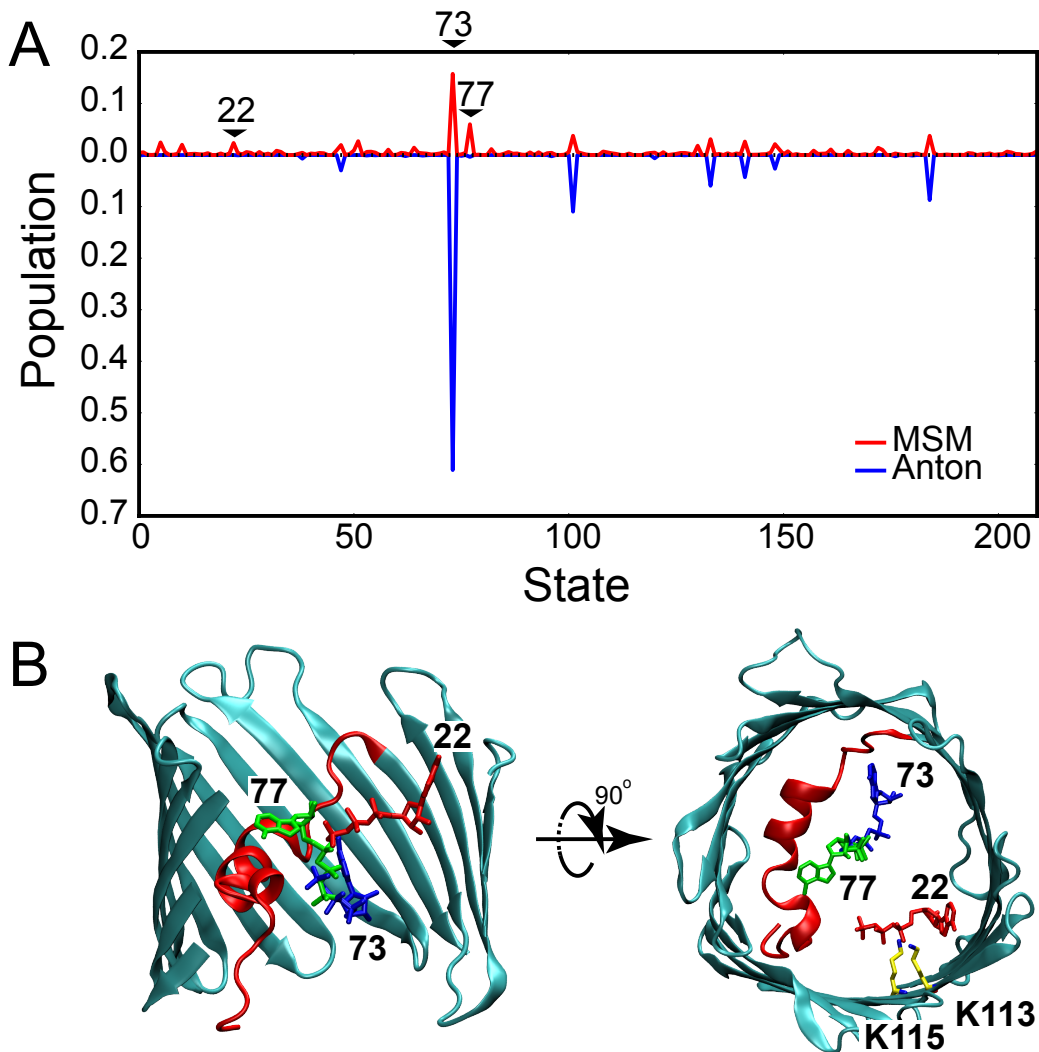The relaxation times, or implied timescales, $\hat{t}$, are computed from the eigenvalues as

Figure 4.4: ATP adopts many conformations in the mVDAC1 pore. (A) Steady state distribution of ATP in the pore domain, defined as -18 Å $\leq$ z $\leq$ 20 Å, calculated from the Markov state model (red). The most stable state (73) has 16% occupancy. The data from the longest Anton simulation (4.8 μs) was projected onto the MSM states (blue), and the time spent in each state was used to determine the relative probability of occupancy (plotted in the negative direction for clarity). All of the most populated states from the Anton simulation are also predicted to be stable based on the MSM. (B) The most stable ATP configurations are structurally diverse. The generators for the 1st (73, blue), 2nd (77, green), and 9th (22, red) most populated states are depicted. The phosphate tails of states 73 and 77 interact with basic residues (K12, R15, and K20) on the N-terminal helix, while the phosphate tail of state 22 interacts with K113 and K115 on the wall of the β-barrel.

follows:

$$\hat{t}_i = -\frac{\tau}{\ln \lambda_i(\tau)}, \tag{4.2}$$

where $\lambda_i(\tau)$ is the ith eigenvalue of $\boldsymbol{P}(\tau)$. In general, the implied timescales plateau and become constant at long lag times in which the model obeys Markovian dynamics. To determine an appropriate lag time to construct our MSM, we computed the implied timescales for τ ranging from 0 to 40 ns, and plotted the 50 slowest transitions in Fig. 4.5.

The resulting spectrum is well behaved, and based on the asymptotic behavior, we chose a lag time of 5 ns to construct our MSM for all subsequent analysis.

4.4.3   Local equilibration times within states

To ensure that there are no kinetic barriers within MSM states, we calculated the time required for ATP to relax within each state. Following the approach in Ref. [90], we partitioned each state into two sub-states using the same hybrid K-centers/K-medoids algorithm used to construct the full MSM, as shown in Fig. 4.6. We then computed the 2×2 transition probability matrix between the two sub-states for a range of lag times. The relaxation time between the two sub-states, $\hat{t}$, is related to the second eigenvalue, λ_2, of the sub-state transition probability matrix:

$$\hat{t}_2 = -\frac{\tau}{\ln \lambda_2(\tau)}. \tag{4.3}$$

As seen in Fig. 4.6, the internal relaxation times of all but 23 of the 210 generators were less than the lag time used in the construction of the full MSM. Since the majority of the states relax faster than the lag time used to construct the MSM, it indicates that there are no internal barriers within states and that they are well equilibrated. Moreover, removing these 23 states from our analysis had no significant change in the MFPT for ATP permeation.

4.4.4   ATP flux through the channel

The rate of ATP flux through the channel was computed from the MSM, allowing us to compare our simulations with experimentally determined rates. To do this, we calculated
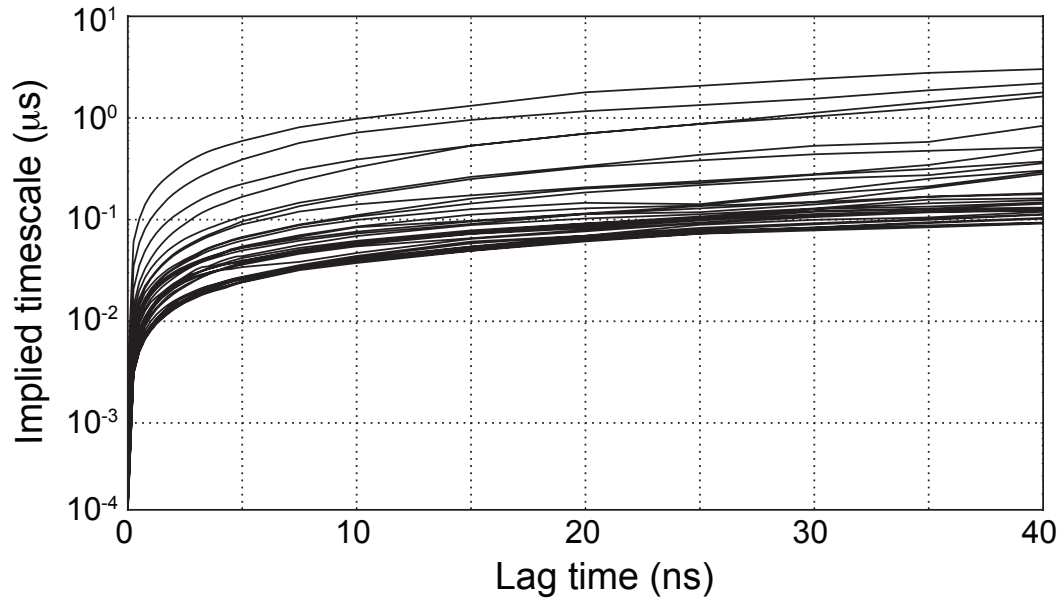
Figure 4.5: Implied timescales for the Markov state model. The 50 slowest implied timescales are plotted on a log scale for a range of lag times, τ. Implied timescales were computed using Eq. 4.2, and the transition probability matrix obtained from the full 40 μs dataset coupled to the continuum bath as described in Sec. 2.4.3. The lag time chosen for all analysis in the main text was 5 ns, since the timescales plateau at this point indicating that the model is Markovian.
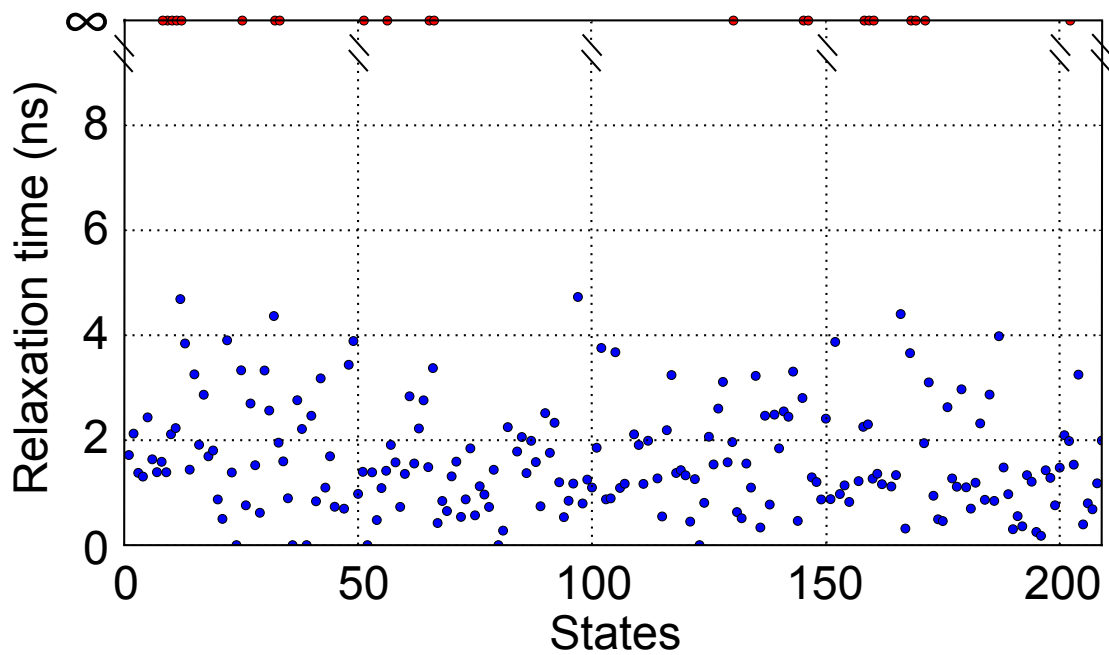
Figure 4.6: States in the Markov state model are well equilibrated. Each state in the MSM was separately analyzed for internal barriers. To do this, all configurations in each state were isolated and clustered into two substates using the k-centers/k-medoids algorithm. The transition probability matrix at a 5 ns lag time was constructed, and the second eigenvalue was used to identify the relaxation time between the substate according to Eq. 4.2. Of the 210 states in the full MSM, 187 states (89%) relaxed faster than the 5 ns lag time used to construct the full MSM (blue circles), while 23 states (11%) relaxed much slower or failed to relax at all (red circles). Dropping these later states from the MFPT analysis had little effect on the results (data not shown).

the mean first passage time (MFPT) between states, the reciprocal of which is the associated reaction rate [121]. The MFPT is determined from the transition probability matrix by solving the following set of linear equations [78]:

$$m_{i,j} = P_{i,j}\tau + \sum_{k \neq j} P_{i,k} \left(m_{k,j} + \tau \right). \tag{4.4}$$

where $m_{i,j}$ is the MFPT from state i to j. Our simulations predict that the MFPT to move from the IMS to the cytoplasm is 27 μs under the 5 mM to 0 mM ATP gradient applied using the continuum coupling model, while the reverse MFPT, with bathing concentrations switched, is 14 μs. While these rates should be identical based on symmetry considerations, the MSM boundaries do not extend fully into bulk solution giving rise to slight asymmetries. The average rate based on these MFPTs is ~49,000 ATP/sec, which is very close to the interpolated, experimental value of 50,000 ATP/sec in 5 mM ATP recorded from VDAC channels extracted from Neurospora crassa [2, 122]. Thus, the ATP flux computed from the MSM designate the crystallographic mVDAC1 structure as an open conformation.

4.4.5   The mechanism of ATP permeation

The mechanism that ATP uses to move from the IMS to the cytoplasm was determined by applying transition path theory (TPT) [90] to the MSM (Fig. 4.8A). Our analysis shows that permeation occurs via a number of distinct and often intersecting routes (Fig. 4.8B) involving 18 of the 28 pore-facing, basic residues (Fig. 4.9 A-C). The top 5 paths, which share the same entry point and account for ~70% of the total flux, are visualized in Fig. 4.8B by plotting the coordinates of the -phosphate along each transition. In all pathways, ATP enters the channel through simultaneous interactions with the positively charged main-chain nitrogen of the N-terminal methionine and K174 on the barrel wall. Along paths 1-4, K174 maintains an interaction with the -phosphate of ATP as the nucleotide is passed to K12 on the helix (Fig. 4.9A,B). ATP then interconverts among a series of stable states that involve interactions with helix-residues K12, R15 and K20 (Fig. 4.9A,B). For paths 1-4, the rate-limiting step is thus dissociation of ATP from the α-helix. Along path 1, carrying 32% of the flux, ATP is released from the helix and transferred to R218 before exiting into the

65

cytosol. Along paths 2-4, the phosphate tail breaks free from the N-terminal helix and moves across the channel to associate with K113 and K115 on the barrel wall. From there, paths 2, 3 and 4 diverge as ATP jumps to different basic residues lining the outer mouth of the channel (Fig. 4.9B). Along path 5, which is unique in that no interactions are seen between the ATP and the α-helix basic residues, translocation involves a sequential move along a series of lysine residues (K174, K119 and K96) on the β-barrel wall, opposite to the α-helix (Fig. 4.9C). The ATP then reaches K61, and the transition to R139, closer to the cytosol, is the rate-limiting step. From here, path 5 joins with path 2 and exits via interactions with K161 and K163. Interestingly, while accounting for only 2% of the calculated flux, path 5 displays one of the highest ATP flux rates (Fig. 4.8A).

4.4.6   VDAC's voltage sensor

The potential difference across a membrane gives rise to an electric field. Charged residues within this region experience a force, and if the electric field is large enough, the residues may move thereby stabilizing the protein in a different conformation. This is the basis of voltage sensing. VDACs primary voltage sensor is thought to be the N-terminal helix [29, 30]. The N-terminal helix is flanked by three glycine residues at the C-terminal end, making it ideally suited to be a mobile element.

However, the degree of motion and exact molecular position of the helix during the gating is hotly debated [24, 34, 123–126]. Based on electrophysiological and biochemical properties of VDAC, colombini and coworkers developed a model for VDAC gating where the N-terminal α-helix slides in and out of the channel during the gating process [34, 127]. Several alternative models are proposed, where the N-terminal helix resides on the membrane surface [24, 124], or is exposed to cytosol [125, 126] in the closed state. Although these models for gating mechanism differ, they propose a large conformational change of the N-terminal α-helix.

To study if the N-terminal α-helix is stable in our MD simulation, we plotted the RMSD of the N-terminal α-helix (residues 1 to 25) for all the seeds. As seen in the figure (4.7), no significant conformational change in the N-terminal segment was observed. This supports the hypothesis that the N-terminal α-helix does not undergo independent dynamic

66
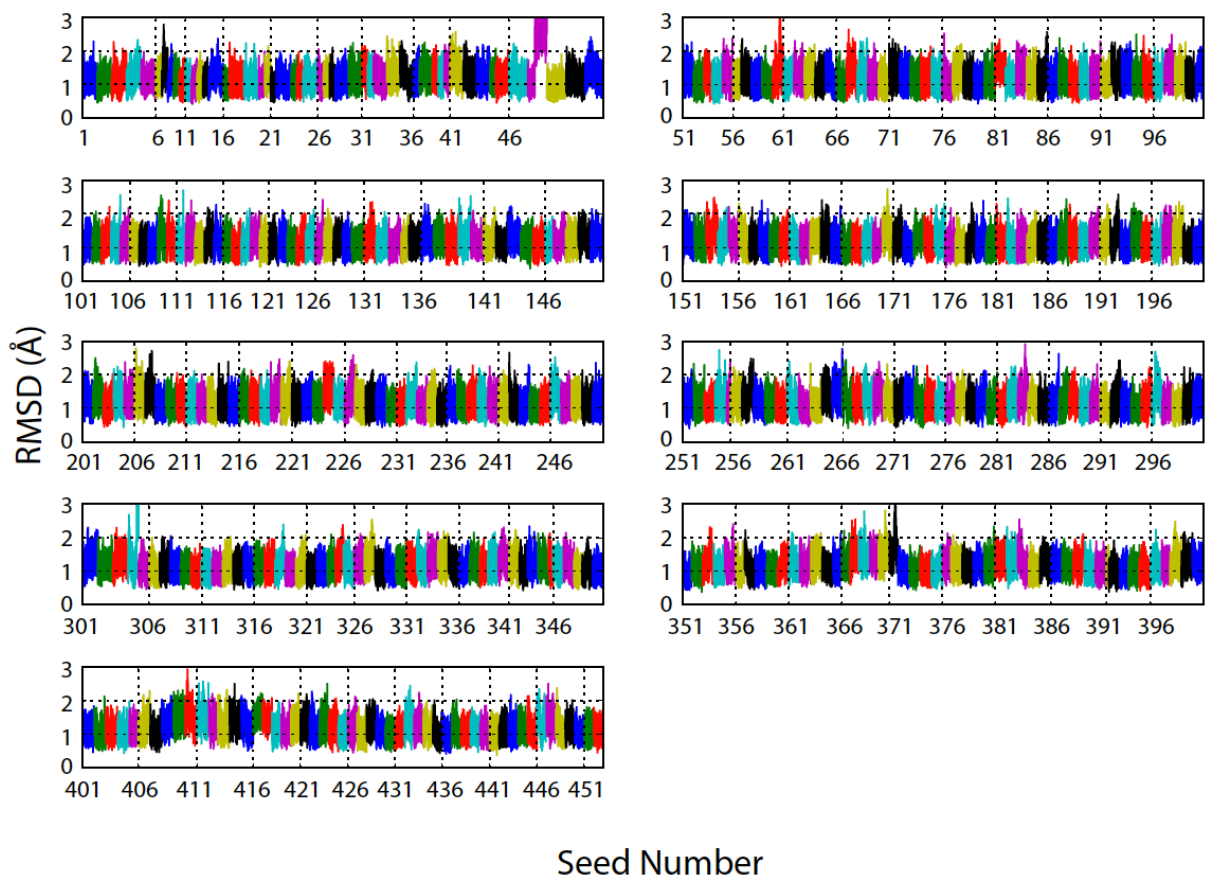
Figure 4.7: RMSD of the N-terminal α-helix for seeds 1 to 453. The RMSD was computed for residues 1 to 25 using heavy-atoms after aligning the β-barrel from all the trajectories to a reference structure.
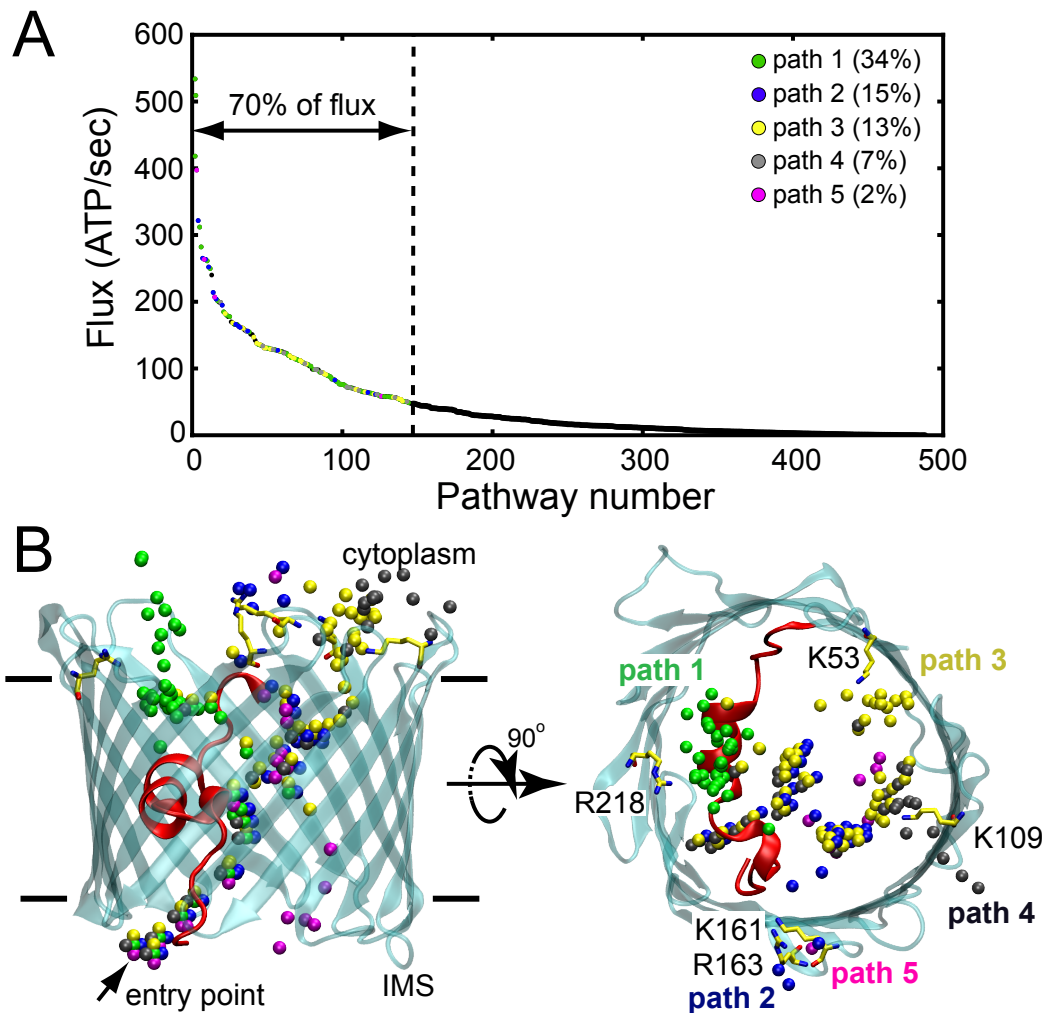
Figure 4.8: A high ATP flux is achieved through multiple, distinct pathways. (A) Transition path analysis was used to calculate the ATP flux for all the possible pathways from the IMS to the cytoplasm. The highest flux pathways giving rise to 70% of the total flux were identified and grouped into 5 primary paths based on spatial analysis in panel B. Paths were numbered according to their probability and color coded. (B) The -phosphate of ATP was plotted for all of the highest flux pathways identified in panel A. All paths enter at the same point, but exit via different locations at the cytoplasmic face. Only path 5 (magenta) avoids interactions with the central portion of the N-terminal α-helix.
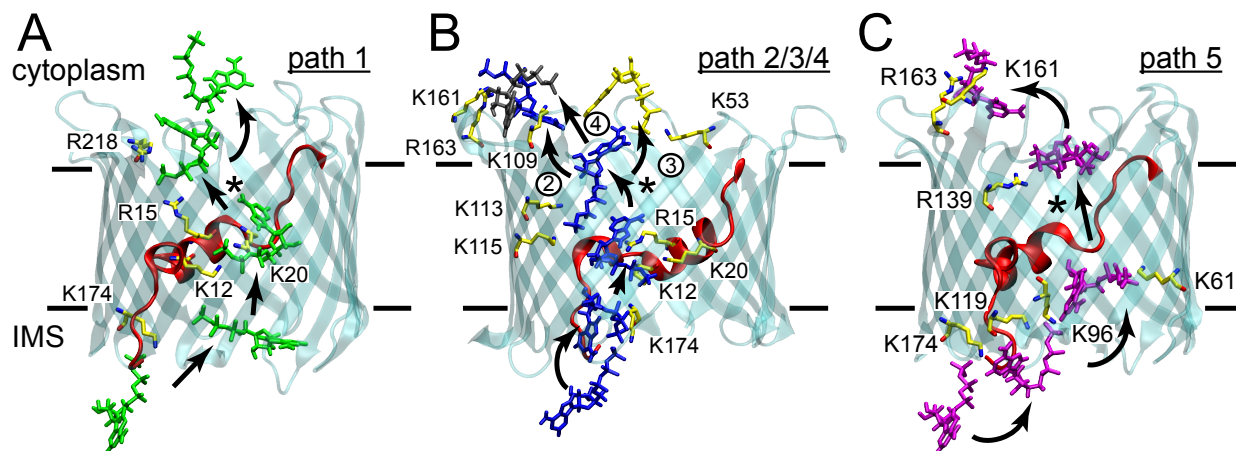
Figure 4.9: ATP permeates via a network of basic residues. (A) The highest probability pathway (34%) from the IMS to the cytoplasm moves up and over the N-terminal helix. In all panels, arrows indicate the direction of flow from the IMS to the cytosol, and * indicate the rate limiting steps. For paths 1-4, the rate-limiting step is release from the patch of basic residues on the helix (K12, R15, and K20). Solid lines delineate the membrane boundaries. (B) Paths 2-4 (15%, 13%, and 7% probability, respectively) share many states in common, except for the final state prior to escape to the cytoplasm. Numbers next to the final transition differentiate the paths. (C) Path 5 (2% probability) is the least probable path, and it avoids the basic residues on the helix.

rearrangement during the gating process (see sec. 5 for further discussion).

5.0 CONCLUSIONS AND FUTURE WORK

Both the PB and the PNP electrostatic calculations show that the mVDAC1 x-ray structure is anion selective, suggesting that the structure is open. There is a weak (2.5 k_BT), but clear minimum in the anion transfer free energy, indicating that the channel favors the permeation of anions more than cations. This might be exactly what we would expect of a channel whose primary role is to quickly shuttle high-valence, negatively charged metabolites across the membrane. Our conductance calculations using PNP suggest that mVDAC1 represents the open structure, and is conducive to anion passage. PNP and MD simulations predict the ratio of the anion to cation conductance to be 1.75 and 2.2, respectively, which are in excellent agreement with the experimental values.

Our MSM approach to small molecule transport bears the potential to overcome the current limitations of MD simulations in terms of providing a statistical description of slow processes. Here, we illustrate how ATP passes through VDAC by hopping along a series of weak binding sites (see composite movies 2 and 3 for full crossing events via paths 1 and 2) respectively, and we faithfully reproduce the experimentally determined flux. The MSM also reveals a myriad of pathways available for ATP permeation, and it provides their relative probabilities along with the occupancy of distinct ATP conformations in the channel. Such a rich statistical description is generally difficult to attain with individual, long MD simulations, but it is tractable using MSM approaches, as we have shown here. The use of MSM generated steady state distributions, along with transition pathway analysis, thus provides a novel framework for studying ligand binding to membrane proteins, as well as ion or small molecule permeation through channels and transporters.

While we believe that high ATP flux can be achieved by ATP molecules permeating one at time, it is possible that a second ATP molecule comes into the channel, and 'bumps' the

first ATP molecule, so that it can more easily diffuse through the channel. This bump may occur through steric contact, but most likely it will be a long-range electrostatic repulsion. While not evident in our MD simulations, anions such as chloride ions could facilitate the diffusion of negatively charged ATP through electrostatic repulsion. This mechanism was first suggested by Hodgkin and Huxley to explain the fast movement of ions through voltage-gated ion channels [128], and Benoit Roux later showed this to be true in a set of simulations carried out on the KcsA channel [129]. It is likely that the same principles are at work here. This could be tested by computing a PMF of ATP permeation using the 2D weighted histogram analysis methods (WHAM) [130]. we will introduce a second ATP molecule into the system and compute a two-dimensional PMF of ATP permeation. Our MSM showed that the ATP permeates through the channel using a network of residues located throughout the channel. Based on this result, we will use a relevant reaction coordinate that involves movement of the first ATP along the top 5 flux pathways identified by the MSM. The most straight-forward second reaction coordinate will the distance between the two ATP molecules along these pathways.

Also what remains to be identified are the structural rearrangements that lead to channel closure, and thereby prevent ATP from transitioning to the cystosol. VDAC1's voltage dependence is more complicated than a simple rearrangement of the N-terminal α-helix because the channel closes at both negative and positive membrane potentials. Energetically, it seems unlikely that both closed states are the same, since the membrane potential would bias the channel in different states at different voltages.

It is suggested in past studies [29,30] that the N-terminal helix, which has four basic and two acidic residues, is the voltage sensor and metabolite gate. The most common model for VDAC gating involves the movement of the N-terminal α-helix in and out of the channel during the gating process [34, 127]. It is entirely possible that the closed states involve a more complex rearrangement than simply moving the N-terminal helix, as suggested by functional work [26,34,123]. Several alternative models are proposed, where the N-terminal helix resides on the membrane surface [24,124] or is exposed to cytosol [125,126]. Although these models for gating mechanism differ, they propose a large conformational change of the N-terminal α-helix.

There is now increasing evidence that the N-terminal α-helix does not undergo independent dynamic rearrangement during the gating process. Tejido and coworkers showed that crosslinking the N-terminal α-helix to the β-barrel wall does not prevent VDAC1 voltage gating [131]. This suggests that the N-terminal α-helix is either not the voltage sensor, or does not move independently of the rest of the voltage sensing domain [131]. Further, recent NMR studies showed that the N-terminal α-helix remains in the same location in both the open and closed conformation [132]. Supporting these data, it was reported that deletion of the N-terminal α-helix did not alter VDAC voltage gating [123].

Thus, there is an ongoing controversy regarding the role of N-terminal α-helix as a voltage sensor, but a new mechanism is emerging where the β-barrel wall elongates, and pore constricts around intact the N-terminal segment [131]. Villinger and coworkers showed pronounced dynamics in β-strands 1-6 and 16-19 [133]. These residues are also found to have increased B-factors in the crystal structure compared to the other β-strand residues around the N-terminal α-helix [131].

Computationally, it is more difficult to determine these alternate conformations if they involve rearrangements of the β-barrel itself, but such a scenario cannot be ruled out. Moreover, if gating involves the N-terminal helix, it is doubtful that it moves as a rigid body. All of these factors complicate the determination of the closed states from the known mVDAC1 structure. Fortunately, the closed and open states exhibit different selectivities and conductances, so future studies can judge the likelihood of any given hypothetical state based on sensor valence calculations as well as selectivity calculations using PNP theory.

To determine the atomic movements involved in gating process, our experimental collaborators from Abramson lab attempted to resolve the structure of the closed conformation by co-crystallization and soaking existing crystals in the presence of factors that promote the closed conformation of the channel, namely NADH and pH of 5.0. However, purified mVDAC1 failed to crystallize in the presence of NADH or at lower pH, and existing crystals when soaked in these agents completely obliterated crystal diffraction. This likely indicates a conformational change from open to closed state, that prevented the crystals from packing in the original structures.

To stabilize the closed conformations and to investigate if the N-terminal α-helix under-

goes structural rearrangement upon gating, Abramson lab is now carrying out functional cross-linking studies. They have created a series of double-Cys mutants of mVDAC1 with the first Cys substitution in the N-terminal voltage sensing domain, and the second Cys mutation at various rationally selected residues (on β-strands facing the α-helix) on the channel. Each construct will be tested for functionality and subjected to cross-linkers of various sizes under conditions known to favor the open and closed conformations. These set of experiments would enable us to track the movements of the N-terminal α-helix region under different voltage and pH conditions.

To analyze changes to the VDAC topology as well as to obtain a more precise measurement of the movement of the N-terminal α-helix upon gating, Abramson lab is using single molecule FRET (smFRET) techniques. They generated eight sets of double-Cys mutants based on the atomic coordinates from the open conformation and labeled them with appropriate donor and acceptor FRET pairs. Monitoring smFRET signals would elucidate the distance changes between the labeled residues associated with the gating process.

In parallel, we will use modeling techniques to generate unbiased, hypothetical models of the closed state. Our long simulations on Anton supercomputer showed that the mVDAC1 structure is stable on the microsecond timescale, and does not undergo big conformational changes. Therefore, we do not expect to see a mVDAC1 transition from open to closed state using standard brute force MD simulations. In order to obtain the putative closed state, we will use the Weighted Emsemble (WE) method to sample the rare-event of transitioning to a closed state. WE has been shown to enhance sampling of rare transitions [134, 135]. While PB and PNP calculations can be used to determine if a model is cation selective (a requirement of the closed state), the models will be filtered based on their satisfaction of FRET and cross-linking data.

We believe that the extensive MD simulations carried out in this study will aid in the interpretation of the experimental data generated in Abramson lab. The correct model of gating process can only be achieved after evaluating all reliable data from experiments, as well as computational models.

BIBLIOGRAPHY

[1] S. J. Schein, M. Colombini, A. Finkelstein, *The Journal of Membrane Biology* **30**, 99 (1976).

[2] T. Rostovtseva, M. Colombini, *Biophysical Journal* **72**, 1954 (1997).

[3] J. M. Adams, S. Cory, *Science* **281**, 1322 (1998).

[4] S. Shimizu, M. Narita, Y. Tsujimoto, *Nature* **399**, 483 (1999).

[5] L. Galluzzi, O. Kepp, N. Tajeddine, G. Kroemer, *Oncogene* **27**, 4633 (2008).

[6] H. Azoulay-Zohar, A. Israelson, S. Abu-Hamad, V. Shoshan-Barmatz, *Biochemical Journal* **377**, 347 (2004).

[7] M. G. Vander Heiden, *et al.*, *Proceedings of the National Academy of Sciences* **97**, 4666 (2000).

[8] G. Beutner, A. Rück, B. Riede, W. Welte, D. Brdiczka, *FEBS Letters* **396**, 189 (1996).

[9] A.-K. Östlund, U. Göhring, J. Krause, D. Brdiczka, *Biochemical Medicine* **30**, 231 (1983).

[10] R. A. Nakashima, P. S. Mangan, M. Colombini, P. L. Pedersen, *Biochemistry* **25**, 1015 (1986).

[11] V. De Pinto, J. A. Al Jamal, F. Palmieri, *Journal of Biological Chemistry* **268**, 12977 (1993).

[12] J. Rasschaert, W. J. Malaisse, *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1015**, 353 (1990).

[13] M. Colombini, *Molecular and Cellular Biochemistry* **256**, 107 (2004).

[14] S. Shimizu, A. Konishi, T. Kodama, Y. Tsujimoto, *Proceedings of the National Academy of Sciences* **97**, 3100 (2000).

[15] T. K. Rostovtseva, *et al.*, *Journal of Biological Chemistry* **279**, 13575 (2004).

[16] M. G. Vander Heiden, *et al.*, *Journal of Biological Chemistry* **276**, 19414 (2001).

[17] M. J. Sampson, *et al.*, *Journal of Biological Chemistry* **276**, 39206 (2001).

[18] K.-D. Hinsch, *et al.*, *Journal of Biological Chemistry* **279**, 15281 (2004).

[19] X. Xu, W. Decker, M. Sampson, W. Craigen, M. Colombini, *The Journal of Membrane Biology* **170**, 89 (1999).

[20] M. d. C. Cesar, J. E. Wilson, *Archives of Biochemistry and Biophysics* **422**, 191 (2004).

[21] M. Colombini, *Journal of Membrane Biology* **111**, 103 (1989).

[22] V. De Pinto, O. Ludwig, J. Krause, R. Benz, F. Palmieri, *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **894**, 109 (1987).

[23] C. A. Mannella, *The Journal of Cell Biology* **94**, 680 (1982).

[24] X. W. Guo, *et al.*, *Journal of Structural Biology* **114**, 41 (1995).

[25] B. W. Hoogenboom, K. Suda, A. Engel, D. Fotiadis, *Journal of Molecular Biology* **370**, 246 (2007).

[26] R. Ujwal, *et al.*, *Proceedings of the National Academy of Sciences* **105**, 17742 (2008).

[27] M. Bayrhuber, *et al.*, *Proceedings of the National Academy of Sciences* **105**, 15370 (2008).

[28] S. Hiller, *et al.*, *Science* **321**, 1206 (2008).

[29] D. A. Koppel, *et al.*, *Journal of Biological Chemistry* **273**, 13794 (1998).

[30] M. Colombini, E. Blachly-Dyson, M. Forte, *Ion Channels* **4**, 169 (1996).

[31] T. K. Rostovtseva, W. Tan, M. Colombini, *Journal of Bioenergetics and Biomembranes* **37**, 129 (2005).

[32] E. Blachly-Dyson, S. Peng, M. Colombini, M. Forte, *Science* **247**, 1233 (1990).

[33] H. Rui, K. I. Lee, R. W. Pastor, W. Im, *Biophysical Journal* **100**, 602 (2011).

[34] J. Song, C. Midson, E. Blachly-Dyson, M. Forte, M. Colombini, *Biophysical Journal* **74**, 2926 (1998).

[35] C. A. Mannella, *Journal of Structural Biology* **121**, 207 (1998).

[36] N. E. Schoppa, K. McCormack, M. A. Tanouye, F. J. Sigworth, *Science* **255**, 1712 (1992).

[37] S. Stanley, J. A. Dias, D. D'Arcangelis, C. A. Mannella, *Journal of Biological Chemistry* **270**, 16694 (1995).

[38] L. Shao, K. W. Kinnally, C. A. Mannella, *Biophysical Journal* **71**, 778 (1996).

[39] P. S. Mangan, M. Colombini, *Proceedings of the National Academy of Sciences* **84**, 4896 (1987).

[40] T. K. Rostovtseva, A. Komarov, S. M. Bezrukov, M. Colombini, *Biophysical Journal* **82**, 193 (2002).

[41] A.-C. Lee, M. Zizi, M. Colombini, *Journal of Biological Chemistry* **269**, 30974 (1994).

[42] T. K. Rostovtseva, S. M. Bezrukov, *Biophysical Journal* **74**, 2365 (1998).

[43] G. Yehezkel, N. Hadad, H. Zaid, S. Sivan, V. Shoshan-Barmatz, *Journal of Biological Chemistry* **281**, 5938 (2006).

[44] J. Song, M. Colombini, *Journal of Bioenergetics and Biomembranes* **28**, 153 (1996).

[45] G. Yehezkel, S. Abu-Hamad, V. Shoshan-Barmatz, *Journal of Cellular Physiology* **212**, 551 (2007).

[46] A.-C. Lee, X. Xu, M. Colombini, *Journal of Biological Chemistry* **271**, 26724 (1996).

[47] T. Sugiyama, S. Shimizu, Y. Matsuoka, Y. Yoneda, Y. Tsujimoto, *Oncogene* **21**, 4944 (2002).

[48] M. G. Vander Heiden, C. B. Thompson, *Nature Cell Biology* **1**, E209 (1999).

[49] M. Y. Liu, M. Colombini, *Journal of Bioenergetics and Biomembranes* **24**, 41 (1992).

[50] N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, *Proceedings of the National Academy of Sciences* **98**, 10037 (2001).

[51] D. Sitkoff, N. Ben-Tal, B. Honig, *The Journal of Physical Chemistry* **100**, 2744 (1996).

[52] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, *Nucleic Acids Research* **32**, W665 (2004).

[53] M. Grabe, H. Lecar, Y. N. Jan, L. Y. Jan, *Proceedings of the National Academy of Sciences* **101**, 17640 (2004).

[54] B. Roux, R. MacKinnon, *Science* **285**, 100 (1999).

[55] M. Grabe, D. Bichet, X. Qian, Y. N. Jan, L. Y. Jan, *Proceedings of the National Academy of Sciences* **103**, 14361 (2006).

[56] M. G. Kurnikova, R. D. Coalson, P. Graf, A. Nitzan, *Biophysical Journal* **76**, 642 (1999).

[57] D. Shaw, *et al.*, *ACM SIGARCH Computer Architecture News* (ACM, 2007), vol. 35, pp. 1–12.

[58] S. Jo, T. Kim, V. Iyer, W. Im, *Journal of Computational Chemistry* **29**, 1859 (2008).

[59] L. Kalé, *et al.*, *Journal of Computational Physics* **151**, 283 (1999).

[60] W. Im, B. Roux, *Journal of Molecular Biology* **322**, 851 (2002).

[61] M. Cohn, T. Hughes, *Journal of Biological Chemistry* **237**, 176 (1962).

[62] J.-C. Liao, S. Sun, D. Chandler, G. Oster, *European Biophysics Journal* **33**, 29 (2004).

[63] S. Jo, T. Kim, W. Im, *PLoS One* **2**, e880 (2007).

[64] A. MacKerell Jr, *et al.*, *The Journal of Physical Chemistry B* **102**, 3586 (1998).

[65] A. MacKerell Jr, M. Feig, C. Brooks III, *Journal of the American Chemical Society* **126**, 698 (2004).

[66] J. Klauda, *et al.*, *The Journal of Physical Chemistry B* **114**, 7830 (2010).

[67] N. Foloppe, A. MacKerell Jr, *Journal of Computational Chemistry* **21**, 86 (2000).

[68] J. Pavelites, J. Gao, P. Bash, A. Mackerell, *Journal of Computational Chemistry* **18**, 221 (1998).

[69] W. Jorgensen, J. Chandrasekhar, J. Madura, R. Impey, M. Klein, *The Journal of Chemical Physics* **79**, 926 (1983).

[70] J. Phillips, *et al.*, *Journal of Computational Chemistry* **26**, 1781 (2005).

[71] J. Ryckaert, G. Ciccotti, H. Berendsen, *Journal of Computational Physics* **23**, 327 (1977).

[72] K. Bowers, *et al.*, *SC 2006 Conference, Proceedings of the ACM/IEEE* (IEEE, 2006), pp. 43–43.

[73] G. Martyna, D. Tobias, M. Klein, *The Journal of Chemical Physics* **101**, 4177 (1994).

[74] Y. Shan, J. L. Klepeis, M. P. Eastwood, R. O. Dror, D. E. Shaw, *The Journal of Chemical Physics* **122**, 54101 (2005).

[75] L. Maragliano, G. Cottone, G. Ciccotti, E. Vanden-Eijnden, *Journal of the American Chemical Society* **132**, 1010 (2010).

[76] H. Vashisth, L. Maragliano, C. F. Abrams, *Biophysical Journal* **102**, 1979 (2012).

[77] L. Maragliano, E. Vanden-Eijnden, *Chemical Physics Letters* **426**, 168 (2006).

[78] N. Singhal, C. D. Snow, V. S. Pande, *The Journal of Chemical Physics* **121**, 415 (2004).

[79] J. Chodera, N. Singhal, V. Pande, K. Dill, W. Swope, *The Journal of Chemical Physics* **126**, 155101 (2007).

[80] X. Huang, G. Bowman, S. Bacallado, V. Pande, *Proceedings of the National Academy of Sciences* **106**, 19765 (2009).

[81] N. Singhal, V. Pande, *The Journal of Chemical Physics* **123**, 204909 (2005).

[82] N. Hinrichs, V. Pande, *The Journal of Chemical Physics* **126**, 244101 (2007).

[83] G. Bowman, X. Huang, V. Pande, *Methods* **49**, 197 (2009).

[84] K. A. Beauchamp, *et al.*, *Journal of Chemical Theory and Computation* **7**, 3412 (2011).

[85] R. Scalco, A. Caflisch, *The Journal of Physical Chemistry. B* **115**, 6358 (2011).

[86] H. C. Berg, *Random Walks in Biology* (Princeton University Press, 1993).

[87] K. Yoshizaki, H. Watari, G. K. Radda, *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research* **1051**, 144 (1990).

[88] O. Choudhary, *et al.*, *Journal of Molecular Biology* **396**, 580 (2010).

[89] P. Metzner, C. Schütte, E. Vanden-Eijnden, *Multiscale Modeling & Simulation* **7**, 1192 (2009).

[90] F. Noé, C. Schütte, E. Vanden-Eijnden, L. Reich, T. Weikl, *Proceedings of the National Academy of Sciences* **106**, 19011 (2009).

[91] V. A. Voelz, G. R. Bowman, K. Beauchamp, V. S. Pande, *Journal of the American Chemical Society* **132**, 1526 (2010).

[92] E. Weinan, E. Vanden-Eijnden, *Journal of Statistical Physics* **123**, 503 (2006).

[93] V. A. Voelz, G. R. Bowman, K. Beauchamp, V. S. Pande, *Journal of the American Chemical Society* **132**, 1526 (2010).

[94] B. E. Cohen, *et al.*, *Science* **296**, 1700 (2002).

[95] C. N. Schutz, A. Warshel, *Proteins: Structure, Function, and Bioinformatics* **44**, 400 (2001).

[96] B. Hille, *Ion Channels of Excitable Membranes*, vol. 507 (Sinauer Sunderland, MA, 2001).

[97] D. A. Doyle, *et al.*, *Science* **280**, 69 (1998).

[98] F. I. Valiyaveetil, M. Leonetti, T. W. Muir, R. MacKinnon, *Science* **314**, 1004 (2006).

[99] S. Y. Noskov, S. Berneche, B. Roux, *Nature* **431**, 830 (2004).

[100] W. Humphrey, A. Dalke, K. Schulten, *Journal of Molecular Graphics* **14**, 33 (1996).

[101] S. Y. Noskov, W. Im, B. Roux, *Biophysical Journal* **87**, 2299 (2004).

[102] P. Graf, M. G. Kurnikova, R. D. Coalson, A. Nitzan, *The Journal of Physical Chemistry B* **108**, 2006 (2004).

[103] G. Báthori, G. Csordás, C. Garcia-Perez, E. Davies, G. Hajnóczky, *Journal of Biological Chemistry* **281**, 17347 (2006).

[104] V. De Pinto, G. Prezioso, F. Thinnes, T. A. Link, F. Palmieri, *Biochemistry* **30**, 10191 (1991).

[105] B. M. McDonald, M. M. Wydro, R. N. Lightowlers, J. H. Lakey, *FEBS Letters* **583**, 739 (2009).

[106] R. Ujwal, D. Cascio, V. Chaptal, P. Ping, J. Abramson, *Channels* **3**, 167 (2009).

[107] S. Törnroth-Horsefield, R. Neutze, *Proceedings of the National Academy of Sciences* **105**, 19565 (2008).

[108] B. Roux, *Biophysical Journal* **73**, 2980 (1997).

[109] L. D. Islas, F. J. Sigworth, *The Journal of General Physiology* **117**, 69 (2001).

[110] B. Chanda, O. K. Asamoah, R. Blunck, B. Roux, F. Bezanilla, *Nature* **436**, 852 (2005).

[111] A. Messina, S. Reina, F. Guarino, V. De Pinto, *Biochimica et Biophysica Acta (BBA)-Biomembranes* **1818**, 1466 (2012).

[112] T. K. Rostovtseva, *et al.*, *Proceedings of the National Academy of Sciences* **105**, 18746 (2008).

[113] S. Geisler, *et al.*, *Nature Cell Biology* **12**, 119 (2010).

[114] R. Benz, *Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes* **1197**, 167 (1994).

[115] H. Abrecht, R. Wattiez, J.-M. Ruysschaert, F. Homblé, *Plant Physiology* **124**, 1181 (2000).

[116] C. Young, *et al.*, *High Performance Computing Networking, Storage and Analysis* (IEEE, 2009), pp. 1–11.

[117] K. L. Sheldon, E. N. Maldonado, J. J. Lemasters, T. K. Rostovtseva, S. M. Bezrukov, *PloS One* **6**, e25539 (2011).

[118] S. B. Ozkan, K. A. Dill, I. Bahar, *Biopolymers* **68**, 35 (2003).

[119] K. A. Beauchamp, *et al.*, *Journal of Chemical Theory and Computation* **7**, 3412 (2011).

[120] T. Elston, H. Wang, G. Oster, *Nature* **391**, 510 (1998).

[121] P. Reimann, G. Schmid, P. Hänggi (1999).

[122] T. Rostovtseva, M. Colombini, *Journal of Biological Chemistry* **271**, 28006 (1996).

[123] B. Popp, R. Benz, W. Neupert, R. Lill, *et al.*, *Journal of Biological Chemistry* **271**, 13593 (1996).

[124] S. Reymann, *et al.*, *Biochemical and Molecular Medicine* **54**, 75 (1995).

[125] V. De Pinto, *et al.*, *Italian Journal of Biochemistry* **52**, 17 (2003).

[126] V. De Pinto, S. Reina, F. Guarino, A. Messina, *Journal of Bioenergetics and Biomembranes* **40**, 139 (2008).

[127] J. Song, C. Midson, E. Blachly-Dyson, M. Forte, M. Colombini, *Journal of Biological Chemistry* **273**, 24406 (1998).

[128] A. L. Hodgkin, A. F. Huxley, *The Journal of Physiology* **116**, 449 (1952).

[129] S. Berneche, B. Roux, *Nature* **414**, 73 (2001).

[130] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, P. A. Kollman, *Journal of Computational Chemistry* **13**, 1011 (1992).

[131] O. Teijido, *et al.*, *Journal of Biological Chemistry* **287**, 11437 (2012).

[132] S. Hiller, G. Wagner, *Current Opinion in Structural Biology* **19**, 396 (2009).

[133] S. Villinger, *et al.*, *Proceedings of the National Academy of Sciences* **107**, 22546 (2010).

[134] J. L. Adelman, *et al.*, *Biophysical journal* **101**, 2399 (2011).

[135] D. Bhatt, D. M. Zuckerman, *Journal of chemical theory and computation* **6**, 3527 (2010).