NONEQUILIBRIUM MOLECULAR DYNAMICS CALCULATION OF THE
CONDUCTANCE OF THE KCSA POTASSIUM ION CHANNEL

A Thesis

Presented to

The Faculty of Graduate Studies

of

The University of Guelph

by

HENDRICK W. DE HAAN

In partial fulfilment of requirements

for the degree of

Doctor of Philosophy in Physics

April, 2007

©Hendrick W. de Haan, 2007

Library and
Archives Canada

Bibliothèque et
Archives Canada

Published Heritage
Branch

Direction du
Patrimoine de l'édition

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-33600-7

Our file Notre référence

ISBN: 978-0-494-33600-7

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

NONEQUILIBRIUM MOLECULAR DYNAMICS CALCULATION OF THE CONDUCTANCE OF THE KCSA POTASSIUM ION CHANNEL

Hendrick W. de Haan

Advisor:

University of Guelph, 2007

Dr. Saul Goldman and Dr. Chris Gray

Ion channels are integral membrane proteins that are involved in many biological processes. Consequently, their medical relevance as well as their ability to conduct ions at near diffusion limits while maintaining a high degree of selectivity make them an interesting and important area of research. We have developed an efficient method for the atomistic determination of the conductance of a biological ion channel model by applying an external field to the conducting ions only. The underlying theory is discussed and demonstrated in a simple test system consisting of ions in a box of water. This approach is then applied to the experimentally determined structure of the KcsA potassium channel from which a conductance in good agreement with the experimental result is predicted. Additionally, results from simulations which investigate the effect of altering the protonation state of key groups,

relaxing the constraints on the transmembrane helices, and substituting the potassium ions with sodium ions are presented.

ACKNOWLEDGEMENTS

First of all, I wish to thank Dr. Saul Goldman and Dr. Chris Gray for their continued guidance and support throughout this project. Additionally, a great deal of thanks goes to Dr. Igor Tolokh for many insightful discussions, technical help, and advice in general. I also wish to acknowledge Dr. Jim Davis and Dr. Don Sullivan for being on my committee and to thank Dr. Règis Pomès for being my external examiner. Thanks also to Dr. Bernhard Nickel for sitting on my examination committee and for providing insight through many useful discussions. Additionally, thanks to Dr. Robert Guy for important advice on setting up the channel models. Also, I would like to acknowledge the help of the excellent support staff at the physics department and in particular Reggi Vallillee for much help over the years. Likewise, thanks to the SHARCNET crew and especially David McCaughan for all his help.

For my scientific peers, I'd like to thank Stewart Hayne, Jeremy Pencer, Karl Martel, George White, Bruno Tomberli, Jeffrey Karn, Christoph Fares, Chris Schultz-Nielsen, and Scott Houliston among many others. While there are too many people to acknowledge everyone by name, I wish to thank the following people in particular: Angela Joosse for her dedicated force and integral support, Mike Degroot and Tymen Devries for being Mike & Tymen,

Chantal and Johanne Paquet, Heather Bishop, Kevin MacKenzie and Kathy Zaborsky, the LLA, and Prapaphan “Toom” Chuensomtoi. Additionally, I owe a special debt of gratitude to all my homes away from home: Al Richardson and Trish Uniac, Eva Joosse and Joel Koops, and Pauline Schaap. Finally, thanks to all my family: my parents; Rob, April, and Oliver; Dan, Erica, Jonah, and Simon; Tracey, Benoy, and Dev; Peggy, Steven, and Chloe. It’s been a long journey so thanks for all the help along the way.

TABLE OF CONTENTS

1	INTRODUCTION	1
1.1	Introduction	1
1.2	Ion Channels	2
1.2.1	Ion Channel Structure	3
1.2.2	Ion Channel Function: Nerve Signal Propagation . . .	11
1.2.3	Medical Relevance	16
1.3	Computer Simulations	17
2	MOLECULAR DYNAMICS COMPUTER SIMULATIONS	19
2.1	Introduction	19
2.2	Basic Algorithm	20
2.3	The Ergodic Hypothesis	23
2.4	Model Force Field	24
2.4.1	Non-Bonded Terms	25
2.4.2	Bonded Terms	29
2.4.3	Special Force Terms	37

2.5	Selected Topics	39
2.5.1	Time Step	40
2.5.2	Periodic Boundary Conditions	40
2.5.3	Long Ranged Interactions	42
2.5.4	Temperature Coupling	44
2.6	Conclusions	46
3	THEORY	48
3.1	Introduction	48
3.2	Derivation	49
3.3	Selective Coupling	56
3.3.1	$\vec{E}$ coupled to all charges	56
3.3.2	$\vec{E}$ coupled to conducting ions only	57
3.4	Equivalence of conductivity expressions	58
3.4.1	water-water term	59
3.4.2	Cross Terms	61
3.5	Conclusion	62
4	THE WATER BOX	64
4.1	Introduction	64
4.2	System Setup	64
4.3	Simulation Details	65
4.4	Data Analysis	66

4.4.1	Equilibrium Runs: Position versus Time	67
4.4.2	Equilibrium Runs: Velocity versus Time	71
4.4.3	Nonequilibrium Runs	79
4.5	Conclusions	87
5	KcsA	88
5.1	Introduction	88
5.2	K ⁺ Channel Protein Structure	89
5.2.1	The Selectivity Filter	91
5.2.2	The Water Pool	93
5.2.3	Charged Groups	94
5.3	System Setup	97
5.4	Simulations Details	102
5.5	System Preparation	104
5.6	Nonequilibrium Runs	106
5.6.1	Applied Voltage $V = 5$ volts	107
5.6.2	Applied Voltage $V = 4$ volts	111
5.6.3	Applied Voltage $V = 3$ volts	112
5.6.4	Applied Voltage $V = 2$ volts	115
5.6.5	Applied Voltage $V = 1$ volt	117
5.6.6	Kinetics	118
5.6.7	Current vs. Voltage	123
5.7	Conclusions	129

6	FOLLOWUP SIMULATIONS	131
6.1	Introduction	131
6.2	Applied Voltage $V = 1$ volts continued simulations	132
6.3	Reversed electric field	135
6.4	Altering the charged state of ARG64	139
6.5	Relaxed constraints of the transmembrane alpha helices	144
6.6	Selectivity: the sodium ion	149
6.7	Conclusions	155
7	CONCLUSION	157
7.1	Introduction	157
7.2	Theory	157
7.3	The Water Box	158
7.4	The KcsA Potassium Ion Channel	160
7.5	Followup Simulations	162
7.6	Future Work	164
7.7	Conclusion	167
8	APPENDIX A	169
9	APPENDIX B	174

LIST OF TABLES

4.1	Diffusion Coefficient (in units of $10^{-9}m^2/s$) for a potassium ion in a box of water molecules.	69
4.2	Diffusion Coefficient (in units of $10^{-9}m^2/s$) for a potassium ion and a chloride ion in a box of water molecules	69
4.3	Diffusion Coefficient (in units of $10^{-9}m^2/s$) of the potassium ion either alone in a box of water molecules or with a chloride ion in a box of water molecules. The nonequilibrium results are labelled according to either having the field applied to the ions only (V on Ion) or applied to all atoms in the system (V on All). The relevant equilibrium values are included as well. For these results, a cutoff radius of 0.99 nm with PME for long ranged interactions is used.	86
5.1	Number of events per 5 nanoseconds at 5 volts. This data gives a current of $2.35 \times 10^{-10} \pm 6.584 \times 10^{-11}$ Amperes.	123
5.2	Number of events per 5 nanoseconds at 4 volts. This data gives a current of $1.73 \times 10^{-10} \pm 4.796 \times 10^{-11}$ Amperes.	124

5.3	Number of events per 5 nanoseconds at 3 volts. This data gives a current of $5.127 \times 10^{-11} \pm 4.104 \times 10^{-11}$ Amperes.	125
5.4	Number of events per 5 nanoseconds at 2 volts. This data gives a current of $3.204 \times 10^{-11} \pm 3.510 \times 10^{-11}$ Amperes.	126
5.5	Number of events per 5 nanoseconds at 1 volt. This data gives a current of $1.602 \times 10^{-11} \pm 1.602 \times 10^{-11}$ Amperes.	127
6.1	Number of events per 5 nanoseconds at 1 volt. This data gives a current of $1.602 \times 10^{-11} \pm 1.602 \times 10^{-11}$ Amperes.	134

LIST OF FIGURES

1.1	The 20 amino acids out of which all proteins are built (taken from [14]).	5
1.2	The process of linking amino acids via the peptide bond to form polypeptide chains (taken from [15]).	6
1.3	The two secondary structures: the alpha helix (a) and the beta sheet (b) (taken from [16]).	8
1.4	Final configuration of the KcsA potassium ion channel (adapted from [17]).	10
1.5	Schematic outlining the process of neuron excitation upon stimulus reception, transmission of the signal down the length of the axon, and transmission of the signal to adjacent neurons via neurotransmitters (taken from [18]).	13
1.6	Generation of the action potential (taken from [19]).	15
2.1	Coulombic potential for 2 potassium ions, with $\epsilon_r = 1$	26

2.2	Lennard-Jones potential for 2 oxygen atoms using parameters from the GROMACS force-field, i.e. $\epsilon = 1.914$ kJ/mol and $\sigma = 0.258$ nm	28
2.3	Bond stretching model and potential for an oxygen atom and carbon atom using parameters from the GROMACS force-field. Here $b_{OO} = 0.123$ nm and $k_{OO} = 502080$ kJ/mol nm ² . .	30
2.4	Rotation angle model and potential for a water molecule using the parameters contained in the GROMACS force-field, with $\theta_O = 109.47$ degrees and $k_{HOH} = 383$ kJ/mol degrees ²	32
2.5	Models of improper dihedral types and resulting potential energy profile when atom 1 is a carbon and atom 4 is an oxygen such that $\xi_o = 0$ and $k_\xi = 0.051$ kJ/mol degrees ²	34
2.6	Model (in the <i>trans</i> form) and resulting potential energy profile for the proper dihedral when atom 1 is a carbon and atom 4 is a nitrogen such that $\phi_0 = 180$, n=2, and $k_\phi = 33.473$ kJ/mol.	36
2.7	Schematic demonstrating the use of periodic boundary conditions in two dimensions. The dashed circle represents the maximum sphere of direct interaction.	41
4.1	Mean square displacement of a potassium ion in a box of water versus time.	68

4.2	Autocorrelation function of the potassium ion (in units of $\frac{nm^2}{ps^2e^2}$ where e is the charge on an electron) for a single potassium ion in a box of water molecules.	72
4.3	Cumulative integration of the autocorrelation function of the potassium ion (in units of $\frac{nm^2}{ps^2e^2}$ where e is the charge on an electron) for a single potassium ion in a box of water molecules.	73
4.4	Autocorrelation function of the electric current (in units of $\frac{nm^2}{ps^2e^2}$ where e is the charge on an electron) averaged over all the water molecules for a single potassium ion in a box of water molecules.	74
4.5	Means square displacement of the net dipole moment of the water (in units of $\frac{nm^2}{e^2}$ where e is the charge on an electron). .	75
4.6	Means square displacement of the net dipole moment of the water divided by time and the cumulative integration of the water current autocorrelation function versus time (in units of $\frac{nm^2}{ps^2e^2}$ where e is the charge on an electron).	76
4.7	The cross correlations function of the electric current (in units of $\frac{nm^2}{ps^2e^2}$ where e is the charge on an electron).	77
4.8	The cumulative integration of the cross correlations function of the electric current (in units of $\frac{nm^2}{ps^2e^2}$ where e is the charge on an electron).	78

4.9	Net position (after removal of PBC) versus time for the potassium ion in a box of water along with a chloride ion. The different trajectories are from various voltages applied across the box to the ions only. The voltages range from 0.5 volts to 7 volts in half a volt steps.	80
4.10	Net position (after removal of PBC) versus time for the potassium ion in a box of water along with a chloride ion. The different trajectories are from various voltages applied across the box to all atoms in the system. The voltages range from 0.5 volts to 7 volts in half a volt steps.	81
4.11	Potassium ion current vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) for a potassium ion in a box of water molecules.	82
4.12	Potassium ion current vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) with both a potassium and a chloride ion in water.	83
4.13	Potassium ion current/voltage vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) for a potassium ion in a box of water molecules.	84

4.14	Potassium ion current/voltage vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) for a potassium ion and a chloride ion in a box of water molecules.	85
5.1	Top view of the KcsA ion channel. The four subunits are indicated by the different colours, with the transmembrane helices as darker ribbons and the other regions as lighter tubes. Note that the helices are at a significant angle and the subunits appear to unfold as the aperture of a camera. The selectivity filter and pore axis are indicated by the dotted orange box. . .	90
5.2	Cross section view displaying the major regions of the KcsA potassium ion channel. The transmembrane helices are represented by purple ribbons. The selectivity filter is indicated by the dotted orange box, with red dots denoting carbonyl oxygens. The orange arrows indicate the dipole generated by the pore helices. Note that these dipoles point to the center of the protein in a region labelled the water cavity. A crude schematic of the lipid bilayer is included to help orient the reader.	91
5.3	The highly conserved TVGYG sequence which constitutes the selectivity filter of KcsA. The binding sites are labelled S0-S4.	92

5.4	Protonation states of charged amino acids for the KcsA model.	
	Note that since the primary sequences of all the subunits are identical, each subunit will contain all the charged groups shown; they are broken up over two subunits simply for clarity. The numbers in black indicate the distance (in Å) from the pore axis while the numbers in green indicate the distance between the atoms containing net charges on pairs of charged groups.	95
5.5	Cross sectional view of the system setup for KcsA showing the placement of the potassium ions (green spheres) and the waters located between subunits (red and gray spheres). The carbonyl oxygens of the selectivity filter are shown to indicate the location of the binding sites.	99
5.6	Top view of the system setup for KcsA showing the placement of the potassium ions (green spheres) and the waters located between subunits (red and gray spheres).	100
5.7	Full system setup of the KcsA protein, potassium ions, and column of water (shown as red and grey balls and sticks). . . .	102
5.8	The final configuration of the system following the equilibration procedure. The different groups are indicated by the coloured regions.	105

5.9 Ion axial trajectories versus time for 5 volts applied across the system.	108
5.10 Close up of the selectivity filter showing the location of the blocking water molecule	110
5.11 Ion axial trajectories versus time for 4 volts applied across the system.	112
5.12 Ion axial trajectories versus time for 3 volts applied across the system. The data from $t=0$ to $t=50$ ns is displayed in the five panels	114
5.13 Ion axial trajectories versus time for 2 volts applied across the system.	116
5.14 Ion axial trajectories versus time for 1 volt applied across the system.	117
5.15 Full table of the observed sequence for all conductance events at 3 volts.	119
5.16 Magnified portion of the trajectory generated at 3 volts. The inset displays the simplified kinetics scheme which is followed by both events shown in the trajectory.	122
5.17 Current vs. Voltage for the nonequilibrium simulations on KcsA128	
6.1 Ion axial trajectories versus time for 1 volts applied across the system.	133

6.2	Full Current vs. Voltage for the nonequilibrium simulations on KcsA	135
6.3	Ion axial trajectories versus time for 3 volts applied across the system in a direction opposite from the results previously presented.	137
6.4	System setup with the external field driving the ions from the intracellular side to the extracellular side.	138
6.5	System setup with a charge on arginine 64. The ARG64 residue is shown for two subunits and all four Chloride counter- ions are shown in light blue. Also shown are all atoms of the water molecules which are within a 4 Å by 4 Å square centered on the conduction pathway.	141
6.6	Trajectories generated from the system setup with a net charge on arginine 64.	142
6.7	Snapshot of the system setup with a net charge on arginine 64.	143
6.8	Trajectories produced from a system setup with the trans- membrane helices free to move around their initial position in a harmonic well with a force constant of $50 \frac{kJ}{nm^2 mol e}$. Three volts are applied across the system.	145
6.9	System configuration for the relaxed transmembrane helices with ions in S0,S2,S4 and water molecules in S1 and S3. . . .	146

6.10	Snapshots outlining the events observed with relaxed trans- membrane helix constraints.	148
6.11	Trajectories generated from the system setup with a replace- ment of potassium ions to sodium ions.	150
6.12	System configuration with a water molecule in S1. Sodium ions are shown in purple.	151
6.13	Snapshots outlining the first two events observed for the sodium ion conductance.	153
7.1	Potassium ion current vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) with both a potassium and a chloride ion in water.	159
7.2	Current vs. Voltage for the nonequilibrium simulations on KcsA. The upper axis displays the voltage as applied across the length of the system while the lower voltage is the cor- rected voltage as determined by the voltage drop across the selectivity filter.	161
8.1	Graphical representation of the domain of integration as de- fined in equations (8.9) and (8.10)	171
9.1	Modified GROMACS position restraint code.	175

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

There are many biological processes that are initiated or regulated by the transfer of ions across the cell membrane. As will be discussed in the following section, the classic example regarding the use of ions as a means of biological signalling is that of nerve signal propagation. Additionally, the transfer of ions across cell membranes is known to be the primary mechanism for controlling the contraction of muscle, regulation of the heartbeat, modulation of cell volume, and a host of other fundamental biological processes [1]. As a consequence, malfunctioning ion channels are implicated in a number of diseases [2]. However, the lipid bilayer which comprises the cell membrane is relatively impermeable to ions and does not allow for the rapid trans-location of ions which is required to mediate these processes [3,4]. Rather, highly specialized proteins called ion channels are embedded within the cell membrane. These proteins allow for the transport of ions across the cell membrane by either forming a hole (or “pore”) through which the ions pass easily (passive transport) or by actively using energy to pump the ions

against a concentration gradient (active transport).

Among the fascinating properties of ion channels is the ability to conduct ions at near diffusion limits (on the order of 10^7 ions per second per channel [5]). Yet, in seeming contradiction, the channels are highly selective to conduct only the ions for which they are “tuned”. For example, the ratio of potassium to sodium conductance is on the order of 1000 to 1 for potassium ion channels even though a potassium ion is *larger* than a sodium ion [6-8]. Hence, the combination of their fundamental biological and medical relevance and the remarkable characteristics of their functioning make ion channels an interesting and important area of research.

1.2 ION CHANNELS

Ion channels are found in the cells of a wide variety of systems including viruses, plant cells, and animals. Consequently, the number and diversity of known ion channels is staggering. However, there are fundamental characteristics shared among groups of ion channels which greatly ease the task of organizing ion channels into separate families. The first such criteria is the propensity to selectively conduct one type of ion over another. This leads to a division among ion channels into superfamilies which conduct sodium, potassium, calcium, or chloride among others. In addition to ion selectivity, further classification can be performed by examining the behaviour of the ion channels. For example, direct observation of the current produced by ions

passing through ion channels indicates that channels tend to open and close. Experiments have demonstrated that this “gating” can be controlled by various environmental factors. Hence, channels which demonstrate gating have been classified as to whether the gating is controlled by the trans-membrane voltage, the pH of the environment, the binding of a certain ligand to the channel, or by the mechanical stressing of the lipid membrane [9,10]. Not surprisingly, ion channels which demonstrate a particular behaviour are found to have similar structures even when occurring in quite different organisms. For example, all potassium ion channels are found to contain a selectivity filter, pore helix, and 2 trans-membrane helices per subunit regardless of the other details of function. Consequently, even though a voltage-gated potassium ion channel will contain an additional 4 trans-membrane helices per subunit, the conduction pathway can be compared to that of the known KcsA structure [11,12], a bacterial ion channel. Hence, the lessons learned from studying one particular ion channel can be applied to a host of other ion channels.

1.2.1 Ion Channel Structure

As for all proteins, ion channels are built from amino acids (shown in Fig. 1.1) linked via peptide bonds to form a polypeptide chain [13]. As shown in Fig. 1.2, the amino acids are joined together by stripping off the hydroxyl (OH) group of the carboxyl group of one amino acid and the hydrogen (H) of

the amino group from an adjacent amino acid producing a water molecule. This allows the backbone nitrogen N of the amino group to covalently bond to the carboxyl carbon to form the peptide bond. Through this process, many amino acids can be linked together to form polypeptide chains. This ordered list of amino acids which comprises the protein is termed the primary sequence.

Amino acids with hydrophobic side groups

Amino acids with hydrophilic side groups

Amino acids that are in between

Figure 1.1: The 20 amino acids out of which all proteins are built (taken from [14]).

FORMATION OF A PEPTIDE BOND

Figure 1.2: The process of linking amino acids via the peptide bond to form polypeptide chains (taken from [15]).

For each primary sequence, the atoms of the amino acids, through their electrostatic and Van der Waals interactions, lead to the formation of local structures known as secondary structures. The two basic forms of secondary structure are the alpha helix and the beta sheet shown in Fig. 1.3. These

structures represent stable configurations of the amino acids and proteins generally contain several alpha helices or beta sheets or even a mix of the two. However, not all amino acids in the primary sequence form secondary structures and there is often a large percentage which form more flexible sections linking the secondary structures together.

Figure 1.3: The two secondary structures: the alpha helix (a) and the beta sheet (b) (taken from [16]).

The final configuration of the protein is then determined by the secondary structures being packed together and associating both with each other and the local environment (such as the lipid bilayer in the case of membrane

embedded proteins). Some proteins are formed from the association of a number of polypeptide chains which may or may not be identical in terms of primary and secondary structure. For example, all potassium ion channels are formed from four separate polypeptide chains (named subunits) which are identical. In particular, each subunit of the KcsA potassium ion channel contains three alpha helices and connecting regions. A schematic of the resolved structure of the KcsA ion channel which shows how the protein is formed by packing these elements together is shown in Fig. 1.4 (the structure of KcsA will be discussed in much more detail in a late chapter).

Figure 1.4: Final configuration of the KcsA potassium ion channel (adapted from [17]).

The structure shown in Fig. 1.4 represents a significant advancement in the understanding of ion channel structure and function. In 1998, the research group under Roderick MacKinnon managed to resolve the structure of the KcsA potassium ion channel via x-ray crystallography [20]. The barrier which needed to be overcome to achieve this breakthrough is that the technique requires a well ordered crystal of the membrane proteins. For membrane bound proteins, this is a difficult task as the proteins must be removed from the membrane and packed into a crystal while keeping the protein struc-

ture intact. Consequently, detergent molecules are used to compensate for the absence of the membrane and allow crystallization. Using these methods, Doyle et al. were able to generate a crystal of the KcsA potassium channel and subsequently resolve the structure to atomic scale (a structure resolved to 2 Å was presented by the same group in 2001 [21, 22]). This result, for which MacKinnon shared the Nobel Prize in chemistry in 2003, represented a significant advancement by essentially providing the scientific community with the first “picture” of ion channels. Subsequent work has resolved the structure of the KirBac1.1 inward-rectifier potassium channel [23], voltage dependent K^+ channels [24, 25], a NaK channel [26], mechanosensitive ion channels [27, 28], the CIC chloride channel (now known as the H⁺/Cl⁻ exchange transporter) [29, 30], gramicidin [31], as well as other ion channels and integral membrane proteins. For our work, the determination of the structure of KcsA represents a starting point as from this structure we can construct a system to study the details of ion channel function.

1.2.2 Ion Channel Function: Nerve Signal Propagation

The generation of the neural pulse (or “action potential”) serves as an excellent illustration of the biological relevance of ion channels. Not only is it an obviously integral physiological process, but the study of the neural pulse is historically significant. In the early 1950’s, Alan Hodgkin and Andrew Huxley studied the action potential of the giant squid axon by the recently

developed voltage clamp technique (it is the axon, not the squid, which is giant as it can be up to 1mm in diameter and extends from the head of the squid to the tail such that it controls the movement of the tail [32]). This technique allowed Hodgkin and Huxley to precisely characterize the action potential and lead them to postulate the existence of ions channels [33–37]. For this work, they shared the Noble Prize in physiology or medicine in 1963. Hence, a brief overview of the generation of the action potential is presented as an example of the biological relevance of ion channels.

In neural cell triggering, the signal received at the dendrites of the neuron must be propagated down the length of the axon at whose terminus it will trigger the release of neurotransmitters. These neurotransmitters will then “excite” adjacent neurons such that the process is repeated and the nerve signal can be transmitted from the brain to the relevant muscle tissue realizing the signal. This process is displayed schematically in Fig. 1.5 below.

Figure 1.5: Schematic outlining the process of neuron excitation upon stimulus reception, transmission of the signal down the length of the axon, and transmission of the signal to adjacent neurons via neurotransmitters (taken from [18]).

In this discussion, the details of the propagation of the signal down the length of the axon are examined. This signal is named the action potential and, as will be seen, it is formed by the coordinated behaviour of sodium and potassium ion channels. Typically, the membrane of the cell is held at a resting potential of -70 millivolts (mV), with the inside negative, which is maintained by a charge imbalance across the cell membrane. At rest, potassium ion concentrations are higher within the cell than outside of the cell while sodium and chloride ion concentrations are higher outside of the cell.

Upon excitation, sodium ion channels will begin to open. This allows an inward flux of sodium ions according to the concentration gradient. This

area of the membrane begins to become depolarized (the potential increases from negative towards positive values) as the positive charges flow into it. If the electric potential reaches a certain potential, a large number of sodium ion channels will open causing rapid further depolarization. This potential is termed the threshold potential and is generally found around -55 mV. The depolarization continues until the potential is in excess of approximately 30 mV. At this point, the sodium channels will close and the potassium channels, which lag behind the sodium channels, will open up. This allows a flux of potassium out of the cell which begins to reverse the depolarization (positive charges flow out and the potential decreases). Consequently, the potential begins to return towards its negative resting value. In fact, the potential overshoots its original value and becomes hyper-polarized at a value below -70 mV. At this point, active transport of the ions across the membrane to restore the original concentration gradients of the ions begins and this area of the membrane slowly relaxes to its resting potential at -70 mV from which point it can be excited again. This process is outlined in Fig. 1.6.

Figure 1.6: Generation of the action potential (taken from [19]).

This sequence of events forms the action potential, which, as is shown in Fig. 1.6 has the shape of a pulse. Since the depolarization of one area of the membrane will trigger adjacent areas of the membrane to begin to depolarize, the action potential will be formed at subsequent locations and the signal will travel down the axon. Hence, the signal represented by the pulse of the action potential will be transmitted down the length of the axon by a cascading sequence of depolarization / re-polarization events as mediated by the behaviour of the ion channels.

Although the description of the propagation of nerve signals as outlined above is relatively brief and neglects many important details which are discussed elsewhere [38,39], it serves to illustrate how the transfer of ions across cell membranes as mediated by ion channels can result in electrochemical

signalling which can carry messages through nerve and muscle tissue. While this process is relatively well understood, there are still many fundamental questions regarding the precise structure and function of the ion channels themselves. In this document, research is presented which serves to elucidate the details of ion conduction through the KcsA potassium ion channel. While this is just one particular channel, we will see that there are many shared characteristics among ion channels and, especially, among the family of potassium ion channels such that insight into the functioning of one channel can shed light upon the nature of ion channel function in general.

1.2.3 Medical Relevance

As illustrated by the description of the generation of the action potential given above, ion channels play an integral role in biological processes. Hence, it is not surprising to learn that there are also a range of diseases associated with ion channels occurring in a malfunctioning state. For example, a set of specific mutations of the sodium channel occurring in muscle tissue are observed to alter the nature of the neural pulse. In an organism expressing this mutated channel, the currents through sodium channels are long lasting and hence the depolarization portion of the action potential is extended. Such mutations result in two muscular disorders known as periodic paralysis and myotonia. The primary symptom of these conditions is muscle weakness or even paralysis often following periods of intense activity [40]. This is but

one example of a “channelopathy”: a disease whose fundamental cause is a malfunctioning ion channel. Other examples include cystic fibrosis, congenital hearing loss, and epilepsy [2, 40]. Clearly, ion channels present not only a fascinating puzzle into the basis of biological functions, but also an important and promising field of medical research.

1.3 COMPUTER SIMULATIONS

While electro-physiology experiments have characterized the functioning of ion channels and techniques such as x-ray crystallography and NMR have yielded the structures of certain ion channels, a great deal of uncertainty concerning the precise details of ion channel function and structure remains. However, ion channels are difficult systems to study as the dynamics of ion transport occur on the atomic scale and the picosecond timescale. Hence, while experiments in the laboratory can provide us with either the static structure or the dynamics of isolated portions of the protein, a complete picture directly solving structure and function concurrently is impossible at the moment. For this reason, computer simulations are an important avenue of research.

In a computer simulation, a researcher starts by building a representative system which likely consists of a protein structure (as given by x-ray crystallography where possible), a lipid bilayer, water, and ions. From this initial structure, the system is evolved in time according to a set of rules derived

from the known behaviour of atoms and molecules. Using this methodology, the details of ion channel function can be measured directly.

In this document, the research conducted by our group into developing a method of efficiently calculating the conductance of ion channels and the application of this methodology to the KcsA potassium ion channel is presented. In attempting to elucidate the details of ion channel function, an accurate calculation of the conductance is a reasonable first criteria as not only is this the primary function of ion channels, but it is a behaviour which has been well characterized in the laboratory. Our goal, in striving to simulate this function, is to produce trajectories which not only yield a reasonable conductance, but whose details reveal the precise nature of the process of ion conductance. For example, how important is the interaction between the ions as they pass through the channel? Are the ions associated with water molecules or do they become completely dehydrated as they enter the channel? What mechanism is responsible for selecting potassium over sodium to such a large degree? While difficult to address in the laboratory, the precise control over the the system and abundance of information produced by computer simulations make them a perfectly suited approach to investigate these and a whole host of other questions.

CHAPTER 2: MOLECULAR DYNAMICS

COMPUTER SIMULATIONS

2.1 INTRODUCTION

Molecular dynamics (MD) simulations attempt to directly simulate a system of N -interacting particles by numerically integrating the equations of motion. In particular, MD simulations are highly suitable for solving systems which describe biological systems on the molecular level. As such, systems appropriate for study via MD generally consist of protein, membrane, and water. The topology of these systems is described according to groups of atoms forming amino acid residues and molecules. The MD package can then read this input information and, given the coordinates of each atom and the relationship between the atoms, construct a molecular model of the system.

Once we have a model of the system, we can begin to “simulate” the system by calculating its time evolution through numerically integrating the equations of motion for each atom. To achieve this, a force-field describ-

ing the interactions is used to calculate the total force on each atom. While some advanced simulations of small systems incorporate quantum mechanical effects and methods [41], these considerations are generally far too computationally expensive to be used on systems containing a full protein. Hence, the force-field is primarily empirical in nature and the simulations evolve according to a classically deterministic description. In addition to a brief overview of the numerical integration procedure, the details of the various terms used in constructing the force field are discussed below. Finally, some details regarding specific features of MD simulations which are particular important for our work are highlighted.

2.2 BASIC ALGORITHM

As stated above, typically MD simulations are purely classical and the core of the simulation lies in numerically solving Newton’s equations of motion given by:

$$\vec{F}_i = m \ddot{\vec{r}}_i \quad (2.1)$$

$$= m \frac{d^2}{dt^2} \vec{r}_i \quad (2.2)$$

where $\vec{F}_i$ is the total force on the i ’th atom and $\vec{r}_i$ is the position vector of the i ’th atom. In order to numerically integrate (2.2), a set of initial conditions is required for each velocity $\vec{v}_i$ and position $\vec{r}_i$. The initial positions are

given by the initial configuration of the system. For example, the structure of a protein solved by x-ray crystallography could be used as the initial configuration. For the velocities, often an initial temperature distribution is assigned to the system according to a Maxwellian distribution. This allows the mean temperature to be set and generates a realistic variation of velocities among the particles. The particular nature of the temperature distribution is controlled by a “seed” number which allows different initial conditions according to different distributions with the same mean temperature to be generated. Obviously, this has uses in generating a number of ensembles to run in parallel in order to calculate an average quantity.

Once the initial conditions are given, the total force $\vec{F}_i$ for each particle must be calculated. As will be discussed in detail later, the force acting on each particle generally has three sources: i) non-bonded interactions, ii) bonded interactions, iii) restraint and external forces. Some of the force terms are directly calculated while others are derived from the interaction energy by:

$$\vec{F}_i = -\frac{\partial V}{\partial \vec{r}_i} \quad (2.3)$$

where V is the potential energy. Each contribution is calculated and the results are summed to yield the total force. The equation of motion given by (2.2) is then solved in two steps by:

$$\frac{d\vec{r}_i}{dt} = \vec{v}_i \quad (2.4)$$

and

$$\frac{d\vec{v}_i}{dt} = \frac{\vec{F}_i}{m_i} \quad (2.5)$$

In particular, GROMACS uses the leap-frog algorithm [43]. In this algorithm, the velocities and positions are calculated at staggered time steps such that:

$$\vec{v}(t + \frac{\Delta t}{2}) = \vec{v}(t - \frac{\Delta t}{2}) + \frac{\vec{F}(t)}{m} \Delta t \quad (2.6)$$

and

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \vec{v}(t + \frac{\Delta t}{2}) \Delta t \quad (2.7)$$

where Δt is the time step, and $\vec{r}(t)$ and $\vec{v}(t - \frac{\Delta t}{2})$ are the position and velocity initial conditions respectively. Using this information, the position-velocity configuration of the system at time $t + \Delta t$ is calculated. The process is then repeated using this generated “frame” as the new starting point. This cyclic process, which continues until the specified number of time steps is reached, constitutes the heart of the MD algorithm. In the following sections, some of the important details of performing MD simulations are discussed.

2.3 THE ERGODIC HYPOTHESIS

Examining the details of the numerical integration algorithm discussed above, it is obvious that there is error inherent in the process. With a finite time step and a limited precision in the specification of the initial conditions, the simulated trajectory will deviate from the “true” trajectory very quickly (It is important to note that “true” trajectory here does not mean the one that would be generated in the lab, but rather the trajectory given from an exact solution of (2.2) for a system of interacting particles. This is impossible except in very special circumstances, but this theoretical solution is here referred to as the “true” trajectory). The finite time step introduces error by breaking the continuous behaviour into discrete jumps. On the other hand, the inability to precisely describe the initial coordinates introduces error as it is well known that small variations in initial conditions yields rapidly and dramatically divergent trajectories [44]. Consequently, rather than the generated trajectory being an evolution of the system in time, it is more a series of relevant “snapshots” constituting a statistically significant ensemble of states. However, this failure to replicate the “true” evolution of the system is not an issue in most simulations due to the ergodic hypothesis. The ergodic hypothesis states that given enough time, a system will sample all of phase-space such that a thermodynamic quantity calculated from averaging over a single sufficiently long trajectory in time is equivalent to averaging over a sufficiently large ensemble of configurations [44]. This allows simulators to

treat the data generated from simulations as trajectories in time and, using equivalent statistical techniques, analyze the data to generate meaningful results. Hence, while this discussion is an important insight into the relevance of simulations, the data produced from MD simulations will be considered time trajectories for simplicity.

2.4 MODEL FORCE FIELD

As shown by (2.2), the details of the time evolution of the system will be governed by the particulars of the force field from which the total force $\vec{F}_i$ acting on each particle is calculated. The nature of the force field is governed by two separate properties: the type and form of the terms included and the numerical parameters describing the interaction given by each term. For example, a chemical bond is modelled by a spring joining two masses and as such, is included in the force calculation as a harmonic potential. The details of the harmonic well are determined by the force constant of the spring and the equilibrium separation between the masses. Hence, the construction of a force field requires both a model for each relevant interaction as well as a full set of parameters quantifying those interactions. As stated, the simulations are purely classical and consequently, the form of the terms comprising the force field are generally empirical. The models and resulting form for the terms included in GROMACS are discussed below in three categories: non-bonded interactions, bonded interactions, and additional “special” terms.

In the following, the notation and nomenclature adopted by the GROMACS simulation package is used for the sake of transparency. For a more elaborate discussion including other options available in the GROMACS force-field, we refer to the GROMACS manual [42]

2.4.1 Non-Bonded Terms

The non-bonded interaction terms in MD simulations are comprised of two terms: the electrostatic Coulombic interaction and the short ranged Lennard-Jones term.

Coulombic Term

The energy of the Coulombic interaction is given by

$$V_c(r_{ij}) = f \frac{q_i q_j}{\epsilon_r r_{ij}} \quad (2.8)$$

where f is the physical constant $\frac{1}{4\pi\epsilon_0}$ which evaluates to $138.935\,485 \frac{\text{kJnm}}{\text{mol e}}$, where e is magnitude of the charge on an electron, in the particular system of units that GROMACS uses [42], q_i and q_j is the charge on atom i and atom j respectively, and r_{ij} is the center to center distance between atoms i and j . Additionally, ϵ_r is the relative dielectric constant which may be specified in the program input. This potential is shown in Fig. 2.1. Using (2.3) we get:

$$\vec{F}_c(\vec{r}_{ij}) = f \frac{q_i q_j}{\epsilon_r r_{ij}^2} \frac{\vec{r}_{ij}}{r_{ij}} \quad (2.9)$$

Figure 2.1: Coulombic potential for 2 potassium ions, with $\epsilon_r = 1$.

This gives the force on an atom due to the Coulombic interaction with another atom. To get the net Coulombic force on a particular atom, the interaction between that atom and all the other atoms are calculated independently and summed. This methodology is termed the pairwise addition and is used for all applicable terms in the force-field.

Lennard-Jones Term

The next non-bonded term is the Lennard-Jones interaction. The Lennard-Jones term is the $C^{(6)} - C^{(12)}$ term which combines the short-range repulsion of atoms being pushed into one another and the mid-range attraction which is due to weak but favourable interactions primarily from mutual electronic

polarization. The precise form of this potential is given by

$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6} \quad (2.10)$$

where $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ are the parameters which give the precise shape of the potential. The force derived from this is

$$\vec{F}_i(\vec{r}_{ij}) = \left(12 \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - 6 \frac{C_{ij}^{(6)}}{r_{ij}^6} \right) \frac{\vec{r}_{ij}}{r_{ij}} \quad (2.11)$$

The energy term given in (2.10) can be rewritten

$$V_{LJ}(r_{ij}) = 4\epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right) \quad (2.12)$$

In this form, the constants have a more physical meaning as ϵ_{ij} is the well depth and σ_{ij} is the effective diameter of the atom as defined by the closest distance at which $V(r)=0$ (that is, any decrease in distance results in a very sharply rising repulsive energy). These parameters are labelled in Fig. 2.2 which displays the Lennard-Jones potential for two oxygen atoms.

Figure 2.2: Lennard-Jones potential for 2 oxygen atoms using parameters from the GROMACS force-field, i.e. $\epsilon = 1.914$ kJ/mol and $\sigma = 0.258$ nm

The constants, whether in the form $C_{ij}^{(6)}$ and $C_{ij}^{(12)}$ or ϵ_{ij} and σ_{ij} , are specific to the pair interaction. Typically, a force field will contain a limited number of specific constants for important interactions. For the rest, parameters are generated from the combination rules which derive the pair interactions from the self interaction constants:

$$C_{ij}^{(6)} = \left(C_{ii}^{(6)} \cdot C_{jj}^{(6)} \right)^{\frac{1}{2}} \quad (2.13)$$

$$C_{ij}^{(12)} = \left(C_{ii}^{(12)} \cdot C_{jj}^{(12)} \right)^{\frac{1}{2}} \quad (2.14)$$

or

$$\sigma_{ij} = \frac{1}{2} (\sigma_{ii} + \sigma_{jj}) \quad (2.15)$$

$$\epsilon_{ij} = (\epsilon_{ii}\epsilon_{jj})^{\frac{1}{2}} \quad (2.16)$$

Hence, in addition to constants for specific interactions, the force-field must contain a pair of constants (either $C_{ii}^{(6)}$ and $C_{ii}^{(12)}$ or σ_{ii} and ϵ_{ii}) for each atom type. In general, $C_{ii}^{(6)}$ and $C_{ii}^{(12)}$ cannot be derived from σ_{ii} and ϵ_{ii} and the force-field will conform to one standard. In GROMACS, the $C_{ii}^{(6)}$ and $C_{ii}^{(12)}$ notation is used and the force-field contains $C_{ii}^{(6)}$ and $C_{ii}^{(12)}$ parameters.

2.4.2 Bonded Terms

The forces comprising the bonded terms of the force field approximate the energy profile of bond stretching and rotation between atoms and groups of atoms which arise from the association of atoms into molecules.

Bond Stretching

The association between two covalently bonded atoms is modelled by two masses connected by a spring. This results in a harmonic potential given by:

$$V(r_{ij}) = \frac{1}{2} k_{ij} (r_{ij} - b_{ij})^2 \quad (2.17)$$

where k_{ij} is the force constant and b_{ij} is the equilibrium separation. A sketch of this model and a plot of the resulting potential is shown in Fig. 2.3.

Figure 2.3: Bond stretching model and potential for an oxygen atom and carbon atom using parameters from the GROMACS force-field. Here $b_{OO} = 0.123$ nm and $k_{OO} = 502080$ kJ/mol nm²

The following force is derived from (2.17):

$$\vec{F}(\vec{r}_{ij}) = k_{ij}(r_{ij} - b_{ij})\frac{\vec{r}_{ij}}{r_{ij}} \quad (2.18)$$

Hence, the force-field must contain a spring constant k_{ij} and equilibrium

separation b_{ij} for each possible covalent bond between atoms.

Rotation Angle

Analogous to the bond stretching potential, the potential of the rotation angle formed by 3 atoms is modelled by a harmonic well given by

$$V(\theta_{ijk}) = \frac{1}{2}k_{ijk}(\theta_{ijk} - \theta_{ijk}^0)^2 \quad (2.19)$$

Here, k_{ijk} is the force constant and θ_{ijk}^0 is the equilibrium angle. Note that the force-field must contain a set of each of these parameters for all relevant ijk combinations. The model and the resulting potential profile are shown in Fig. 2.4.

Figure 2.4: Rotation angle model and potential for a water molecule using the parameters contained in the GROMACS force-field, with $\theta_O = 109.47$ degrees and $k_{HOH} = 383$ kJ/mol degrees².

The forces on each atom as derived from (2.19) are:

$$\vec{F}_i = -\frac{dV(\theta_{ijk})}{d\vec{r}_i} \quad (2.20)$$

$$\vec{F}_k = -\frac{dV(\theta_{ijk})}{d\vec{r}_k} \quad (2.21)$$

$$\vec{F}_j = -\vec{F}_i - \vec{F}_k \quad (2.22)$$

where

$$\theta_{ijk} = \arccos \frac{(\vec{r}_{ij} \cdot \vec{r}_{kj})}{r_{ij}r_{kj}} \quad (2.23)$$

Improper Dihedrals

Improper dihedrals extend the potential beyond comparing the position of atoms to comparing the orientation of planes of atoms in an effort to model these extended structures. For example, improper dihedrals will serve to maintain the planar nature of a benzene ring. Again, this is modelled using a harmonic potential:

$$V(\xi_{ijkl}) = k_\xi(\xi_{ijkl} - \xi_0)^2 \quad (2.24)$$

where ξ_{ijkl} is the angle between the plane formed by atoms ijk and the plane formed by atoms jkl . The models for a number of improper dihedral situations (taken from [42]) and the potential for them are shown in Fig. 2.5.

Figure 2.5: Models of improper dihedral types and resulting potential energy profile when atom 1 is a carbon and atom 4 is an oxygen such that $\xi_o = 0$ and $k_\xi = 0.051 \text{ kJ/mol degrees}^2$.

Proper Dihedrals

The proper dihedrals are implemented into the force-field to model the preference of a molecule to assume a *cis* or a *trans* configuration. The potential has the following form:

$$V(\phi_{ijkl}) = k_\phi(1 + \cos(n\phi - \phi_s)) \quad (2.25)$$

where ϕ is the angle between the ijk and jkl planes and $\theta = 0$ corresponds to the *cis* configuration. The model (taken from [42]) and potential are shown in Fig. 2.6.

Figure 2.6: Model (in the *trans* form) and resulting potential energy profile for the proper dihedral when atom 1 is a carbon and atom 4 is a nitrogen such that $\phi_0 = 180$, $n=2$, and $k_\phi = 33.473$ kJ/mol.

2.4.3 *Special Force Terms*

In addition to the terms arising from the bonded and non-bonded interactions, there are extra terms which might be added to the net force calculation depending on the nature of the simulation. In the following, two examples relevant to our work are discussed: i) a force arises from the application of an external electric field ii) force arises from the implementation of position restraints.

External Electric Field

GROMACS allows an external electric field $\vec{E}$ to be applied to the system as a whole. In this case, an additional energy is added to the system given by:

$$V = \sum_i q_i \vec{E} \cdot \vec{r}_i \quad (2.26)$$

and the following term is added to the force field for each atom:

$$\vec{F}_i = q_i \vec{E} \quad (2.27)$$

where q_i is the charge on the i 'th atom. In addition to applying a force to the entire system, GROMACS allows us to apply an external force directly to a subset of atoms. Using this option, it is actually the “applied acceleration” which is specified in the input file such that a single number can be used for a group of atoms. For example, one could apply a gravitational field to all the atoms of a system or selected part of the system simply by specifying

the acceleration of the field. The respective forces are then calculated within GROMACS by multiplying the given acceleration by the individual atomic masses. The resulting term is added to the net force for all the atoms specified to be affected by the field.

Position Restraints

As will be discussed in further detail in the chapters concerning the ion channel simulations, it is sometimes desirable to constrain parts of a system to either their initial configurations or to stay within a specified region. For example, using such constraints, one can allow a protein to fluctuate around its initial position such that there is thermal motion without losing the overall configuration due to a solvent which hasn't been fully energy minimized or equilibrated.

In our simulations, position restraints will be used in the ion channel simulations for two purposes. The first is to constrain the water in our system to remain within a cylinder of 15 Å. As will be discussed later in further detail, this option was not actually available in the native GROMACS package and was added manually. In the second instance, some simulations were performed where significant portions of the protein were restrained to their initial position yet free to move (that is, each atom is allowed to oscillate around its initial position). These simulations were done to compare the results to simulations where these parts of the protein were fixed (or “frozen”)

in place. Here, the position restraints are used to allow thermal fluctuations while preventing the channel from disintegrating due to the lack of a lipid bilayer.

In either case, the restraining force is modelled by a harmonic well where the force constant is supplied by the user in the input files. In the case of restraining the water, the force term is added only if the water molecule strays outside of the specified cylinder. If this happens, the distance to the nearest point on the cylinder is calculated and the force is applied back along this vector. For the case of restraining the protein, the distance of each atom from its initial position is used as the separation vector with the equilibrium separation for the harmonic well being set to zero.

2.5 SELECTED TOPICS

The field of molecular dynamics is a large and ever-growing one. Consequently, it is beyond the scope of this document to discuss in even cursory detail the various options and methods which are available to a simulator using a modern MD package. However, there are some details which are particularly pertinent to our research. These are discussed in the sections below.

2.5.1 Time Step

In order to simulate a system on significant timescales in a reasonable amount of computation time, we would like to use as large a time step as possible. However, if the time step is too large, some of the relevant molecular motions may not be properly sampled. Hence, we are limited to a time step which is smaller than the fastest motion. In our systems (and in general), this is the bond stretching vibrations of covalently bonded hydrogen atoms. To appropriately account for this motion, a timescale not greater than 1 femtosecond should be used. However, as this is such a fast motion, in most cases we can freeze these bonds at the mean separation without a significant effect on the results. In GROMACS, this constraint is implemented using the LINCS algorithm [45]. With these bond lengths fixed, we can safely simulate with a time step of 2 femtoseconds. With the exception of some water box simulations, 2 femtoseconds is the time step used in our simulations.

2.5.2 Periodic Boundary Conditions

To achieve maximum computer efficiency, it is an essential consideration when constructing a system for study via molecular dynamics simulations to minimize the number of atoms. However, in limiting the system size, artifacts due to the finite system size are introduced [46]. For example, if a system of pure water is simulated without periodic boundary conditions, a significant percentage of the water will form a border with what is essentially a vacuum

and the water at the edge will form distinct structures [47].

The standard method of reducing these effects is to implement periodic boundary conditions (PBC). In PBC's, the system is replicated in all three dimensions around the central system. Hence, atoms forming the right hand side of the system will interact with the “image” of atoms on the left hand side of the system rather than vacuum. This is demonstrated schematically in Fig. 2.7.

Figure 2.7: Schematic demonstrating the use of periodic boundary conditions in two dimensions. The dashed circle represents the maximum sphere of direct interaction.

In GROMACS, the *minimum image convention* is used. In this convention, only the nearest image of any atom is used in the direct calculation of the forces. Hence, a sphere is drawn around each atom and interactions are calculated only with atoms within this sphere (this is shown in Fig. 2.7 by the dashed circle which shows that atom i interacts with an image of j). To avoid an atom seeing multiple instances of the same atom or itself, this sphere is limited to a radius which cannot exceed one-half of the smallest box length.

With most of our simulations being performed with an external field applied, periodic boundary conditions offer an additional benefit by maintaining the number of ions in the simulation. Through PBC's, an ion exiting the bottom of the system will appear at the top. Hence, a system can be run for long periods of time with the same ions cycling through.

2.5.3 Long Ranged Interactions

As discussed above, we introduce PBC's to correct for the finite system size effects. However, in their implementation, it is necessary to limit the radius of the sphere of interaction. Hence, a similar limited system size is introduced as there will essentially be vacuum outside of this sphere for any particular atom. While this is not an issue for the short ranged Lennard-Jones interactions or the bonded interactions, the electrostatic interaction is quite long ranged in comparison and truncating this force can have a noticeable impact on the

results of a simulation.

There are several methods to correct for the truncation of these long ranged interactions. For example, one can implement a *reaction field* where the space beyond the direct interactions is assigned a dielectric constant and the interaction of the atom with the medium is added in a continuum approximation approach [48]. While this is a reasonable solution to the problem, this approach is limited by having only a single dielectric constant implemented. Hence, the medium outside of the system should be homogeneous for the reaction field approach to be reasonable.

In our simulations, we use the particle-mesh Ewald (PME) methodology to handle the long ranged interaction. This approach is an extension of the Ewald summation [49]. In the Ewald summation, the forces are calculated beyond the limited sphere of interaction and are in fact summed over the images generated by the periodic boundary conditions. A direct implementation of this approach would be prohibitively expensive in terms of computer cycles. The trick used to resolve this problem is to break up the sum over the images into two sums: one in direct space and the other in reciprocal space. For technical reasons, this computation is much quicker and allows for an effective extension of the sphere of interaction to distances which are satisfactory for the Coulombic interaction as well as the shorter range interactions. Given this methodology, rather than having a sharp demarcation of the space each atom sees, the cut-off distance determines the portion of the

calculation which is done in real space with the other interactions being sums over a reasonable number of wave vectors in reciprocal space (for example, 10).

The particle-mesh Ewald is an extension of the Ewald summation approach [50, 51]. Here, for the purposes of efficiency, the charges beyond the direct interaction sphere are assigned to a grid [42]. This allows for the a fast Fourier transform of the grid from which the reciprocal space contribution can be calculated from a single sum rather than summing over individual wave vectors. This approach scales much better with system size and is more efficient in all but the smallest simulation systems.

Throughout our research on this project, the importance of correctly handling long-range interactions has become evident. Whether studying a simple system of only water and ions or a full system with a protein, the effect of “turning on” an algorithm to simulate long-ranged interactions was dramatic. As a consequence, the bulk of the simulations discussed in this document were performed using the particle-mesh Ewald algorithm for long-ranged interactions.

2.5.4 Temperature Coupling

In a molecular dynamics simulation, the issue of controlling the temperature of the system requires special attention. In the absence of temperature control, the mean temperature of the system can deviate from the desired

temperature (for example, physiological temperature) due to numerical error. In a more dramatic example, energy added to the system via the application of an external force will not be properly dispersed due to the finite size of the system. Hence, we aim to artificially couple the system to a heat bath.

One option that GROMACS offers to achieve this was developed Berendsen et al. [52]. Essentially, the Berendsen approach periodically scales the velocities of the atoms to achieve the user specified system temperature. The temperature is derived from the instantaneous velocities of the atoms according to the equipartition theorem:

$$\langle \frac{1}{2}mv^2 \rangle = \frac{3}{2}kT \quad (2.28)$$

where m is the mass of the atom, v is the velocity, k is the Boltzmann constant, and T is the temperature. As shown above, the velocity of each atom is calculated as a necessary part of the MD algorithm. Thus, we can scale the mean temperature of the system by scaling the velocities. This results in a numerical algorithm such that:

$$\frac{dT}{dt} = \frac{T_0 - T}{\tau} \quad (2.29)$$

where T is the temperature of the system, T_0 is the temperature of the heat bath (ie, desired temperature of the system), and τ is a constant which determines the rate of correcting the temperature. Hence, the user specifies the desired system temperature and an appropriate time constant in the

input files for the simulation.

In GROMACS, the system can actually be coupled to several baths at separate temperatures and with different time constants by splitting the system as a whole into groups. Temperature coupling is an important consideration for our project as we will be primarily running non-equilibrium simulations. Additionally, as will be discussed later, the ability to couple the system in separate groups will allow us to maintain a more realistic temperature profile.

2.6 CONCLUSIONS

In this chapter, a basic outline of how molecular dynamics simulations function has been given. In particular, the various terms which comprise the force field which is used to calculate the total force on each atom was described in some detail. For all these terms in the force field, parameters quantifying the interaction are needed. The subject of appropriate force field and force field parameters is a large and evolving subject and most simulation packages will come with various options for both the form of certain interactions as well as different sets of parameters to use in the calculation. Consequently, the user must make careful decisions regarding the force field when initially setting up the system for study.

Beyond details concerning the force-field, several issues which have a direct impact on the simulations which will be discussed in this document were covered. These included the use of periodic boundary conditions, the

treatment of long-ranged interactions, and methods of controlling the temperature of the system. Again, all of these techniques have options for their implementation which should be investigated when setting up a simulation.

CHAPTER 3: THEORY

3.1 INTRODUCTION

In this section, the theory behind our methodology of applying the electric field to the conducting ions only is given. In order to calculate the conductance of the channel, an electric field is applied to the system. Due to the relatively long simulation times required to observe a conductance event, we would like to apply as large a field as possible. However, applying a field so large that it is in the nonlinear regime could drastically affect the conductance as the protein may be distorted. Hence, we would like to apply the field to the conducting ions only. In order to do this, we need to show that the results in the linear regime are still relevant. Consequently, we go back to the Kubo theory of conductance and show that the results as obtained for applying the field to all charged species in the system or to the conducting ions only yield the same results. The following derivation primarily follows that outlined by R. Balescu [53], but there are many other excellent sources [54, 55].

3.2 DERIVATION

We begin with a system of N interacting particles which we assume has been undisturbed for a sufficient amount of time to be at equilibrium and thus can be described by the canonical phase space distribution :

$$\rho^0(\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N) = \frac{1}{N!h^{3N}Z} e^{-\beta H(\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N)} \quad (3.1)$$

where h is Planck's constant and β is the familiar $\frac{1}{k.T}$ factor where k is the Boltzmann constant and T is the temperature. Here, the q 's represent the position coordinate vectors and the p 's represent the momenta. Z is the partition function.

To begin the derivation, we start with the results of the Liouville theory [56] which states that:

$$\frac{\partial \rho}{\partial t} = L\rho \quad (3.2)$$

where L is Liouvillian operator defined by:

$$L = \sum_i \left(\frac{\partial H}{\partial \vec{p}_i} \frac{\partial}{\partial \vec{q}_i} - \frac{\partial H}{\partial \vec{q}_i} \frac{\partial}{\partial \vec{p}_i} \right) \quad (3.3)$$

When the external field is turned on at time $t=0$, there will be an additional term in the Liouvillian L^E so that:

$$L^T = L + L^E \quad (3.4)$$

We introduce the propagator $W(t)$ given by:

$$\rho(t) = W(t)\rho(0) \quad (3.5)$$

Hence, from (3.2) and (3.4) we have

$$\frac{\partial \rho(t)}{\partial t} = (L + L^E)\rho(t) \quad (3.6)$$

Using (3.5) gives

$$\frac{\partial W(t)}{\partial t} = (L + L^E)W(t) \quad (3.7)$$

We also introduce the unperturbed propagator defined by

$$\frac{\partial U(t)}{\partial t} = LU(t) \quad (3.8)$$

The solution to (3.7) is given by

$$W(t) = U(t) + \int_0^t d\tau U(\tau)L^E W(t - \tau) \quad (3.9)$$

To see that this is indeed the solution, we take the partial derivative of (3.9) with respect to time:

$$\begin{aligned}
\frac{\partial}{\partial t}W(t) &= \frac{\partial}{\partial t}U(t) + \int_0^t d\tau \frac{\partial}{\partial t}U(t-\tau)L^EW(\tau) + U(t-\tau)L^EW(\tau)|_{\tau=t} \\
&= LU(t) + \int_0^t d\tau LU(t-\tau)L^EW(\tau) + U(0)L^EW(t) \\
&= L\left(U(t) + \int_0^t d\tau U(t-\tau)L^EW(\tau)\right) + L^EW(t) \\
&= (L + L^E)W(t)
\end{aligned} \tag{3.10}$$

in agreement with (3.7). Examining (3.9), we make an approximation by taking the first iteration of (3.5) (linear response) and letting $W(t) \rightarrow U(t)$ such that

$$W(t) \doteq U(t) + \int_0^t d\tau U(\tau)L^EU(t-\tau) \tag{3.11}$$

Recall that

$$\rho(t) = W(t)\rho^0 \tag{3.12}$$

and hence

$$\rho(t) = U(t)\rho^0 + \int_0^t d\tau U(\tau)L^EU(t-\tau)\rho^0 \tag{3.13}$$

ρ^0 is the equilibrium solution of the unperturbed Liouville equation and hence:

$$U(t)\rho^0 = \rho^0 \tag{3.14}$$

Thus (3.13) becomes

$$\rho(t) = \rho^0 + \int_0^t d\tau U(\tau) L^E \rho^0 \quad (3.15)$$

We are interested in the average current $\vec{J}(t)$ which will be given by

$$\vec{J}(t) = \int d\vec{x} \vec{j} \rho(t) \quad (3.16)$$

where, for convenience, all of the position and momentum variables are collected into a single vector $\vec{x} = (\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N)$ which is termed the *phase space vector* [44] and $\vec{j}$ is the microscopic current defined by

$$\vec{j} = \sum_{i=1}^N e_i \vec{v}_i \quad (3.17)$$

where e_i is the charge of the i 'th atom and v_i is the velocity of the i 'th atom.

Using (3.15), we get from (3.16)

$$\vec{J}(t) = \int d\vec{x} \vec{j} \rho^0 + \int_0^t d\tau \int d^N x \vec{j} U(\tau) L^E \rho^0 \quad (3.18)$$

Now $\vec{j}$ is an odd function of the velocities since $\vec{J}(-\vec{v}) = -\vec{J}(\vec{v})$ where $\vec{v} = (\frac{\vec{p}_1}{m}, \frac{\vec{p}_2}{m}, \dots)$. However, for ρ^0 we have

$$\rho^0 \propto e^{-\beta H(\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N)} \quad (3.19)$$

$$\propto e^{-\beta \sum_i \frac{\vec{p}_i^2}{2m_i}} \quad (3.20)$$

$$\propto e^{-\frac{\beta}{2} \sum_i m_i v_i^2} \quad (3.21)$$

where the part of the Hamiltonian dependent on the velocities is the kinetic energy term and the other terms are omitted (this is shown in more detail below). Therefore, ρ^0 is an even function of the velocities as $\rho^0(-\vec{v}) = \rho^0(\vec{v})$. Hence, the first term in (3.18) is the product of an odd and an even function over a symmetric range (all of phase space) and thus evaluates to zero. Consequently we find

$$\vec{J}(t) = \int_0^t d\tau \int d\vec{x} \vec{j} U(\tau) L^E \rho^0 \quad (3.22)$$

Thus far, the perturbation to the equilibrium has been represented in general by L^E . However, we are concerned with the results of applying an external electric field to the system. Hence L^E is given by

$$L^E = \sum_i \left(\frac{\partial H^E}{\partial \vec{p}_i} \frac{\partial}{\partial \vec{q}_i} - \frac{\partial H^E}{\partial \vec{q}_i} \frac{\partial}{\partial \vec{p}_i} \right) \quad (3.23)$$

where H^E is the term added to the Hamiltonian by turning on the electric field $\vec{E}$. Specifically we have

$$H^E = \sum_i e_i \vec{E} \cdot \vec{q}_i \quad (3.24)$$

and thus (3.23) becomes:

$$L^E = \sum_i \frac{\partial}{\partial \vec{p}_i} (e_i \vec{E} \cdot \vec{q}_i) \frac{\partial}{\partial \vec{q}_i} - \frac{\partial}{\partial \vec{q}_i} (e_i \vec{E} \cdot \vec{q}_i) \frac{\partial}{\partial \vec{p}_i} \quad (3.25)$$

$$= - \sum_i e_i \vec{E} \cdot \frac{\partial}{\partial \vec{p}_i} \quad (3.26)$$

$$= - \vec{E} \cdot \sum_i \frac{e_i}{m_i} \frac{\partial}{\partial \vec{v}_i} \quad (3.27)$$

Therefore we have

$$L^E \rho^0 = - \vec{E} \cdot \sum_i \frac{e_i}{m_i} \frac{\partial}{\partial \vec{v}_i} \rho^0 \quad (3.28)$$

Recall that:

$$\rho^0 = \frac{1}{N! h^{3N} Z} e^{-\beta H(\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N)} \quad (3.29)$$

where H is the equilibrium Hamiltonian given by the sum of the kinetic energies and a general potential energy term:

$$H(x_1, \dots, x_N) = \sum_i \frac{\vec{p}_i^2}{2m_i} + U(\vec{q}_1, \dots, \vec{q}_N) \quad (3.30)$$

$$= \frac{1}{2} \sum_i m_i \dot{\vec{q}}_i^2 + U(\vec{q}_1, \dots, \vec{q}_N) \quad (3.31)$$

$$= \frac{1}{2} \sum_i m_i \vec{v}_i^2 + U(\vec{q}_1, \dots, \vec{q}_N) \quad (3.32)$$

and hence

$$\rho^0 = \frac{1}{N! h^{3N} Z} e^{\frac{-\beta}{2} \sum_i m_i \vec{v}_i^2} e^{-\beta U(\vec{p}_1, \dots, \vec{p}_N)} \quad (3.33)$$

Therefore we have

$$\frac{\partial}{\partial \vec{v}_i} \rho^0 = \left(\frac{1}{N! h^{3N} Z} \right) \beta m_i \vec{v}_i e^{-\frac{\beta}{2} \sum_i m_i \vec{v}_i^2} e^{-\beta U(\vec{p}_1, \dots, \vec{p}_N)} \quad (3.34)$$

$$= \beta m_i \vec{v}_i \rho^0 \quad (3.35)$$

Using this result in (3.28) gives

$$L^E \rho^0 = -\vec{E} \cdot \sum_i \frac{e_i}{m_i} (\beta m_i \vec{v}_i \rho^0) \quad (3.36)$$

$$= -\beta \vec{E} \cdot \sum_i e_i \vec{v}_i \rho^0 \quad (3.37)$$

$$(3.38)$$

and substituting this in (3.22) gives

$$\vec{J}(t) = \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \vec{j} U(\tau) \left(\sum_i e_i \vec{v}_i \right) \rho^0 \quad (3.39)$$

Formally, $U(\tau) = e^{\tau L}$ (from solving (3.8)) and hence:

$$\vec{J}(t) = \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \vec{j} e^{\tau L} \left(\sum_i e_i \vec{v}_i \right) \rho^0 \quad (3.40)$$

To continue the derivation, we now look at two separate cases.

3.3 SELECTIVE COUPLING

Equation (3.40) is general; we have not specified to which charges the electric field is coupled. We now take two cases:

i) $\vec{E}_{all}$ where the field is coupled to all the charges

ii) $\vec{E}_{ion}$ where the field is coupled only to the ions

3.3.1 $\vec{E}$ coupled to all charges

In (3.40), the sum will run over all charges. Hence, by definition:

$$\sum_{i=1}^N e_i \vec{v}_i = \vec{j} \quad (3.41)$$

Thus, the expression for the current becomes:

$$\vec{J}(t) = \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \vec{j} e^{\tau L} \vec{j} \rho^0 \quad (3.42)$$

Manipulating this we obtain

$$\vec{J}(t) = \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \rho^0 \vec{j} e^{\tau L} \vec{j} \quad (3.43)$$

$$= \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \rho^0 \vec{j} \vec{j}(\tau) \quad (3.44)$$

$$= \beta \vec{E} \cdot \int_0^t d\tau < \vec{j} \vec{j}(\tau) > \quad (3.45)$$

This has the form of a linear response equation - specifically it is the standard Ohmic relation:

$$\vec{J}(t) = \vec{\sigma}(t) \cdot \vec{E} \quad (3.46)$$

where the conductivity tensor $\vec{\sigma}$ is given by:

$$\vec{\sigma} = \lim_{t \rightarrow \infty} \vec{\sigma}(t) \quad (3.47)$$

$$= \beta \int_0^\infty d\tau \langle \vec{j} \vec{j}(\tau) \rangle \quad (3.48)$$

This is the Green-Kubo formula for the conductivity [57]. As with other fluctuation-dissipation relations, it links a nonequilibrium transport coefficient (here the conductivity $\vec{\sigma}$) to the autocorrelation function of equilibrium fluctuations of another property - in this case the microscopic current $\vec{j}$. A consequence of this theorem is the Onsager hypothesis [58, 59] which states that *the regression of microscopic thermal fluctuations at equilibrium follows the macroscopic law of relaxation of small non-equilibrium disturbances*. In other words, the response of a system after being perturbed from equilibrium by an external influence is statistically equivalent to the system relaxing to equilibrium following an internal fluctuation. There are many other examples of such relations. For example the thermal conductivity can be related to the autocorrelation function of the microscopic heat fluxes [53].

3.3.2 $\vec{E}$ coupled to conducting ions only

Going back to (3.40), if we apply the field to the ions only, then the sum will run only over the ions. Hence, we define:

$$\sum_{i=1}^{N_k} e_i \vec{v}_i = \vec{j}_k \quad (3.49)$$

where N_k is the number of ions and $\vec{j}_k$ is the microscopic current of the ions only.

Thus, the expression for the current becomes

$$J(t) = \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \vec{j} e^{\tau L} \vec{j}_k \rho^0 \quad (3.50)$$

Again, we manipulate this

$$\vec{J}(t) = \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \rho^0 \vec{j} e^{\tau L} \vec{j}_k \quad (3.51)$$

$$= \beta \vec{E} \cdot \int_0^t d\tau \int d\vec{x} \rho^0 \vec{j} \vec{j}_k(\tau) \quad (3.52)$$

$$= \beta \vec{E} \cdot \int_0^t d\tau \langle \vec{j} \vec{j}_k(\tau) \rangle \quad (3.53)$$

Hence, we find the conductivity is given by

$$\vec{\sigma} = \beta \int_0^\infty d\tau \langle \vec{j} \vec{j}_k(\tau) \rangle \quad (3.54)$$

3.4 EQUIVALENCE OF CONDUCTIVITY EXPRESSIONS

To end up with the same conductance for both cases, we need to show that (3.54) is equivalent to (3.48). To do this, note that the total microscopic current $\vec{j}$ can be broken up into contributions from the ion charges and the water charges

$$\vec{j} = \sum_j e_j \vec{v}_j \quad (3.55)$$

$$= \sum_{i=1}^{N_k} e_i \vec{v}_i + \sum_{j=1}^{3N_w} e_j \vec{v}_j \quad (3.56)$$

$$\equiv \vec{j}_k + \vec{j}_w \quad (3.57)$$

where N_K is the number of ions and N_w is the number of water molecules. We neglect other partial charges of atoms (e.g. those of the ion channel protein) as they are relatively immobile and hence will not contribute to the observed current.

Using this in (3.48), we obtain for the case of the electric field coupled to all charges (noting the equivalence of cross terms)

$$\begin{aligned}
\vec{\sigma}_{all} &= \beta \int_0^\infty d\tau \langle \vec{j} \vec{j}(\tau) \rangle \\
&= \beta \int_0^\infty d\tau \langle (\vec{j}_k + \vec{j}_w)(\vec{j}_k(\tau) + \vec{j}_w(\tau)) \rangle \\
&= \beta \int_0^\infty d\tau \langle \vec{j}_k \vec{j}_k(\tau) \rangle + \beta \int_0^\infty d\tau \langle \vec{j}_w \vec{j}_w(\tau) \rangle + \\
&\quad 2\beta \int_0^\infty d\tau \langle \vec{j}_k \vec{j}_w(\tau) \rangle
\end{aligned} \tag{3.58}$$

For the case of the field applied only to the ions, we obtain

$$\begin{aligned}
\vec{\sigma}_{ion} &= \beta \int_0^\infty d\tau \langle \vec{j} \vec{j}_k(\tau) \rangle \\
&= \beta \int_0^\infty d\tau \langle (\vec{j}_k + \vec{j}_w) \vec{j}_k(\tau) \rangle \\
&= \beta \int_0^\infty d\tau \langle \vec{j}_k \vec{j}_k(\tau) \rangle + \beta \int_0^\infty d\tau \langle \vec{j}_w \vec{j}_k(\tau) \rangle
\end{aligned} \tag{3.59}$$

To show the equivalence between (3.59) and (3.58) we will show that the water-water term and the ion-water term are negligible.

3.4.1 water-water term

The water-water term is given by

$$\int_0^\infty d\tau \langle \vec{j}_w \vec{j}_w(\tau) \rangle \tag{3.60}$$

where

$$\vec{j}_w = \sum_{j=1}^{3N_w} e_j \vec{v}_j \quad (3.61)$$

$$= \sum_{j=1}^{3N_w} e_j \frac{d}{dt} \vec{q}_j \quad (3.62)$$

$$= \frac{d}{dt} \sum_{j=1}^{3N_w} e_j \vec{q}_j \quad (3.63)$$

Hence we define:

$$\vec{j}_w = \dot{\vec{\mu}} \quad (3.64)$$

where $\vec{\mu}$ is the net dipole moment of the water given by summing the product of charge and position of each water atom. We can recast this using the following identity (for a derivation of this identity, see Appendix A):

$$\int_0^\infty dt \langle \dot{\vec{A}} \dot{\vec{A}}(t) \rangle = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle |\vec{A}(t) - \vec{A}(0)|^2 \rangle \quad (3.65)$$

Hence, the water-water term becomes:

$$\int_0^\infty d\tau \langle \dot{\vec{\mu}} \dot{\vec{\mu}}(\tau) \rangle = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle |\vec{\mu}(t) - \vec{\mu}(0)|^2 \rangle \quad (3.66)$$

Now, the right hand side of the above is simply the fluctuation of the net dipole moment growth of the water divided by t. Since the water charges are bound the growth is bounded, and hence for $t \rightarrow \infty$ the ratio (3.65) vanishes. This should be true whether or not conducting ions are present (at equilibrium) in the system.

3.4.2 Cross Terms

So far we have considered a general case with an electric field applied in any direction and hence considered the conductivity tensor $\vec{\sigma}$. For the following discussion, it is convenient to take the electric field applied in the z direction such that σ_{zz} is the relevant component of the tensor. Thus, since the normal conductivity is given by the zz component of the tensor and the conductance g will be given by

$$g = \frac{A}{L} \sigma_{zz} \quad (3.67)$$

where A is the cross sectional area and L is the axial length of the system. Hence, from this point forward, we use the notation that $j \equiv j_z$ and $\sigma \equiv \sigma_{zz}$ etc.

To show the cross terms are negligible, we will use the Schwarz inequality:

$$(A, B)^2 \leq (A, A)(B, B) \quad (3.68)$$

where (A, B) denotes any scalar product. Since we want to relate this to the cross terms given in (3.58) and (3.59), we define the scalar product here as:

$$(A, B) = \int_0^\infty d\tau < A(0)B(\tau) > \quad (3.69)$$

Hence:

$$(A, A) = \int_0^\infty d\tau < A(0)A(\tau) > \quad (3.70)$$

and

$$(B, B) = \int_0^\infty d\tau < B(0)B(\tau) > \quad (3.71)$$

A requirement for a scalar product (and the Schwarz inequality) is $(A,A) \geq 0$, with equality if and only if $A=0$. This is true for (3.70) since the integral is proportional to the relaxation time for $A(t)$. Using these in the Schwarz inequality (3.68) gives

$$\left(\int_0^\infty d\tau \langle A(0)B(\tau) \rangle \right)^2 \leq \left(\int_0^\infty d\tau \langle A(0)A(\tau) \rangle \right) \left(\int_0^\infty d\tau \langle B(0)B(\tau) \rangle \right) \quad (3.72)$$

Specifically we have

$$\int (\langle j_k(0)j_w(t) \rangle^2) \leq \left(\int \langle j_k(0)j_k(t) \rangle \right) \left(\int \langle j_w(0)j_w(t) \rangle \right) \quad (3.73)$$

Since the water-water term approaches zero and the ion-ion term is finite, we conclude that the cross term is also negligible as it is bounded by the above relationship.

3.5 CONCLUSION

We have shown that the water-water term is negligible by showing that

$$\int_0^\infty d\tau \langle j_w j_w(\tau) \rangle = 0 \quad (3.74)$$

In addition, we have shown that

$$\left(\int \langle j_k(0)j_w(t) \rangle^2 \right) \leq \left(\int \langle j_k(0)j_k(t) \rangle \right) \left(\int \langle j_w(0)j_w(t) \rangle \right) \quad (3.75)$$

and since the ion-ion term remains finite while the water-water term goes to zero, the cross terms are negligible as well. Hence, if we re-examine the

expressions for the two cases as given by (3.58) and (3.59) we find that both are given by:

$$\vec{\sigma} = \beta \int_0^\infty d\tau < \vec{j}_k(0) \vec{j}_k(\tau) > \quad (3.76)$$

Thus, provided we remain in the linear regime, we expect to find the same conductance regardless of whether we apply the field to all charged species in the system or to the conducting ions alone.

CHAPTER 4: THE WATER BOX

4.1 INTRODUCTION

To directly test the theory, we constructed a simplified system consisting of either one or two ions in a box of water molecules. Two types of runs were performed on this system. The first set were equilibrium runs in which no field was applied and the water and ions were free to diffuse. In the second set, an electric field was applied to either all the atoms (with their partial charges) in the system or the ions only. All data is produced using the GROMACS simulation package [60,61].

4.2 SYSTEM SETUP

The system was constructed by putting either a single potassium ion or a potassium and chloride ion in a 20 Å cubic box. The system was then overlaid with water using the GROMACS utility program genbox [42]. The water model we used was the simple point charge (SPC) model [62]. After overlaying the water, the resulting system consists of either one potassium

ion and 215 SPC model water molecules or one potassium ion, one chloride ion, and 214 SPC model water molecules. Due to the details of the overlay procedure, these densities are below the standard 1000 grams/litre. To correct the low densities, water molecules were inserted using the GROMACS program `edit_conf` [42] until each system contained their respective ions and 256 water molecules which brings the densities closer to 1000 grams/litre. This system is then energy minimized using a steepest descent algorithm until the changes in energies converge to machine precision [42]. The energy minimized system is then equilibrated for 10 nanoseconds according to the details giving in the following section. All runs, both equilibrium and nonequilibrium simulations, are started in the configuration given by the last frame of the equilibration simulation.

4.3 SIMULATION DETAILS

For all the water box simulations, periodic boundary conditions (PBC) were used in all dimensions and the temperature was controlled by coupling the system as a whole to an external bath at 298 Kelvin using the Berendsen algorithm [52]. After equilibration for 10 nanoseconds, all simulations are set to produce 10 nanoseconds of data with all the atom coordinates being written every 1 picosecond. In the equilibrium simulations performed for the purpose of analyzing the velocity autocorrelation functions, the velocity was written to disk at every step. A time step of 2 femtoseconds was used.

For the equilibrium runs concerned with calculating the diffusion coefficient, simulations were performed with a sharp cutoff at 7 Å and 9.9 Å as well as simulations in which these cutoffs are used for short-range interactions, but long ranged interactions are handled using a particle mesh ewals (PME) algorithm. In the case of the autocorrelation equilibrium simulations, the cutoff is set at 9.9 Å and PME is always used for the long ranged interactions. These parameters are also used for the nonequilibrium simulations, but here an electric field varying from 0.5 to 7 volts across the 20 Å cubic box was applied to either only the ions or to all atoms in the system.

4.4 DATA ANALYSIS

To analyze the generated trajectories, the effect of the periodic boundary conditions must be accounted for. For example, due to PBC's, an ion exiting the right side of the box will jump over and enter the left side of the box. However, in order to calculate things such as the diffusion coefficient or the ion current, a continuous trajectory is needed. Hence, a program was written which adds the box length to the relevant coordinate every time the ion exits and re-enters the box. This program is run on all the trajectories before analysis.

4.4.1 *Equilibrium Runs: Position versus Time*

For the equilibrium runs, the diffusion of the ions was monitored. After removing the effects of the periodic boundary conditions, the position versus time data is output in all three dimensions. From this data, the diffusion coefficient D is calculated using the Einstein relation [63]:

$$D = \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |\vec{r}(t) - \vec{r}(0)|^2 \rangle \quad (4.1)$$

where t is time and $\vec{r}$ is the position. Figure 4.1 displays a typical result for the averaged mean square displacement versus time. In this case, the data is taken from having both a potassium and a chloride ion in the box of water and the long ranged forces are handled via particle mesh Ewald (PME). The graph presents the mean displacement over 10 picoseconds averaged from 10 nanoseconds of data.

Figure 4.1: Mean square displacement of a potassium ion in a box of water versus time.

To test the importance of long-ranged forces, runs were conducted with a cut-off of 9.9 Å or 7.5 Å beyond which the interactions were either ignored (cut-off) or handled via the particle mesh algorithm (PME) [49]. All results are presented in Tables 4.1 and 4.2. In each case, a 9.8 nanosecond run was used and the resulting average mean square displacement for the first 10 picoseconds is used to calculate the diffusion coefficient. For this calculation, points within 10 picoseconds of each other are not used to avoid correlation between points. Consequently, each data point on the graph is the average

of 980 data points (9.8 nanoseconds of data over 10 picosecond windows).

One Ion	R=9.9 Å	R=7.5 Å
	K ⁺	K ⁺
Cut-off	3.600±0.036	4.119±0.034
PME	2.600±.021	2.540±0.023

Table 4.1: Diffusion Coefficient (in units of $10^{-9}m^2/s$) for a potassium ion in a box of water molecules.

Two Ions	R=9.9 Å		R=7.5 Å	
	K ⁺	Cl ⁻	K ⁺	Cl ⁻
Cut-off	2.933±0.016	2.783±0.025	3.248±0.046	2.843±0.047
PME	2.484±0.013	2.180±0.02	2.619±0.01	2.146±0.028

Table 4.2: Diffusion Coefficient (in units of $10^{-9}m^2/s$) for a potassium ion and a chloride ion in a box of water molecules

Table 4.1 shows the calculated diffusion coefficient for a potassium ion in a box of water. Examining Table 4.1, note the relatively large difference between the results obtained using sharp cutoffs and PME. Going from a sharp cutoff of R=9.9 Å to using a PME algorithm, D drops from $3.600 \pm 0.036 \times 10^{-9}m^2/s$ to $2.600 \pm .021 \times 10^{-9}m^2/s$. Similarly, at a cutoff of R=7.5 Å to using PME, D drops from $4.119 \pm 0.034 \times 10^{-9}m^2/s$ to $2.540 \pm 0.023 \times 10^{-9}m^2/s$. This is

the first of several indications that the handling of long-ranged interactions has an important impact on the results. Furthermore, notice that the impact of changing the cutoff from $R=9.9 \text{ \AA}$ to $R=7.5 \text{ \AA}$ in the case of sharp cutoffs causes D to change from $3.600 \pm 0.036 \times 10^{-9} m^2/s$ to $4.119 \pm 0.034 \times 10^{-9} m^2/s$. However, when PME is implemented, changing the size of the cutoff radius affects D to a much smaller degree as it changes from $2.600 \pm 0.021 \times 10^{-9} m^2/s$ to $2.540 \pm 0.023 \times 10^{-9} m^2/s$. This is to be expected as implementing PME effectively eliminates the sharp limited sphere of interaction (all interactions within the cutoff radius are explicitly handled while interactions outside are handled via PME). Table 4.2 shows the same results for the case of two ions in a box of water. The conclusions drawn from Table 4.2 are similar to the above, but the differences between the cases are less dramatic. This seems to indicate that although GROMACS can handle a system setup with a non-neutral net charge, it may be preferable to have a net neutral system. Additionally, it appears that it is best to have as many interactions handled explicitly as computation time allows. From this rather unsurprising result, we decided that using a cutoff of $R=9.9 \text{ \AA}$ (as large as possible in a box of length $20 \text{ \AA}$) as well as PME for long ranged interactions is the best setup. This is the system which is used for the nonequilibrium runs described in the following section.

4.4.2 Equilibrium Runs: Velocity versus Time

In addition to the position versus time data, an attempt was made to directly analyze the velocity autocorrelation functions given in the theory. Recall the ion-ion contribution to the conductivity tensor was given by:

$$\sigma_{ion}^{\vec{z}} = \beta \int_0^\infty d\tau \langle \vec{j}_k(0) \vec{j}_k(\tau) \rangle \quad (4.2)$$

where

$$\vec{j}_k = \sum_{j=1}^{Nk} q_j \vec{v}_j \quad (4.3)$$

or in the case of a single ion in a box of water:

$$\vec{j}_k = q\vec{v} \quad (4.4)$$

where q is the charge on the ion and $\vec{v}$ is the velocity of the ion. To analyze this result from the equilibrium runs, runs were performed with a time step of 1 femtosecond with the velocity being written every step. Although this resulted in very large data files (there were problems reading the trajectory files in older versions of the linux kernel), this was necessary in order to obtain sufficient resolution. The resulting ion autocorrelation curve for a potassium ion in a box of water is shown in figure Figure 4.2. The shape of this curve is well-known for an ion velocity autocorrelation function. The curve starts at a maximum and quickly falls off. A small shoulder is reached at

around 0.1 picoseconds beyond which the curve dips below zero to become negatively correlated before finally trailing away approaching zero. This curve is qualitatively correct when comparing to known results [64].

Figure 4.2: Autocorrelation function of the potassium ion (in units of $\frac{\text{nm}^2}{\text{ps}^2 e^2}$ where e is the charge on an electron) for a single potassium ion in a box of water molecules.

As equation 4.2 displays, the ion-ion contribution to the conductance is given by the integral of the curve shown in Fig. 4.2. However, there are difficulties in quantitatively analyzing this result. Even with the high data output, the error in the curve shown in Fig. 4.2, especially at longer times, is significant. Hence, we are limited to analysing the behaviour over the first 5 to 10 picoseconds at most. However, it is well known that the long

tails of the autocorrelation functions are important and should not to be neglected [64]. Thus, while the integration of the auto-correlation functions indicate behaviour consistent with our predictions, it is impossible to exactly verify our theoretical predictions from this data. The cumulative integration curve for the data shown in Fig. 4.2 is displayed below in Fig. 4.3. Note that the curve appears to saturate at a finite value as expected.

Figure 4.3: Cumulative integration of the autocorrelation function of the potassium ion (in units of $\frac{nm^2}{pse^2}$ where e is the charge on an electron) for a single potassium ion in a box of water molecules.

Figure 4.4 shows the autocorrelation of the electric current averaged over all the water molecules for a single potassium in a box of water.

Figure 4.4: Autocorrelation function of the electric current (in units of $\frac{\text{nm}^2}{\text{ps}^2 e^2}$ where e is the charge on an electron) averaged over all the water molecules for a single potassium ion in a box of water molecules.

In the above figure, note that the qualitative behaviour of rapid oscillations damping towards zero is qualitatively correct. However, the period of the oscillation is about 40 femtoseconds. If this rapid oscillation arises from the the bond bending of the water molecule, one would expect a frequency on the order of 20 femtoseconds. We are unsure if this discrepancy is a result a limitation of the SPC water model or the calculation. Recall the relationship:

$$\int_0^\infty d\tau \langle \dot{\vec{\mu}} \dot{\vec{\mu}}(\tau) \rangle = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle |\vec{\mu}(t) - \vec{\mu}(0)|^2 \rangle \quad (4.5)$$

where $\vec{\mu}$ is the net dipole moment of the water given by the sum of the charge

times position of all the atoms of the water molecules. The plot of the mean square deviation of the net dipole moment is included below.

Figure 4.5: Means square displacement of the net dipole moment of the water (in units of $\frac{nm^2}{e^2}$ where e is the charge on an electron).

Note that the dipole moment appears to be saturating at a finite value as predicted. To demonstrate that the water-water term goes to zero, the cumulative integration curve for the water current autocorrelation function as well as the mean square displacement of the net water dipole divided by the time is plotted below.

Figure 4.6: Means square displacement of the net dipole moment of the water divided by time and the cumulative integration of the water current autocorrelation function versus time (in units of $\frac{nm^2}{pse^2}$ where e is the charge on an electron).

In Fig. 4.6, note that there is good agreement between the curves at a sufficiently long time (the equivalence between the two forms is a long time limit). Furthermore, both curves appear to be asymptotically tending towards zero. However, a great deal more data at long times would be necessary to map out the long tails sufficiently to confirm this. As a final test of the theory, the cross correlation curves are included. The data for $\langle j_k(0)j_w(t) \rangle$ and $\langle j_w(0)j_k(t) \rangle$ is plotted below. Note that these curves are equivalent within the error range of the data.

Figure 4.7: The cross correlations function of the electric current (in units of $\frac{\text{nm}^2}{\text{ps}^2 e^2}$ where e is the charge on an electron).

In Fig. 4.7, both curves osciallate around zero. Hence, integrating the curves we get the results shown below.

Figure 4.8: The cumulative integration of the cross correlations function of the electric current (in units of $\frac{nm^2}{ps^2 e^2}$ where e is the charge on an electron).

Note that again these curves oscillate around zero and the fluctuations appear to be dampening as time goes on. Again, the curves are equivalent within the large error within the data due to the significant oscillations. Consequently, while a great deal more data would be necessary to verify this, we expect both of these curves to approach zero as required by theory. Hence, analysis of the data we currently have appears to verify our theory that the ion-ion term is the only term which contributes to the conductance. To directly test this prediction, nonequilibrium runs were performed.

4.4.3 *Nonequilibrium Runs*

For all the nonequilibrium runs, long ranged forces were handled via the particle mesh-Ewald PME algorithm. The currents were plotted against the applied voltage to yield current versus voltage graphs. Figure 4.9 shows the potassium ion trajectories for the case of a potassium and a chloride ion in a box of water. Here, the voltage is applied only to the ions and ranges from 0.5 volts to 7.0 volts in half a volt steps.

Figure 4.9: Net position (after removal of PBC) versus time for the potassium ion in a box of water along with a chloride ion. The different trajectories are from various voltages applied across the box to the ions only. The voltages range from 0.5 volts to 7 volts in half a volt steps.

Figure 4.10 shows similar data, but in this case the voltage is applied to all atoms in the system.

Figure 4.10: Net position (after removal of PBC) versus time for the potassium ion in a box of water along with a chloride ion. The different trajectories are from various voltages applied across the box to all atoms in the system. The voltages range from 0.5 volts to 7 volts in half a volt steps.

These trajectories are included primarily to indicate the small error which results from running the simulations for 10 ns. Consequently, the error for the current at each voltage lies within the data point itself and is not included on the graphs. The corresponding ion current can be calculated from the slope of the above plots. From this data, the current versus voltage (I vs. V) plots can be generated. Figs. 4.11 and 4.12 shows the I vs. V plots

for the field applied both to all charges and to the conducting ions only. In figure 4.11, there is only a potassium in a box of water molecules while figure 4.12 displays the results for the potassium ion current when there is also a chloride ion present.

Figure 4.11: Potassium ion current vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) for a potassium ion in a box of water molecules.

Figure 4.12: Potassium ion current vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) with both a potassium and a chloride ion in water.

In Fig. 4.11 and 4.12, note the relative overlap of the linear regimes when the field is applied to all atoms and applied to the ions only. This is particularly striking in the case of a potassium and chloride ion in the water (Fig. 4.12) where the linear regimes lie almost completely on top of each other. Additionally, after non-linearity sets in (at around 2.5 - 3 volts in both cases), note that the case of the field applied to all charges deviates from linearity much more quickly than in the case of the field applied to the conducting ions alone. This is the primary benefit of our approach: straying

into the nonlinear regime is less dramatic when the field is applied to the ions alone. We believe this effect will be even more important in the case of the ion channel, where any distortion of the selectivity filter due to non-linearity could have a large impact on the conductance.

To determine the conductance, plots of (I/V) vs. V were generated for the above data.

Figure 4.13: Potassium ion current/voltage vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) for a potassium ion in a box of water molecules.

Figure 4.14: Potassium ion current/voltage vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) for a potassium ion and a chloride ion in a box of water molecules.

The resulting conductance can be calculated from the mean of the values up to 3 volts (linear regime). For convenience, the conductance can be converted into a diffusion coefficient D using the Nernst-Einstein relation [64]:

$$D = \sigma / q^2 \beta \rho \quad (4.6)$$

where σ is the conductivity, ρ is the density of the charged particles, and all other terms are as previously defined. Note that the relation between the

conductance g and the conductivity σ is given by:

$$\sigma = gL/A \quad (4.7)$$

where L is the length of the system and A is the cross-sectional area. The results for all cases is shown in Table 4.3 along with the relevant equilibrium results.

	Nonequilibrium		Equilibrium
	V on Ions	V on All	
K^+ alone	2.173 ± 0.156	2.756 ± 0.214	2.600 ± 0.021
K^+ and Cl^-	2.138 ± 0.131	2.333 ± 0.102	2.484 ± 0.013

Table 4.3: Diffusion Coefficient (in units of $10^{-9}m^2/s$) of the potassium ion either alone in a box of water molecules or with a chloride ion in a box of water molecules. The nonequilibrium results are labelled according to either having the field applied to the ions only (V on Ion) or applied to all atoms in the system (V on All). The relevant equilibrium values are included as well. For these results, a cutoff radius of 0.99 nm with PME for long ranged interactions is used.

In Table 4.3, we see that the conductances derived from having the field applied to the ions only or to all the atoms are indeed similiar. Again, having a net neutral system seems to have an impact as the results are particularly close in this case (the K^+ and Cl^- case). Here, the diffusion coefficients are

$2.138 \pm 0.131 \times 10^{-9} m^2/s$ for the field applied to the ions only and $2.333 \pm 0.102 \times 10^{-9} m^2/s$ for the field applied to all atoms. Furthermore, both of these results compare well with the equilibrium results of $2.484 \pm 0.013 \times 10^{-9} m^2/s$.

4.5 CONCLUSIONS

The water box served as an excellent test case in which to check our methodology before proceeding. Recall that our basic idea is to apply an external force to the ions only while leaving the rest of the system uncoupled. Our primary motivation is to avoid distortions of the protein which may result from straying into the non-linear regime. The water box allowed us to directly test this approach by comparing the resulting conductances from having an external field applied to either the ions only or to all the atoms. As seen in Fig. 4.12, the results are very similar in the linear regime and diverge from linearity at different rates. This is quantified in Table 4.3 where the two values lie within error of each other. Additionally, both results compare well with the equilibrium value of $D = 2.484 \pm 0.013 \times 10^{-9} m^2/s$. We found this result very encouraging as a fundamental ratification of our theory and we now endeavour to apply this approach to a more complicated system containing the protein KcsA.

CHAPTER 5: KCSA

5.1 INTRODUCTION

The primary application of our methodology will be the efficient elucidation of the details of ion conduction through potassium ion channels. Specifically, we endeavor to determine the conductance of the KcsA protein. As discussed in the introduction, the determination of the structure of this membrane spanning potassium channel via xray crystallography was a major step in the understanding of ion channel structure and function in general and allowed us to take an experimentally determined model as a starting point for building our simulation system. Additionally, KcsA has a relatively high conductance (100-200 pS) depending on the concentration [65]) and a relatively low molecular weight (76 KDa [66]). This means that we can limit the size of the system and generate meaningful data within a reasonable time frame. Consequently, KcsA was an excellent choice for our study. We aim to produce the first molecular dynamics (MD) simulations of ion currents through a realistic ion channel model done at atomic scale. Previous non-equilibrium (field applied to the system) MD simulations of

ion currents [67,68] were based on simplified models of membrane pores, or Brownian Dynamics simulations with adjustable parameters [69,70].

5.2 K^+ CHANNEL PROTEIN STRUCTURE

In general, potassium ion channels are comprised of four identical subunits. These individual polypeptides associate to form a tetramer with a pore running through the center. This structure is displayed in Fig. 5.1. Each subunit contains 2 transmembrane helices (voltage gated K^+ channels will contain an additional 4 TM helices per subunit [71,72]), a short pore helix, a turret region extending into the extracellular medium, and a flexible section lined with carbonyl oxygens which forms the selectivity filter. These different regions are outlined in Fig. 5.2. In this image, the subunits above and below the plane of the page have been removed for clarity. Images of the protein were generated using the Visual Molecular Dynamics (VMD) package [73].

Figure 5.1: Top view of the KcsA ion channel. The four subunits are indicated by the different colours, with the transmembrane helices as darker ribbons and the other regions as lighter tubes. Note that the helices are at a significant angle and the subunits appear to unfold as the aperture of a camera. The selectivity filter and pore axis are indicated by the dotted orange box.

Figure 5.2: Cross section view displaying the major regions of the KcsA potassium ion channel. The transmembrane helices are represented by purple ribbons. The selectivity filter is indicated by the dotted orange box, with red dots denoting carbonyl oxygens. The orange arrows indicate the dipole generated by the pore helices. Note that these dipoles point to the center of the protein in a region labelled the water cavity. A crude schematic of the lipid bilayer is included to help orient the reader.

5.2.1 *The Selectivity Filter*

Among all potassium ion channels, the selectivity filter (or “P-loop”) is a highly conserved sequence TVGYG (threonine - valine - glycine - tyrosine

- glycine) which is implicated in the remarkable selection of potassium ions over other ions (most often, in comparison to sodium). As shown in Fig. 5.2, this region is the narrowest stricture through which the ion must pass on its path through the protein. The details of the selectivity filter are shown in Fig. 5.3.

Figure 5.3: The highly conserved TVGYG sequence which constitutes the selectivity filter of KcsA. The bindings sites are labelled S0-S4.

The backbone carbonyl oxygens of the amino acid residues lining the selectivity filter create binding sites. As seen in Fig. 5.3, the site labelled S1 is defined by the backbone carbonyl oxygens from the tyrosine residue above and the backbone carbonyl oxygens from glycine below. S2 is similarly defined with these charges being given by glycine above and valine below; for S3 it is valine above and threonine below. S4 is a little bit different

in that it has the backbone carbonyl of threonine above and the oxygen from the reactive group of threonine below. Additionally, there is a partial binding site above S1 which is defined by the backbone carbonyl oxygens of glycine above and tyrosine below. However, as can be seen in Fig. 5.3, the glycine oxygens are further apart and this site is more of a precursor to being in the selectivity filter than a part of the selectivity filter itself. Recall that there are two additional subunits which are not shown, so that each site actually contains 4 partial negative charges above and 4 below. These binding sites help compensate for the energy cost of at least partially dehydrating the potassium ions as they enter the protein. As will be seen in the simulation data, ions jump through these sites in discrete steps to traverse the channel. The details of binding and correlation between the ions in the kinetics of these jumps is of primary interest to understanding the details of conduction.

5.2.2 *The Water Pool*

Another notable structural feature is the water cavity which lies between the intracellular mouth and the selectivity filter comprising most of the center of the pore. Recall that the membrane is composed of a lipid bilayer with polar/charged head groups and neutral fatty acid tails. This means that the interior of the membrane is a primarily hydrophobic environment with a low dielectric constant. Consequently, there is an energy cost for an ion to

cross the membrane as it is unable to polarize its environment as it can in the bulk water. The presence of a significant amount of water in the “water pool” allows for some compensation of this energy cost. In addition, the pore helices contain a net dipole (a net partial negative charge on the water pool side and a net positive charge on the opposite end) which also serves to stabilize ions within the water pool.

5.2.3 *Charged Groups*

Like all proteins, the primary sequence contains amino acids which may occur in a charged state. Recall that, in terms of charged amino acids, glutamic acid and aspartic acid have a net negative charge unless neutralized by a proton (protonated) while lysine and asparagine have a net positive charge unless stripped of their extra proton (deprotonated). Considering that the passage of an ion crossing the channel will be heavily influenced by electrostatic forces (see for example the binding sites formed by the carbonyl oxygens of the selectivity filter), the protonation state of these groups is of significant importance. However, it can be difficult to know the protonation states of these groups as the local pH of the local environment at the atomic level is a hard quantity to calculate. To resolve this, the general location of the groups is used as an indication of what the charge state should be. Fig. 5.4 shows all the charged groups that are part of the protein.

Figure 5.4: Protonation states of charged amino acids for the KcsA model. Note that since the primary sequences of all the subunits are identical, each subunit will contain all the charged groups shown; they are broken up over two subunits simply for clarity. The numbers in black indicate the distance (in Å) from the pore axis while the numbers in green indicate the distance between the atoms containing net charges on pairs of charged groups.

In Fig. 5.4, all the charged groups on the left subunit are neutralized while the charges on the right are in a charged state with either a net positive/negative charge of one (with the exception of ASP80). Note that all four subunits contain all of these charges and they are separated in the figure for clarity only. The groups on the left (ARG27, ARG52, ARG64, ARG122) tend to be exposed to the extracellular solution, the intracellular solution, or extending parts of the protein not shown. As will be discussed later, the amount of solution in our system setup will be limited in order to perform the simulations in a reasonable time. Consequently, in the absence of a full ion/counter-ion solution which would provide counter charges for these groups to associate with, they are instead exactly neutralized by choosing a neutral topology when setting up the system.

On the right, the groups come in charged pairs. Examining Fig. 5.4 near the inner mouth, we see that the oxygen of GLU118, which contains a partial negative charge, is 2.25 Å from the hydrogen groups of ARG121 which carry a net partial positive charge. At this distance, this effectively forms a hydrogen bond and hence the groups are left with these charges. A similar situation is evident between GLU120 and ARG117 where the charged atoms are separated by approximately 2.58 Å. Looking at the figure near the turret region, we see that hydrogens of ARG89 are approximately 7.35 Å from the oxygens of GLU51. Although this separation is too large to be considered a bond, we consider these charges to cancel out and choose the charged state

when setting up the system.

The final charged groups constitute a special situation. Just behind the selectivity filter, there are two negatively charged groups: GLU71 and ASP80. The distance between the end of the reaction groups of these two residues is only 1.88 Å. Considering the packing of these groups, it is evident that one of them must be protonated and that that hydrogen is shared between the two groups. There is some debate about which group contains the hydrogen and which is left in negative state [78–80]. Based upon the findings of Roux et al., we decided to protonate ASP80 and have GLU71 carry a net negative charge.

This is the rationale which was used for choosing the different charged states for the charged residues. Note that with these choices, there is a net charge of negative four for the protein. Once we add in the four potassium ions, the system will have a net charge of zero. Although a net charge of zero is not required to perform the simulations (GROMACS can handle a non-zero net charge even when using PME), there have been some indications that a net zero system is preferable (see the water box results).

5.3 SYSTEM SETUP

We start with the structure for KcsA as determined by Mackinnon et al. in 2001 [22] via x-ray crystallography. One issue with this model that the inner mouth is in a closed state such that the pore formed by the transmembrane

helices on the intracellular side is too narrow to freely allow a potassium ion into or out of the channel [81–86]. Since we are interested in the conductance of ions through the channel, we must open the inner mouth before proceeding. To accomplish this, a force was applied to the groups near the inner mouth which pulled them radially away from the pore axis. Specifically, a linear force was applied to the C_α atoms of the backbone. After simulating for 100 picoseconds, the results were analysed and the subunit in the most stable position was isolated and replicated and rotated 3 times to construct a new tetramer. This system setup is taken as the starting configuration for the protein.

Once we have the protein in a state with the inner mouth in a conducting state, potassiums were added to the system. Two ions were placed in the selectivity filter; one in position S1 and the other in position S3. A third ion was placed in the water cavity beneath the selectivity filter and the fourth ion was placed above the selectivity filter. These positions are shown in Fig. 5.5.

Figure 5.5: Cross sectional view of the system setup for KcsA showing the placement of the potassium ions (green spheres) and the waters located between subunits (red and gray spheres). The carbonyl oxygens of the selectivity filter are shown to indicate the location of the binding sites.

In addition to the ions, note that Fig. 5.5 shows the location of 16 water molecules. The location of these water molecules was determined along with the protein structure [22]. These water molecules are located between adjacent subunits and serve as bridges between the subunits linking partial charges

through hydrogen bonds. The location of these waters can be seen more clearly from the top view which is shown in 5.6.

Figure 5.6: Top view of the system setup for KcsA showing the placement of the potassium ions (green spheres) and the waters located between subunits (red and gray spheres).

The final step in setting up the system is to solvate the system by filling the space with water. This step results in a very large number of waters being added to the system. If we define the system with a cell size of a box 80 Å per side, with the protein at the middle, filling the empty space with water results in the numbers of atoms going from 4016 to over 40,000 atoms.

Clearly, this would create a massive jump in calculation time. For this reason, we decided to limit the water to be a cylinder of radius 15 Å centered on the pore axis. This allows for a fully hydrated conduction pathway which axially links adjacent images via periodic boundary conditions (PBC). Thus, an ion can travel from the top of the system, through the protein, and out the bottom to re-enter through the solution. Making these cuts reduces the system size to 7148 atoms - a much more manageable system for the simulations we wish to perform.

There is an additional detail which needs to be addressed at this point. It is necessary to apply constraints to the system in order to maintain the water within a cylindrical structure centered on the pore axis. This possibility was not present in GROMACS, and hence the code was altered to include this option. Specifically, a harmonic constraint pulling the oxygens of the water molecules back into the cylinder if they strayed outside of the 15 Å radius was implemented (see appendix B for details). This procedure was implemented by altering the code present within GROMACS for constraining atom coordinates to their original position. Fig. 5.7 shows the full system setup including the cylinder of water centered on the pore axis.

Figure 5.7: Full system setup of the KcsA protein, potassium ions, and column of water (shown as red and grey balls and sticks).

5.4 SIMULATIONS DETAILS

For all runs, equilibrium and nonequilibrium, a time step of 2 femtoseconds is used (the LINCS algorithm is used to fix the hydrogen bonds allowing

a time step of 2 fs [45]). Interactions within 15 Å are calculated explicitly while interactions beyond this distance are handled via particle mesh Ewald (PME) [49]. Periodic boundary conditions (PBC) are used in all three dimensions but are relevant only in the axial direction where ions can exit and re-enter the system to have a continuous current. To maintain a reasonable temperature distribution, the system is coupled in several independent groups to an external bath at 298 Kelvin. Each segment per subunit that is free to move is coupled individually. The water is coupled in three separate groups of approximately even numbers of water molecules.

The reason for coupling the water in three groups instead of one requires a brief discussion on the nature of the temperature control with MD programs. As discussed in the chapter regarding molecular dynamics simulations in general, the temperature is calculated from the instantaneous velocity of the atom. The velocities are of course calculated as a part of the numerical integration which is the heart of MD simulations. Consequently, the temperature of atoms and average temperature of a group of atoms is readily available. To maintain the temperature within range of a constant value, the velocities of the atoms within the defined groups can be scaled up or down to meet the given value. The rate at which the velocities are adjusted can be controlled by setting the time constant (more details are given in the GROMACS manual for more details [42]). If the water was coupled as one large group, it would be possible to have part of the water heating

up and the velocities of all the water molecules being scaled down before the excess heat was dissipated throughout the water. For this reason, the water is coupled as three separate groups of about 320 water molecules each so that the groups can mix and local heating is avoided. In order for GROMACS to function properly, it was necessary to couple the ions to an external bath as well. However, to maintain the relative freedom of the ions, a very large time constant is used. This ensures that the coupling does not hinder the motion of the ions.

5.5 SYSTEM PREPARATION

Once we have the system setup, it is necessary to energy minimize and equilibrate the system so that it is in a state from which we can start the simulations. Energy minimization is done in GROMACS and performed to machine precision. The equilibration procedure is much more involved. To properly prepare the system, the system was equilibrated in three distinct steps.

For the first step, the selectivity filter is fixed in place while the other regions of the protein, the ions, and the water molecules are free to move. Basically, this allows the transmembrane helices to relax after they were pulled apart to achieve an open state at the inner mouth. This run is performed for only 1 ps. This allows the relaxation without having the helices resume a closed configuration. The different groups into which the protein was segmented are shown in Fig. 5.8.

Figure 5.8: The final configuration of the system following the equilibration procedure. The different groups are indicated by the coloured regions.

In the second step, The inner transmembrane helix was also frozen in place. During this 1 picosecond run, the outer transmembrane helix, the pore helix and turret region, the potassium ions, and water are free to move. This allows for further relaxation of the outer helices while maintaining an open

configuration of the inner helices. In the final stage of equilibration, the two transmembrane helices are frozen in place while the pore helix, turret region, and selectivity filter are free to move in addition to the ions and the water molecules. This is the major equilibration step as the final configuration of this run will be used as the starting point for the nonequilibration runs. Consequently, we perform this run for 1 nanosecond to allow sufficient time for the system to relax. The final configuration is shown in 5.8.

5.6 NONEQUILIBRIUM RUNS

Once the system is properly equilibrated, the nonequilibrium simulations can be started. For these runs, a voltage ranging from 1 volt to 5 volts across the system cell (80 Å) is applied to the conducting ions. For the initial set of runs, the voltage is applied such that the ions travel from the top of the system, through the protein, and out of the bottom (from the extracellular side to the intracellular side). Since our system does not explicitly include a membrane, the two transmembrane helices are frozen in place to maintain the overall protein structure. The pore helix, turret region, and selectivity filter are all free to move. The water is again constrained to remain with a 15 Å radius centered on the pore axis. The ions are also subject to this constraint such that they remain within the conducting pathway. The runs are setup to produce 5 nanoseconds of data with the coordinates being written every picosecond. The bulk of the initial data was generated on the first generation

of SHARCNET. At best, it took one week to generate a 5 nanosecond run per voltage. When there were heavy loads on the system, it could take up to 3 weeks to generate this data. These times have been significantly improved with the second generation of SHARCNET currently in operation upon which our followup simulations were conducted.

5.6.1 Applied Voltage $V = 5$ volts

The data generated from the nonequilibrium runs at 5 volts is shown in Fig. 5.9. The spikes in the plot indicate an ion exiting the bottom of the system and re-entering from the top via periodic boundary conditions. Due to the external field being applied to the ions, an ion exiting the selectivity filter will relatively quickly make its way through the water cavity, exit the protein, and travel through the water cylinder to bottom of the system. Hence, the spikes in the graphe each indicate a conductance “event”.

Figure 5.9: Ion axial trajectories versus time for 5 volts applied across the system.

In Fig. 5.9, notice that there are a significant number of events (i.e. 17) in the first 10 nanoseconds. The approximate locations of the binding sites are indicated by shading and labelled S0-S4. Since the selectivity filter, turret region and pore helix are free to move, there will actually be some variation in the exact location of these sites. However, given the general location, one can see how the ions pass through the selectivity filter by jumping through the sites in discrete steps as expected. Furthermore, one can see the importance of cooperativity between the ions as it is readily apparent that the

movement of one ion between sites is often correlated with the movement of a different ion between sites. For example, looking at the events in the first 10 nanoseconds, note that the movement of an ion from S0 to S1 (that is, into the selectivity filter) is usually accompanied by the movement of an ion from S2 to S3 as well as an ion going from S3 to S4 and then often quickly out of the channel. This correlated behaviour is due to the ion-ion repulsion and the details of the kinetics will be examined more closely at 3 volts. At this point, it is sufficient to merely draw the reader's attention to the highly correlated behaviour demonstrated in the events in Fig. 5.9.

After the first 10 nanoseconds, events are seen over the next 7 nanoseconds (the 10-17 nanoseconds portion of Fig. 5.9). However, after this time, the system seems to lock down and no further events are observed. This was a curious result and examination of the ion and protein behaviour over this stretch yielded little insight. It was not until the water in the vicinity of the selectivity filter was examined that things became clear. While for the rest of the events, the potassium ion traverses the channel fully dehydrated, Fig. 5.10 shows that a water molecule gets dragged with a potassium ion into the selectivity filter right before the events stop.

Figure 5.10: Close up of the selectivity filter showing the location of the blocking water molecule

The water molecule gets dragged (or pushed) into the S1 binding site and, once there, seems to stop the conductance. In fact, the system becomes locked into a configuration containing an ion in site S0, an ion in S2, and an ion in S4. As Fig. 5.9 shows, this configuration does not change over the next 18 nanoseconds of data collection. This is remarkable not only as the water seems to block the channel, but also as this configuration (0,2,4) is previously an unstable one with ions spending very little time in S4 over all.

The role of water in the conduction process is an open issue. While there is certainly water conduction through ion channels [74–76], the mechanism for water-ion co-transport or the ratio of water to ion conduction is not certain.

In the next chapter, the role of water in conduction is discussed further since in these followup simulations, water was prominent in the observed dynamics. However, in this case, the result of having the water block the channel may be an indication that we are in the non-linear regime. Consequently, further simulations were performed at lower voltages.

5.6.2 Applied Voltage $V = 4$ volts

Fig. 5.11 shows all the data collected with a voltage of four volts applied across the system. This data is similar to that shown for 5 volts. There is a significant number of events in the first 10 nanoseconds (10 events) and in the second 10 nanoseconds (14). However, after 3 events from 20 to 24 nanoseconds, the system again locks up with no events witnessed for the next 16 nanoseconds. As observed at 5 volts, the system seems to get stuck in a configuration consisting of ions in S0, S2, and S4. Closer inspection reveals that indeed, there is a water molecule in S1 in this “blocked” configuration. Again, while this result is puzzling, we continued to collect data at lower voltages.

Figure 5.11: Ion axial trajectories versus time for 4 volts applied across the system.

5.6.3 Applied Voltage $V = 3$ volts

The data from the simulations with 3 volts applied across the system is shown in Fig. 5.12. The 3 volt data agrees with many of the characteristics seen at 4 and 5 volts. For example, note the strong correlation between ions moving between binding sites as seen previously. However, there are also some differences. There appears to be a larger variation in the time between events. This may indicate a more natural distribution of events

as driven by the natural fluctuations required to achieve events. Another indication of this is an increase in the times an ion is observed to fluctuate between two adjacent binding sites (for example, examine the first 40 - 60 nanoseconds in which the ion in S2 will jump to S1 and then back down to S2 several times). Finally, the long pause in events which begins at 17 nanoseconds is not a blockage and conductance is observed to resume at 29 nanoseconds. Events are then seen sporadically until data collection is ended at 50 nanoseconds. This type of behaviour seems to indicate that we may have moved into the linear regime and to further test this, we ran simulations at even lower voltages.

Figure 5.12: Ion axial trajectories versus time for 3 volts applied across the system. The data from $t=0$ to $t=50$ ns is displayed in the five panels

5.6.4 *Applied Voltage $V = 2$ volts*

Fig. 5.11 shows all the data collected with a voltage of two volts applied across the system. The data collected here is very similiar to that observed at 3 volts. However, there are fewer events and the fluctuations and oscillations are even more pronounced (note the ion in S2 going back to S1 and down again between 17 and 19 nanoseconds). To fill out the eventual current versus voltage plot, we began to collect data at one volt as well.

Figure 5.13: Ion axial trajectories versus time for 2 volts applied across the system.

5.6.5 Applied Voltage $V = 1$ volt

The data generated at 1 volt is shown in Fig. 5.14. When viewing Fig. 5.14, note that some of the spikes are not events, but an ion oscillating across the boundary (specifically, the spikes observed at 8.5 and 13.7 nanoseconds). This is previously unobserved as these oscillations are suppressed at higher voltages. In fact, during the 20 nanoseconds of data shown, only one true event is observed. Clearly, much more data is required to make any conclusions about the current at this voltage. However, as the computer system we were using, SHARCNET, was being dismantled for a complete upgrade, we decided to continue with the project by calculating the current versus voltage plot for the data we had so far managed to collect.

Figure 5.14: Ion axial trajectories versus time for 1 volt applied across the system.

5.6.6 *Kinetics*

To examine the process of conductance in detail, the sequence of ion motion through the selectivity filter for all of the conductance events observed at 3 volts was examined. The full table is shown in Fig. 5.15. The n values represent the number of times that pathway was observed. The numbers are the average time, in picoseconds, between the initial and final state. The uncertainty is given by the standard deviation between the times when possible.

Figure 5.15: Full table of the observed sequence for all conductance events at 3 volts.

While this figure is comprehensive, it is actually too detailed to give a feel for the order of events. Hence, a simplified scheme using the dominant conduction pathway was constructed. This is shown in the inset of Fig. 5.16. The main part of Fig. 5.16 shows a magnified portion of the trajectory at 3 volts. This is actually the second 5 nanosecond run and corresponds to the 5 to 10 nanosecond portion of the trajectory shown in Fig. 5.12. The events shown here follow the kinetics scheme shown in the inset.

The order of events is as follows. Starting with ions in sites S1 and S3, an ion approaches from above the selectivity filter. As this ion enters S0, the ion in S1 almost simultaneously moves into S2 (as seen in the full analysis given in Fig. 5.15, 7 times out of 15 events an external ion enters S1 before the S2 to S3 transition, 3 out of 15 times the S2 to S3 transition preceeds the ion entering S0, and 5 out of 15 times the external ion entering S0 and the S2 to S3 transition happen simultaneously within a time resolution of 1 picosecond). This state is relatively long lived, with the next step being the transition of the ion in S3 to S4. Once the configuration S0,S2,S4 is achieved, there is a relatively rapid progression of the ion in S2 moving to S3 and quickly the ion in S4 exiting the selectivity filter. After a brief time in the water cavity, this ion will exit the protein and leave through the bottom of the system producing a spike the plot which is counted as an event.

It is interesting to compare these results to other findings. In particular, the x-ray crystallography results from Morais-Cabral et al. [21] predict an occupancy of 0.5 ions in each of the main binding sites (S1-S4) at sufficient ion concentration (200 mM). These results also predict that there will be 2 ions in the selectivity filter at any one time and consequently ions are generally found in either S1 and S3 or in S2 and S4. Results from energetic analysis performed by Berneche and Roux also find that there is a minimal energy barrier for ions moving between (S1,S3) and (S2,S4) in the selectivity filter [77]. However, in our simulations there is a preference for the ions to

occupy S1 and S3 in the selectivity filter. Furthermore, we do not observe any significant oscillation of the ions between (S1,S3) and (S2,S4). Rather, S4 has a relatively low occupancy as ions reaching this site are seen to exit the selectivity filter shortly thereafter. These results can be seen in Fig. 5.16.

Figure 5.16: Magnified portion of the trajectory generated at 3 volts. The inset displays the simplified kinetics scheme which is followed by both events shown in the trajectory.

5.6.7 Current vs. Voltage

To calculate the current corresponding to the various voltages, the number of events per 5 nanosecond run was calculated. This is then converted into a current by calculating the charge per unit time. The uncertainty in the current is derived from considering the variation of the number of events per 5 nanoseconds. For 5 volts, only the first 15 nanoseconds are used as after this the system becomes blocked as described above. This gives 3 blocks of 5 nanoseconds and the number of events are shown in Table 5.1.

Time Period	Number of Events
0 - 5 ns	10
5 - 10 ns	7
10 - 15 ns	5

Table 5.1: Number of events per 5 nanoseconds at 5 volts. This data gives a current of $2.35 \times 10^{-10} \pm 6.584 \times 10^{-11}$ Amperes.

At 4 volts, the first 25 nanoseconds are used to calculate the current. After this time, the sytem becomes blocked similarly to the situation at 5 volts. The data for the resulting 5 blocks of 5 nanoseconds is shown in Table 5.2.

Time Period	Number of Events
0 - 5 ns	5
5 - 10 ns	5
10 - 15 ns	7
15 - 20 ns	7
20 - 25 ns	3

Table 5.2: Number of events per 5 nanoseconds at 4 volts. This data gives a current of $1.73 \times 10^{-10} \pm 4.796 \times 10^{-11}$ Amperes.

At 3 volts, the full 50 nanoseconds of data is used to calculate the current. This data is shown in Table 5.3.

Time Period	Number of Events
0 - 5 ns	2
5 - 10 ns	3
10 - 15 ns	4
15 - 20 ns	1
20 - 25 ns	0
25 - 30 ns	1
30 - 35 ns	1
35 - 40 ns	0
40 - 45 ns	3
45 - 50 ns	1

Table 5.3: Number of events per 5 nanoseconds at 3 volts. This data gives a current of $5.127 \times 10^{-11} \pm 4.104 \times 10^{-11}$ Amperes.

The full 50 nanoseconds of data is used to calculate the current the current at 2 volts as well. This data is shown in Table 5.4.

Time Period	Number of Events
0 - 5 ns	2
5 - 10 ns	0
10 - 15 ns	0
15 - 20 ns	2
20 - 25 ns	1
25 - 30 ns	3
30 - 35 ns	0
35 - 40 ns	0
40 - 45 ns	0
45 - 50 ns	2

Table 5.4: Number of events per 5 nanoseconds at 2 volts. This data gives a current of $3.204 \times 10^{-11} \pm 3.510 \times 10^{-11}$ Amperes.

At one volt, we were only able to collect 20 nanoseconds of data in our first attempt. The number of events for this time is shown below.

Time Period	Number of Events
0 - 5 ns	1
5 - 10 ns	0
10 - 15 ns	0
15 - 20 ns	1

Table 5.5: Number of events per 5 nanoseconds at 1 volt. This data gives a current of $1.602 \times 10^{-11} \pm 1.602 \times 10^{-11}$ Amperes.

Using the data above, we can generate the current versus voltage graph. This is shown in Fig. 5.17. There is one subtle point which must be addressed before we can continue the data analysis. Recall that in our approach, we apply the field to the ions only. To calculate the conductance it is necessary to convert our field, applied to the ions only, to an equivalent trans-membrane voltage. As it is seen in Fig. 5.16, the ions spend most of their time in the selectivity filter region. Thus, this region is the most resistant part of the open channel. This is in agreement with the estimation of the applied voltage profile across KcsA made by Berneche and Roux [69] which showed that for an open KcsA potassium channel, virtually all the voltage drop occurs across the filter; the rearrangement of charges cancels out the field in all areas except across the selectivity filter. Consequently, the field is effectively magnified by having a voltage drop over a limited region in comparison to our system where the voltage drops linearly over the length of the entire system. Therefore,

to calculate the KcsA conductance, we use the electrical potential difference for the ion in the applied potential at the beginning and at the end of the selectivity filter (12 Å length) as an estimation for the equivalent voltage as opposed to the full system length (80 Å) and scale our voltage by (12/80). This voltage is plotted in Fig. 5.17 on the lower horizontal axis while the upper axis is labelled by the applied field over the axial length of the box (80 Å). For convenience, the latter voltages will be used to refer to the different data points.

Figure 5.17: Current vs. Voltage for the nonequilibrium simulations on KcsA

Note that the data point at one volt is a tentative point and hence is shown in paranthesis. Examining Fig. 5.17, we see that the linear regime extends up to 3 volts (applied) and the resulting conductance from fitting over these points is 113 pS. To compare this to experimental results [65], it is necessary to estimate the effective concentration of the bathing solutions. At 3 volts, we estimate there are on average 2.1 ions in the channel. Morais-Cabral et al. [21] found an average of 2.1 ions in the selectivity filter at a concentration of 200 mM. This corresponds to a conductance of 125 ps [65]. The dependence of the conductance on the concentration at these values is not very steep and the range of concentration from 100-300 mM corresponds to a conductance range of 100-150 pS. Hence, even with the uncertainties due to the system setup and the estimation of concentration, it is clear that the results are in good agreement with laboratory measurements.

This result exceeded our expectations and a paper was published which outline these results [87]. To continue the project, we devised several simulations to test the limits of the system setup. These experiments are discussed in the next chapter.

5.7 CONCLUSIONS

Starting with the structure determined by MacKinnon et al. in 2001 [22], a simulation system consisting of the protein, 4 potassium ions, and a column of water centered on the pore axis was constructed. After opening the inner

mouth into a conducting configuration, the system was prepared by a multi-step equilibration process to relax all parts of the system without losing the overall “open” state. Nonequilibrium runs were then performed by applying a voltage to the ions only. The trajectories from these runs indicate the high correlation between ions as they pass through the channel. In the runs performed with 4 and 5 volts across the system, a water molecule being dragged into binding site S1 stopped the conductance. As later seen in the full current versus voltage plot, these points are in the non-linear regime and this effect may just be due to non-linearity. At 2 and 3 volts, a significant number of events are observed and plotting the resulting current versus voltage yields a conductance of 113 pS which is in good agreement with the known value. Furthermore, careful analysis of the order of events at 3 volts yields a rough kinetics scheme of the correlated passage of ions through the selectivity filter.

CHAPTER 6: FOLLOWUP SIMULATIONS

6.1 INTRODUCTION

The results presented in the previous chapter gave not only a reasonable value for the conductance, but illuminated several details of the mechanisms of ion conduction. However, for purposes of putting these results in the appropriate context, several simulations designed to test the limits of the system setup and methodology were conducted. The first set of simulations simply continue the runs performed with one volt across the system in order to collect more data at this voltage. Beyond this, simulations were designed to test the effect of reversing the direction of the applied field, altering the protonation state of a charged group, or relaxing the constraints that kept the alpha helices fixed. Additionally, simulations investigating ion selectivity were performed.

6.2 APPLIED VOLTAGE $V = 1$ VOLTS CONTINUED SIMULATIONS

In the first set of simulations we performed with this system setup, only 18 ns of data was collected at an applied voltage of one volt. As this is the lowest voltage we tested, there were only 2 observed events in this time. Once SHARCNET came back online (after undergoing a complete overhaul), we continued the runs with an applied voltage of one volt. The resulting full 50 ns of data is shown in Fig. 6.1.

Figure 6.1: Ion axial trajectories versus time for 1 volts applied across the system.

When examining Fig. 6.1, note that not all the spikes are events as some as simply ions fluctuating across the system boundary. Recall that previously 2 events in 20 ns were observed. The additionally data contains 3 more

events such that we observe 5 events in 50 ns total. Hence, the calculated conductance remains identical but now is calculated from more events to improve the statistics. The complete table of events is shown below.

Time Period	Number of Events
0 - 5 ns	1
5 - 10 ns	0
10 - 15 ns	0
15 - 20 ns	1
20 - 25 ns	1
25 - 30 ns	0
30 - 35 ns	1
35 - 40 ns	0
40 - 45 ns	0
45 - 50 ns	1

Table 6.1: Number of events per 5 nanoseconds at 1 volt. This data gives a current of $1.602 \times 10^{-11} \pm 1.602 \times 10^{-11}$ Amperes.

Thus, as the conductance at one volt is identical to that which was calculated earlier, the same current versus voltage plot can be presented but now with proper error bars for the 1 volt data point.

Figure 6.2: Full Current vs. Voltage for the nonequilibrium simulations on KcsA

6.3 REVERSED ELECTRIC FIELD

The direction of the electric field used for the simulations is relatively arbitrary. In most cells containing potassium ion channels, the inside of the cell is negative with respect to the outside. However, there are ion channels which are observed to conduct potassium ions from outside the cell to inside (inward rectifiers) or ion channels which open at a certain voltage to release

potassium ions to the outside (voltage gated ion channels) or indeed ion channels which seem to conduct potassium ion channels in both directions rather freely (the 2P “leak current” ion channels) [9]. Furthermore, studying the ion channel in isolation, the behaviour should be linear through the origin in the Ohmic regime. Consequently, the choice of applying the field such that the ions flowed from outside the cell to inside in our simulations was a pragmatic decision: simulations conducted at higher voltages indicated that the geometry of the channel was more amenable to this direction of ion flow.

Hence, one of the first tests we conducted was to reverse this voltage. Thus, simulations identical to those discussed in the previous chapter were conducted except that the applied voltage was now -3 volts instead of 3 volts. The trajectories from these simulations are shown in Fig. 6.3.

Figure 6.3: Ion axial trajectories versus time for 3 volts applied across the system in a direction opposite from the results previously presented.

Examining Fig. 6.3, it is clear that the conductance as calculated in the reverse direction from previous results (discussed in the previous chapter) will be significantly less. While there are 11 events observed in the first 40 ns with the field driving the ions in the extracellular to intracellular direction, there is only one event when the field is driving the ions in the opposite direction. In the aims of finding a difference between these two cases, the system setup is shown in Fig. 6.4.

Figure 6.4: System setup with the external field driving the ions from the intracellular side to the extracellular side.

As shown in Fig. 6.4, there is an ion in S1 and S3. The other ions are in the water pool waiting to enter S4. However, there is a water molecule in S4 and the system stays in this configuration for 8 ns. Shortly after 8 ns, the ion in S3 moves up to S2. Simultaneously, the water molecule in S4 moves to S3 and the ion previously in S1 begins to oscillate between S1 and S0. At 11.3 ns, an ion from the water pool enters S4 and once this occurs, the ion oscillating between S1 and S0 stays in S0. Hence, for the 2 ns, there are ions in S0, S2, and S4 with a water molecule in S3. At 13.5 ns, the ion in S0 exits the selectivity filter yielding the only event. After this point, there are ions in S2 and S4 with a water molecule in S3. As there are no more events observed, this appears to be a blocked configuration. Hence, while we were unable to obtain the same conductance in the reverse direction, this may be part of a larger puzzle concerning the nature of water-ion interaction as they pass through the channel. Recall that with the voltage in the opposite direction, even when driving the ions at a high voltage yielding many events of fully dehydrated ions, a configuration with a water molecule in the selectivity filter stopped conduction altogether. This mechanism may explain the results seen in Fig. 6.4 as well.

6.4 ALTERING THE CHARGED STATE OF ARG64

As discussed in the chapter introducing the simulations concerned with KcsA, most of the charged amino acids were exactly neutralized. The justification of

this approach is that the system setup lacked a full ion / counter-ion solution and lipid bilayer with which these charged groups would form associations. Of course, exactly neutralizing these groups does not precisely mimic this effect. In reality, a charged amino acid which is exposed to the a solution containing ions will perturb the local environment to create a cloud of loose associations with oppositely charged ions. The distance over which this happens is called the Debye radius [64]. Hence, exactly neutralizing the charged groups in the topology over-neutralizes them as in reality there is still a finite separation between the effective charge centers.

To test the impact of this effect, a system was constructed in which an arginine residue near the selectivity filter was left in a charged state. Chloride ions were placed 10 Å (on the order of the Debye radius for a net positive charge in solution) directly above these four arginine groups (one from each subunit). to approximate the average position of an ion from solution cancelling the net charge. This setup is displayed in Fig. 6.5.

Figure 6.5: System setup with a charge on arginine 64. The ARG64 residue is shown for two subunits and all four Chloride counter-ions are shown in light blue. Also shown are all atoms of the water molecules which are within a 4 Å by 4 Å square centered on the conduction pathway.

Using this setup, the same equilibration method as employed for previous runs was used and data was collected by applying 3 volts across the system in the original direction. The resulting trajectories are displayed below.

Figure 6.6: Trajectories generated from the system setup with a net charge on arginine 64.

As is evident from Fig. 6.6, the majority of the data displays no events with a cluster of events taking place from 25 - 30 ns. An examination of the system setup confirmed our suspicions; there are water molecules in the selectivity filter which seem to prevent events for the first 25 ns of data. A typical configuration is shown in Fig. 6.5 with water molecules in positions

S1 and S3. As can be seen in the trajectory, the ion in S0 moves towards the outer mouth at around 24.5 ns. As this occurs, the water previously in S1 leaves the selectivity filter, passing through S0 and into the extracellular bulk solution. Subsequently, at around 25 ns, the ion previously in S0 returns to the S0 and passes quickly into the S1 position. Thus, there are now ions in S1, S2, and S4 with a water molecule occupying S3. This setup is shown in Fig. 6.7.

Figure 6.7: Snapshot of the system setup with a net charge on arginine 64.

At 25.4 ns, the ion in S4 exits the selectivity filter. Then, in what appears to be a concerted motion at 25.7 ns, the ion in S2 moves to S3 as the water molecule in S3 moves to S4. When in the S4 site, the water molecule is free to exchange with the water in the water pool. At 26 ns, the S3 ion moves to S4 and from this point on, several events are observed. In all of these events, the ions pass through the filter fully dehydrated. However, at 30 ns, another water molecule is pushed into the S1 position and the events stop.

This data displays interesting behaviour regarding the interplay between

ions and water occupying sites in the selectivity filter. However, due to the relatively short portion of trajectory which yielded events comparable to the data previously produced at 3 volts (discussed in the previous chapter), no quantitative analysis was possible. Rather, a more detailed simulation study which investigated both hydrated and bare ion passage through the channel while changing the charge state of key groups would be necessary. This undertaking is beyond the scope of this project and is left as future work.

6.5 RELAXED CONSTRAINTS OF THE TRANSMEMBRANE ALPHA HELICES

For the purposes of minimizing simulation time, an explicit lipid bilayer was not implemented. Hence, to keep the protein together, constraints to fix the atoms in place were applied to the transmembrane helices. However, even a minimal amount of flexibility may produce quite different results [88]. Hence, a system was constructed in which the transmembrane helices, previously frozen in place, were now allowed to oscillate around their original positions. These simulations were run with the atoms of the transmembrane helices being restrained to their original positions within a harmonic well with a force constant of $50 \frac{kJ}{nm^2 mol e}$. The trajectory is shown in Fig. 6.8.

Figure 6.8: Trajectories produced from a system setup with the transmembrane helices free to move around their initial position in a harmonic well with a force constant of $50 \frac{kJ}{nm^2 mol e}$. Three volts are applied across the system.

In the trajectories shown in Fig. 6.8, the runs start with an ion in S3 and an ion oscillating between S1 and S2. Around 1.3 ns, the ion in S3 moves to S4. At 1.5 ns, an ion enters the S0 and pushes a water molecule into S1. Thus, the system is in the configuration we have observed several times: an ion in S0, S2, and S4 with a water molecule in S1. However, unlike previous observations of this state for potassium ions where the alpha-helices are frozen, events are now observed for this configuration. At 7 ns, the ion in

S4 exits the selectivity filter and subsequently, the ions in S0 and S2 move into S1 and S3 while the water molecule moves from S1 to S2. This configuration persists for the next 2 ns. At 9 ns, as an ion moves into S0 and pushes a water molecule into S1, the ions in S1 and S3 move into S2 and S4 while the water molecule in S2 moves into S3. The resulting configuration is shown in Fig. 6.9.

Figure 6.9: System configuration for the relaxed transmembrane helices with ions in S0,S2,S4 and water molecules in S1 and S3.

This is a short lived state and within 0.1 ns, the ion in S4 exits the selectivity filter and the other ions move down so that the system then contains an ion in S1 and S3 with a water molecule between them in S2. The water molecule that moved from S3 to S4 is observed to exchange with the waters in the water pool. This state is relatively stable and lasts for over 18 ns.

Eventually, the system leaves this configuration as the ions move from S1 and S3 to S2 and S4 and the water molecule between them moves from S2 to S3 at 27.7 ns. Then, at 28.1 ns, an ion enters S0 and pushes a water

molecule into S1. Shortly thereafter, at 28.5 ns, there is a concerted motion as the ion in S4 exits the filter and the ions in S0 and S2 move into S1 and S3. At the same time, the water molecule in S1 moves into S2 and the water in S3 moves into S4 where it is free to exchange with waters from the water pool. The system stays in this configuration for the next 2 nanoseconds. At this point, the ions shift from S1 and S3 to S2 and S4 and the water between them moves into S3. There are no further changes observed by the time the trajectory ends at 35 ns. Snapshots from the event sequences are shown in Fig 6.10.

Figure 6.10: Snapshots outlining the events observed with relaxed transmembrane helix constraints.

Thus, for relaxed constraints on the transmembrane helices, we encounter the water blockage issue earlier than in previous simulations for the potassium ion. However, with a greater degree of flexibility, the filter seems to be able to accommodate the water molecule to the degree that eventually there is concurrent conductance of an ion and water molecule across the channel. While there are not enough events to produce a statistically significant result, the trajectories indicate that altering the constraints on the transmembrane helices could have a measurable effect on the conductance results. Again, it is left to future work to pursue these details.

6.6 SELECTIVITY: THE SODIUM ION

As discussed at several points in this document, one of the remarkable features of potassium ion channels is their ability to selectively conduct potassium over sodium in ratios on the order of 100 to 1. While the precise mechanism for selectivity is still unknown, there have been many simulations [89–91], calculations [7, 92–95], and experiments [6, 96–98] which shed light on various aspects of potassium versus sodium in K^+ ion channels. In this section, we attempt to simulate this selectively by replacing the potassium ions with sodium ions. To set up the system, the configuration used to produce the potassium ion currents is taken with a direct replacement of the four potassium ions by sodium ions. This system is then equilibrated for 10 ns and the nonequilibrium runs are started from this point. The trajectories

of the sodium ions are shown in Fig. 6.11.

Figure 6.11: Trajectories generated from the system setup with a replacement of potassium ions to sodium ions.

As shown in Fig. 6.11, the run starts with ions in S1 and S3. Within several picoseconds, the ion in S1 moves to S2 and oscillates between S1 and S2 several times. At 0.15 ns, an ion from the extracellular side moves into S0 and brings a water molecule along with it which occupies S1. Within 20 picoseconds, the ion in S3 moves into S4 and thus at 200 picoseconds, there are ions in S0, S2, and S4 with a water molecule in S1. This system setup is shown in Fig. 6.12.

Figure 6.12: System configuration with a water molecule in S1. Sodium ions are shown in purple.

The system configuration shown above appears to be relatively stable and the system remains in this state for over 1 ns. At 1.4 ns, the ion in S4 exits the selectivity filter into the water pool. The system then remains with an ion in S1 and S3 with a water molecule in between in S2 for the following 2 ns. At 3.5 ns, an ion enters S0. Within a couple of picoseconds, the ions in S1 and S3 move to S2 and S4 respectively and the water molecule in S2 moves to S3. At 3.6 ns, the ion in S0 moves to S1. This yields a selectivity filter with ions in S1,S2, and S4 with a water molecule in S3. Within 100 picoseconds, the ion in S4 exits the selectivity filter and the ions in S1 and S2 move to sites S2 and S3. Soon after, at 3.8 ns, an ion enters from the

extracellular side and occupies S0. This ion pushes a water molecule into S1 and at the same time, the ion in S3 moves into S4.

This results in the previously observed semi-stable state of having ions in S0, S2, and S4 with a water molecule in S1. In fact, the system remains in this configuration for over 17 ns. Then, at 22 ns, the ion in S0 moves laterally and the water in S1 passes through S0 and exits the selectivity filter. Snapshots from the succession of events as outlined above is shown in Fig. 6.13.

Figure 6.13: Snapshots outlining the first two events observed for the sodium ion conductance.

Continuing from the last frame shown in Fig. 6.13, five quick events are then observed. In each case, the ions pass through the filter fully dehydrated. Then, at 23.5 ns, an ion enters S0 and at the same time a water molecule moves into S1. This results in the relative blockage of the channel as seen before; no events are observed for the remaining 16 ns of data.

The sodium trajectories described above yield some interesting results. The first and second events are produced by the simultaneous passage of ions and water molecules across the selectivity filter. In contrast, for the potassium ion simulations performed with the transmembrane helices fixed in place (as they are for these sodium ion simulations), water is observed to block the channel and does not pass through the filter. Additionally, in the absence of water molecules in the selectivity filter, the channel appears to conduct sodium ions relatively well (this is seen in the trajectory from 22-24 ns). Although this segment is too short for quantitative analysis, 5 events within 2 ns is certainly comparable to the rate of events observed for potassium ions. However, the rate at which water was pushed into the channel is higher than was seen for potassium. This raises the possibility that the stronger interaction between sodium and water causes water to be dragged or pushed into the selectivity filter and consequently this lowers the conductance. Hence, this data suggests a promising set of simulation experiments investigating the conduction of sodium through the KcsA channel both concurrently with and in the absence of water. These investigations are left for

future work.

6.7 CONCLUSIONS

As the results from the simulations presented in this chapter indicate, there are many details of conduction left to be explored. The first set of simulations continued the runs at one volt, the lowest voltage which we simulated. This data confirmed the tentative results we obtained earlier and thus strengthened our previous conclusions. However, other simulation results presented in this chapter indicate that the conductance mechanism may be more complicated than we previously had observed. When applying the voltage in the direction opposite to that which produced a reasonable conductance, we immediately run into the water blockage issue. Altering the charged state of ARG64 yielded interesting results regarding the involvement of water as well. Here, once the waters originally in S1 and S3 leave the filter, there is a period of significant activity as 6 events are observed in less than 5 ns. However, another water molecule is pushed into the selectivity filter at this point and the events stop once again. Water also played a pivotal role in understanding the results of the simulations run under the conditions of relaxed constraints on the transmembrane helices. In these runs, water and ions were observed to traverse the selectivity filter together. Finally, the simulations regarding sodium conduction produced some surprising results. In the first part of these trajectories, sodium ions are observed passing through

the selectivity filter with water. However, when the water is expelled from the filter, a period of rapid conductance of fully dehydrated sodium ions is observed. Again, this period of high conductance comes to an end when a water molecule is dragged or pushed into the channel.

All of these simulations were designed to test the robustness of the system setup. The periods of high conductances observed in the sections of the trajectories involving fully dehydrated ions indicate that indeed the conductance behaviour outlined in the previous chapter can be replicated under slightly varying circumstances. However, the portions of the trajectories which show a minimum or absence of events due to the involvement of water molecules indicate that there are more subtleties which have yet to be explored.

CHAPTER 7: CONCLUSION

7.1 INTRODUCTION

This thesis presents the results of a research project designed to investigate the conduction mechanism of the KcsA potassium ion channel by directly simulating ions traversing the channel. To achieve this goal, a unique methodology of applying an external field to only the conducting ions was employed. The project thus breaks up into four distinct sections: the theory justifying the methodology, application of the methodology to the “water box” test system, application of the methodology to the KcsA ion channel, and followup simulations testing the system setup under varying conditions.

7.2 THEORY

In the third chapter of this document, the theory outlining the theoretical justification for applying the electrical field to only the conducting ions is presented. Recall that the primary motivation for using this approach is to avoid gross distortions of the protein when a significantly large external field

is applied to the system. Starting from the general theory of conductance in the Kubo formalism, the conductivity tensor is derived for two cases: i) applying the field to all charged species and ii) applying the field to the conducting ions only. The resulting expressions are then determined to be equivalent by showing that the water-water term is negligible and then eliminating the cross term by using the Schwarz inequality. Consequently, it is demonstrated that if one is within the linear regime, applying an external field to the conducting species only will yield the same conductance results as for applying the external field to all charged species. The next step was to test this theory in a simple setup.

7.3 THE WATER BOX

The water box was constructed as a simple test system in order to demonstrate the results of our theoretical prediction. This system was constructed by placing either a single potassium ion or a potassium and chloride ion in a cubic simulation cell filled with water molecules (simple point charge i.e., SPC model). Simulations were performed applying various voltages across the system cell to either all atoms or to the ions alone. The resulting current versus voltage plots for the potassium ion are shown below (in this graph, there are two ions in the cell (potassium and chloride) and particle-mesh Ewald (PME) is used for long ranged forces).

Figure 7.1: Potassium ion current vs. voltage for the water box system with an external field applied to all charged species (boxes) and to the ions alone (crosses) with both a potassium and a chloride ion in water.

The plot shown in 7.1 confirms what our theory predicts; in the linear regime, the conductance calculated from applying the field to the entire system is nearly identical to that calculated from applying the field to the conducting ions only. Furthermore, the results diverge when entering the nonlinear regime with the results for applying the field to only the conducting ions diverging less rapidly from linearity as opposed to the case of applying the field to all atoms. We believe this effect may be enhanced when applied to the protein as any distortion of the key elements of the protein (ie, the

selectivity filter) could significantly impact the conductance. Hence, with the theory validated by the test system, this methodology was then applied to the KcsA ion channel.

7.4 THE KCSA POTASSIUM ION CHANNEL

Perhaps the single most significant event in our current understanding of ion channel structure and function was the determination of the structure of KcsA in 1998 [20]. This remarkable experimental accomplishment provided the first complete picture of ion channel structure via x-ray crystallography. However, while the structure is now known, there are still many questions regarding the details of the structure-function relationship. For example, by what mechanism can ion channels demonstrate both high conductance and high selectivity? In order to elucidate these details, we endeavoured to directly simulate the conductance process by driving the ions through the channel. In particular, we employed our methodology of applying the external field to only the conducting ions to avoid gross distortions of the protein. The simulations were performed with a voltage ranging from 1-5 volts across the simulation cell. The resulting current versus voltage plot is shown below.

Figure 7.2: Current vs. Voltage for the nonequilibrium simulations on KcsA. The upper axis displays the voltage as applied across the length of the system while the lower voltage is the corrected voltage as determined by the voltage drop across the selectivity filter.

Note the linearity of the plot up to the three volt (applied voltage) point. The conductance as fit from these data points is 113 pS. While it is difficult to determine exactly what conductance we would expect with our system setup and methodology, we can determine that it would be in the range of

100-150 pS. Hence, the results of our simulations yields a conductance for the KcsA channel that is in good agreement with what has been measured experimentally. Additionally, the trajectories of the ions as they pass through the channel display some interesting behaviour. For example, the importance of ion-ion interaction is evident from the strong correlation of events as ions jump between binding sites. While these results exceeded our expectations, the next step was to test our system setup by altering key properties and measuring the effect.

7.5 FOLLOWUP SIMULATIONS

In the chapter regarding the followup simulations, the results from several runs which were used to further investigate the system were presented. These simulations included setups in which the direction of the electric field was reversed, the protonation state of ARG43 was changed, the constraints of the transmembrane helices were relaxed, and the potassium ions were changed to sodium ions. While the results of these simulations proved difficult to analyse quantitatively, some interesting behaviour was observed, primarily with respect to the interaction between the ions and water in the conduction process.

For the simulations with the reversed electric field and the charged state of ARG43, water molecules in the selectivity filter appear to block the conduction process. While this state persists throughout the runs with the electric

field in the opposite direction, the water molecules which are initially in the selectivity filter for the charged ARG43 runs are eventually cleared (they do not pass across the length of the selectivity filter but rather leave through the nearest entrance/exit). Once this occurs, a number of conduction events with fully dehydrated ions are observed until another water molecule enters the filter and the conduction events stop.

The simulations with the relaxed transmembrane helix constraints and the sodium ions also displayed interesting ion-water interaction. For these setups, water was again observed in the selectivity filter. Here, however, the water and ions pass across the length of the selectivity filter together and thus simultaneous water and ion conduction is observed. From these results, it appears that the additional motion provided by loosening the constraints on the transmembrane helices allows the selectivity filter to accommodate a water molecule more easily and allows enough flexibility for the water molecule to pass through the selectivity filter. Similarly, when sodium ions are substituted for the potassium ions, the selectivity filter is able to conduct water molecules in concert with sodium ions. This is contrast to the results for a system setup with potassium ions and frozen transmembrane helices where a water molecule was never observed to pass through the entire selectivity filter and rather seemed to block channel altogether; this blockage was observed even at four and five volts applied across the system. Hence, while these simulations did not replicate the high selectivity that is seen in the

lab, the details of the ion conduction was quite different from that observed for potassium and clues as to the selectivity mechanism may be found by exploring these differences.

7.6 FUTURE WORK

As the section discussing the results from the followup simulations indicates, there is still much to be learned from MD simulations of ion channels. However, in devising future experiments, there are many improvements which can be implemented. While the computational facilities available to our group when this project began were impressive at the time, increases in the speed of computer processors as well as the construction of computing facilities such as SHARCNET have increased our computational power many times over. In fact, while it used to take a week or two to generate one nanosecond of data for the typical system setup used for this project, we are now able to generate a 5 nanosecond trajectory overnight. Clearly, this increase in computational power allows for significant improvements to the system setup and simulation details.

An increase in computational power allows for a more complete simulation procedure. For example, the lack of electronic polarizability in the force-field is a well known limitation of many simulations [99,100]. Consequently, there are various methods being developed to incorporate polarizability into molecular dynamics simulations [101,102] and to generate polarizable force-

fields [103]. For example, the latest version of GROMACS includes the option to treat polarization by attaching shell particles to the atoms [42]. While the addition of a polarizable term to the force-field will increase the computation time, there may be a significant affect on the results [104]. Beyond the inclusion of polarizability, there are other improvements which can be made to the force-field. Of course, a simulation is only as good as the force-field which is used to define the interactions. However, these parameters are fit from varying scenarios and ultimately can only approximate the interactions for specific scenarios. Hence, refinements of the force-field which improve the description of important interactions will have a significant impact on the results [105]

Of the improvements which can be made to the system setup, perhaps the most important will be the inclusion of an explicit lipid membrane. With an appropriately constructed lipid membrane into which the protein can be inserted, there will be no need for any constraints on the protein itself. This will certainly allow for a closer approximation of the laboratory setup and, as the results from the simulations with relaxed constraints indicate, will have a significant impact on the results. Additionally, rather than limiting the water in the system to a cylinder centered on the pore axis, the full area above and below the membrane can be filled with bulk water containing both the conducting ions and counter-ions. This will generate a complete ion/counter-ion solution which has many benefits including: i) a direct control over the

concentration of ions in the solution allowing for simulations testing the conductance mechanism at varying concentrations ii) allowing charged amino acids exposed to the bulk to be left in a charged state to be “naturally” neutralized by association with ions rather than by explicitly neutralizing it iii) eliminating the pockets of vacuum generated by cutting away all water not within 15 Å of the conduction pathway.

When the system is expanded to include the membrane and the ion/counterion solution, the number of atoms in the simulation cell will increase by at least an order of magnitude. However, the increased computational costs are no longer prohibitively expensive and the improvement in approaching the laboratory setup will be well worth the cost. Additionally, with the improvement in computing speed, the need to apply the external field to the conducting ions may also be unnecessary. While this approach allowed us to perform the runs at voltages much larger than those applied in the laboratory, it is now possible to run simulations at lower voltages comparable to the voltages used experimentally. Hence, one may be able to apply the external field to all the atoms without fear of distorting the protein and destroying the conductance mechanism.

Implementing these changes will result in the “next generation” of system setup. While the simplified setup has generated some interesting and illuminating results, the followup runs demonstrate that it is time to broaden the simulations with these enhancements. With this system, a similar study

investigating the conduction of potassium and sodium ions at a range of voltages and concentrations would generate interesting results which could be compared against both experimental results [65,106] as well as theoretical predictions of the ion conduction kinetics [107]. Additionally, simulations attempting to replicate the selectivity as well as probing the effect of changing the protonation state of key groups or mutating a particular amino acid would help illuminate the details of the conduction mechanism.

7.7 CONCLUSION

The goal of this project was to directly simulate the conductance of potassium ions through the KcsA ion channel. In pursuit of this objective, a unique methodology of applying the field to the conducting ions only was developed. After validating this approach with a simple test system, simulations were performed on a system containing the experimentally determined KcsA structure. The results of these simulations yielded a conductance in very good agreement with the experimental results. Furthermore, the trajectories gave insight into the details of ion conductance and demonstrated the importance of such things as ion correlation. Subsequent to these results, followup simulations were performed to test the system under varying conditions. While these results proved to be ambiguous in terms of quantitative analysis, the behaviour which was observed, particularly with regard to the ion-water interactions, indicate that there is much left to be discovered about

the detailed operation of ion channels.

Consequently, the next step will be the construction of a more robust system setup including a membrane and a full ion/counter-ion solution upon which similar non-equilibrium simulations can be performed. A systematic study of this system under varying conditions (applied voltage, ion concentration, protonation states of charged groups) is sure to yield data which provides new information about the conduction process, the importance of the ion-water relationship (i.e., the fully dehydrated conduction mechanism versus the correlated ion-water transfer), and general insight into the structure-function relationship for integral membrane proteins.

CHAPTER 8: APPENDIX A

In this appendix the following identity is derived:

$$\int_0^\infty \langle \dot{\vec{A}} \dot{\vec{A}}(t) \rangle = \lim_{t \rightarrow \infty} \frac{1}{2t} \langle |\vec{A}(t) - \vec{A}(0)|^2 \rangle \quad (8.1)$$

For convenience, we will use the position vector $\vec{r}$ as the general vector $\vec{A}$ and hence $\dot{\vec{A}}$ is represented by the velocity vector $\vec{v}$. We begin the derivation with the following:

$$\vec{r}(t) - \vec{r}(0) = \int_0^t dt' \vec{v}(t') \quad (8.2)$$

Squaring (8.2) and taking the average gives

$$\langle [\vec{r}(t) - \vec{r}(0)]^2 \rangle = \int_0^t dt' \int_0^t dt'' \langle \vec{v}(t'') \cdot \vec{v}(t') \rangle \quad (8.3)$$

Using the fact that the velocity correlation is invariant under a shift in time, we obtain

$$\langle [\vec{r}(t) - \vec{r}(0)]^2 \rangle = \int_0^t dt' \int_0^t dt'' \langle \vec{v}(t'' - t') \vec{v}(0) \rangle \quad (8.4)$$

We now define a new time variable τ as follows:

$$\tau = t'' - t' \quad (8.5)$$

$$d\tau = dt'' \quad (8.6)$$

Converting the limits of intergration we have

$$t'' = 0 \quad , \quad \tau = -t' \quad (8.7)$$

$$t'' = t \quad , \quad \tau = t - t' \quad (8.8)$$

and hence (8.4) becomes

$$< [\vec{r}(t) - \vec{r}(0)]^2 > = \int_0^t dt' \int_{-t'}^{t-t'} d\tau < \vec{v}(\tau) \cdot \vec{v}(0) > \quad (8.9)$$

Equation (8.9) can be recast as

$$< [\vec{r}(t) - \vec{r}(0)]^2 > = 2 \int_0^t d\tau \int_0^{t-\tau} dt' < \vec{v}(\tau) \cdot \vec{v}(0) > \quad (8.10)$$

To see that (8.9) is equivalent to (8.10), examine Fig. 8.1 which displays a graphical representation of the domain of integration across τ and t' .

Figure 8.1: Graphical representation of the domain of integration as defined in equations (8.9) and (8.10)

In equation (8.9), the interior integral over τ runs from $-t'$ to $t - t'$. Hence, when $t = 0$, the integral runs from 0 to t along the τ axis. As t' increases in the outer integral, this line of integration along τ shifts upwards and to the left as represented by the horizontal lines in Fig. 8.1. At the upper limit, when $t' = t$, the line runs from $\tau = -t$ to $\tau = 0$. Consequently, the area enclosed is the parallelogram shown in the left hand side of Fig. 8.1. The right hand side of Fig. 8.1 shows one half of this parallelogram and here the order of integration is switched such that the inner integration runs over t' while the outer integration runs over τ .

Taking just one half of the parallelogram, the limits of integration will run from 0 to $t - \tau$ over t' in the inner integration and from $\tau = 0$ to $\tau = t$ in the outer integration (this is represented by the vertical lines in the right hand side of Fig. 8.1 which indicate the lines of integration across t' as τ

increase from 0 to t). Note that we have to multiply this result by 2 since we are only integrating over $\frac{1}{2}$ of the original area. Hence, the integration in (8.9) as shown in the right hand side of Fig. 8.1 is equivalent to that in (8.10) which is shown in the right hand side of Fig. 8.1.

Examining (8.10), we see that the argument is independent of the inner integration over t' and hence:

$$\langle [\vec{r}(t) - \vec{r}(0)]^2 \rangle = 2 \int_0^t d\tau \langle \vec{v}(\tau) \vec{v}(0) \rangle \int_0^{t-\tau} dt' \quad (8.11)$$

$$= 2 \int_0^t d\tau \langle \vec{v}(\tau) \vec{v}(0) \rangle (t - \tau) \quad (8.12)$$

$$= 2t \int_0^t d\tau \left(1 - \frac{\tau}{t}\right) \langle \vec{v}(\tau) \vec{v}(0) \rangle \quad (8.13)$$

$$= \lim_{t \rightarrow \infty} 2t \int_0^t d\tau \langle \vec{v}(\tau) \vec{v}(0) \rangle \quad (8.14)$$

Thus, we end up with the result that

$$\lim_{t \rightarrow \infty} \frac{1}{2t} \langle |\vec{r}(t) - \vec{r}(0)|^2 \rangle = \int_0^\infty \langle \dot{\vec{v}}(t) \dot{\vec{v}}(0) \rangle \quad (8.15)$$

as required.

As an interesting and important use of this identity, consider the following connection to the diffusion coefficient D :

$$D = \frac{1}{3} \int_0^\infty dt \langle \vec{v}_i(t) \vec{v}_i(0) \rangle \quad (8.16)$$

$$= \lim_{t \rightarrow \infty} \frac{1}{6t} \langle |\vec{r}(t) - \vec{r}(0)|^2 \rangle \quad (8.17)$$

Hence, the diffusion coefficient can be calculated either from the mean square displacement of the position or from the autocorrelation function of the velocity.

CHAPTER 9: APPENDIX B

As discussed in the chapter regarding MD simulations, the ability to restrain a group of particles according to a geometric surface was not available in GROMACS. Hence, for the simulations with the protein, it was necessary to alter the GROMACS source code such that the water molecules in the simulations were restrained to be within 15 Å of the conducting pore. This appendix describes the details of this change.

The relevant file is `bondfree.c` which is found under the path *gromacs* – 3.3.1/*src/gmxlib/* using the standard installation. In particular, the function that needs to be altered is the *posres* function. The code for this function, with the changes, is shown below (some superfluous code/comments have been edited out for clarity).

```

real posres(int nbonds,
            const t_iatom forceatoms[], const t_iparams forceparams[],
            const rvec x[], rvec f[], rvec fshift[],
            const t_pbc *pbc, const t_graph *g,
            real lambda, real *dvdlambda,
            const t_mdatoms *md, t_fcdata *fcd)
{
    int i, ai, m, type, ki;
    const t_iparams *pr;
    real vtot, fm, fc;
    rvec dx;

    real r_hdh; /*Distance from pore axis*/
    real ro_hdh=1.5; /*Set axial distance*/
    rvec base; base[0] = 4; base[1] = 4; base[2] = 0;
    /*vector given system center*/
    int fctest; /*value of force constant*/

    vtot = 0.0;
    for(i=0; i<nbonds; i++) {
        type = forceatoms[i+1];
        ai = forceatoms[i+2];
        pr = &forceparams[type];
        fctest=(int)pr->posres.fcA[0]; /*get the force constant value*/

        if(fctest == 999) { /*if water molecule*/
            ki = pbc_rvec_sub(pbc, x[ai], base, dx);
            dx[2]=0; /*Ensure the difference in the Z-direction is zero*/
            r_hdh = sqrt( (dx[0]*dx[0] + (dx[1]*dx[1]) ); /*Planar distance*/
            dx[0]=dx[0]*(1-ro_hdh/r_hdh); /*distance from cylinder in x*/
            dx[1]=dx[1]*(1-ro_hdh/r_hdh); /*distance from cylinder in y*/
            v=0;
            if( r_hdh > ro_hdh ) { /*Do restraint if atom is outside the defined*/
                for (m=0; (m<DIM); m++) { /*distance from the pore axis*/
                    *dvdlambda += harmonic(pr->posres.fcA[m], pr->posres.fcB[m],
                                            0, pr->posres.posOB[m]-pr->posres.posOA[m],
                                            dx[m], lambda, &v, &fm);
                    vtot += v;
                    f[ai][m] += fm;
                }
            }
        }
        else {
            if(fctest != 0) { /*Do regular constraint*/
                ki = pbc_rvec_sub(pbc, x[ai], forceparams[type].posOA, dx);
                v=0;
                for (m=0; (m<DIM); m++) {
                    *dvdlambda += harmonic(pr->posres.fcA[m], pr->posres.fcB[m],
                                            0, pr->posres.posOB[m]-pr->posres.posOA[m],
                                            dx[m], lambda, &v, &fm);
                    vtot += v;
                    f[ai][m] += fm;
                }
            }
        }
    }
    return vtot;
}

```

Figure 9.1: Modified GROMACS position restraint code.

While the code shown in the above figure contains comments which ex-

plain the process, a couple of additional notes are required. When modifying this function, we wanted to add the ability to restrain the waters to a cylinder centered on the pore axis while maintaining the ability to restrain other atoms to their initial position. The water restraint is used in all protein simulations while the “regular” position restraint is used only for the simulations where the constraint on the atoms of the transmembrane helices are relaxed from being fixed in place to being restrained to their initial positions. Hence, in the code, we discriminate between water molecules (specifically, the oxygen atom of the water molecule) and other atoms by assigning all oxygen atoms of the water molecule a force constant of $999 \frac{kJ}{molnm^2} e$ (the default value is $1000 \frac{kJ}{molnm^2} e$). The value of the force constant is defined in the run input files and is stored as ‘pr->posres.fcA[m]’ where m is the dimension (i.e., 0 for x, 1 for y, and 2 for z).

If we are dealing with an atom other than a water molecule, the function `pbcrvec_sub(pbc,x[ai],forceparams[type].posres.pos0A,dx)` is called which calculates the difference between the atom position ‘x[ai]’ and the initial position of the atom given by ‘forceparams[type].posres.pos0A’ and returns the difference as the vector ‘dx’. This distance, along with the force constant given by ‘pr->posres.fcA[m]’, is used to calculate the force ‘fm’ which is then added to the total force on the atom. This is the regular position restraint as included in the GROMACS package.

If it is a water molecule, the function `ki = pbcrvec_sub(pbc,x[ai],base,dx)`

is called. This function compares the atom position given by 'x[ai]' to the center of the system as defined by the vector 'base' and returns the difference as the vector 'dx'. The planar distance is then calculated as r_hdh, and if the water molecule lies outside of the radius of the cylinder defined by 'ro_hdh', the constraint is applied. This is the constraint procedure used for the water molecules.

BIBLIOGRAPHY

- [1] B. Hille, *Ion Channels of Excitable Membranes*, (Sinauer Associates, Sunderland, MA, 2001), 3rd ed.
- [2] F.M. Ashcroft, *Ion Channels and Diseases*, (Academic Press, San Diego, CA, 2000).
- [3] A. Parsegian, “Energy of an ion crossing a low dielectric membrane: Solution to four relevant electrostatic problems”, *Nature*, **221**, 844-846, (1969).
- [4] B. Roux, S. Berneche, and W. Im, “Ion Channels, Permeation, and Electrostatics: Insight into the Function of KcsA”, *Biochemistry*, **39(44)**, 13296-13306 (2000).
- [5] D. P. Tieleman, P.C. Biggin, G.R. Smith, and M.S.P. Sansom, “Simulation approaches to ion channel structure-function relationships”, *Q. Rev. of Biophysics*, **34**, 473-561 (2001).
- [6] E. Gouaux and R. MacKinnon, “Principles of Selective Ion Transport in Channels and Pumps”, *Science*, **310**, 1461-1465 (2005).

- [7] S.Y. Noskov and B. Roux, "Ion selectivity in potassium ion channels", *Biophysical Chemistry*, **124**(3), 279-291 (2006).
- [8] I.H. Shrivastava, D.P. Tieleman, P.C. Biggin, and M.S.P. Sansom, " K^+ versus Na^+ ions in a K channel selectivity filter: A simulation study", *Biophysical Journal*, **83**, 633-645 (2002).
- [9] S.J. Korn and J.G. Trapani, "Potassium Channels", *Transactions on Nanobioscience*, **4**, 21-33 (2005).
- [10] F. Hucho and C. Weise, "Ligand-gated ion channels", *Angew. chem. Int. Ed*, **40**, 3100-3116 (2001).
- [11] M.S.P. Sansom, I.H. Shrivastava, J.N. Bright, J. Tate, C.E. Capener, P.C. Biggin, "Potassium channels: structures, models, simulations", *Biochimica et Biophysica Acta*, **1565**, 294-307 (2002).
- [12] I.H. Shrivastava, C.E. Capener, L.R. Forrest, and M.S.P. Sansom, "Structure and dynamics of K channel pore-lining helices: a comparative simulation study", *Biophysical Journal*, **78**, 79-92 (2000).
- [13] C. Branden and J. Tooze, *Introduction to Protein Structure*, (Garland Publishing, Inc., NY, NY, 1999), 2nd ed.
- [14] <http://web.mit.edu/esgbio/www/lm/proteins/aa/aminoacids.html>
- [15] <http://www.science.siu.edu/microbiology/micr302/figure%207.07.jpg>

- [16] <http://sosnick.uchicago.edu/precpscstru.html>
- [17] <http://calcium.ion.ucl.ac.uk/3d-structure.html>
- [18] <http://hyperphysics.phy-astr.gsu.edu/hbase/biology/nervecell.html>
- [19] <http://faculty.washington.edu/chudler/ap.html>
- [20] D.A. Doyle, J.M. Cabral, R.A. Pfuetzner, A. Kuo, J.M. Gulbis, S.L. Cohen, B. T. Chait, and R. MacKinnon, "The structure of the potassium channel: molecular basis of K^+ conduction and selectivity", *Science* **280**, 69 (1998).
- [21] J.H. Morais-Cabral, Y. Zhou, and R. MacKinnon, "Energetic optimization of ion conduction rate by the K^+ selectivity filter", *Nature* **414**, 37 (2001).
- [22] Y. Zhou, J.H. Morais-Cabral, A. Kaufman, and R. MacKinnon, "Chemistry of ion coordination and hydration revealed by a K^+ channel-FAB complex at 2.0 Å resolution", *Nature* **414**, 43 (2001).
- [23] A. Kuo, J.M. Gulbis, J.F. Antcliff, T. Rahman, E.D. Lowe, J. Zimmer, J. Cuthbertson, F.M Ashcroft, T. Ezaki, and D.A. Doyle, "Crystal structure of the potassium channel KirBac1.1 in the closed state", *Science*, **300**, 1922-1926, (2003).

- [24] Y. Jiang, A. Lee, J. Chen, V. Ruta, M. Cadene, B.T. Chait, and R. MacKinnon, “X-ray structure of a voltage-dependent K^+ channel”, *Nature*, **423(6935)**, 33-41, (2003).
- [25] S.B. Long, E.B. Campbell, and R. Mackinnon, “Crystal structure of a mammalian voltage-dependent Shaker family K^+ channel”, *Science*, **309(5736)**, 867 (2005).
- [26] N. Shi, S. Ye, A. alam, L. Chen, and Y. Jiang, “Atomic structure of a Na^+ - and K^+ -conducting channel”, *Nature*, **440(7083)**, 427-429, (2006).
- [27] G. Chang, R.H. Spencer, A.T. Lee, M.T. Barclay, and D.C. Rees, “Structure of the MscL homolog from Mycobacterium tuberculosis: a gated mechanosensitive ion channel”, *Science*, **282(5397)**, 2220-2226, (1998).
- [28] , R.B. Bass, P. Strop, M. Barclay, and D.C. Rees, “Crystal structure of Escherichia coli MscS, a voltage-modulated and mechanosensitive channel”, *Science*, **298(5598)**, 1582-1587, (2002).
- [29] R. Dutzler, E.B. Campbell, M. Cadene, B.T. Chait, and R. MacKinnon, “X-ray structure of a ClC chloride channel at 3.0 Å reveals the molecular basis of anion selectivity”, *Nature*, **415(6869)**, 287-294, (2002).

- [30] R. Dutzler, E.B. Campbell, and R. MacKinnon, "Gating the selectivity filter in ClC chloride channels", *Science*, textbf300(5616), 108-112, (2003).
- [31] B.A. Wallace, "Common structural features in gramicidin and other ion channels", *Bioessays*, textbf22(3), 227-34, (2000).
- [32] D.J. Aidley and P.R. Stanfield, *Ion Channels: Molecules in Action*, (Cambridge University Press, Cambridge, 1996).
- [33] A.L. Huxley and A.F. Hodgkin, "Measurement of Current-Voltage Relations in the Membrane of the Giant Axon of Loligo", *Journal of Physiology*, **1**, 424-448 (1952).
- [34] A.L. Huxley and A.F. Hodgkin, "Currents Carried by Sodium and Potassium Ions Through the Membrane of the Giant Axon of Loligo", *Journal of Physiology*, **1**, 449-472 (1952).
- [35] A.L. Huxley and A.F. Hodgkin, "The Dual Effect of Membrane Potential on Sodium Conductance in the Giant Axon of Loligo", *Journal of Physiology*, **1**, 497-506 (1952).
- [36] A.L. Huxley and A.F. Hodgkin, "The Components of Membrane Conductance in the Giant Axon of Loligo", *Journal of Physiology*, **1**, 473-496 (1952).

- [37] A.L. Huxley and A.F. Hodgkin, “A Quantitative Description of Membrane Current and Its Application to Conduction and Excitation in Nerve”, *Journal of Physiology*, **1**, 500-544 (1952).
- [38] J.G. Nicholls, A.R. Martin, B.G. Wallace, and P.A. Fuchs. *From Neuron to Brain*, Sinauer Associates, Sunderland, MA, 2002.
- [39] Y. Pichon, L. Prime, P. Benquet, and F. Tiaho, “Some aspects of the physiological role of ion channels in the nervous system”, *Eur. Biophys. J.*, **33**, 211-226 (2004).
- [40] B. Dworakowska and K. Dolowy, “Ion channels - related diseases”, *Acta Biochimica Polonica*, **47(3)**, 685-793, (2000).
- [41] W.H. Miller, “The Semiclassical Initial Value Representation: A Potentially Practical Way for Adding Quantum Effects to Classical Molecular Dynamics Simulations”, *J. Phys. Chem. A*, **105(13)**, 2942-2955, (2001).
- [42] D. van der Spoel, E. Lindahl, B. Hess, A.R. van Buuren, E. Apol, P.J. Meulenhoff, D.P. Tieleman, A.L.T.M. Sijbers, K.A. Feenstra, R. van Drunen, and H.J.C. Berendsen, “Gromacs users manual version 3.3”, *www.gromacs.org*, (2005).
- [43] R.W. Hockney, S.P. Goel, and J. Eastwood, “Quiet high resolution computer models of a plasma”, *J. Comp. Phys.*, **14**, 148-158, (1974).

- [44] M.E. Tuckerman, and G.J. Martyna, "Understanding modern molecular dynamics: techniques and applications", *J. Phys. Chem. B*, **104**, 159-178, (2000).
- [45] B. Hess, H. Bekker, H.J.C. Berendsen, and J.G.E.M. Fraaije, "LINCS: A linear constraint solver for molecular simulations", *J. Comp. Chem.*, **18**, 1463-1472, (1997).
- [46] M.P. Allen and D.J. Tildesley, *Computer simulations of liquids*, (Oxford Science Publications, Oxford, UK, 2002).
- [47] H.L. Gordon and S. Goldman, "The glassy wall boundary for simulations of inhomogeneous systems", *Molec. Sim.*, **3**, 315-323, (1989).
- [48] W.F. van Gunsteren, H.J.C. Berendsen, and J.A.C. Rullmann, "Inclusion of reaction fields in molecular dynamics. Application to liquid water", *Faraday Discuss. Chem. Soc.*, **66**, 58-70, (1978).
- [49] P.P. Ewald, "Die berechnung optischer und elektrostatischer gitterpotentiale", *Ann. Phys.*, **64**, 253-287, (1921).
- [50] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An N-log(N) method for Ewald sum in large systems", *J. Chem. Phys.*, **98**, 10089-10092, (1993).

- [51] U. Essmann, L. Perera, M.L. Berkowitz, T. Darden, H. Lee, and L.G. Pedersen, “A smooth particle mesh Ewald potential”, *J. Chem. Phys.*, **103**, 8577-8592, (1995).
- [52] H.J.C. Berendsen, J.P.M. Postma, A. DiNola, and J.R. Haak, “Molecular dynamics with coupling to an external bath”, *J. Chem. Phys.*, **81**, 3684-3690, (1984).
- [53] R. Balescu, *Statistical dynamics: matter out of equilibrium*, (Imperial College Press, London, 1997).
- [54] R. Kubo, “The fluctuation-dissipation theorem”, *Rep. Prog. Phys.*, **29**, 255-284, (1966).
- [55] R. Zwanzig, “Time-correlation functions and transport coefficients in statistical mechanics”, *Ann. Rev. of Phys. Chem.*, **16**, 67-102, (1965).
- [56] M. Thornton, *Classical Dynamics*, (Saunders College Publishing, Toronto, ON, 1995), 1st ed.
- [57] R. Kubo, “Statistical-mechanical theory of irreversible processes. I.: General theory and simple applications to magnetic conduction problems”, *J. of the Phys. Soc. of Japan*, **12(6)**, 570-586, (1957).
- [58] L. Onsager, “Reciprocal Relations in Irreversible Processes. I.”, *Phys. Rev.*, **37**, 405-426 (1931).

- [59] L. Onsager, “Reciprocal Relations in Irreversible Processes. II.”, *Phys. Rev.*, **38**, 2265-2279 (1931).
- [60] H.J.C. Berendsen, D. van der Spoel, and R. van Drunen, “GROMACS: a message-passing parallel molecular dynamics implementation”, *Comp. Phys. Comm.* **91** 43-56, (1995).
- [61] E. Lindahl, B. Hess, and D. van der Spoel, “GROMACS 3.0: a package for molecular simulation and trajectory analysis”, *J. Mol. Mod.* **7** 306-317, (2001).
- [62] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren and J. Hermans, in B. Pullman (ed.), *Intermolecular Forces*, (Reidel, Dordrecht, 1981).
- [63] A. Einstein, “Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen”, *Ann. Phys., Lpz.* **17**, 549-560, (1905).
- [64] D.A. McQuarrie, *Statistical Mechanics*, (Harper and Row, New York, 1976).
- [65] M. LeMasurier, L. Heginbotham, and C. Miller, “KcsA: it’s a potassium channel”, *J. Gen. Physiol.*, **18**, 303-314, (2001).
- [66] M.L. Molina, F.N. Barrera, A.M. Fernández, J.A. Poveda, M.L. Renart, J.A. Encinar, G. Riquelme, and J.M. González-Ros, “Clustering

and coupled gating modulate the activity in KcsA, a potassium channel model”, *J. Biol. Chem.*, published online prior to print, (2006).

- [67] P.S. Crozier, R.L. Rowley, N.B. Holladay, D. Henderson, and D.D. Busath, “Molecular dynamics simulation of Continuous current flow through a model biological membrane channel”, *Phys. Rev. Lett.* **86**, 2467-2470 (2001).
- [68] Y. Yang, D. Henderson, and D. Busath, *J. Chem. Phys.*, “Applied-field molecular dynamics study of a model calcium channel selectivity filter”, **118**, 4213 (2003).
- [69] S. Berneche and B. Roux, “A microscopic view of ion conduction through the K⁺ channel”, *PNAS* **100**, 8644-8648 (2003).
- [70] S. Kuyucak, O.S. Andersen, and S.H. Chung, “Models of permeation in ion channels”, *Rep. Prog. Phys.* **64**, 1427 (2001).
- [71] F. Bezanilla, “Voltage-Gated Ion Channels”, *IEEE Transactions on Nanobioscience*, **4**(1), 34-48, (2005).
- [72] C. Miller, “An overview of the potassium channel family”, *Genome Bio.*, **1**(4), reviews0004.10004.5, (2000).
- [73] W. Humphrey, A. Dalke, and K. Schulten, “VMD: visual molecular dynamics”, *J. Molec. Graphics* **14**, 33-38 (1996).

- [74] C. Miller, "Coupling of water and ion fluxes in a K^+ -selective channel of sarcoplasmic reticulum", *Biophys. J.*, **38**, 227-230, (1982).
- [75] S.M. Saparov and P. Pohl, "Beyond the diffusion limit: Water flow through the empty bacterial potassium channel", *PNAS*, **101**(14), 4805-4809, (2004).
- [76] H. Ando, M. Kuno, H. Shimizu, I. Muramatsu, and S. Oiki, "Coupled K⁺ Water Flux through the HERG Potassium Channel Measured by an Osmotic Pulse Method", *J. Gen. Physiol.*, **126**(5), 529-538, (2005).
- [77] S. Berneche and B. Roux, "Energetics of ion conduction through the K^+ channel", *Nature* **414**, 73-77, (2001).
- [78] S. Berneche and B. Roux, "The ionization state and the conformation of Glu-71 in the KcsA K^+ channel", *Biophys. J.*, **82**, 772-780, (2002).
- [79] K.M. Ranatunga, I.H. Shrivastava, G.R. Smith, and M.S.P. Sansom, "Side-chain ionization states in a potassium channel", *Biophys. J.*, **80**, 1210-1219, (2001).
- [80] H. Choi and L. Heginbotham, "Functional influence of the pore helix glutamate in the KcsA K^+ channel", *Biophys. J.*, **86**, 2137-2144, (2004).
- [81] L.G. Cuello, J. G. Romero, D. M. Cortes, and E. Perozo, "pH-dependent gating in the *Streptomyces lividans* K^+ channel", *Biochemistry*, **37**, 3229-3236, (1998).

- [82] E. Perozo, D. M. Cortes, and L. G. Cuello, "Three-dimensional architecture and gating mechanism of a K^+ channel studied by epr spectroscopy", *Nat. Struct. Biol.*, **5**, 459-469, (1998).
- [83] E. Perozo, D. M. Cortes, and L. G. Cuello, "Structural rearrangements underlying K^+ -channel activation gating", *Science*, **285**, 73-78, (1999).
- [84] P. Sompornpisut, Y.S. Liu, and E. Perozo, "Calculation of Rigid-Body Conformational Changes Using Restraint-Driven Cartesian Transformations", *Biophys. J.*, **81**, 2530-2546, (2001).
- [85] S. Haider, A. Grottesi, B.A. Hall, F.M. Ashcroft, and M.S.P. Sansom, "Conformational Dynamics of the Ligand-Binding Domain of Inward Rectifier K Channels as Revealed by Molecular Dynamics Simulations: Toward an Understanding of Kir Channel Gating", *Biophys. J.*, **88**, 3310-3320, (2005).
- [86] V. Jogini and B. Roux, "Electrostatics of the Intracellular Vestibule of K^+ Channels", *J. Mol. Biol.*, **354**, 272288, (2005).
- [87] H.W. de Haan, I.S. Tolokh, C.G. Gray, and S. Goldman, "Nonequilibrium molecular dynamics calculation of the conductance of the KcsA potassium ion channel", *Phys. Rev. E.*, **74**, 030905-1-030905-4, (2006).

- [88] T.W. Allen, O.S. Andersen, and B. Roux, "On the importance of atomic fluctuations, protein flexibility, and solvent in ion permeation", *J. Gen. Physiol.*, **124**, 679-690, (2004).
- [89] T.W. Allen, M. Hoyles, S. Kuyucak, and S.H. Chung, "Molecular and Brownian dynamics study of ion selectivity and conductivity in the potassium channel", *Chem. Phys. Lett.*, **313**, 358-365, (1999).
- [90] I.H. Shrivastava, D.P. Tieleman, P.C. Biggin, and M.S.P. Sansom, "K⁺ versus Na⁺ ions in a K channel selectivity filter: a simulation study", *Biophys. J.*, **83**, 633-645, (2002).
- [91] D. Asthagiri and Lawrence R. Pratt, "Role of fluctuations in a snug-fit mechanism of KcsA channel selectivity", *J. Chem. Phys.*, **125**, 024701-1-024701-6, (2006).
- [92] George Eisenman and Richard Horn, "Ionic Selectivity Revisited: The Role of Kinetic and Equilibrium Processes in Ion Permeation Through Channels", *J. Membrane Biol.*, **76**, 197-225, (1983).
- [93] V.B. Luzhkov and J. Åqvist, "K⁺/Na⁺ selectivity of the KcsA potassium channel from microscopic free energy perturbation calculations", *Biochimica et Biophysica Acta*, **1548**, 194-202, (2001).
- [94] S. Berneche and B. Roux, "A gate in the selectivity filter of potassium channels", *Structure*, **13**, 591-600, (2005).

- [95] B. Corry, T. Vorab, and S.H. Chung, "Electrostatic basis of valence selectivity in cationic channels", *Biochimica et Biophysica Acta*, **1711**, 72-86, (2005).
- [96] L. Kiss, D. Immke, J. LoTurco, and S.J. Korn, "The Interaction of Na^+ and K^+ in Voltage-gated Potassium Channels: Evidence for Cation Binding Sites of Different Affinity", *J. Gen. Physiol.*, **111**, 195-206, (1998).
- [97] N. Shin, H. Soh, S. Chang, D.H. Kim, and C. Park, "Sodium Permeability of a Cloned Small-Conductance Calcium-Activated Potassium Channel", *Biophys. J.*, **89**, 3111-3119, (2005).
- [98] F.I. Valiyaveetil, M. Sekedat, R. Mackinnon, and T.W. Muir, "Structural and Functional Consequences of an Amide-to-Ester Substitution in the Selectivity Filter of a Potassium Channel", *J. Am. Chem. Soc.*, **128**, 11591-11599, (2006).
- [99] B.T. Thole, "Molecular polarizabilities with a modified dipole interaction", *Chem. Phys.*, **59**, 341345, (1981).
- [100] Karen A. Duca and Peter C. Jordan, "Comparison of Selectively Polarizable Force Fields for Ion-Water-Peptide Interactions: Ion Translocation in a Gramicidin-like Channel", *J. Phys. Chem. B*, **102**, 9127-9138, (1998).

- [101] G. Lamoureux and B. Roux, “Modeling induced polarization with classical drude oscillators: Theory and molecular dynamics simulation algorithm”, *J. Phys. Chem. A.*, **119**, 30253039, (2004).
- [102] G. Lamoureux, A.D. MacKerell, and B. Roux, “A simple polarizable model of water based on classical drude oscillators”, *J. Phys. Chem. A.*, **119**, 51855197, (2004).
- [103] M. Swart and P.Th. van Duijnen, “DRF90: a polarizable force field”, *Mol. Simul.*, **32(6)**, 471-484, (2006).
- [104] S.Y. Noskov, G. Lamoureux, and B. Roux, “Molecular dynamics study of hydration in ethanol-water mixtures using a polarizable force field”, *J. Phys. Chem. B*, **109(14)**, 6705-13, (2005).
- [105] S. Bernèche and B. Roux, “On the Potential Functions used in Molecular Dynamics Simulations of Ion Channels”, *Biophys. J.*, **82(3)**, 1677-1684, (2002).
- [106] I.S. Tolokh, I.I. Tolokh, H.C. Cho, N. D’Avanzo, P.H. Backx, S. Goldman and C.G. Gray, “Non-Michaelis-Menten kinetics model for conductance of low-conductance potassium ion channels”, *Phys. Rev. E*, **71**, 021906-1-021906-9.

- [107] I.S. Tolokh, S. Goldman and C.G. Gray, “Unified modeling of conductance kinetics for low- and high-conductance potassium ion channels”, *Phys. Rev. E*, **74**, 011902-1-011902-12 (2006).