A Comprehensive Dynamic Model for KCNQ1/KCNE1 Ion Channel: Structural & Functional Studies

by

Horia Jalily Hasani

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Pharmaceutical Sciences

Faculty of Pharmacy and Pharmaceutical Sciences
University of Alberta

ABSTRACT

The voltage-gated KCNQ1/KCNE1 potassium ion channel plays a key role in maintaining the heart rhythm. An active channel generates the slow delayed rectifier (I_{Ks}) current in the heart. Both loss-of-function and gain-of-function mutations in KCNQ1 or KCNE1 are linked to many heart-related diseases, including long QT syndromes, congenital atrial fibrillation, and short QT syndrome. On the other hand, the KCNQ1/KCNE1 channel is also found to be an off-target for many non-cardiovascular drugs, leading to fatal cardiac irregularities. This Thesis aims at understanding the structure and function of the KCNQ1/KCNE1 ion channel at the atomistic level.

To accomplish this task, we used several state-of-the-art molecular modeling approaches to build a structural model for KCNQ1 protein. This model was tested against available experimental data which ensured its accuracy. Following that, we included the KCNE1 protein component using a data-driven protein-protein docking simulation. The presence of KCNE1 associated with KCNQ1 produced profound effects on the dynamics of the channel. More importantly, through potassium ion permeation studies using Steered Molecular Dynamics, it was found that KCNE1 causes changes in the topology of the pore. This change translates into a slight blockade of ion permeation and is in agreement with the known effect of KCNE1, which is a slowing of activation. To our knowledge, this Thesis represents the first study to highlight the effect of the KCNE1 protein on the structure of the KCNQ1 pore domain as well as on ion permeation.

Next, we docked a panel of compounds consisting of Chromanol 293B and its 8 derivatives, in the KCNQ1/KCNE1 channel. The small molecule docking simulations provided us with a predicted binding affinity score which correlated well with the

experimental activity of the compounds. This indicated that our model has been capable of discriminating between blockers of differential activity. Furthermore, Steered Molecular Dynamic simulations were performed for the ligand bound channel complexes. Through this study we were able to directly investigate the effect of the blockers on ion permeation. Such that strong blockers had a profound effect on the passage of potassium ion which was evident from their binding mode and interactions with the binding site residues, force profiles as well as pore topology analysis. The weak blockers on the other hand, did not have direct interference with the normal passage of ions. The structure activity relationships of the ligands revealed the pharmacophoric features responsible for the degree of their effect on the channel and allowed us to get more insights into how small molecule blockers can affect the orientation of specific residues in the protein.

Overall, the findings from this Thesis are important and investigated novel aspects of the KCNQ1/KCNE1 channel complex. We believe this Thesis marks a good starting point for further studies to investigate different drug scaffolds and different mechanisms by which they can affect ion permeation in KCNQ1/KCNE1 ion channels. Eventually, allowing the identification of potential cadiotoxicity of certain drug molecules early in the drug development stages, and therefore to prevent the unfortunate consequences of designing a drug molecule which is toxic to the heart.

PREFACE

A version of Chapter 2 has been submitted to The Journal of Molecular Graphics and Modelling as *Jalily Hasani H, Ahmed M, Barakat K. A Comprehensive Structural Model for the Human KCNQ1/KCNE1 Ion Channel.*

A version of Chapter 3 has been submitted to The Journal of Molecular Modeling as *Jalily Hasani H, Ganesan A, Ahmed M, Barakat K. Ion Permeation Studies in the Human KCNQ1/KCNE1 Ion Channel.*

Appendix B of this thesis has been published as *Jalily Hasani H, Barakat K. Homology Modeling: an Overview of Fundamentals and Tools. Int Rev Model Simulations (IREMOS); Vol 10, No 2. 201.*

A version of Appendix C has been published as *Jalily Hasani H, Barakat K. Protein-Protein Docking: Are We There Yet?, Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery. Dastmalchi S, Hamzeh-Mivehroud M, Sokouti B, editors. IGI Global; 173-195 p. 2016.*

I dedicate this thesis to my parents Ozra and Mostafa, and my beloved brother Pouria, who have given me the courage and the support I needed to pursue my dreams and, to make the impossible "possible".

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor, Dr. Khaled Barakat for his generous support and encouragement throughout the course of my degree. He was always supportive of my ideas and his expertise, understanding and patience added considerably to my graduate experience. I also wish to thank my co-supervisor Dr. Ayman El-Kadi and committee members; Dr. Arno Siraki and Dr. Carlos A. Velázquez-Martínez.

A very special thanks goes out to Dr. Marawan Ahmed for the assistance he provided at all levels of my research project. I couldn't have hoped for a better mentor! I also acknowledge my other lab mates Dr. Aravindhan Ganesan, Dr. Subha Kalyaanamoorthy, Fitch Feng and Nawreen Hena.

I am short of words to show my gratitude to my dearest parents Ozra and Mostafa. They were the ones who helped me get through the toughest moments during this time and I surely couldn't have made it this far without them. My biggest thanks go to my brother Pouria, without whose motivation and encouragement, I would not have considered a graduate career in pharmaceutical research. I doubt that I will ever be able to convey my appreciation fully, but I owe him my eternal gratitude. I'd like to express my gratitude and love to Amir, whose companionship kept me going through hard days.

And lastly, I recognize that this research was supported by the financial assistance from the University of Alberta, the Natural Sciences and Engineering Research Council of Canada (NSERC), as well as support from the Li Ka Shing Applied Virology Institute (LKSAVI).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APD	Action Potential Duration
CaM	Calmodulin
Ca_V	Voltage-Gated Calcium Ion Channel
CT	Carboxy Terminal
I_{Ks}	Slow Delayed Rectifier potassium current
JLNS	Jervell and Lange-Nielsen syndrome
K_{ir}	Inward Rectifier Potassium Ion Channel
K_V	Voltage-Gated Potassium Ion Channel
LGIC	Ligand-Gated Ion Channel
LQTS	Long QT Syndrome
MD	Molecular Dynamics
Na_V	Voltage-Gated Sodium Ion Channel
NMR	Nuclear Magnetic Resonance
NT	Amino Terminal
PD	Pore Domain
PDB	Protein Data Bank
PIP_2	Phosphatidylinositol 4,5-bisphosphate
POPC	Phosphatidylcholine
REMD	Replica Exchange Molecular Dynamics
RMSD	Relative Mean Square Deviation
RMSF	Relative Mean Square Fluctuations
SMD	Steered Molecular Dynamics
SQTS	Short QT Syndrome
Tdp	Torsade de Points
TMD	Transmembrane Domain
VGIC	Voltage-Gated Ion Channel
VSD	Voltage-Sensing Domain

CHAPTER 1: INTRODUCTION

The heart; this muscular beating organ is the source of our life. From the first beat on our third week to the last moment of our lives, its continuous and unresting movement is controlled by the passage of ions. Any disturbance or interruption to this precise ionic flow may damage this machine or stop it from beating (Magder, 1998). The flow of ions in the heart is controlled by a set of proteins, called cardiac ion channels. They allow ions to travel back and forth across the cardio cellular membrane lipid bilayers (Grant, 2009). Despite the accumulated knowledge in the literature, the details of these mechanisms have been a mystery for decades.

The three main types of ion channels that aid in maintaining the heart rhythm are sodium, calcium and potassium channels (Roden, Balser, George, & Anderson, 2002). A change in the membrane voltage (potential) controls when and which cardiac channel should work. It also controls the opening and closing of these channels. This in turn can change their action potentials, which are the elementary units of electrical signals in biology. These dynamic action potentials are further translated into contractions and relaxations of the heart muscle (Silva et al., 2009). The normal functioning of ion channels involves numerous intricate details that have been the subject of research for many years. The important role of ion channels in the normal functioning of the heart is further emphasized by the number of diseases that result from mutations within these proteins (Dworakowska & Dolowy, 2000). This makes it crucial to completely understand their working mechanisms.

Thusly motivated, this thesis will focus on this complicated process, taking one of these ion channels as an example, namely the $K_V7.1$ potassium channel. We have made use of computational modelling as a tool to understand the structure, function and dynamics of this channel.

The $K_V7.1$ channel is a voltage-gated ion channel, meaning that it is directly linked to the change of the membrane potential. It is also called the KCNQ1 ion channel and generates the slow activating delayed rectifier (I_{KS}) potassium current, which has a significant role in controlling repolarization phase of cardiac action potential, its duration and eventually the normal pace of the heart (Geoffrey W. Abbott, 2014). Furthermore, mutations in cardiac KCNQ1 ion channel represent a leading cause of heritable long QT

syndrome (LQTS), Short QT syndrome (SQTS), Jervell and Lange-Nielsen syndrome (JLNS) and atrial fibrillations associated with life-threatening arrhythmias, deliquium seizures and sudden death (Modell & Lehmann, 2006; Mousavi Nik, Gharaie, & Jeong Kim, 2015; Tester & Ackerman, 2014). Altogether, the malfunction of this ion channel is a leading cause of cardiac abnormalities making it crucial to accurately study its function and develop drugs for the treatment of ion channel induced diseases.

The primary goal of this Thesis is to develop a structural model for the KCNQ1 ion channel in its active form as well as to understand the ion conduction mechanism and modes of drug binding. Our literature review identified many limitations and gaps in previous research, which we aimed at solving in the current model. With that in mind, the model for KCNQ1 ion channel was developed with extensive inclusion of available experimental information. Further, we modeled the interaction of this ion channel with its auxiliary subunit KCNE1 protein and other cellular components. This dissertation should assist in overcoming some of the controversies in the field and provide additional pieces of information to the current dilemmatic puzzle of this biological assembly.

Our methodologies combined several cutting-edge modeling techniques together. These included homology modeling to develop the initial KCNQ1 protein structure, classical as well as advanced molecular dynamic (MD) simulation techniques to capture the dynamicity of the proteins under study, protein-protein docking simulations to assemble KCNQ1 and KCNE1 together, ligand-docking to obtain drug-bound structures and ion permeation studies.

This Thesis is organized as follows. Chapter 1 is divided into three main parts. In part 1, background information about ion channels, with an emphasis on cardiac ion channels in particular is provided. Part 2 focuses on the KCNQ1 ion channel, describing its structural properties and its role in controlling the action potential and its regulation by other cellular components. Part 3 includes a brief literature review on the current state of research in this area. This is then followed by a detailed description of the research objectives and questions of this Thesis. Chapter 2 presents our efforts in developing a comprehensive homology model for the KCNQ1 protein and the incorporation of the KCNE1 accessory protein via state-of-the-art computational techniques. Chapter 3 includes the essential application and studies using our validated KCNQ1/KCNE1 model.

The reported findings in this Chapter are novel information related to the mechanisms of ion permeation through the channel pore with and without the presence of the accessory KCNE1 proteins. This is followed by a presentation of the data generated through the testing of KCNQ1/KCNE1 small molecule blockers as well as their effect on the ion permeation through the channel. The overall findings and the links between the different chapters are provided in Chapter 4. And finally, Chapter 5 includes the implications for future studies. Appendix A of the Thesis encompasses a detailed review of Molecular Dynamic (MD) simulations and its enhanced variants which are also employed in the studies. Appendix B presents the manuscript titled: "Homology Modeling: An Overview of Fundamentals and Tools", describing one of the main techniques that has been employed in this Thesis. Appendix C describes another main modeling technique used in this study, namely protein-protein docking in the form of a publication; "Protein-Protein Docking: Are We There Yet?". Appendix D, contains the supplementary data from Chapter 2.

1.1: ION CHANNELS

1.1.1. Introduction

All cells are surrounded by membrane bilayers. These protective barriers, not only separate the cells interior from extracellular fluid, but also significantly block the transfer of organic molecules and ions from and into the cells. Although these entities are critical for many biological processes, their uncontrolled transfer across the cells can lead to significant adverse and, in many cases, deleterious effects. For ions' passage, ion channels are the membrane spanning gates to enter or exit the cell. They allow the flux of ions across the cell membrane through opening and closing of nanoscopic pores at the center of their structure (Ashcroft, 2000). Ion channels are superior to other types of cellular carrier proteins due to their extreme selectivity and their exclusive capability of transporting up to 100 million ions per second. This rate is 100,000 times greater than the fastest known carrier proteins in the cell (Aidley & Stanfield, 1996).

By regulating a precise ion concentration, ion channels can control the electrostatic environment of the cells. Electrical signals, as will be discussed later in this

chapter, are essential for many cells and eventually for a whole body organ such as the heart, muscles and brain to survive and function normally (Yu & Catterall, 2004). At the cellular level, these tissues rely highly on ion channels to maintain a particular ion concentration at the right time and in the right amount. Taking the heart as an example, it requires a harmonic coordination of ionic flux for contractility and, thus, supplying blood to all other parts of the body (Grant, 2009).

In 1952, Hodgkin and Huxley made the first step to understand the complex mechanisms of ion channels in the plasma membrane. They developed a conductance-based theoretical model to describe how action potentials are produced (Hodgkin & Huxley, 1990). Today and after almost six decades, hundreds of ion channel genes have been cloned and many of them have been expressed and purified. Furthermore, the currents conducted by many of these channels have been clearly characterized. Given this accumulated knowledge, one can associate particular ion channels with certain physiological functions or human diseases (Dworakowska & Dolowy, 2000). While this represents an astonishing progress, yet there is much more to be revealed about the function and structure of these complex macromolecules, given the complexity in determining their structures experimentally.

This section serves as the first part of the introduction Chapter and begins with a general classification of the different ion channels and describes their main types and subtypes. This is followed by a brief description of the ion channels distribution throughout the human body. The chapter then focuses on cardiac ion channels; cardiac action potential, the types of cardiac currents and their associated ion channels in the heart. This chapter lays the background for the next chapters.

1.1.2. Classification of Ion Channels

Conventionally, the ion channel superfamily is broadly classified into: 1) Voltage-gated ion channels (VGICs) and 2) Ligand-gated ion channels (LGICs). This classification solely relies on the way these proteins are stimulated to perform their function. VGICs respond to changes in the electrical properties of their environment, while LGICs are stimulated by their interaction with selective intrinsic ligands. Figure 1.1.1, illustrates these two main types of ion channels and their different mechanisms of activation.

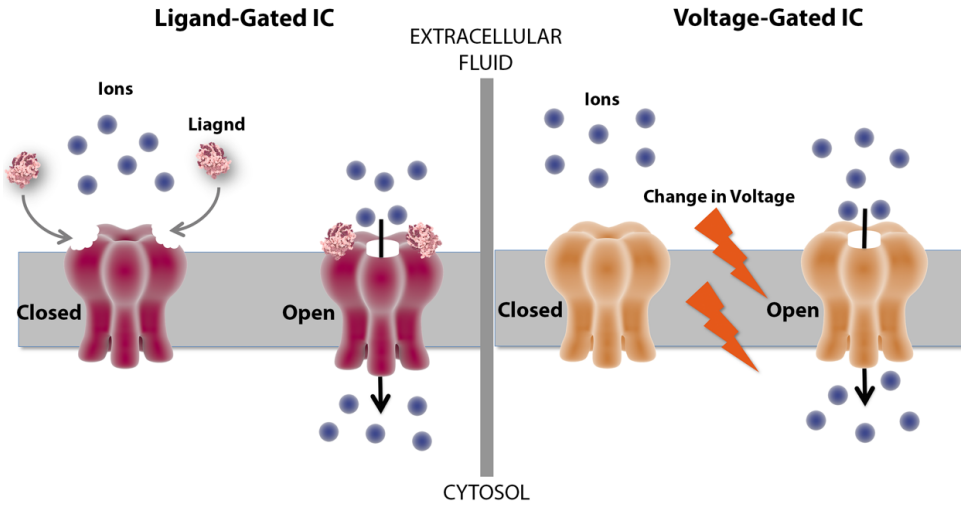


Figure 1.1.1. The activation mechanisms of ligand-gated (left) and voltage-gated (right) ion channels. Ligand-gated ion channels are transformed from a closed state to an open state upon binding of a ligand to specific binding sites. Voltage-gated ion channels, on the other hand, are activated by changes that take place in the membrane potential. Upon activation, both types of ion channels open their pore for ions to pass from the extracellular space to the cytosol or vice versa.

1.1.2.1. Ligand-gated ion channels

The state transitions of ligand-gated ion channels are driven by their specific interactions with particular ligands (see Figure 1.1.1) (Hucho & Weise, 2001). These multi-functional proteins are essential for synaptic transmission and other forms of cellular signalling and communications. Their main exceptional characteristic is their ability to convert intracellular chemical signals into electrical information that can spread over to other cells. LGICs are less selective than VGICs and can allow the passage of two or even more types of ions at the same time (Lemoine et al., 2012) .

The most important ligand-gated ion channels are those activated by neurotransmitters in the nervous system. Others are activated by circulating hormones and locally released substances. Nicotinic acetyl choline receptors (nAchR), GABA-A glycine and 5-HT3 serotonin receptors are all classical examples of LGICs that form a family termed as the Cys loop channels (Galligan, 2002). These proteins are composed of three main regions (see Figure 1.1.2 for an example): a) an extracellular ligand binding domain, b) a transmembrane cylinder, forming the pore, and c) an intracellular region

responsible for trafficking, localization and regulation by secondary messengers (Zagotta, 2015).

LGICs are pentameric assemblies each possessing 350-500 amino acids in length, containing four transmembrane alpha helices. Most of them contain a signature sequence of 13 amino acids in their extracellular N-terminal region flanked by cysteine residues, and so the name Cys loop family (Cascio, 2004). The other families of LGICs are glutamate-gated receptors (iGluRs) and ATP-gated P2X receptors, which are structurally different from the Cys loop channels family. iGluRs are usually tetramers formed by dimers of dimers whereas the P2X channels are homo or hetero tetramers (Hucho & Weise, 2001). Since cardiac ion channels, which are the focus of this thesis, are voltage dependant in nature, the rest of this section will be devoted mainly to VGICs and describe the three main VGICs sub-types, namely sodium, calcium and potassium ion channels.

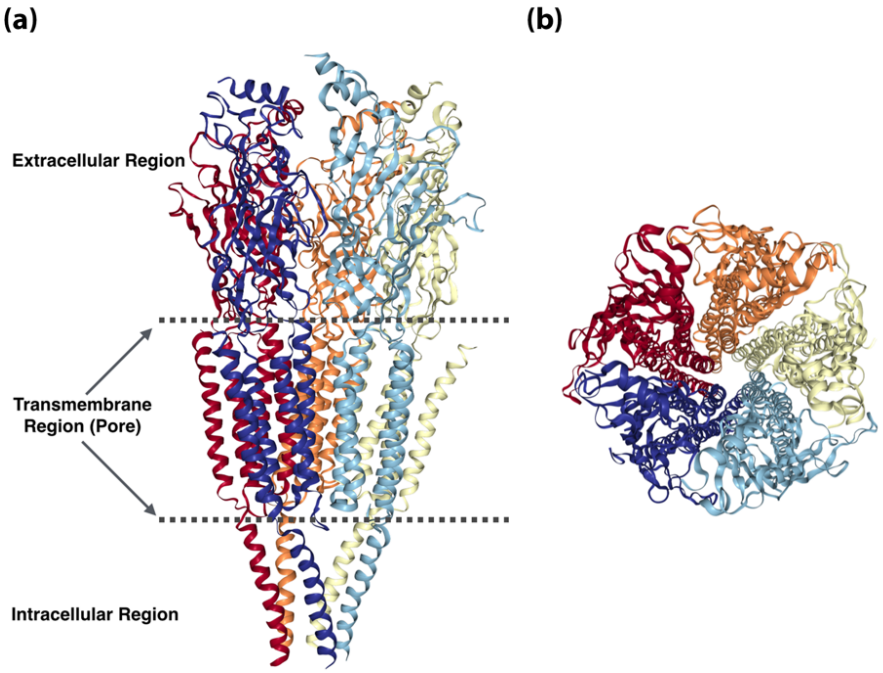


Figure 1.1.2. The 3D structure of a nicotinic acetylcholine receptor (PDB ID: 2BG9). (a) Lateral view, showing the three main regions: extracellular, transmembrane and intracellular parts. (b) Top view, showing the five chains forming a pentameric channel with the pore in the centre (the structure is coloured by chain).

1.1.2.2. Voltage-gated Ion Channels

Voltage-gated ion channels are membrane proteins that catalyze the transport of ions across the membrane when triggered by changes in membrane potential. The variations in the membrane voltage induce significant conformational changes in the channel (S. Chowdhury & Chanda, 2015). When in the "closed state", channel pores are impermeable to ions. Membrane depolarization induces a conformational change in the ion channel, which causes the channel to change to an "open conformation", allowing the ions to move in or out of the cell (G Yellen, 1998). Some VGICs such as sodium channels possess a third conformation known as the "inactivation state" which is a non-conducting state (see Figure 1.1.3), refractory to the successive activation (Ahern, Payandeh, Bosmans, & Chanda, 2016).

The main types of voltage gated ion channels in the human body are: Na^{+}, $K^{+,}$ Cl^{-} and Ca^{2+} ion channels. They are distributed in different tissues and organs, mainly the muscles, brain, heart and nervous system (Bezanilla, 2007). Voltage-dependent ion channels are responsible for the generation of an action potential, which is the elementary unit of electrical signals in biology. Action potentials eventually translate into different functions in the body, e.g. muscle contraction & relaxation, neuronal transmission, hormone secretion, cell division and gene expression (Bean, 2007; Rudy, 2008).

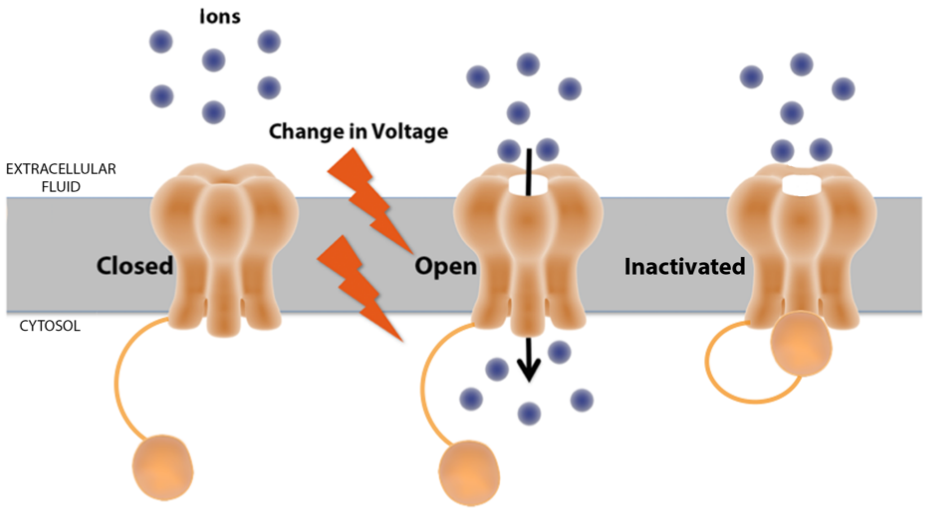


Figure 1.1.3. A voltage-gated ion channel shown in the three states: closed, open, and inactivated. The channel opens in response to changes in membrane voltage, allowing ions to pass. After activation, they become inactivated for a brief period and do not open in response to a signal.

Voltage-dependent ion channels share several common recognizable features. Figure 1.1.4 displays a general schematic of voltage-gated ion channels with their common features. Structurally, they all possess an integral membrane part known as the voltage-sensing domain (VSD). More or less, the VSDs are all composed of a few alpha subunits (S. Chowdhury & Chanda, 2015), one or more pore domains (PD) and occasionally associated with an auxiliary beta subunit (Isom, De Jongh, & Catterall, 1994). The alpha subunit segments are named as S1, S2, S3 and so forth depending on the number of segments in the respective protein. Generally, S1-S4 segments act as the voltage sensing components of the ion channels. The S4 segment is unique in its composition, such that it has an unusual abundance of positively charged residues that play a significant role as voltage sensors. These positive charges lead to the sensitive movement of S4 as a result of any changes in the charge of the membrane on the inner and/or outer surface. This movement is then translated into conformational changes, which initiate the state transitions. The S5 and S6 segments along with a pore-loop (P-loop) from all domains join together and form the central PD of the channel (Tien, Young, Jan, & Jan, 2014).

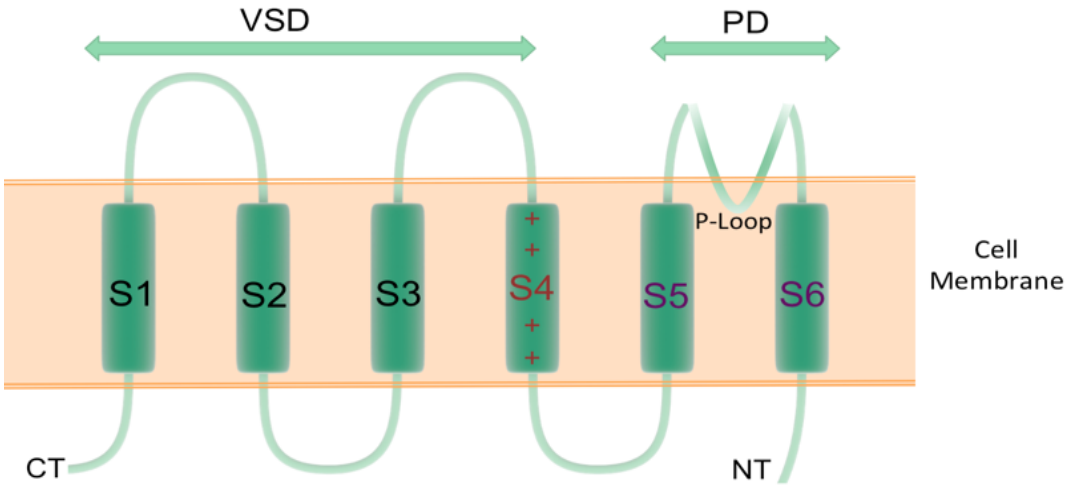


Figure 1.1.4. General Topology of a voltage-gated ion channel. A VGIC is composed of several segments (S1 to S6) embedded in the cell membrane. The voltage-sensing domain (VSD) is composed of S1 to S4, the pore domain (PD) consists of S5, S6 and the P-loop. There are loops that join the segments together, as well as N-terminal (NT) and C-terminal (CT) tails.

One of the main properties of all ion channels is their selective permeability to a certain type of ion over the other types. That is, each ion channel opens to a particular type of ion and is entirely or partially impermeable to the others. This property is mainly attributed to a conserved sequence of amino acids located in the pore of the channel, called the selectivity filter. In addition to the structure of the selectivity filter, this property is also controlled by other factors including size, valency and hydration energy. The latter factor is extensively discussed in (Ashcroft, 2000; D. Kim, McCoy, & Nimigean, 2015; Roux et al., 2011).

Needless to say, various ion channels differ in their structural composition from one class to another or even within the same class. These variations are highlighted in the following sections, describing the three main classes of VGICs, i.e. voltage-gated sodium, calcium and potassium ion channels. As this thesis is focused on a cationic (potassium) ion channel, the anionic chloride channels (CLCs), which are structurally and biophysically distinct from the other cationic ion channels will not be discussed here. However, interested readers are referred to (Accardi & Pusch, 2000; Ashcroft & Ashcroft, 2000a; Thomas J Jentsch, Stein, Weinreich, & Zdebik, 2002; Pusch, 2007) for detailed information on CLCs.

i) Voltage-Gated Sodium Channels

Voltage-gated sodium channels' family, known as the Na_v ion channels, comprises of nine alpha-subunit SCNnA genes (See Table 1.1.1). The first five members, SCN1A-SCN5A encode $Na_v1.1$-$Na_v1.5$ and SCN8A-SCN11A encode $Na_v1.6$-$Na_v1.9$. The sodium ion channels trigger the initiation and propagation of action potentials in excitable cells in different cell types (Ashcroft & Ashcroft, 2000d). The sites of expression and localization of Na_v ion channels range from brain neurons (CNS, PNS and DRG), to skeletal and heart muscles. Sodium channel subtypes are differentially expressed and localized within individual cells, in addition to their differential expression in different cell types (WilliamA Catterall, 2015). $Na_v1.5$ ion channel is specifically expressed in the cardiac muscle tissue, whereas $Na_v1.4$ is present in the skeletal muscles. The other seven members ($Na_v1.1$, $Na_v1.2$, $Na_v1.3$, $Na_v1.6$, $Na_v1.7$, $Na_v1.8$, $Na_v1.9$) are mainly localized in the brain and in the nervous system (WilliamA Catterall, 2015).

The alpha-subunit of sodium ion channels is a long polypeptide composed of 1,700-2,000 amino acids. Structurally, the protein is divided into four homologous domains (domains I-IV), connected by cytosolic loops (L1-L3), which are less conserved in their sequence compared to the transmembrane regions (see Figure 1.1.5). Each domain comprises of six transmembrane segments (S1-S6) connected by several cytosolic and extracellular linkers (Dib-Hajj & Priestley, 2010). The alpha subunit of sodium channels is sufficient to produce a functional protein. However, when bound to a beta subunit and/or to other partners, several of the channel's functions and properties are modified and regulated. These include channel stability, density of the current produced, phosphorylation and glycosylation (Brackenbury & Isom, 2011).

Table 1.1.1. Voltage-gated sodium channel family. Source: (Abriel, 2007; William a Catterall, 2012; Lehmann-Horn & Jurkat-Rott, 1999; Trezise, Dale, & Main, 2010)

CHANNEL	GENE	PRIMARY LOCATION	DISEASE
$Na_v1.1$	SCN1A	CNS	• Generalized epilepsy with febrile seizures • Dravet syndrome (Severe myoclonic epilepsy in infancy) • Benign neonatal convulsions • Familial hemiplegic migraine type III
$Na_v1.2$	SCN2A	CNS	• Benign familial neonatal-infantile seizures
$Na_v1.3$	SCN3A	Embryonic CNS	
$Na_v1.4$	SCN4A	Skeletal Muscle	• Hypokalemic periodic paralysis type II • Normokalemic periodic paralysis • Hyperkalemic periodic paralysis • Paramyotonia congenta
$Na_v1.5$	SCN5A	Heart	• Long QT syndrome type III • Brugada syndrome
$Na_v1.6$	SCN7A	CNS	•
$Na_v1.7$	SCN9A	PNS	• Erythromelalgia • Paroxysmal extreme pain disorder • Congenital indifference to pain
$Na_v1.8$	SCN10A	DRG	• Small fibre neuropathy
$Na_v1.9$	SCN11A	DRG	

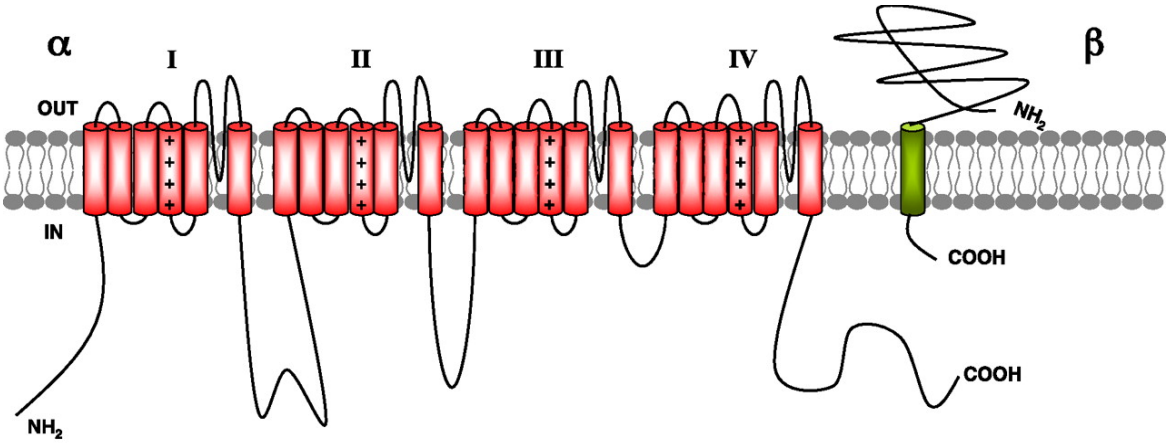


Figure 1.1.5. Schematic representation of the α and β-subunits of a complete Na_v ion channel structure. The figure shows the four domains (Domain I – Domain IV), with each domain including a voltage-sensing domain (VSD) and a pore domain (PD). The positively charged S4 helical segments of the VSDs, which interact with the respective PD, are shown as green helices and the location of inactivation gate (IFMT) is also marked. Source: (Vacher, Mohapatra, & Trimmer, 2008)

Mutations in sodium channels have been associated with a variety of neural, muscular and cardiac disorders such as periodic paralysis, cardiac arrhythmia, epilepsy, and chronic pain (see Table 1.1.1 for more details). Moreover, these proteins are targeted by a wide array of drugs including local anaesthetics, anti-arrhythmics, analgesics and anticonvulsants (Bagal, Marron, Owen, Storer, & Swain, 2015). Over the past few decades, Na_v ion channels have opened doors to extensive research and investigation of sodium selectivity and ion conductance, voltage-dependent activation, fast inactivation and drug block. Recent and elaborated reviews on sodium ion channels are provided in (Ahern et al., 2016; William a Catterall, 2012).

ii) Voltage-Gated Calcium Channels

Ca^{2+} ions regulate a wide range of functions such as cardiac action potential propagation, neurotransmission and hormone release, skeletal muscle contraction, and calcium-dependent gene transcription (Minor & Findeisen, 2010; Simms & Zamponi, 2014). The activation of voltage-gated calcium channels causes a transitory rise in the intracellular Ca^{2+} concentrations and is mainly stimulated by membrane depolarization (Clapham, 2007). Voltage-gated calcium ion channels (also called as Ca_v ion channels) are subdivided into two categories; 1) low-voltage activated channels (LVA), 2) high-voltage

activated (HVA) channels. The former type of Ca_v ion channels is activated by modest membrane depolarizations and rapidly inactivated. The latter class are the ion channels that require larger membrane depolarizations to transform into their activated or open state (Petegem & Minor, 2006).

Another classification of Ca_v ion channels is based on their subunit gene subfamily. In this context, three subfamilies have been identified so far, namely Ca_v1, Ca_v2, and Ca_v3, with several members for each subtype. The Ca_v1 and Ca_v2 ion channels are mainly of HVA type, whereas Ca_v3 is of the LVA category of calcium ion channels (Dolphin, 2009). Table 1.1.2 lists all the Ca_v ion channels along with their encoding genes, site of localization and associated diseases. Ca_v1 ion channels (Ca_v 1.1-1.4) play a significant role in contractility activities, secretion, regulation of gene expression and integration of synaptic input in neuronal synaptic transmission in specialized sensory cells. The Ca_v2 subfamily members, on the other hand, are more dominant in neuroendocrine cells and facilitate synaptic transmission at fast synapses. The Ca_v3 subfamily primarily act as pacemakers in neuronal cell bodies and dendrites and also serve to maintain the rhythmic firing in cardiac and smooth muscle myocytes (W. A. Catterall, 2005).

Structurally, Ca_v ion channels highly resemble the structure of the Na_v ion channels in terms of their homologous composition, and similar to Na_v ion channels, a single continuous gene encodes the Ca_v types (Zhorov & Tikhonov, 2004). Each alpha subunit monomer is composed of 6 segments (S1 to S6), and the main structural components are the selectivity filter, VSD, and P-loop (See Figure 1.1.6). An important structural feature of Ca_v ion channels is their association with calmodulin (CaM), which binds to the intracellular N-terminal (Halling, Aracena-Parks, & Hamilton, 2005). CaM is occasionally considered as a distinct subunit of Ca_v ion channels (called as gamma subunit), due to its profound effect on the function of these ion channels. There are also several other beta subunits associated with these proteins (Buraei & Yang, 2010) , which modify and control their function to a great extent, highlighted in (Hofmann, Lacinová, & Klugbauer, 1999; Petegem & Minor, 2006). Voltage-gated calcium channels are also targeted by several drugs for pain control in severe chronic refractory pain and cardiac diseases like hypertension and angina. Other possible indications of calcium channel

blockers include migraine, movement disorders, hearing and vision problems (Belardetti & Zamponi, 2012; Mcdonough, 2013).

Table 1.1.2. The voltage-gated calcium channel family, along with their encoding gene, localization sites and associated diseases. Sources: (Bidaud, Mezghrani, Swayne, Monteil, & Lory, 2006; Heyes et al., 2015; Lehmann-Horn & Jurkat-Rott, 1999; Simms & Zamponi, 2014)

CHANNEL	GENE	PRIMARY LOCATION	DISEASE
$Ca_v1.1$	CACNA1S	Skeletal muscle, transverse tubules	• Hypokalemic periodic paralysis • Malignant hyperthermia
$Ca_v1.2$	CACNA1C	CNS, smooth and cardiac muscle, neuronal cell bodies	• Timothy syndrome • Cardiac arrhythmia • Autism spectrum disorders
$Ca_v1.3$	CACNA1D	CNS, endocrine cells, pacemaker cells, cochlear hair cells	
$Ca_v1.4$	CACNA1F	Retinal rod and bipolar cells	• Stationary night blindness
$Ca_v2.1$	CACNA1A	CNS, heart pituitary dendrites, presynaptic terminal, skeletal muscle (end plate)	
$Ca_v2.2$	CACNA1B	CNS, dendrites, presynaptic terminal, some cell bodies	• Familial hemiplegic migraine • Cerebellar ataxia
$Ca_v2.3$	CACNA1E	CNS, dendrites, presynaptic terminal	
Ca_v 3.1	CACNA1G	CNS, heart, placenta, lung, kidney	• Absence seizures
Ca_v 3.2	CACNA1H		
Ca_v 3.3	CACNA1I		

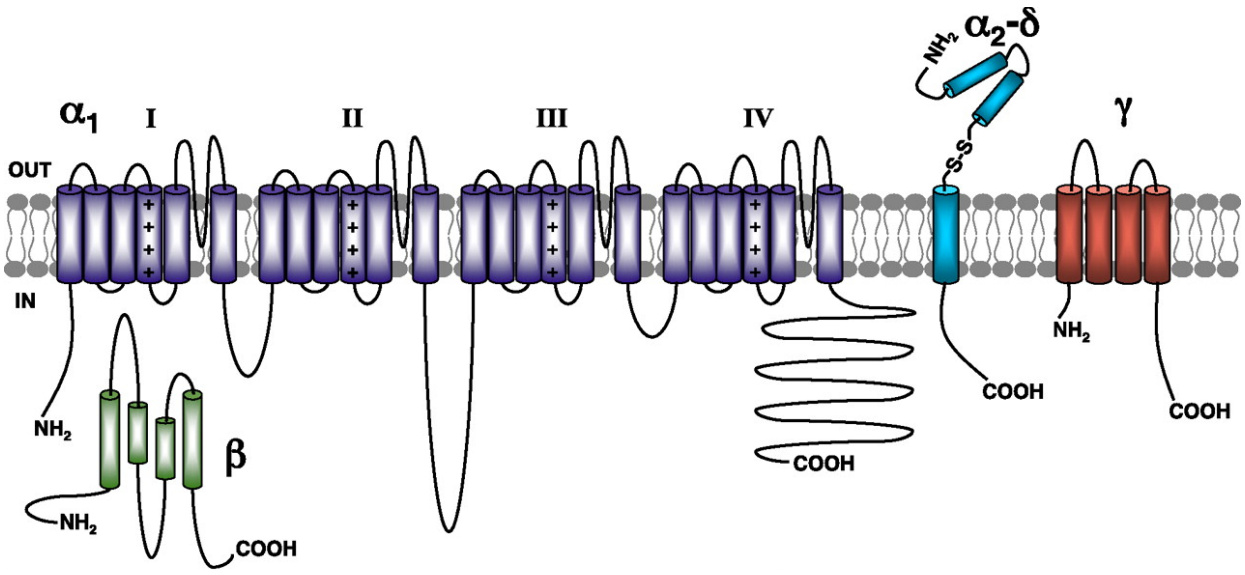


Figure 1.1.6. Schematic representation of a Ca_v ion channel, depicting the α1, Ca_vβ Ca_vα2δ, and Ca_vγ subunits. Source: (Vacher et al., 2008).

iii) Voltage-Gated Potassium Channels

The voltage-gated potassium channels form the largest family of VGICs and are named as K_V ion channels. More than 40 members of this diverse group of ion channels have been identified and named K_V1 up to K_V12 with numerous subtypes to each class (Gutman et al., 2005; Miller, 2000). Their expression sites include the heart, brain, auditory and vestibular organs and epithelia (T J Jentsch, 2000), where they perform central roles (see Table 1.1.3).

This superfamily of closely related tetrameric membrane proteins are mainly composed of homomeric structural components, i.e. the alpha subunits are identical in their sequence and structure (Kuang, Purhonen, & Hebert, 2015). However, they may rarely become heterotetramers when different subunits are formed within the same family as in the case of the K_V1, K_V7, and K_V10 families (Po, Roberds, Snyders, Tamkun, & Bennett, 1993).

Table 1.1.3. The voltage-gated potassium channel family, along with their encoding gene, localization sites and most common diseases associated with them. Source: (Arévalo, 2015; Miceli et al., 2015; Pletscher-Frankild, Pallejà, Tsafou, Binder, & Jensen, 2015; Uhlén et al., 2015; Vacher et al., 2008), www.genecards.org

CHANNEL	GENE	PRIMARY LOCATION	DISEASE
Voltage-gated potassium channel			
K_V1 subfamily: K_V1.1 K_V1.2 K_V1.3 K_V1.4 K_V1.5 K_V1.6	KCNA1 KCNA2 KCNA3 KCNA4 KCNA5 KCNA6A	CNS, PNS, Brain, Neurons Eyes Heart endothelium Colon Kidney Vascular smooth muscle	• Episodic • Ataxia/Myokymia • Syndrome Isolated • Hypomagnesemia • Seizures • Tremor • Neuropathic pain • Atrial fibrillation and arrhythmia
K_V2 subfamily: K_V2.1 K_V2.2	KCNB1 KCNB2	Retina of the eyes. Retinal neurons Brain Liver	• Infantile epilepsy • Encephalopathy • Brain ischemia
K_V3 subfamily: K_V3.1 K_V3.2 K_V3.3 K_V3.4	KCNC1 KCNC2 KCNC3 KCNC4	Brain Skeletal muscle	• Epilepsy • Spinocerebellar ataxia • Hereditary spastic paraplegia • Joubert syndrome
K_V4 subfamily: K_V4.1 K_V4.2 K_V4.3	KCND1 KCND2 KCND3	Heart Brain Eyes	• Atrial fibrillation • Spinocerebellar ataxia • Corneal ectasia
K_V5 subfamily: K_V5.1	KCNF1	Heart Brain Liver Skeletal muscle Kidney Pancreas	• Fragile X syndrome • Intellectual disability • Pancreatitis • LQTS
K_V6 subfamily: K_V6.1 K_V6.2 K_V6.3 K_V6.4	KCNG1 KCNG2 KCNG3 KCNG4	Brain Heart	• Fragile X syndrome • LQTS • Pulmonary hypertension
K_V7 subfamily: K_V7.1 K_V7.2 K_V7.3 K_V7.4 K_V7.5	KCNQ1 KCNQ2 KCNQ3 KCNQ4 KCNQ5	Heart Inner ear Stomach and colon Pancreas Prostate Kidney	• LQTS and cardiac arrhythmia • Jervell and Lange-Nielsen syndrome • Benign familial neonatal seizure • Migraine • Neuropathic pain
K_V8 subfamily: K_V8.1 K_V8.2	KCNV1 KCNV2	CNS Brain	• Schizophrenia • Epilepsy • Autistic disorder

Table 1.1.3. Continued.

CHANNEL	GENE	PRIMARY LOCATION	DISEASE
K_V9 subfamily: K_V9.1 K_V9.2 K_V9.3	KCNS1 KCNS2 KCNS3	CNS Brain Eyes	• Sensory peripheral neuropathy • Migraine • Neuroblastoma
K_V10 subfamily: K_V10.1	KCNH1	Heart Brain, CNS Lungs	• Congenital myasthenic syndrome • Epilepsy • LQTS
K_V11 subfamily: K_V11.1 K_V11.2	KCNH2 KCNH3	Heart Brain Ovary, testis	• LQTS, SQTS • Brugada syndrome • SIDS • drug-induced Torsade de Pointes
Inward-rectifier potassium ion channel			
K_{ir}1 subfamily: K_{ir}1.1	KCNJ1	Kidneys Pancreas	• Hypokalemia • Hyperaldosteronism • Bartter syndrome • Nephrocalcinosis and osteopenia
K_{ir}2 subfamily: K_{ir}2.1 K_{ir}2.2 K_{ir}2.3 K_{ir}2.4 K_{ir}2.6	KCNJ2 KCNJ12 KCNJ4 KCNJ14 KCNJ18	Granulocytes Eyes Brain Skeletal muscle Heart Corneal dystrophy	• Smith-Magenis syndrome • Long QT syndrome • Atrial fibrillation
K_{ir}3 subfamily: K_{ir}3.1 K_{ir}3.2 K_{ir}3.3 K_{ir}3.4	KCNJ3 KCNJ6 KCNJ9 KCNJ5	CNS Dopaminergic neurons Eyes	• Down syndrome • Pain agnosia • Corneal dystrophy • Conn's syndrome • Adenoma
K_{ir}4 subfamily: K_{ir}4.1 K_{ir}4.2	KCNJ10 KCNJ15	Brain Spinal cord Eyes	• EAST syndrome • Pendred Syndrome • Epilepsy • Down syndrome • Familial periodic paralysis • Corneal dystrophy
K_{ir}5 subfamily: K_{ir}5.1	KCNJ16	Kidney Brain, neurons	• EAST syndrome • Sesame syndrome • Body dysmorphia
K_{ir}6 subfamily: K_{ir}6.1	KCNJ8	Brain Cardiac endothelium	• Cantu syndrome • SIDS • J-wave syndromes
K_{ir}7 subfamily: K_{ir}7.1	KCNJ13	Retina of the eyes	• Vitreous syneresis • Stickler syndrome

Table 1.1.3. Continued.

CHANNEL	GENE	PRIMARY LOCATION	DISEASE
$K_{ir}1$ subfamily: $K_{ir}1.1$	KCNJ1	Kidneys Pancreas	• Hypokalemia • Hyperaldosteronism • Bartter syndrome • Nephrocalcinosis and osteopenia
$K_{ir}2$ subfamily: $K_{ir}2.1$ $K_{ir}2.2$ $K_{ir}2.3$ $K_{ir}2.4$ $K_{ir}2.6$	KCNJ2 KCNJ12 KCNJ4 KCNJ14 KCNJ18	Granulocytes Eyes Brain Skeletal muscle Heart Corneal dystrophy	• Smith-Magenis syndrome • Long QT syndrome • Atrial fibrillation
$K_{ir}3$ subfamily: $K_{ir}3.1$ $K_{ir}3.2$ $K_{ir}3.3$ $K_{ir}3.4$	KCNJ3 KCNJ6 KCNJ9 KCNJ5	CNS Dopaminergic neurons Eyes	• Down syndrome • Pain agnosia • Corneal dystrophy • Conn's syndrome • Adenoma
$K_{ir}4$ subfamily: $K_{ir}4.1$ $K_{ir}4.2$	KCNJ10 KCNJ15	Brain Spinal cord Eyes	• EAST syndrome • Pendred Syndrome • Epilepsy • Down syndrome • Familial periodic paralysis • Corneal dystrophy
$K_{ir}5$ subfamily: $K_{ir}5.1$	KCNJ16	Kidney Brain, neurons	• EAST syndrome • Sesame syndrome • Body dysmorphia
$K_{ir}6$ subfamily: $K_{ir}6.1$	KCNJ8	Brain Cardiac endothelium	• Cantu syndrome • SIDS • J-wave syndromes
$K_{ir}7$ subfamily: $K_{ir}7.1$	KCNJ13	Retina of the eyes	• Vitreous syneresis • Stickler syndrome

Unlike Ca_v and Na_v ion channels, which are encoded by a single gene with slightly different repeating sequences, K_V ion channels are encoded by four copies of the same gene. Accordingly, the alpha subunit is formed cotranslationally, consisting of six putative segments, termed S1 to S6 (See Figure 1.1.7). Segments S1-S4 form the VSD that undergoes a conformational change upon voltage changes, i.e. membrane depolarization. S4 is the primary mobile element of VSD and its displacement in the membrane plays a significant role in the activation and opening of the pore (Geoffrey W. Abbott, 2014). The pore domain (PD) in the center of the protein and the selectivity filter are formed by S5 and S6 of the four identical copies of alpha subunits. The VSD and PD work together, such that the displacement of S4 in the VSD causes a conformational change in the PD, opening the pore and allowing ions to pass selectively (Peroz et al.,

2008). The selectivity of K^+ channels for potassium ions is associated with a conserved signature sequence motif; TxGYG in the P-loop. This sequence motif creates the selective behavior of the channel protein (Robbins, 2001; Sansom et al., 2002).

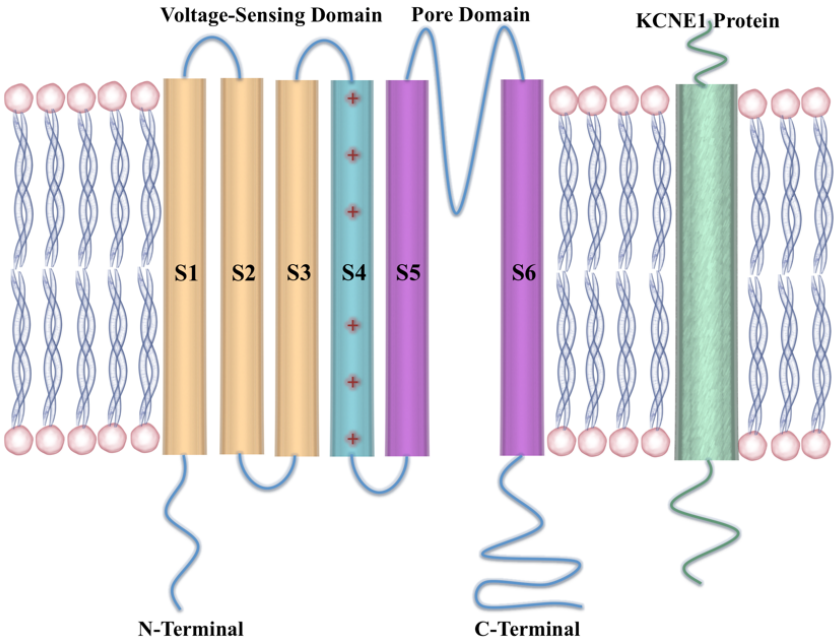


Figure 1.1.7. Schematic representation of a single K_V α subunit and its Beta subunit. Four of such α subunits assemble to form a functional K_V channel. The KCNE1 beta subunit plays an important role in controlling and modifying the actions of the channel.

Similar to Na_v and Ca_v ion channels, the voltage-gated K^+ channels are also associated with several beta subunits of diverse structure and function. Some notable ones include the KCNE family of proteins (minK and MiRPs) (G W Abbott & Goldstein, 2001), calmodulin (CaM) (Wen & Levitan, 2002)(Halling et al., 2005), protein kinase C (PKC) zeta-interacting proteins (Boland & Jackson, 1999; Thomas et al., 2003), KCHip (K_V4.x-channel-interacting proteins) and KCHap (K_V1.x- and K_V2.x-channel-associated proteins) (An et al., 2000). Phosphorylation or dephosphorylation (Park, Yang, Seikel, & Trimmer, 2008), ubiquitinylation (Boehmer et al., 2008), SUMOylation (H. Wu, Chen, Cheng, & Qi, 2016) and palmitoylation (Shipston, 2011) may also modify the K_V channel properties.

KCNQ channels play a significant role in human disease and harbour the highest number of mutations that are closely linked to cardiac arrhythmias, deafness and benign familial neonatal epilepsy (T J Jentsch, 2000; Shieh, Coghlan, Sullivan, &

Gopalakrishnan, 2000). Due to the vast number of functions and processes affected by these proteins, they are also considered as potential drug targets for the treatment of cancer, autoimmune diseases, metabolic, neurological and cardiovascular disorders (Wulff, Castle, & Pardo, 2009). However, another surprising fact is that K_V channels, in particular, the K_V11.1 (hERG) ion channel, has become an obstacle in the journey of developing a drug. Such that drugs which are designed to target other organs and proteins, block the cardiac hERG ion channel and account for cardiotoxicity effects (Anwar-Mohamed et al., 2014; FLORESCU, CINTEZA, & VINEREANU, 2013). This further adds to the ambiguities of studying ion channels and remains an active area of research.

As mentioned above, the K_V ion channels are the most abundant and diverse members of VGICs superfamily, from both structural and functional points of view. There has been countless number of research done and massive amount of knowledge has been generated concerning these channels. However, the physiological function and structural characteristics of many of the K_V family members has remained to be conclusively solved.

1.1.3. Ion Channel-Related Diseases

Combining the wealth of data obtained from ion channels' structures with their biophysical properties enabled the rationalization of diseases aetiology and phenotype. The number of such "channelopathies" are constantly increasing. This includes muscular and neurological diseases such as movement disorders, migraine, epilepsy, arrhythmias, deafness and metabolic disorders such as diabetes and many others (Kass, 2005). Many of these diseases and disorders stem from mutations in the coding region of ion channel genes, leading to the gain or loss of channel function, either of which may have detrimental consequences. In addition, defective regulation of channel activity by intracellular or extracellular ligands or modulators is responsible for certain diseases. This can be due to the mutations in the genetic encoding of the regulatory molecules themselves, or because of defects in the pathways leading to their synthesis as in the case of some forms of diabetes mellitus (Proks & Lippiat, 2006). Auto-antibodies to ion channel proteins may produce disease either by down-regulating or by enhancing channel

function (Martinez-Martinez et al., 2013). Furthermore, ion channels may act as lethal agents, being secreted by cells and inserted into the membrane of a target cell to form large nonselective pores that cause cell lysis and death. Complement and the hemolytic toxin produced by the bacterium *Staphylococcus aureus* are examples of this type of ion channels (Ashcroft & Ashcroft, 2000c).

The frequency of most channelopathies in the general population is very low. However, the insight they provide regarding the link between ion channel structure and function, and into the physiological role of the different ion channels has been invaluable (Ashcroft & Ashcroft, 2000b). This complexity and involvement in several diseases further lead to ion channels frequently becoming the therapeutic targets for correcting and curing many disorders and diseases (Bagal et al., 2015; Belardetti & Zamponi, 2012; Kaczorowski, McManus, Priest, & Garcia, 2008; Wulff et al., 2009). Tables 1.1.1, 1.1.2 and 1.1.3 include a list of diseases associated with sodium, calcium and potassium ion channels, respectively.

1.1.4. Distribution of Ion Channels in the Human Body

Ion channels are present all over the human body in various organs. The most important sites of expression of ion channels are the brain, nervous system, skeletal and vascular muscles and the heart. However, they are also present in other organs such as the gastrointestinal tract (Feranchak, 2003), epithelial tissue of kidney (Kuo & Ehrlich, 2012), pancreas (Braun et al., 2008), lungs (Hollenhorst, Richter, & Fronius, 2011) audio-vestibular organs (Gabashvili, Sokolowski, Morton, & Giersch, 2007), sweat and salivary glands (Gao et al., 2009; Palmer, 2007).

The functions of ion channels vary depending on their differential localization. Neuronal ion channels trigger the nerve impulse generation and ultimately control the signalling characteristics of a neuron (Lai & Jan, 2006). In skeletal muscles, ion channels control the excitability required for contractions. Conversely, in the cardiac tissue, these proteins serve to regulate the excitation of myocytes and thus the contractility of the heart muscle. In vascular muscles, ion channels regulate the vascular tone in the walls of resistance arteries and arterioles (W. F. Jackson, 2000). Other functions of ion channels

include controlling hormonal secretions, sensation of the environment, gene expression and cell division (Yu & Catterall, 2004).

Although all the functions of ion channels in different parts of the body are vital, the cardiac ion channels have been regarded to be of utmost significance. This is mainly due to their high prevalence in cardiac tissue as well as their high impact over the function of the heart. A cordial association of these proteins with their currents generates the action potential of the heart. Action potential is the main signal required for the contraction of the heart muscle and thereby the main function of the heart; i.e. to supply blood to the body. The following section aims at highlighting and describing cardiac ion channels and their significance in more depth.

1.1.5. Cardiac Ion Channels

Regulation of the contractile activity of muscle cells in the heart is dependent on a complex interplay of events, jointly called the excitation contraction coupling (ECC). ECC is the process during which a stimulatory electrical signal, such as an action potential from the neurons is transformed into muscle movement (Bers, 2002). However, the formation and propagation of cardiac action potential is highly dependent on ion channel proteins present in the cardiac myocytes. More than 20 ion channels are expressed in the cardiac tissue (Grant, 2009), giving rise to several types of currents, each of which possess specific characteristics.

Cardiac ion channels and more specifically the cardiac voltage-gated potassium KCNQ1 ion channel, serve as the focus of this thesis. To provide more background, the following section describes the cardiac action potential in more details. This is followed by a comprehensive description of the various ion channels and the currents that they produce in the heart.

1.1.5.1. Cardiac Action Potential

As mentioned above, the normal function of the heart relies on the accurate coordination of cardiac ion channels. The electrical charge on both sides of the cardiac cell membrane is altered even when a single ion channel is activated. This difference in the membrane charge is called membrane potential, which creates an activation signal and propagates

throughout the heart. To grasp the basics of cardiac electrophysiology, one must have a thorough understanding of the cardiac action potential. Therefore, this section describes the cardiac action potential followed by highlighting the main cardiac currents generated by cardiac VGICs.

Cardiac action potential represents the changes in the voltage of a single cardiac cell plotted over time. The cardiac action potential shown in Figure 1.1.8 is a series of events associated with changes in the membrane voltage. It can be broken down into different phases that are the result of atomic level changes in the ion channels.

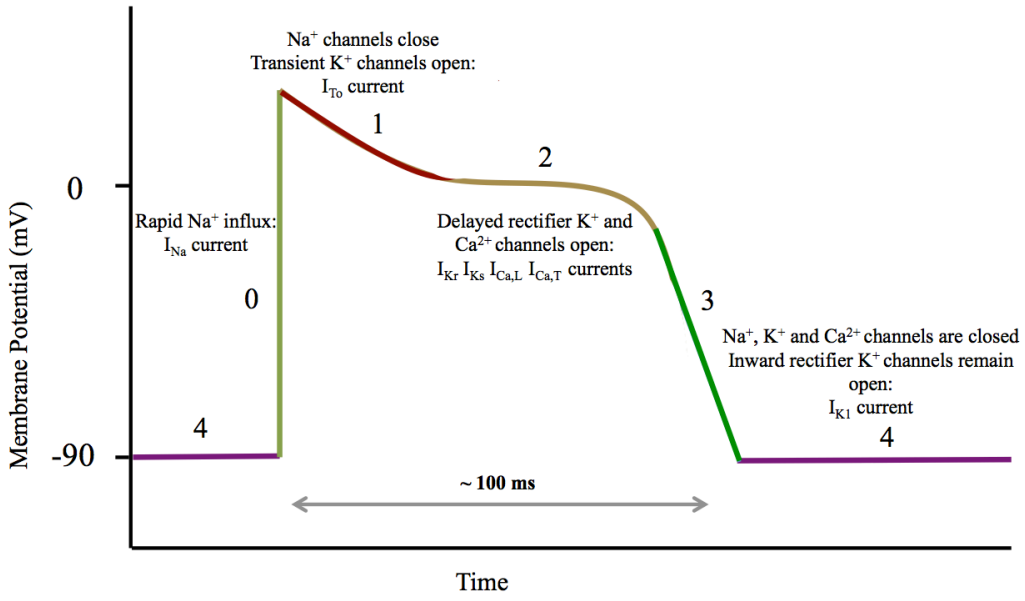


Figure 1.1.8. Cardiac Action Potential. Phase 0: depolarization due to the opening of fast sodium channels. Phase 1: Early repolarizations because of a rapid decrease in sodium ion influx as fast sodium channels are inactivated. Phase 2: plateau phase in which the influx (Ca^{2+}) and efflux (K^{+}) of positive charge become equal. Phase 3: repolarization phase, sodium and calcium channels all close, delayed rectifier potassium channels remain open and membrane potential returns to its baseline at about -90 mV. Phase 4: resting membrane potential, inwardly rectifying K^{+} channels remain open to maintain the membrane potential in a stable condition.

At resting membrane potential (phase 4) most Na^{+}, K^{+} and Ca^{2+} ion channels are in the closed/inactivated state. Voltage gated potassium rectifier ion channels are the exceptions that remain open to maintain the negative membrane resting potential. However, the rate of ion exchange across the membrane is almost zero, sustaining the

membrane potential at about -90 mV. Phase 4 represents both the beginning and ending of a complete action potential cycle. At Phase 0, the Na^+ channels are rapidly activated when triggered by a neighbouring cardiomyocyte or pacemaker cell. This allows a huge influx of Na^+ ions, making the intracellular voltage, positive and causing a rapid depolarization. The L-type Ca^{2+} channels then open simultaneously during Phase 0 at about -40 mV. This causes a small but steady influx of Ca^{2+} down its concentration gradient (Nerbonne & Kass, 2005).

After the depolarization phase is complete at about +30 mV, the cell tends to return to its resting membrane potential. This is achieved in the consecutive three phases. In Phase 1, which is called the initial/early repolarization, the Na^+ channels close as rapidly as they opened. This leads to a short and quick decline of action potential. The potassium channels then open slowly, causing K^+ ions to move out of the cell and return the membrane voltage to approximately 0 mV. At a certain stage the influx of the positive charges by L-type Ca^{2+} channels and the efflux of positive charge by the delayed rectifier K^+ channels create a balance, leading to the formation of a plateau phase. This is called Phase 2 of the action potential, which remains just below 0 mV throughout phase 2 (Grant, 2009; Nerbonne & Kass, 2005; Pinnell, Turner, & Howell, 2007).

During Phase 3, which is the main repolarization stage, the calcium ion channels are gradually inactivated but the potassium ion channels remain open, now dominating over the influx of Ca^{2+}. This brings the membrane voltage back to its resting membrane potential at −90 mV to prepare the cell for a new cycle of depolarization. To return the ionic concentration gradients back to its primary form, Na^+ and Ca^{2+} ions have to return to the extracellular compartment, and K^+ ions to the cell interior. This is when the sarcolemmal Na^+- Ca^{2+} exchanger, Ca^{2+}-ATPase and Na^+-K^+-ATPase pumps come into play and restore the normal intracellular and extracellular ionic concentrations (Rudy, 2008).

The action potential duration (APD) differs from one region of the heart to another, for example the atrial cardiomyocytes have a shorter action potential than the ventricular cardiomyocytes due to the differences in the underlying currents. The majority of the lethal events happening at the cardiac level are due to abnormalities in the APD. These include LQTS, SQTS, arrhythmias, ventricular/atrial fibrillations and sudden

cardiac death, which all have an underlying effect on one or the other factors that control the APD (Campuzano et al., 2010).

1.1.5.2. Currents in the Heart

The main ionic currents, which play significant roles in the formation of the cardiac action potential, are discussed in this section. Table 1.1.4 lists these currents along with type of ion channels that generate them and their auxiliary beta subunit. Needless to say, cardiac ion channels are complicated biological systems and a comprehensive discussion on them is beyond the scope of this thesis. However excellent reviews can be found in (Grant, 2009) and (Roden et al., 2002). The main ionic current types: sodium, potassium and calcium are described in the following section.

Table 1.1.4. The principal ionic currents in the heart. The alpha subunit ion channel that generates them along with the encoding gene. The last column lists the auxiliary beta subunits associated with the ion channel proteins.

Current	Ion Channel (Pore forming α-subunit)	Gene	β-subunit
I_{Na}	$Na_v1.5$	SCN5A	$Na_v\beta$
I_{Kur}	$K_V1.5$	KCNA5	-
I_{To}	$K_V4.2$-4.3	KCND2/3	KChIP2
I_{KS}	$K_V7.1$/KCNQ1	KCNQ1	KCNE1
I_{Kr}	$K_V11.1$/HERG	KCNH2	KCNE2/KCNE3
I_{K1}	$K_{ir}2.1$-2.2	KCNJ2/12	-
$I_{Ca,L}$	$Ca_v1.2$	CACNA1C	$Ca_v\alpha_2/\beta_2/\delta$
$I_{Ca,T}$	$Ca_v3.1$-3.2	CACNA1G	$Ca_v\alpha_2/\beta_2/\delta$

Sodium Current (I_{Na})

As mentioned in the previous section, it is the sodium ion channels' rapid activation and the resulting sodium current, I_{Na} that causes the depolarisation of the cardiac cells. The voltage-gated sodium channel $Na_v1.5$ is the primary component in generating the cardiac sodium current. This is demonstrated by the fact that several cardiac syndromes,

including long QT syndrome and Brugada Syndrome, have been linked to mutations in SCN5A, which is the gene encoding $Na_v1.5$ (Ahern et al., 2016).

During the resting membrane potential (Phase 4), the majority of channels are in the closed state. When the $Na_v1.5$ channels receive the depolarising signal, they undergo a very fast transition to the open state (Abriel, 2007). During Phase 4, $Na_v1.5$ recovers from an inactivated state to initiate a new cycle of the action potential. Despite the fact that $Na_v1.5$ channels are the most prominent sodium channels in the heart, other voltage-gated sodium channels may also play a role in generating the cardiac I_{Na}. These include the neuronal sodium channels that are suggested to be involved in electrical-chemical coupling and thus propagation of the cardiac action potential (Maier et al., 2002).

L-type ($I_{Ca,L}$) and T-type ($I_{Ca,t}$) Calcium Currents

In a cardiac muscle, two types of calcium channels are present. The L-type (low-threshold) and the T-type (transient-type). While the L-type calcium ion channels ($Ca_v1.2$) are present in all cardiac cells, the T-type ($Ca_v3.1$-3.2) is only present in specific types of cells (pacemaker, Purkinje, atrial). The L-type $Ca_v1.2$ and the T-type $Ca_v3.1/3.2$ channels generate the $I_{Ca,L}$ and $I_{Ca,T}$ respectively. They are dominantly open during Phase 2 (plateau phase), allowing the Ca^{2+} ions to enter and balance the efflux of K^+ ions (Grant, 2009).

Moreover, the entry of Ca^{2+} ions into the cytosol triggers the sarcoplasmic reticulum to release additional calcium ions. This process is essential for the excitation-contraction coupling to happen. The released calcium interacts with troponin-C, causing myosin to interact with actin and eventually leading to contraction of the cardiomyocytes (Gez, Hagalili, Shainberg, & Atlas, 2012). Therefore, Ca_v ion channels are additionally important for the normal sequence of events that eventually translate into the physical contraction of the cardiac muscle.

Transient Outward (I_{to}), Inward Rectifier (I_{K1}), Delayed Rectifier (I_K) Currents

Potassium ion channels are the largest and most variant group of ion channels in the heart. The major currents produced by K^+ ion channels are the transient outward (I_{to}), the inward rectifier (I_{K1}) and the delayed rectifier (I_K) currents. The contribution of each of these currents to the action potential falls into the different phases, described below (Ashcroft & Ashcroft, 2000e).

The transient outward (I_{to}) is generated by the $K_V4.2$, $K_V4.3$ pore-forming α-subunit, and to some degree by the $K_V1.4$ subunits, in association with ß-subunits, such as KChiP2. It is of two components: the fast ($I_{to,f}$) and slow ($I_{to,s}$) that are mainly prominent during Phase 1 of the action potential to initiate the repolarization process. Any change in the density and duration of these currents can have strong influences on the normal formation and propagation of the action potential (Roden et al., 2002).

Inwardly rectifying currents are also of several types; including the quasi-instantaneous rectifier I_{K1}, the ATP inhibited I_{KATP} and the muscarinic receptor stimulated I_{KACh}. However, of specific interest to us is the inward rectifier K^+ current which sets the resting membrane potential to be stable at ~ -90 mV. During Phase 4, when all the other ion channels are closed or inactivated, the $K_{ir}2.1$ and $K_{ir}2.2$ remain open allowing a minimal influx of K^+ ions into the cell. This steady influx is also regarded as a slow depolarization (Ashcroft & Ashcroft, 2000f).

The delayed-rectifier K^+ currents (I_K) are of three types: slow (I_{KS}), fast (I_{Kr}) and ultrarapid (I_{Kur}), distinguished by their kinetics of activation and deactivation. The slow delayed rectifier I_{KS} current is produced by $K_V7.1$ (KCNQ1) ion channel with its association of its beta subunit, KCNE1 (Jost, Papp, & Varro´, 2007). The fast-delayed rectifier I_{Kr} current is generated by $K_V11.1$ (hERG) ion channel and the auxiliary beta subunit KCNE2/3. Both the slow and fast delayed rectifier currents are crucial for the repolarization (Phase 3) of action potential to take place. The ultrarapid I_{Kur} current is supplied by the $K_V1.5$ and to a lesser extent by $K_V3.1$ ion channels during Phase 1 and Phase 2 of action potential (Roden et al., 2002).

As stated earlier, K^+ currents in the heart have a fundamental role in shaping the cardiac action potential at different phases and stages. Many cardiac abnormalities and diseases are associated with disruptions in the normal formation and propagation of the

K^+ currents. In addition, many existing drugs can inhibit or activate these currents both as a therapeutic target as well as a side effect (cardiotoxicity).

Of specific interest to us was the slow delayed rectifier I_{KS} current that is produced by the $K_V7.1$ (KCNQ1) ion channel and its auxiliary beta subunit, KCNE1 protein. I_{KS} is a slowly activating, voltage-dependent current and its role is most dominant during Phase 2 (i.e. the plateau phase) of the action potential. It serves by offsetting the inward calcium flux and repolarization of cardiac action potential (Jost, Papp, & Varro, 2007). Any disturbance in the normal generation or propagation of the I_{KS} current may lead to instabilities in the heart and result in Torsade de pointes arrhythmias (Jost, Papp, & Varro, 2007). The latter effect is also observed when the I_{KS} current is defective due to genetic mutations or down regulation of KCNQ1/KCNE1 and when blocked by pharmacological agents (Roden & Yang, 2005). There are numerous evidences for the significance of this current in the heart (Jost, Papp, & Varro, 2007; Sanguinetti et al., 1996; Veerman et al., 2013; Youshan Yang & Sigworth, 1998) and yet a lot more remains to be understood about the mechanisms of generation of I_{KS} current by the KCNQ1/KCNE1 protein. In this Thesis, we aim at revealing more atomistic details about the KCNQ1/KCNE1 ion channel. The next chapter will elaborate more on this ion channel and the resultant current and its significant role in the heart.

1.1.6. Summary

The present section of Chapter 1 served as an introductory gate to the world of ion channels. It showed that ion channels are versatile proteins of different classes and subtypes, which are spread throughout the human body. Their localization ranges from brain, heart to muscles, kidneys and eyes where they perform vital functions. Any disruption in the normal function of ion channels can result in a syndrome or disease such as heart irregularities in the form of arrhythmias, seizures, blindness, deafness, etc. Of utmost importance are the ion channels present in the heart, which control the normal rhythm of this beating muscle. The focus of this thesis narrows down to a mysterious potassium ion channel in the heart that plays a fundamental role in the formation of the cardiac action potential, as is introduced in the next section.

1.2: HUMAN KCNQ1 ION CHANNEL

1.2.1. Introduction

The human KCNQ1 (K_V7.1 or K_VLQT1) ion channel belongs to the large family of voltage gated potassium ion channels. The KCNQ1 protein is expressed in the heart, muscles, pancreas, kidney, brain and inner ear (Geoffrey W. Abbott, 2014). The cardiac KCNQ1 ion channel is one of the important proteins in the heart, mainly due to the crucial role that it plays in controlling the repolarization phase of the action potential and secondly due to the diseases related to mutations with the channel. KCNQ1 ion channel was first identified in a study by Wang et al. to reveal some of the genetic causes of sudden death from cardiac arrhythmia (Q. Wang et al., 1996). Since its identification in 1996, KCNQ1 has become one of the most studied ion channels.

This section is focused on providing a general background about the KCNQ1/KCNE1 ion channel, with an emphasis on the important physiologic aspects of this ion channel, such as pathophysiology and regulation.

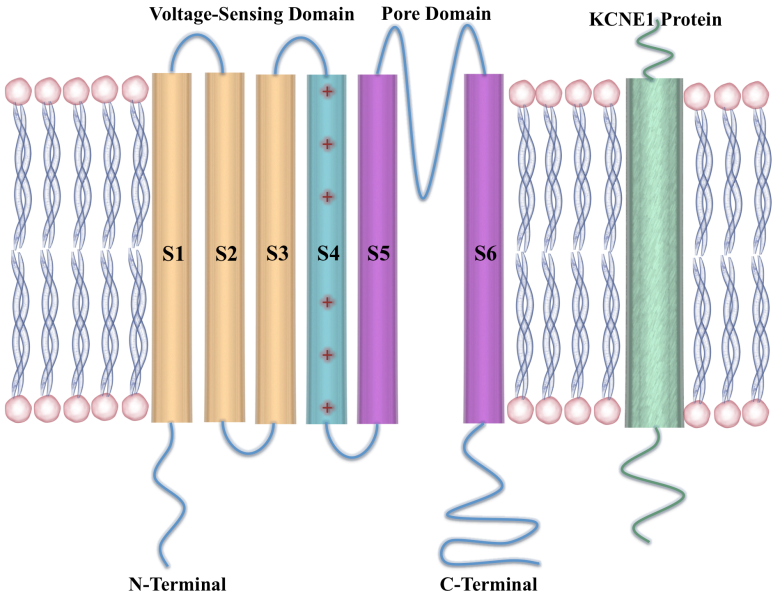
1.2.2. KCNQ1 Channel Properties and Structure

K_V7.1 ion channel protein is encoded by the human KCNQ1 gene (also called K_VLQT1). This gene is composed of 16 exons and the full-length gene spans ~400 kb (Splawski et al., 1998), which is primarily translated into 676 amino acids (isoform 1), forming a functional alpha subunit in the structure of the ion channel. The sequence of KCNQ1 isoform 1 is shown in Figure 1.2.1 with the specific structural components shown. An alternatively spliced variant of the channel (isoform 2) is also expressed in the heart (Demolombe et al., 1998). Comparatively, this isoform lacks almost the whole N terminus and is composed of 549 residues. Isoform 2 of KCNQ1 is non-functional alone but acts as a negative modulatory unit when co-expressed with the functional isoform (Pereon et al., 2000). From this point on, it is the isoform 1 of KCNQ1 that is being discussed in this thesis.

Figure 1.2.1. Amino acid sequence of KCNQ1 (isoform 1) with structural annotation. The figure illustrates the sequence of KCNQ1 with the different structural components highlighted; NT (amino terminal) and CT (carboxy terminal) are shown in red colour, S1-S3 (yellow), S4 (blue), S5-S6 (purple), SF (selectivity filter) shown in green.

Each alpha subunit of KCNQ1 is structurally composed of six alpha-helical transmembrane segments (S1-S6) and a single pore loop (P-loop). This gives rise to a 6TMD-1P configuration i.e. six transmembrane domains and one P-loop segment, shown in Figure 1.2.2.a. Four of such KCNQ1 polypeptides (alpha subunits) assemble to form the functional homomeric tetramer of the KCNQ1 ion channel (Wiener, Haitin, Shamgar, Fernández-Alonso, et al., 2008; W.-P. Yang et al., 1997). Figure 1.2.2.b shows the arrangement of the alpha subunit segments in a 3D tetrameric arrangement. The first four segments (S1-S4) in the alpha subunit are designated as the voltage-sensing domain (VSD) of the ion channel, whereas S5-S6 and their linker form the pore domain (PD). The S4-membrane spanning segment contains regularly spaced positively charged residues (4 amino acids) and acts as the driving force during voltage-dependent gating (Cui, 2016). The S5-S6 pore domain and the intervening P-loop are responsible for the ion conduction and selectivity of the ion channel (Jespersen, Grunnet, & Olesen, 2005).

a)



b)

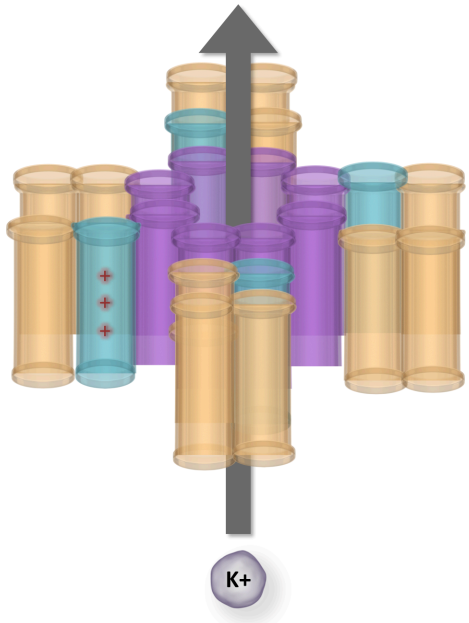


Figure 1.2.2. (a) General topology of the KCNQ1/KCNE1 ion channel. The voltage sensing domain (VSD) and the pore domain (PD) constitute the alpha subunit. The beta subunit of this ion channel is the KCNE1 protein. (**b) The 3D four-fold assembly of the subunits forming the pore in the centre of the channel.** The four subunits of KCNQ1 assemble together and form the pore in the centre of the channel. The S5-S6 and their loops are the main constituents that form the channel's pore.

Like all other VG potassium ion channels, KCNQ1 also follows the trend of possessing a conserved amino acid sequence (GYGD) in its selectivity filter (Aiyar, Rizzi, Gutman, & Chandy, 1996). The selectivity filter is the structural component of a voltage-gated ion channel that infers which ions would have a higher probability of passing through the channel pore. In KCNQ1, the GYGD signature motif is formed by the loop connecting the S5 and S6 segments (Jespersen et al., 2005). The cytoplasmic N- and C-terminal domains consist of 122 and 322 residues, respectively. The C-terminus is exceptionally functional in KCNQ1 and is responsible for its tetramerization, surface expression, assembly and protein-protein interactions (Lvov, Gage, Berrios, & Kobertz, 2010; Wiener, Haitin, Shamgar, Fernandez-Alonso, et al., 2008). Many cardiac abnormalities result from the mistrafficking and impaired assembly as a result of deleting a part of the C-terminal (Schmitt et al., 2000).

1.2.3. KCNQ1 and the I_{KS} Current

The KCNQ1/KCNE1 channel complex conducts the outward slow delayed rectifier potassium current (I_{KS}) in the heart. It is a slowly activating, voltage-dependent current and its role is most obvious during Phase 2 (plateau phase) of the action potential, during which it helps in offsetting the inward calcium flux and repolarization of cardiac action potential (Jost, Papp, & Varro, 2007).

The discovery of the I_{KS} current has an interesting history. Back in 1991, it was believed that the KCNE1 (mink) protein alone is responsible for the generation of the I_{KS} current (Hausdorff, Goldstein, Rushin, & Miller, 1991; Honoré et al., 1991). Around the same time, extensive research was done to identify the underlying causes of the LQTS. As such, the KCNQ1 protein was identified and its mutations were found to be associated with the LQTS and cardiac arrhythmias (Q. Wang et al., 1996). This was the result of defects in a specific type of current that was characterised to be very similar to the previously identified I_{KS} current. Putting the pieces of the puzzle together, in 1996, two groups reported that KCNQ1 and KCNE1 are co-expressed and generate the native slow I_{KS} current (Barhanin et al., 1996; Sanguinetti et al., 1996). With that in mind, the KCNE1 protein was coined as an auxiliary beta subunit that is co-expressed with

KCNQ1, and that it is not able to generate the I_{KS} current by itself (Sanguinetti et al., 1996).

During abnormal conditions of prolonged repolarization in the heart or when there is an elevation in the sympathetic tone, the I_{KS} current plays a vital role through a "safety mechanism". This mainly involves preventing excessive and dangerous lengthening of repolarization, which can otherwise lead to instabilities in the heart and result in Torsade de pointes arrhythmias (Jost, Papp, & Varro, 2007). The latter effect is also observed when the I_{KS} current is defective due to genetic mutations, down regulation or blocked by pharmacological agents (Roden & Yang, 2005).

Despite the fact that the I_{KS} current has never been a successful candidate for blockade and drug design purposes (Roden & Yang, 2005), it is occasionally targeted by non-cardiac drugs (Veerman et al., 2013). To put it in a simpler way, the drugs that are not intended to affect the heart physiology, bind to cardiac ion channels and disturb the generation of the vital current in the heart, mainly the IKr and I_{KS} currents generated by hERG and KCNQ1/KCNE1, respectively. As a consequence these drugs cause unintentional drug-induced cardiotoxicity in the form of QT prolongation and Torsade de Pointes (Tdp), which eventually leads to their discontinuation of development and market withdrawal (Yap & Camm, 2003).

1.2.4. KCNQ1 Pathophysiology

In the heart, mutations in KCNQ1 or the KCNE1 gene that affect trafficking, assembly, or regulation lead to the life-threatening long QT syndrome (LQTS). Congenital LQTS is characterized by delayed repolarization in the action potential of the heart that leads to QT prolongation and T-wave abnormalities on the ECG. Consequent effects of QT prolongation are TdP; a polymorphic ventricular arrhythmia, seizures and sudden cardiac death (Tester & Ackerman, 2014).

LQTS is categorized into two forms: the Romano-Ward syndrome (RWS) and the more fatal Jervell and Lange-Nielsen syndrome (JLNS). The former is inherited in autosomal dominant fashion and is not associated with hearing defects. RWS is more dominant and accounts for the majority of LQTS cases worldwide (Schwartz, Crotti, & Insolia, 2012). JLNS on the other hand is a rare, autosomal recessive form of LQTS. It is

characterized by QT prolongation as well as bilateral sensory-neural deafness and a higher incidence of sudden death (Modell & Lehmann, 2006; Mousavi Nik et al., 2015). Both variations of LQTS are associated with the KCNQ1/KCNE1 ion channel, as discussed below.

The result of loss-of-function mutations in the KCNQ1 or KCNE1 gene is a reduction in the outward I_{KS} current during action potential repolarization. This further leads to a prolonged opening of the channel, delay in the repolarization of the membrane and finally the QT interval is prolonged, emerging as LQTS. In addition, gain-of-function mutations in cardiac KCNQ1 increase current flow and lead to shortening of the cardiac action potential; as seen in a number of cardiac rhythm disorders such as Short QT syndrome (SQTS) and atrial fibrillation (Tester & Ackerman, 2014; Wulff et al., 2009).

KCNQ1 channels also play critical roles in the inner ear, where they mediate K^+ flux across the apical membrane of the marginal cells (MCs) of the stria vascularis into the scala media of the cochlear duct (Rivas & Francis, 2005). The hearing loss in JLNS is due to potassium level disturbances in the inner ear fluid, as a result of defects in the endolymph-producing stria vascularis of the cochlea (Modell & Lehmann, 2006; W. Wang et al., 2015). KCNQ1 ion channels co-expressed with KCNE proteins are also found in the epithelial tissues of lung, stomach, cochlea, intestine and kidney (Jespersen et al., 2005). Their role in the kidney is not yet explicitly ruled out, but animal studies have shown that KCNE1 knockout mice were found to suffer from hypokalemia, urinary and fecal salt wasting and volume depletion. This indicates that the I_{KS} current could possibly play a major role in normal renal function (Vallon et al., 2001).

1.2.5. KCNQ1 Ion Channel Regulation

Like all other signalling proteins and ion channels, KCNQ1 channels are also regulated by several cytosolic as well as intracellular and extracellular entities. The significant regulatory components that highly affect the function, current properties and expression of these proteins, are discussed below.

i) KCNE proteins

KCNQ1 ion channels possess a broad range of current characteristics and functionalities

mainly due to their unique ability to associate with the five members of the KCNE beta subunit family (KCNE1-KCNE5) (Jespersen et al., 2005; Melman, Um, Krumerman, Kagan, & McDonald, 2004). Once bound to the channel, each of the KCNE family members can have a dramatically different consequence on the channel conductance, gating and overall nature of the channel. KCNE proteins are small single transmembrane proteins (1-TM topology) composed of a long alpha helix with an extracellular NH2 terminal and intracellular COOH terminal (Geoffrey W. Abbott, 2016). See Figure 1.2.3 for an example of KCNE proteins, which is KCNE1 structure (PDB ID: 2K21).

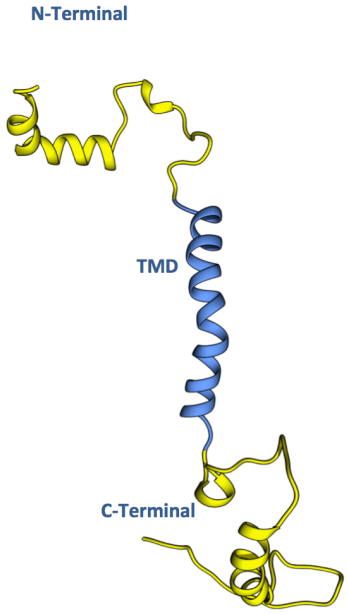


Figure 1.2.3. Structure of the KCNE1 protein (PDB ID: 2K21). The structure has been modified to have a clear distinction of the transmembrane domain (TMD) amino terminal (N-Terminal) and the carboxy terminal (C-Terminal) as opposed to the original NMR structure, which is bent in micellar form.

In the heart, KCNQ1 is co-expressed with KCNE1 (also called minK) and the result is a positive shift in voltage dependence of activation, slow activation and overall modulation of physiology (Geoffrey W. Abbott, 2016; Barro-Soria et al., 2014; J. Chen, Zheng, Melman, & McDonald, 2009; Melman, Krummerman, & McDonald, 2002; Melman et al., 2004; Nakajo & Kubo, 2015; X. Xu, Jiang, Hsu, Zhang, & Tseng, 2008). Although the cardiac KCNQ1 channel on its own is functional, its biophysical properties are dramatically transformed once associated with its beta subunit, KCNE1. These transformations include an increase in channel conductance, slower activation and opening at lower voltages, positive shift in voltage dependence and slowing the time course of deactivation. Overall, the presence of KCNE1 is essential for the native I_{KS} current formation (Osteen, Sampson, & Kass, 2010).

Other KCNE beta subunits i.e. KCNE2-5 also associate with KCNQ1. KCNE2 has a very similar effect in the myocytes to KCNE1, i.e. it modulates the current properties in the same way. Mutations in KCNE1 and KCNE2 proteins can also lead to the fatal LQTS by affecting current characteristics, voltage dependence and channel deactivation (Splawki, Tristani-Firouzi, Lehmann, Sanguinetti, & Keating, 1997; Splawski et al., 2000). Although these mutations are rare, they shed light on the significance of these auxiliary proteins in the heart physiology and KCNQ1 functioning. KCNE3 is associated with KCNQ1 in the colonic epithelia and produces an interesting effect of constitutive activation regardless of the changes in membrane potential (Schroeder et al., 2000). KCNE4 is suggested to have an inhibitory effect on the function of KCNQ1 channels (Grunnet et al., 2002). On the other hand, KCNE5 is shown to have an effect similar to KCNE1 on the channel properties, but in different sites (Angelo et al., 2002).

ii) PIP_2 regulation

Phosphatidylinositol-4,5-bisphosphate (PIP_2) is another regulatory component, necessary for maintaining the activity of KCNQ1/KCNE1 complex. PIP_2 is the major polyphosphoinositide in mammalian cells, which is known to interact with proteins of different kind. It has several regulatory roles in many cellular processes and occasionally acts as a second messenger or cofactor in the cells (McLaughlin, Wang, Gambhir, & Murray, 2002). Other voltage-gated as well as ligand-gated ion channels are also regulated and modulated by their interactions with PIP_2 (Hansen, 2015). Some examples include the inwardly rectifying K^+ (K_{ir}) channels, hERG ion channel, transient receptor potential (TRP) channels, and ion transporters such as the Na^+- Ca^{2+} exchanger (Suh & Hille, 2008).

In KCNQ1 channels, PIP_2 stabilizes the open state of the pore domain (PD), thereby slowing the deactivation kinetics and shifting the activation curve to more negative potentials (Eckey et al., 2014). PIP_2 is also identified as a required factor for coupling and communication between the VSD and PD in KCNQ1 (Zaydman et al., 2013). Furthermore, several of the reported LQTS-associated mutations and loss of function mutations, affect basic residues in the PIP_2 binding pocket, implicating the importance of

this lipid in the normal functioning of the KCNQ1/KCNE1 complex (Dvir, Peretz, Haitin, & Attali, 2014; Sachyani et al., 2013).

iii) Calmodulin regulation

Calmodulin (CaM), a Ca^{2+} binding protein is recognized as an obligate regulatory subunit for many ion channels. Ca^{2+} being a fundamental second messenger of electrical activity in excitable cells, creates a feedback modulation mediated by CaM (S. Ghosh, Nunziato, & Pitt, 2006).

The KCNQ1 C-terminal comprises of amphipathic α helices, coiled-coils and clusters of basic amino acids that form the CaM binding sites (Haitin & Attali, 2008). CaM acts as a regulator of channel gating, modulates the current amplitude and highly contributes to the folding, trafficking and assembly of these tetrameric proteins (Roden, 2006; Shamgar et al., 2006; Yus-Nájera, Santana-Castro, & Villarroel, 2002). LQTS mutants that disrupt CaM interaction and binding to KCNQ1 protein, prevent functional assembly of channels in a dominant-negative manner and thereby alter function (S. Ghosh et al., 2006).

1.2.6. Summary

This section provided the basic background related to the human cardiac voltage-gated KCNQ1 potassium channel, which serves as the focus of this thesis. A complete understanding of the complex KCNQ1 properties, function and structure is necessary to understand its importance in controlling the cardiac action potential. KCNQ1 generates the I_{KS} current in the heart. The disruption of I_{KS} can lead to several forms of irregularities in the heart's function, as seen in congenital JLNS, LQTS and cardiac arrhythmias. The pathogenesis of several heart diseases in which KCNQ1 plays a role, are still lacking an explanation and cure. This adds to the significance of this ion channel and emphasizes the need for further studies. KCNQ1 is regulated by several cellular components including accessory KCNE proteins, calmodulin (CaM) and PIP_2 lipids. For a complete understanding of KCNQ1 function and working mechanism, one needs to consider each of these regulatory components as well. The following chapters will

describe how the different properties and aspects of this mysterious ion channel are addressed in this thesis, through the modeling of the KCNQ1KCNE1 ion channel.

1.3: CURRENT STATE OF KCNQ1 MODELLING AND STRUCTURES

In this Thesis, we modeled the human KCNQ1 potassium ion channel using several computational tools and techniques, based on the most recent experimental data. Prior to presenting the findings of this thesis, a brief literature review on what has been accomplished to date in this area is presented in this chapter from both modelling and structural perspectives. The objective of this chapter is to provide the reader with a comprehensive literature review on this area of research and illustrate what this Thesis attempts to fulfill.

1.3.1. Structural Information: A Prerequisite

Atomic resolution three-dimensional (3D) structure of an ion channel is the key to understanding its fundamental structural and functional properties. These include information about the structural components, possible interacting proteins, evolutionary relationships, drug-binding sites, molecular mechanisms of ion transport and the overall function (Y. Jiang et al., 2002; Sokolova, Kolmakova-Partensky, & Grigorieff, 2001; Wiener, Haitin, Shamgar, Fernandez-Alonso, et al., 2008). When detailed structural data for an ion channel are available, computational techniques and algorithms can further be used to predict several physiologically relevant mechanisms (Beckstein et al., 2003; Corry, 2015; Sansom et al., 2002). However, there is a scarcity of structural data for almost all ion channels with very few known 3D structures.

With ion channels and membrane proteins in general, 3D structural determination using the main experimental techniques of X-ray crystallography, NMR spectroscopy and recently, electron microscopy has always been a big challenge (Carpenter, Beis, Cameron, & Iwata, 2008). KCNQ1 ion channel suffers from a similar scenario; such that it has no complete experimentally derived 3D structure. However, three PDB entries exist for this protein in the PDB database. These are 3BJ4 (Wiener, Haitin, Shamgar, Fernández-Alonso, et al., 2008), 3HFC (Q. Xu & Minor, 2009), 4UMO (Sachyani et al., 2013), which belong to partial sections of the C-terminal intracellular domains only and do not represent any part of the membrane spanning domains (see Figure 1.3.1). Although they provide useful information related to trafficking and assembly (Haitin et al., 2009; Haitin & Attali, 2008; Howard, Clark, Holton, & Minor, 2007; Shamgar et al.,

2006; Wiener, Haitin, Shamgar, Fernández-Alonso, et al., 2008), they are not enough to study the main properties of this protein, i.e. ion transport mechanisms, selectivity behaviour, the coupling of membrane components, protein interactions, drug binding, etc. In June 2017, a cryo-EM structure for the closed conformation of frog KCNQ1 has been published (PDB ID:5VMS), which can offer a great advancement in this area for future research. We will discuss this structure and its relevance to this work later on in this Thesis.

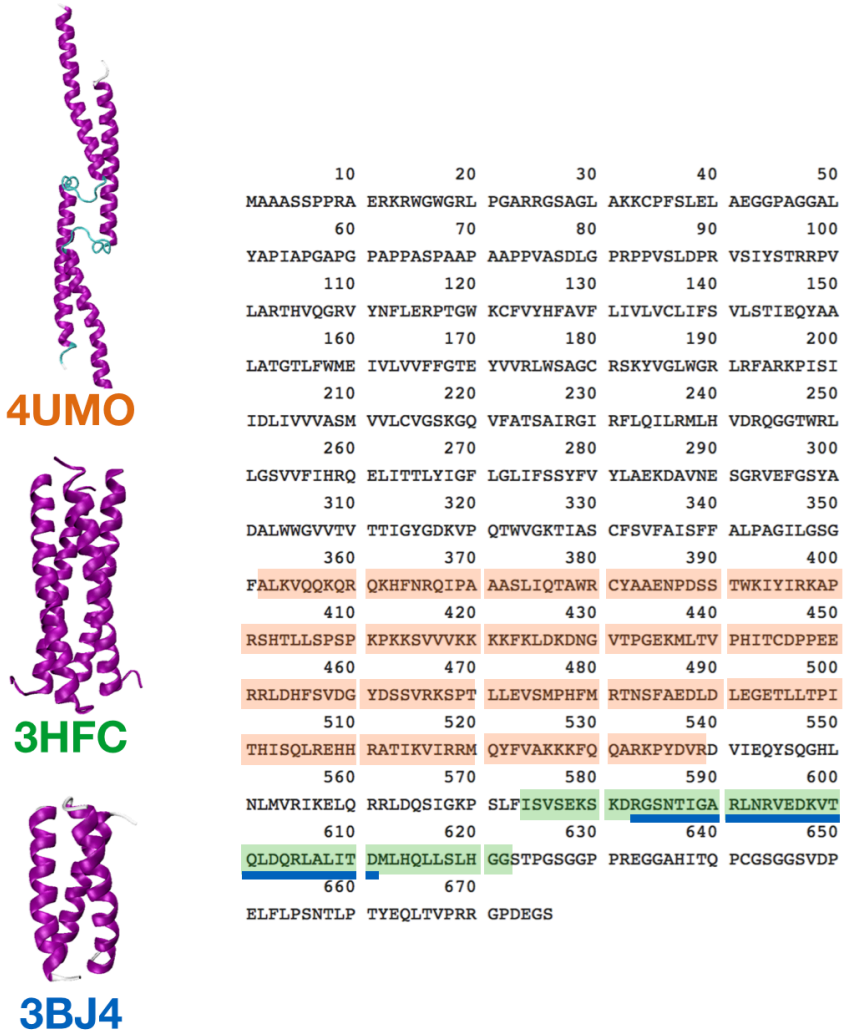


Figure 1.3.1. The structure and sequence of the three PDB entries (PDB IDs: 4UMO, 3HFC, 3BJ4) related to KCNQ1 protein. The three structures belong to partial sections of the C-terminal and are highlighted on the sequence of KCNQ1 to show the coverage.

To compensate the need for 3D structures, the scientific community have applied other techniques of protein structure determination. The homology modeling approach has become one of the crucial techniques for obtaining 3D structures of carrier proteins and ion channels (see Appendix B for more details on Homology Modeling), when the complete experimental structure of the protein of interest is not available (Ravna & Sylte, 2012). Homology modeling involves the prediction of the 3D structure of a given protein based primarily on its sequence similarity to one or more proteins of already known structures (Krieger, Nabuurs, & Vriend, 2003). Similarly, the major structural studies done so far on KCNQ1 ion channel also have relied on *in silico* structure prediction techniques. Some of these modeling attempts will be briefly reviewed in the subsequent section.

1.3.2. Modeling Studies

Modeling attempts for the KCNQ1 ion channel began with the emergence of the very first X-ray crystal structure of the KcsA potassium ion channel (PDB ID: 1K4C (Zhou, Morais-Cabral, Kaufman, & MacKinnon, 2001)), which could be used as a template for building a model. In 2005, Du et al. (Du, Li, Tsai, You, & Xia, 2005) used this structure as a template (36% sequence similarity) for homology modeling of KCNQ1. Using this model, they were able to characterize the binding of small molecule blockers and build a pharmacophore model for KCNQ1 blockers. As can be seen from Figure 1.3.2, despite the 36% sequence similarity, this template is structurally and topologically different from KCNQ1. Furthermore, neither of the regulatory components of KCNQ1, i.e. KCNE1 and PIP_2 were included in their model.

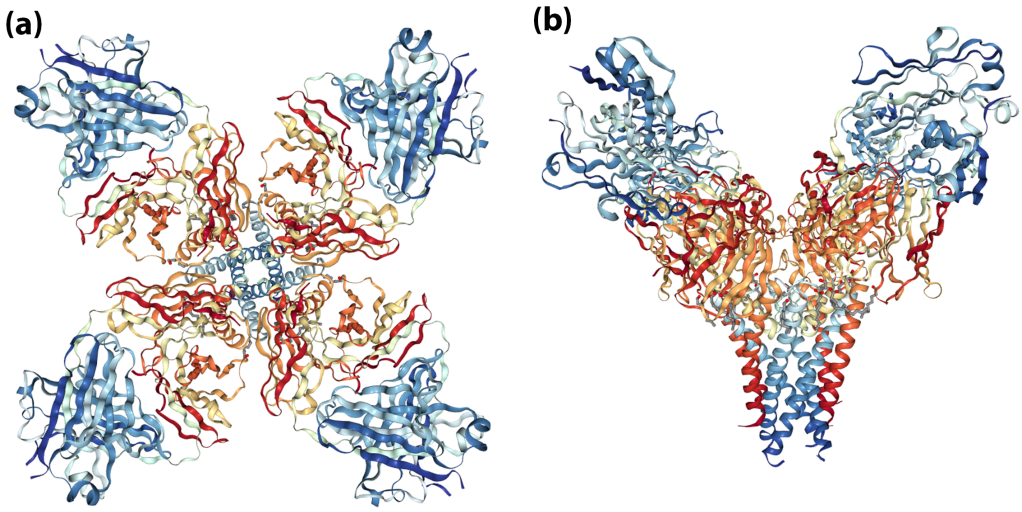


Figure 1.3.2. Structure of Potassium Channel KcsA-Fab complex in high concentration of K^+ (PDB ID: 1K4C). (a) top view, (b) lateral view.

The tendency of modeling KCNQ1 using homology modeling has continued through years as other crystal or NMR structures with higher resolution and more similar sequences became available. Currently, the prototype modeling template for the open conformation of KCNQ1 is the $K_V1.2$-$K_V2.1$ paddle chimera channel (PDB ID: 2R9R (Long, Tao, Campbell, & MacKinnon, 2007)). For the closed state, there was no experimentally derived structure to be used as the template. However, the closed $K_V1.2$ model built by Yarov-Yarovy et al. (Yarov-Yarovoy, Baker, & Catterall, 2006) has frequently been employed as the template for building models for the closed conformation of KCNQ1 (Lvov et al., 2010; J. a Smith, Vanoye, Jr, Meiler, & Sanders, 2007; Van Horn, Vanoye, & Sanders, 2011; Y. Xu et al., 2013a). This model was developed using the Rosetta membrane method, combining homology and de novo structure prediction techniques (Yarov-Yarovoy et al., 2006).

The models and corresponding studies of Smith et al. (J. a Smith et al., 2007) in 2007, mark the subsequent fundamental work towards accurate structural studies of KCNQ1. In this study, models for both open and closed channel conformations were built, using the open $K_V1.2$ structure (PDB ID: 2A79 (Long, Campbell, & MacKinnon, 2005)) and the $K_V1.2$ closed state model of Yarov-Yarovy et al. (Yarov-Yarovoy et al., 2006), as templates. The models were then used to investigate disease-related mutations,

their location and their correlation with gating properties of the channel. However, one cannot regard their models of KCNQ1 as complete ones due to the reasons stated here. First, a template better than PDB ID 2A79 has been published after their work (i.e. PDB ID 2R9R) for the open state of KCNQ1. Their closed model, has been built using an *ab initio* protein modelling approach and it remains vague as to how accurate is this modeling strategy. Also, this study did not implicate any details related to the presence or absence of KCNE1 and its influence on the models for KCNQ1.

In 2008, along with the publication of the NMR structure of KCNE1 (Kang et al., 2008) for the first time, a new trend of structural studies for the KCNQ1/KCNE1 complex began. These included several *in silico* modeling combined with experimental studies to shed more light on the interaction between the two proteins (Chan, 2011; Lvov et al., 2010; Van Horn et al., 2011). (Du et al., 2005; Kang et al., 2008; Kasimova, Zaydman, Cui, & Tarek, 2015; T. Yang et al., 2013) are several other studies that have investigated KCNE1 interaction as well as drug binding properties of KCNQ1 ion channel. However, many other related aspects remained vague. These included the precise effect of KCNE1 on KCNQ1 ion channel in terms of structure and function as well as role of PIP_2 on KCNQ1.

One of the most outstanding modeling studies was done by, Xu et al. (Y. Xu et al., 2013b) in 2013. They built homology models of the KCNQ1 in both open and closed conformations. Using molecular dynamics simulations, they were able to trace the interactions between KCNQ1 and KCNE1. Several pieces of experimental information were included in their modeling strategies and their findings were confirmed by voltage-clamp experiments (Y. Xu et al., 2013a). In a subsequent study, they used their models to study drug binding properties in the KCNQ1/KCNE1 complex (Y. Xu et al., 2015). However, these studies lacked several important aspects; they did not include the PIP_2 lipid in the modeled systems, PIP_2 has a significant influence on the structural and functional properties of the KCNQ1/KCNE1 channel complex; also, for their closed state model, they used an *ab initio* model which definitely lacks the precision of an X-ray or NMR structure and questions the accuracy of the structural findings. And lastly, the experimental structural information, which was offered by several NMR studies, were not addressed and included in their study.

In 2015, Kasimova et al. (Kasimova et al., 2015) attempted to investigate the effect of PIP_2 on the KCNQ1. Their findings were interesting as their structural modeling and simulations confirmed that the effect of PIP_2 on KCNQ1 is quite substantial. However, one large gap in this study was that there was no mention of KCNE1 and whether the influence of PIP_2 on KCNQ1 remains the same if it is in complex with KCNE1 (Kasimova et al., 2015).

The information provided by these studies has been invaluable in answering many questions related to the structure and function of the KCNQ1/KCNE1 ion channel complex. However, as outlined above there are gaps associated with each of them. Furthermore, any implementation of modeling techniques, are associated with the assumptions that may lead to uncertainties in the final results. One way of ensuring an accuracy of structural models is to continually update them by including reliable experimental data as they become available. These pieces of information are essential to supplement and guarantee the reliability of *in silico* research. Accordingly, in a search of such studies, we came across NMR studies (Gayen, Li, & Kang, 2015; D. Peng et al., 2014), in which structural studies of the KCNQ1 VSD had been done. In this paper, the authors reported novel structural properties for the KCNQ1 ion channel which is unique amongst the homologous members of the K_V channel superfamily (D. Peng et al., 2014). This information was not extensively included and investigated in any of the modeling attempts, so far.

An inclusive description of the gaps that we identified in the literature and addressed in the thesis is discussed in the following section.

1.4. References

Abbott, G. W. (2014). Biology of the KCNQ1 Potassium Channel. *New Journal of Science*, *2014*, 1–26. http://doi.org/10.1155/2014/237431

Abbott, G. W. (2016). KCNE1 and KCNE3: The yin and yang of voltage-gated K+ channel regulation. *Gene*, *576*(1), 1–13. http://doi.org/10.1016/j.gene.2015.09.059

Abbott, G. W., & Goldstein, S. A. (2001). Potassium channel subunits encoded by the KCNE gene family: physiology and pathophysiology of the MinK-related peptides (MiRPs). *Molecular Interventions*, *1*(2), 95–107.

Abriel, H. (2007). Roles and regulation of the cardiac sodium channel Nav1.5: Recent insights from experimental studies. *Cardiovascular Research*, *76*(3), 381–389. http://doi.org/10.1016/j.cardiores.2007.07.019

Accardi, A., & Pusch, M. (2000). Fast and Slow Gating Relaxations in the Muscle Chloride Channel Clc-1. *The Journal of General Physiology*, *116*(3), 433–444. http://doi.org/10.1085/jgp.116.3.433

Ahern, C. A., Payandeh, J., Bosmans, F., & Chanda, B. (2016). The hitchhiker's guide to the voltage-gated sodium channel galaxy. *The Journal of General Physiology*, *147*(1), 1–24. http://doi.org/10.1085/jgp.201511492

Aidley, D. J., & Stanfield, P. R. (1996). *Ion Channels—Molecules in Action*. *Cambridge University Press*.

Aiyar, J., Rizzi, J. P., Gutman, G. A., & Chandy, K. G. (1996). The signature sequence of voltage-gated potassium channels projects into the external vestibule. *The Journal of Biological Chemistry*, *271*(49), 31013–31016.

An, W. F., Bowlby, M. R., Betty, M., Cao, J., Ling, H.-P., Mendoza, G., … Rhodes, K. J. (2000). Modulation of A-type potassium channels by a family of calcium sensors. *Nature*, *403*(6769), 553–556. Retrieved from http://dx.doi.org/10.1038/35000592

Angelo, K., Jespersen, T., Grunnet, M., Nielsen, M. S., Klaerke, D. A., & Olesen, S.-P. (2002). KCNE5 induces time- and voltage-dependent modulation of the KCNQ1 current. *Biophysical Journal*, *83*(4), 1997–2006. http://doi.org/10.1016/S0006-3495(02)73961-1

Anwar-Mohamed, A., Barakat, K. H., Bhat, R., Noskov, S. Y., Tyrrell, D. L., Tuszynski, J. a., & Houghton, M. (2014). A human ether-á-go-go-related (hERG) ion channel atomistic model generated by long supercomputer molecular dynamics simulations and its use in predicting drug cardiotoxicity. *Toxicology Letters*, *230*(3), 382–392. http://doi.org/10.1016/j.toxlet.2014.08.007

Arévalo, J. C. (2015). Nedd4-2 regulation of voltage-gated ion channels: An update on structure, function relationships and the pathophysiological consequences of dysfunction. *Journal of Receptor, Ligand and Channel Research*, *8*, 53–63. http://doi.org/10.2147/JRLCR.S52534

Ashcroft, F. M. (2000). Chapter 3 – How ion channels work. In *Ion Channels and Disease* (pp. 21–41). http://doi.org/10.1016/B978-012065310-2/50004-7

Ashcroft, F. M., & Ashcroft, F. M. (2000a). Chapter 10 – Voltage-gated Cl− channels. In *Ion Channels and Disease* (pp. 185–198). http://doi.org/10.1016/B978-012065310-2/50011-4

Ashcroft, F. M., & Ashcroft, F. M. (2000b). Chapter 1 – Introduction. In *Ion Channels and Disease* (pp. 1–2). http://doi.org/10.1016/B978-012065310-2/50002-3

Ashcroft, F. M., & Ashcroft, F. M. (2000c). Chapter 2 – From gene to protein. In *Ion Channels and Disease* (pp. 3–19). http://doi.org/10.1016/B978-012065310-2/50003-5

Ashcroft, F. M., & Ashcroft, F. M. (2000d). Chapter 5 – Voltage-gated Na+ channels. In *Ion Channels and Disease* (pp. 67–96). http://doi.org/10.1016/B978-012065310-2/50006-0

Ashcroft, F. M., & Ashcroft, F. M. (2000e). Chapter 6 – Voltage-gated K+ channels. In *Ion Channels and Disease* (p. 97–II). http://doi.org/10.1016/B978-012065310-2/50007-2

Ashcroft, F. M., & Ashcroft, F. M. (2000f). Chapter 8 – Inwardly rectifying K+ channels. In *Ion Channels and Disease* (pp. 135–159). http://doi.org/10.1016/B978-012065310-2/50009-6

Bagal, S. K., Marron, B. E., Owen, R. M., Storer, R. I., & Swain, N. A. (2015). Voltage gated sodium channels as drug discovery targets. *Channels*, *9*(6), 360–366. http://doi.org/10.1080/19336950.2015.1079674

Barhanin, J., Lesage, F., Guillemare, E., Fink, M., Lazdunski, M., & Romey, G. (1996). K(V)LQT1 and lsK (minK) proteins associate to form the I(Ks) cardiac potassium current. *Nature*, *384*(6604), 78–80. http://doi.org/10.1038/384078a0

Barro-Soria, R., Rebolledo, S., Liin, S. I., Perez, M. E., Sampson, K. J., Kass, R. S., & Larsson, H. P.

(2014). KCNE1 divides the voltage sensor movement in KCNQ1/KCNE1 channels into two steps. *Nature Communications*, *5*, 3750. http://doi.org/10.1038/ncomms4750

Bean, B. P. (2007). The action potential in mammalian central neurons. *Nat Rev Neurosci*, *8*(6), 451–465. Retrieved from http://dx.doi.org/10.1038/nrn2148

Beckstein, O., Biggin, P. C., Bond, P., Bright, J. N., Domene, C., Grottesi, A., … Sansom, M. S. P. (2003). Ion channel gating: Insights via molecular simulations. *FEBS Letters*, *555*(1), 85–90. http://doi.org/10.1016/S0014-5793(03)01151-7

Belardetti, F., & Zamponi, G. W. (2012). Calcium channels as therapeutic targets. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, *1*(4), 433–451. http://doi.org/10.1002/wmts.38

Bers, D. M. (2002). Cardiac excitation-contraction coupling. *Nature*, *415*(6868), 198–205. Retrieved from http://dx.doi.org/10.1038/415198a

Bezanilla, F. (2007). Voltage-Gated Ion Channels. In S.-H. Chung, O. S. Andersen, & V. Krishnamurthy (Eds.), *Biological Membrane Ion Channels: Dynamics, Structure, and Applications* (pp. 81–118). New York, NY: Springer New York. http://doi.org/10.1007/0-387-68919-2_3

Bidaud, I., Mezghrani, A., Swayne, L. A., Monteil, A., & Lory, P. (2006). Voltage-gated calcium channels in genetic diseases. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1763*(11), 1169–1174. http://doi.org/https://doi.org/10.1016/j.bbamcr.2006.08.049

Boehmer, C., Laufer, J., Jeyaraj, S., Klaus, F., Lindner, R., Lang, F., & Palmada, M. (2008). Modulation of the voltage-gated potassium channel Kv1.5 by the SGK1 protein kinase involves inhibition of channel ubiquitination. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *22*(5–6), 591–600. http://doi.org/10.1159/000185543

Boland, L. M., & Jackson, K. A. (1999). Protein kinase C inhibits Kv1.1 potassium channel function. *The American Journal of Physiology*, *277*(1 Pt 1), C100-10.

Brackenbury, W., & Isom, L. (2011). Na+ Channel β Subunits: Overachievers of the Ion Channel Family . *Frontiers in Pharmacology* . Retrieved from http://journal.frontiersin.org/article/10.3389/fphar.2011.00053

Braun, M., Ramracheya, R., Bengtsson, M., Zhang, Q., Karanauskaite, J., Partridge, C., … Rorsman, P. (2008). Voltage-gated ion channels in human pancreatic beta-cells: electrophysiological characterization and role in insulin secretion. *Diabetes*, *57*(6), 1618–1628. http://doi.org/10.2337/db07-0991

Buraei, Z., & Yang, J. (2010). The ss subunit of voltage-gated Ca2+ channels. *Physiological Reviews*, *90*(4), 1461–1506. http://doi.org/10.1152/physrev.00057.2009

Campuzano, O., Beltran-Alvarez, P., Iglesias, A., Scornik, F., Perez, G., & Brugada, R. (2010). Genetics and cardiac channelopathies. *Genet Med*, *12*(5), 260–267. Retrieved from http://dx.doi.org/10.1097/GIM.0b013e3181d81636

Carpenter, E. P., Beis, K., Cameron, A. D., & Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*, *18*(5), 581–586. http://doi.org/10.1016/j.sbi.2008.07.001

Cascio, M. (2004). Structure and function of the glycine receptor and related nicotinicoid receptors. *The Journal of Biological Chemistry*, *279*(19), 19383–19386. http://doi.org/10.1074/jbc.R300035200

Catterall, W. (2015). Voltage-Gated Sodium Channels. In *Handbook of Ion Channels* (pp. 213–226). CRC Press. http://doi.org/doi:10.1201/b18027-20

Catterall, W. a. (2012). Voltage-gated sodium channels at 60: structure, function and pathophysiology. *The Journal of Physiology*, *590*(11), 2577–89. http://doi.org/10.1113/jphysiol.2011.224204

Catterall, W. A. (2005). Voltage-Gated Calcium Channels. *Molecular Biology Intelligence Unit*, 1–373. http://doi.org/10.1016/B978-0-12-374145-5.00112-1

Chan, P. J. (2011). Functional and Biochemical Characterization of KCNQ1/KCNE1 Subunit Interactions in the Cardiac I.

Chen, J., Zheng, R., Melman, Y. F., & McDonald, T. V. (2009). Functional interactions between KCNE1 C-terminus and the KCNQ1 channel. *PLoS ONE*, *4*(4), 1–9. http://doi.org/10.1371/journal.pone.0005143

Chowdhury, S., & Chanda, B. (2015). Basic mechanisms of voltage sensing. *Handbook of Ion Channels*, 25–40. http://doi.org/doi:10.1201/b18027-5

Clapham, D. E. (2007). Calcium Signaling. *Cell*, *131*(6), 1047–1058. http://doi.org/http://dx.doi.org/10.1016/j.cell.2007.11.028

Corry, B. (2015). Long Timescale Molecular Simulations for Understanding Ion Channel Function. In *Pumps, Channels, and Transporters* (pp. 411–441). John Wiley & Sons, Inc. http://doi.org/10.1002/9781119085126.ch15

Cui, J. (2016). Voltage-Dependent Gating: Novel Insights from KCNQ1 Channels. *Biophysical Journal*, *110*(1), 14–25. http://doi.org/10.1016/j.bpj.2015.11.023

Demolombe, S., Baro, I., Pereon, Y., Bliek, J., Mohammad-Panah, R., Pollard, H., … Escande, D. (1998). A dominant negative isoform of the long QT syndrome 1 gene product. *The Journal of Biological Chemistry*, *273*(12), 6837–6843.

Dib-Hajj, S., & Priestley, T. (2010). Chapter 2.3: Voltage-gated sodium channels. In J. N. C. Kew & C. H. Davies (Eds.), *Ion channels: from structure to function*. New York: Oxford University Press.

Dolphin, A. C. (2009). Calcium channel diversity: multiple roles of calcium channel subunits. *Current Opinion in Neurobiology*, *19*(3), 237–244. http://doi.org/http://dx.doi.org/10.1016/j.conb.2009.06.006

Du, L. P., Li, M. Y., Tsai, K. C., You, Q. D., & Xia, L. (2005). Characterization of binding site of closed-state KCNQ1 potassium channel by homology modeling, molecular docking, and pharmacophore identification. *Biochemical and Biophysical Research Communications*, *332*(3), 677–687. http://doi.org/10.1016/j.bbrc.2005.04.165

Dvir, M., Peretz, A., Haitin, Y., & Attali, B. (2014). Recent molecular insights from mutated IKS channels in cardiac arrhythmia. *Current Opinion in Pharmacology*, *15*, 74–82. http://doi.org/http://dx.doi.org/10.1016/j.coph.2013.12.004

Dworakowska, B., & Dolowy, K. (2000). Ion channels-related diseases. *Acta Biochimica Polonica*, *47*(3), 685–703.

Eckey, K., Wrobel, E., Strutz-Seebohm, N., Pott, L., Schmitt, N., & Seebohm, G. (2014). Novel Kv7.1-Phosphatidylinositol 4,5-Bisphosphate Interaction Sites Uncovered by Charge Neutralization Scanning. *Journal of Biological Chemistry*, *289*(33), 22749–22758. http://doi.org/10.1074/jbc.M114.589796

Feranchak, A. P. (2003). Ion Channels in Digestive Health and Disease. *Journal of Pediatric Gastroenterology and Nutrition*, *37*(3), 230–241. http://doi.org/10.1097/00005176-200309000-00006

FLORESCU, M., CINTEZA, M., & VINEREANU, D. (2013). Chemotherapy-induced Cardiotoxicity. *Mædica*, *8*(1), 59–67. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3749765/

Gabashvili, I. S., Sokolowski, B. H. A., Morton, C. C., & Giersch, A. B. S. (2007, September). Ion Channel Gene Expression in the Inner Ear. *JARO: Journal of the Association for Research in Otolaryngology*. New York. http://doi.org/10.1007/s10162-007-0082-y

Galligan, J. J. (2002). Ligand-gated ion channels in the enteric nervous system. *Neurogastroenterology and Motility : The Official Journal of the European Gastrointestinal Motility Society*, *14*(6), 611–623.

Gao, N., Lu, M., Echeverri, F., Laita, B., Kalabat, D., Williams, M. E., … Moyer, B. D. (2009). Voltage-gated sodium channels in taste bud cells. *BMC Neuroscience*, *10*(1), 20. http://doi.org/10.1186/1471-2202-10-20

Gayen, S., Li, Q., & Kang, C. (2015). Structural analysis of the S4–S5 linker of the human KCNQ1 potassium channel. *Biochemical and Biophysical Research Communications*, *456*(1), 410–414. http://doi.org/10.1016/j.bbrc.2014.11.097

Gez, L. S., Hagalili, Y., Shainberg, A., & Atlas, D. (2012). Voltage-driven Ca(2+) binding at the L-type Ca(2+) channel triggers cardiac excitation-contraction coupling prior to Ca(2+) influx. *Biochemistry*, *51*(48), 9658–9666. http://doi.org/10.1021/bi301124a

Ghosh, S., Nunziato, D. a., & Pitt, G. S. (2006). KCNQ1 assembly and function is blocked by long-QT syndrome mutations that disrupt interaction with calmodulin. *Circulation Research*, *98*(8), 1048–1054. http://doi.org/10.1161/01.RES.0000218863.44140.f2

Grant, A. O. (2009). Cardiac ion channels. *Circulation: Arrhythmia and Electrophysiology*, *2*(2), 185–194. http://doi.org/10.1161/CIRCEP.108.789081

Grunnet, M., Jespersen, T., Rasmussen, H. B., Ljungstrøm, T., Jorgensen, N. K., Olesen, S.-P., & Klaerke, D. A. (2002). KCNE4 is an inhibitory subunit to the KCNQ1 channel. *The Journal of Physiology*, *542*(Pt 1), 119–130. http://doi.org/10.1113/jphysiol.2002.017301

Gutman, G. A., Chandy, K. G., Grissmer, S., Lazdunski, M., McKinnon, D., Pardo, L. A., … Wang, X. (2005). International Union of Pharmacology. LIII. Nomenclature and molecular relationships of voltage-gated potassium channels. *Pharmacological Reviews*, *57*(4), 473–508. http://doi.org/10.1124/pr.57.4.10.1

Haitin, Y., & Attali, B. (2008). The C-terminus of Kv7 channels: a multifunctional module. *The Journal of*

Physiology, *586*(7), 1803–1810. http://doi.org/10.1113/jphysiol.2007.149187
Haitin, Y., Wiener, R., Shaham, D., Peretz, A., Cohen, E. B.-T., Shamgar, L., … Attali, B. (2009). Intracellular domains interactions and gated motions of I(KS) potassium channel subunits. *The EMBO Journal*, *28*(14), 1994–2005. http://doi.org/10.1038/emboj.2009.157
Halling, D. B., Aracena-Parks, P., & Hamilton, S. L. (2005). Regulation of voltage-gated Ca2+ channels by calmodulin. *Science's STKE : Signal Transduction Knowledge Environment*, *2005*(315), re15. Journal Article, Review. http://doi.org/10.1126/stke.3152005re15
Hansen, S. B. (2015). Lipid agonism: The PIP2 paradigm of ligand-gated ion channels. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1851*(5), 620–628. http://doi.org/http://dx.doi.org/10.1016/j.bbalip.2015.01.011
Hausdorff, S. F., Goldstein, S. A. N., Rushin, E. E., & Miller, C. (1991). Functional characterization of a minimal potassium channel expressed from a synthetic gene. *Biochemistry*, *30*(13), 3341–3346. http://doi.org/10.1021/bi00227a025
Heyes, S., Pratt, W. S., Rees, E., Dahimene, S., Ferron, L., Owen, M. J., & Dolphin, A. C. (2015). Genetic disruption of voltage-gated calcium channels in psychiatric and neurological disorders. *Progress in Neurobiology*, *134*, 36–54. http://doi.org/10.1016/j.pneurobio.2015.09.002
Hodgkin, A. L., & Huxley, A. F. (1990). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Bulletin of Mathematical Biology*, *52*(1–2), 25–71. http://doi.org/10.1007/BF02459568
Hofmann, F., Lacinová, L., & Klugbauer, N. (1999). Voltage-dependent calcium channels: From structure to function BT - Reviews of Physiology, Biochemistry and Pharmacology, Volume 139 (pp. 33–87). CHAP, Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/BFb0033648
Hollenhorst, M. I., Richter, K., & Fronius, M. (2011). Ion transport by pulmonary epithelia. *Journal of Biomedicine and Biotechnology*, *2011*. http://doi.org/10.1155/2011/174306
Honoré, E., Attali, B., Romey, G., Heurteaux, C., Ricard, P., Lesage, F., … Barhanin, J. (1991). Cloning, expression, pharmacology and regulation of a delayed rectifier K+ channel in mouse heart. *The EMBO Journal*, *10*(10), 2805–2811. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC452989/
Howard, R. J., Clark, K. a., Holton, J. M., & Minor, D. L. (2007). Structural Insight into KCNQ (Kv7) Channel Assembly and Channelopathy. *Neuron*, *53*(5), 663–675. http://doi.org/10.1016/j.neuron.2007.02.010
Hucho, F., & Weise, C. (2001). Ligand-Gated Ion Channels. *Angewandte Chemie International Edition*, *40*(17), 3100–3116. http://doi.org/10.1002/1521-3773(20010903)40:17<3100::AID-ANIE3100>3.0.CO;2-A
Isom, L. L., De Jongh, K. S., & Catterall, W. A. (1994). Auxiliary subunits of voltage-gated ion channels. *Neuron*, *12*(6), 1183–1194. http://doi.org/10.1016/0896-6273(94)90436-7
Jackson, W. F. (2000, January). Ion Channels and Vascular Tone. *Hypertension*.
Jentsch, T. J. (2000). Neuronal KCNQ potassium channels: physiology and role in disease. *Nature Reviews. Neuroscience*, *1*(1), 21–30. http://doi.org/10.1038/35036198
Jentsch, T. J., Stein, V., Weinreich, F., & Zdebik, A. A. (2002). Molecular structure and physiological function of chloride channels. *Physiological Reviews*, *82*(2), 503–568. http://doi.org/10.1152/physrev.00029.2001
Jespersen, T., Grunnet, M., & Olesen, S.-P. (2005). The KCNQ1 potassium channel: from gene to physiological function. *Physiology (Bethesda, Md.)*, *20*, 408–416. http://doi.org/10.1152/physiol.00031.2005
Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T., & MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. *Nature*, *417*(6888), 515–522. Retrieved from http://dx.doi.org/10.1038/417515a
Jost, N., Papp, J. G., & Varro´, A. (2007). Slow Delayed Rectifier Potassium Current (I Ks) and the Repolarization Reserve. *?Ane*, *12*(36), 64–78.
Jost, N., Papp, J. G., & Varro, A. (2007). Slow delayed rectifier potassium current (IKs) and the repolarization reserve. *Annals of Noninvasive Electrocardiology : The Official Journal of the International Society for Holter and Noninvasive Electrocardiology, Inc*, *12*(1), 64–78. http://doi.org/10.1111/j.1542-474X.2007.00140.x
Kaczorowski, G. J., McManus, O. B., Priest, B. T., & Garcia, M. L. (2008). Ion Channels as Drug Targets: The Next GPCRs. *The Journal of General Physiology*, *131*(5), 399–405.

http://doi.org/10.1085/jgp.200709946
Kang, C., Tian, C., Sönnichsen, F. D., Smith, J. A., Meiler, J., George, A. L., … Sanders, C. R. (2008). Structure of KCNE1 and Implications for How It Modulates the KCNQ1 Potassium Channel. *Biochemistry*, *47*(31), 7999–8006. http://doi.org/10.1021/bi800875q
Kasimova, M. a, Zaydman, M. a, Cui, J., & Tarek, M. (2015). PIP2-dependent coupling is prominent in Kv7.1 due to weakened interactions between S4-S5 and S6, 2–10. http://doi.org/10.1038/srep07474
Kass, R. S. (n.d.). The channelopathies: novel insights into molecular and genetic mechanisms of human disease. *The Journal of Clinical Investigation*, *115*(8), 1986–1989. http://doi.org/10.1172/JCI26011
Kim, D., McCoy, J., & Nimigean, C. (2015). Ion selectivity and conductance. *Handbook of Ion Channels*, 13–24. http://doi.org/doi:10.1201/b18027-4
Krieger, E., Nabuurs, S. B., & Vriend, G. (2003). Homology Modeling. In *Structural Bioinformatics* (pp. 509–523). John Wiley & Sons, Inc. http://doi.org/10.1002/0471721204.ch25
Kuang, Q., Purhonen, P., & Hebert, H. (2015). Structure of potassium channels. *Cellular and Molecular Life Sciences*, *72*, 3677–3693. http://doi.org/10.1007/s00018-015-1948-5
Kuo, I. Y., & Ehrlich, B. E. (2012, December). Ion Channels in Renal Disease. *Chemical Reviews*. http://doi.org/10.1021/cr3001077
Lai, H. C., & Jan, L. Y. (2006). The distribution and targeting of neuronal voltage-gated ion channels. *Nat Rev Neurosci*, *7*(7), 548–562. Retrieved from http://dx.doi.org/10.1038/nrn1938
Lehmann-Horn, F., & Jurkat-Rott, K. (1999). Voltage-Gated Ion Channels and Hereditary Disease. *Physiological Reviews*, *79*(4), 1317–1372. Retrieved from http://physrev.physiology.org/content/79/4/1317
Lemoine, D., Jiang, R., Taly, A., Chataigneau, T., Specht, A., & Grutter, T. (2012). Ligand-Gated Ion Channels: New Insights into Neurological Disorders and Ligand Recognition. *Chemical Reviews*, *112*(12), 6285–6318. http://doi.org/10.1021/cr3000829
Long, S. B., Campbell, E. B., & MacKinnon, R. (2005). Crystal Structure of a Mammalian Voltage-Dependent Shaker Family K+ Channel. *Science*, *309*(5736), 897–903. http://doi.org/10.1126/science.1116269
Long, S. B., Tao, X., Campbell, E. B., & MacKinnon, R. (2007). Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature*, *450*(7168), 376–382. http://doi.org/10.1038/nature06265
Lvov, A., Gage, S. D., Berrios, V. M., & Kobertz, W. R. (2010). Identification of a protein-protein interaction between KCNE1 and the activation gate machinery of KCNQ1. *The Journal of General Physiology*, *135*(6), 607–618. http://doi.org/10.1085/jgp.200910386
Maier, S. K. G., Westenbroek, R. E., Schenkman, K. A., Feigl, E. O., Scheuer, T., & Catterall, W. A. (2002). An unexpected role for brain-type sodium channels in coupling of cell surface depolarization to contraction in the heart. *Proceedings of the National Academy of Sciences* , *99*(6), 4073–4078. http://doi.org/10.1073/pnas.261705699
Martinez-Martinez, P., Molenaar, P. C., Losen, M., Stevens, J., Baets, M. H. De, Szoke, A., … Rutten, B. P. F. (2013). Autoantibodies to neurotransmitter receptors and ion channels: from neuromuscular to neuropsychiatric disorders. *Frontiers in Genetics*, *4*, 181. http://doi.org/10.3389/fgene.2013.00181
Mcdonough, S. I. (2013). Calcium ion channels: Challenges and successes in drug discovery. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, *2*(2), 85–104. http://doi.org/10.1002/wmts.71
McLaughlin, S., Wang, J., Gambhir, A., & Murray, D. (2002). PIP2 and Proteins: Interactions, Organization, and Information Flow. *Annual Review of Biophysics and Biomolecular Structure*, *31*(1), 151–175. http://doi.org/10.1146/annurev.biophys.31.082901.134259
Melman, Y. F., Krummerman, A., & McDonald, T. V. (2002). KCNE regulation of KvLQT1 channels: Structure-function correlates. *Trends in Cardiovascular Medicine*, *12*(4), 182–187. http://doi.org/10.1016/S1050-1738(02)00158-5
Melman, Y. F., Um, S. Y., Krumerman, A., Kagan, A., & McDonald, T. V. (2004). KCNE1 binds to the KCNQ1 pore to regulate potassium channel activity. *Neuron*, *42*(6), 927–937. http://doi.org/10.1016/j.neuron.2004.06.001
Miceli, F., Soldovieri, M. V., Ambrosino, P., De Maria, M., Manocchio, L., Medoro, A., & Taglialatela, M. (2015). Molecular pathophysiology and pharmacology of the voltage-sensing module of neuronal ion channels . *Frontiers in Cellular Neuroscience* . Retrieved from http://journal.frontiersin.org/article/10.3389/fncel.2015.00259

Miller, C. (2000). An overview of the potassium channel family. *Genome Biology*, *1*(4), reviews0004.1-reviews0004.5. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC138870/

Minor, D. L. J., & Findeisen, F. (2010). Progress in the structural understanding of voltage-gated calcium channel (CaV) function and modulation. *Channels (Austin, Tex.)*, *4*(6), 459–474. http://doi.org/10.4161/chan.4.6.12867

Modell, S. M., & Lehmann, M. H. (2006). The long QT syndrome family of cardiac ion channelopathies: a HuGE review. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *8*(3), 143–155. http://doi.org/10.109701.gim.0000204468.85308.86

Mousavi Nik, A., Gharaie, S., & Jeong Kim, H. (2015). Cellular mechanisms of mutations in Kv7.1: auditory functions in Jervell and Lange-Nielsen syndrome vs. Romano–Ward syndrome. *Frontiers in Cellular Neuroscience*, *9*, 32. http://doi.org/10.3389/fncel.2015.00032

Nakajo, K., & Kubo, Y. (2015). KCNQ1 channel modulation by KCNE proteins via the voltage-sensing domain. *The Journal of Physiology*, *593*(12), 2617–2625. http://doi.org/10.1113/jphysiol.2014.287672

Nerbonne, J. M., & Kass, R. S. (2005). Molecular physiology of cardiac repolarization. *Physiological Reviews*, *85*(4), 1205–1253. http://doi.org/10.1152/physrev.00002.2005

Osteen, J. D., Sampson, K. J., & Kass, R. S. (2010). The cardiac IKs channel, complex indeed. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(44), 18751–18752. http://doi.org/10.1073/pnas.1014150107

Palmer, L. G. (2007). Ion Channels in Epithelial Cells. In S.-H. Chung, O. S. Andersen, & V. Krishnamurthy (Eds.), *Biological Membrane Ion Channels: Dynamics, Structure, and Applications* (pp. 425–445). New York, NY: Springer New York. http://doi.org/10.1007/0-387-68919-2_12

Park, K.-S., Yang, J.-W., Seikel, E., & Trimmer, J. S. (2008). Potassium Channel Phosphorylation in Excitable Cells: Providing Dynamic Functional Variability to a Diverse Family of Ion Channels. *Physiology*, *23*(1), 49–57. Retrieved from http://physiologyonline.physiology.org/content/23/1/49.abstract

Peng, D., Kim, J., Kroncke, B. M., Law, C. L., Xia, Y., Droege, K. D., … Sanders, C. R. (2014). Purification and Structural Study of the Voltage-Sensor Domain of the Human KCNQ1 Potassium Ion Channel. *Biochemistry*, *53*(12), 2032–2042. http://doi.org/10.1021/bi500102w

Pereon, Y., Demolombe, S., Baro, I., Drouin, E., Charpentier, F., & Escande, D. (2000). Differential expression of KvLQT1 isoforms across the human ventricular wall. *American Journal of Physiology. Heart and Circulatory Physiology*, *278*(6), H1908-15.

Peroz, D., Rodriguez, N., Choveau, F., Baró, I., Mérot, J., & Loussouarn, G. (2008). Kv7.1 (KCNQ1) properties and channelopathies. *The Journal of Physiology*, *586*(7), 1785–1789. http://doi.org/10.1113/jphysiol.2007.148254

Petegem, F. Van, & Minor, D. L. (2006). The structural biology of voltage-gated calcium channel function and regulation. *Biochemical Society Transactions*, *34*(Pt 5), 887–893. http://doi.org/10.1042/BST0340887

Pinnell, J., Turner, S., & Howell, S. (2007). Cardiac muscle physiology. *Continuing Education in Anaesthesia Critical Care & Pain*, *7*(3), 85–88. Retrieved from http://dx.doi.org/10.1093/bjaceaccp/mkm013

Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J. X., & Jensen, L. J. (2015). DISEASES: Text mining and data integration of disease–gene associations. *Methods*, *74*, 83–89. http://doi.org/https://doi.org/10.1016/j.ymeth.2014.11.020

Po, S., Roberds, S., Snyders, D. J., Tamkun, M. M., & Bennett, P. B. (1993). Heteromultimeric assembly of human potassium channels. Molecular basis of a transient outward current? *Circulation Research*, *72*(6), 1326–1336. http://doi.org/10.1161/01.RES.72.6.1326

Proks, P., & Lippiat, J. D. (2006). Membrane ion channels and diabetes. *Current Pharmaceutical Design*, *12*(4), 485–501. Journal Article, Review.

Pusch, M. (2007). Chloride Transporting CLC Proteins1. In S.-H. Chung, O. S. Andersen, & V. Krishnamurthy (Eds.), *Biological Membrane Ion Channels: Dynamics, Structure, and Applications* (pp. 301–333). New York, NY: Springer New York. http://doi.org/10.1007/0-387-68919-2_8

Ravna, A. W., & Sylte, I. (2012). Homology modeling of transporter proteins (carriers and ion channels). *Methods in Molecular Biology (Clifton, N.J.)*, *857*, 281–299. http://doi.org/10.1007/978-1-61779-588-6_12

Rivas, A., & Francis, H. W. (2005). Inner ear abnormalities in a Kcnq1 (Kvlqt1) knockout mouse: a model

of Jervell and Lange-Nielsen syndrome. *Otology & Neurotology : Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *26*(3), 415–424.

Robbins, J. (2001). KCNQ potassium channels : physiology , pathophysiology , and pharmacology, *90*, 1–19.

Roden, D. M. (2006). A new role for calmodulin in ion channel biology. *Circulation Research*, *98*(8), 979–981. http://doi.org/10.1161/01.RES.0000221822.22971.8c

Roden, D. M., Balser, J. R., George, A. L., & Anderson, M. E. (2002). Cardiac ion channels. *Annual Review of Physiology*, *64*, 431–75. http://doi.org/10.1146/annurev.physiol.64.083101.145105

Roden, D. M., & Yang, T. (2005, September). Protecting the heart against arrhythmias: potassium current physiology and repolarization reserve. *Circulation*. United States. http://doi.org/10.1161/CIRCULATIONAHA.105.562777

Roux, B., Bernèche, S., Egwolf, B., Lev, B., Noskov, S. Y., Rowley, C. N., & Yu, H. (2011). Ion selectivity in channels and transporters. *The Journal of General Physiology*, *137*(5), 415–426. http://doi.org/10.1085/jgp.201010577

Rudy, Y. (2008). Molecular basis of cardiac action potential repolarization. *Annals of the New York Academy of Sciences*, *1123*, 113–118. http://doi.org/10.1196/annals.1420.013

Sachyani, D., Dvir, M., Strulovich, R., Tria, G., Pongs, O., Svergun, D., … Hirsch, J. a. (2013). Structural Basis of a Kv7 . 1 (KCNQ1) Potassium Channel Gating Module : Studies of the Intracellular C-terminal Domain in Complex with Calmodulin. *Structure/Folding and Design*, *1*(11), 1582–1594. http://doi.org/10.1016/j.str.2014.07.016

Sanguinetti, M. C., Curran, M. E., Zou, A., Shen, J., Spector, P. S., Atkinson, D. L., & Keating, M. T. (1996). Coassembly of K(V)LQT1 and minK (IsK) proteins to form cardiac I(Ks) potassium channel. *Nature*, *384*(6604), 80–83. http://doi.org/10.1038/384080a0

Sansom, M. S. P., Shrivastava, I. H., Bright, J. N., Tate, J., Capener, C. E., & Biggin, P. C. (2002). Potassium channels : structures , models , simulations, *1565*, 294–307.

Schmitt, N., Schwarz, M., Peretz, A., Abitbol, I., Attali, B., & Pongs, O. (2000). A recessive C-terminal Jervell and Lange-Nielsen mutation of the KCNQ1 channel impairs subunit assembly. *The EMBO Journal*, *19*(3), 332–340. http://doi.org/10.1093/emboj/19.3.332

Schroeder, B. C., Waldegger, S., Fehr, S., Bleich, M., Warth, R., Greger, R., & Jentsch, T. J. (2000). A constitutively open potassium channel formed by KCNQ1 and KCNE3. *Nature*, *403*(6766), 196–199. http://doi.org/10.1038/35003200

Schwartz, P. J., Crotti, L., & Insolia, R. (2012). Long-QT Syndrome. *Circulation: Arrhythmia and Electrophysiology*, *5*(4), 868 LP-877. Retrieved from http://circep.ahajournals.org/content/5/4/868.abstract

Shamgar, L., Ma, L., Schmitt, N., Haitin, Y., Peretz, A., Wiener, R., … Attali, B. (2006). Calmodulin is essential for cardiac IKS channel gating and assembly: Impaired function in long-QT mutations. *Circulation Research*, *98*(8), 1055–1063. http://doi.org/10.1161/01.RES.0000218979.40770.69

Shieh, C.-C., Coghlan, M., Sullivan, J. P., & Gopalakrishnan, M. (2000). Potassium Channels: Molecular Defects, Diseases, and Therapeutic Opportunities. *Pharmacological Reviews*, *52*(4), 557 LP-594. Retrieved from http://pharmrev.aspetjournals.org/content/52/4/557.abstract

Shipston, M. J. (2011). Ion Channel Regulation by Protein Palmitoylation. *The Journal of Biological Chemistry*, *286*(11), 8709–8716. http://doi.org/10.1074/jbc.R110.210005

Simms, B. A., & Zamponi, G. W. (2014). Neuronal Voltage-Gated Calcium Channels: Structure, Function, and Dysfunction. *Neuron*, *82*(1), 24–45. http://doi.org/http://dx.doi.org/10.1016/j.neuron.2014.03.016

Smith, J. a, Vanoye, C. G., Jr, a L., Meiler, J., & Sanders, C. R. (2007). Structural Models for the KCNQ1 Voltage-Gated Potassium Channel. *Biochemistry*, 14141–14152. http://doi.org/10.1021/bi701597s

Sokolova, O., Kolmakova-Partensky, L., & Grigorieff, N. (2001). Three-Dimensional Structure of a Voltage-Gated Potassium Channel at 2.5 nm Resolution. *Structure*, *9*(3), 215–220. http://doi.org/http://dx.doi.org/10.1016/S0969-2126(01)00578-0

Splawki, I., Tristani-Firouzi, M., Lehmann, M. H., Sanguinetti, M. C., & Keating, M. T. (1997). Mutations in the hminK gene cause long QT syndrome and supppress Iks function. *Nature Genetics*, *17*, 338–340.

Splawski, I., Shen, J., Timothy, K. W., Lehmann, M. H., Priori, S., Robinson, J. L., … Keating, M. T. (2000). Spectrum of Mutations in Long-QT Syndrome Genes. *Circulation*, *102*(10), 1178 LP-1185. Retrieved from http://circ.ahajournals.org/content/102/10/1178.abstract

Splawski, I., Shen, J., Timothy, K. W., Vincent, G. M., Lehmann, M. H., & Keating, M. T. (1998). Genomic Structure of Three Long QT Syndrome Genes:KVLQT1, HERG,andKCNE1. *Genomics*, *51*(1), 86–97. http://doi.org/http://dx.doi.org/10.1006/geno.1998.5361

Suh, B.-C., & Hille, B. PIP2 is a necessary cofactor for ion channel function: how and why?, 37Annual review of biophysics 175–195 (2008). United States. http://doi.org/10.1146/annurev.biophys.37.032807.125859

Tester, D. J., & Ackerman, M. J. (2014). GENETICS OF LONG QT SYNDROME. *Methodist DeBakey Cardiovascular Journal*, *10*(1), 29–33. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4051331/

Thomas, D., Zhang, W., Wu, K., Wimmer, A.-B., Gut, B., Wendt-Nordahl, G., … Karle, C. A. (2003). Regulation of HERG potassium channel activation by protein kinase C independent of direct phosphorylation of the channel protein. *Cardiovascular Research*, *59*(1), 14–26. Retrieved from http://dx.doi.org/10.1016/S0008-6363(03)00386-9

Tien, J., Young, D. M., Jan, Y. N., & Jan, L. Y. (2014). Chapter 11 - Molecular Properties of Ion Channels BT - From Molecules to Networks (Third Edition) (pp. 323–348). Boston: Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-0-12-397179-1.00011-7

Trezise, D., Dale, T., & Main, M. (2010). Chapter 2.3: Voltage-gated sodium channels. In J. N. C. Kew & C. H. Davies (Eds.), *Ion channels : from structure to function*. New York: Oxford University Press.

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., … Pontén, F. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220). Retrieved from http://science.sciencemag.org/content/347/6220/1260419.abstract

Vacher, H., Mohapatra, D. P., & Trimmer, J. S. (2008). Localization and Targeting of Voltage-Dependent Ion Channels in Mammalian Central Neurons. *Physiological Reviews*, *88*(4), 1407 LP-1447. Retrieved from http://physrev.physiology.org/content/88/4/1407.abstract

Vallon, V., Grahammer, F., Richter, K., Bleich, M., Lang, F., Barhanin, J., … Warth, R. (2001). Role of KCNE1-dependent K+ fluxes in mouse proximal tubule. *Journal of the American Society of Nephrology : JASN*, *12*(10), 2003–2011. Journal Article, Research Support, Non-U.S. Gov't.

Van Horn, W. D., Vanoye, C. G., & Sanders, C. R. (2011). Working model for the structural basis for KCNE1 modulation of the KCNQ1 potassium channel. *Current Opinion in Structural Biology*, *21*(2), 283–291. http://doi.org/10.1016/j.sbi.2011.01.001

Veerman, C. C., Verkerk, A. O., Blom, M. T., Klemens, C. A., Langendijk, P. N. J., van Ginneken, A. C. G., … Tan, H. L. (2013). Slow delayed rectifier potassium current blockade contributes importantly to drug-induced long QT syndrome. *Circulation. Arrhythmia and Electrophysiology*, *6*(5), 1002–1009. http://doi.org/10.1161/CIRCEP.113.000239

Wang, Q., Curran, M. E., Splawski, I., Burn, T. C., Millholland, J. M., VanRaay, T. J., … Keating, M. T. (1996). Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nature Genetics*, *12*(1), 17–23. http://doi.org/10.1038/ng0196-17

Wang, W., Flores, M. C. P., Sihn, C.-R., Kim, H. J., Zhang, Y., Doyle, K. J., … Yamoah, E. N. (2015). Identification of a key residue in K(v)7.1 potassium channel essential for sensing external potassium ions. *The Journal of General Physiology*, *145*(3), 201–212. http://doi.org/10.1085/jgp.201411280

Wen, H., & Levitan, I. B. (2002). Calmodulin is an auxiliary subunit of KCNQ2/3 potassium channels. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(18), 7991–8001. http://doi.org/22/18/7991 [pii]

Wiener, R., Haitin, Y., Shamgar, L., Fernandez-Alonso, M. C., Martos, A., Chomsky-Hecht, O., … Hirsch, J. A. (2008). The KCNQ1 (Kv7.1) COOH terminus, a multitiered scaffold for subunit assembly and protein interaction. *The Journal of Biological Chemistry*, *283*(9), 5815–5830. http://doi.org/10.1074/jbc.M707541200

Wiener, R., Haitin, Y., Shamgar, L., Fernández-Alonso, M. C., Martos, A., Chomsky-Hecht, O., … Hirsch, J. a. (2008). The KCNQ1 (Kv7.1) COOH terminus, a multitiered scaffold for subunit assembly and protein interaction. *Journal of Biological Chemistry*, *283*(9), 5815–5830. http://doi.org/10.1074/jbc.M707541200

Wu, H., Chen, X., Cheng, J., & Qi, Y. (2016). SUMOylation and Potassium Channels: Links to Epilepsy and Sudden Death. *Advances in Protein Chemistry and Structural Biology*, *103*, 295–321. http://doi.org/10.1016/bs.apcsb.2015.11.009

Wulff, H., Castle, N. A., & Pardo, L. A. (2009). Voltage-gated Potassium Channels as Therapeutic Drug Targets. *Nature Reviews. Drug Discovery*, *8*(12), 982–1001. JOUR. http://doi.org/10.1038/nrd2983

Xu, Q., & Minor, D. L. (2009). Crystal structure of a trimeric form of the KV7.1 (KCNQ1) A-domain tail coiled-coil reveals structural plasticity and context dependent changes in a putative coiled-coil trimerization motif. *Protein Science*, *18*(10), 2100–2114. http://doi.org/10.1002/pro.224

Xu, X., Jiang, M., Hsu, K.-L., Zhang, M., & Tseng, G.-N. (2008). KCNQ1 and KCNE1 in the IKs channel complex make state-dependent contacts in their extracellular domains. *The Journal of General Physiology*, *131*(6), 589–603. http://doi.org/10.1085/jgp.200809976

Xu, Y., Wang, Y., Meng, X. Y., Zhang, M., Jiang, M., Cui, M., & Tseng, G. N. (2013a). Building KCNQ1/KCNE1 channel models and probing their interactions by molecular-dynamics simulations. *Biophysical Journal*, *105*(11), 2461–2473. http://doi.org/10.1016/j.bpj.2013.09.058

Xu, Y., Wang, Y., Meng, X., Zhang, M., Jiang, M., Cui, M., & Tseng, G. (2013b). Building KCNQ1 / KCNE1 Channel Models and Probing their Interactions by Molecular-Dynamics Simulations, *105*(December), 2461–2473. http://doi.org/10.1016/j.bpj.2013.09.058

Xu, Y., Wang, Y., Zhang, M., Jiang, M., Rosenhouse-Dantsker, A., Wassenaar, T., & Tseng, G.-N. (2015). Probing Binding Sites and Mechanisms of Action of an IKs Activator by Computations and Experiments. *Biophysical Journal*, *108*(1), 62–75. http://doi.org/10.1016/j.bpj.2014.10.059

Yang, T., Smith, J. a, Leake, B. F., Sanders, C. R., Meiler, J., & Roden, D. M. (2013). An allosteric mechanism for drug block of the human cardiac potassium channel KCNQ1. *Molecular Pharmacology*, *83*(2), 481–9. http://doi.org/10.1124/mol.112.081513

Yang, W.-P., Levesque, P. C., Little, W. a, Conder, M. L., Shalaby, F. Y., & Blanar, M. a. (1997). KvLQT1, a voltage-gated potassium channel responsible for human cardiac arrhythmias. *Proceedings of the National Academy of Sciences*, *94*(April), 4017–4021. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=20560&tool=pmcentrez&rendertype=abstract

Yang, Y., & Sigworth, F. J. (1998). Single-Channel Properties of I(Ks) Potassium Channels . *The Journal of General Physiology*, *112*(6), 665–678. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2229447/

Yap, Y. G., & Camm, A. J. (2003). Drug induced QT prolongation and torsades de pointes. *Heart*, *89*(11), 1363–1372. http://doi.org/10.1136/heart.89.11.1363

Yarov-Yarovoy, V., Baker, D., & Catterall, W. a. (2006). Voltage sensor conformations in the open and closed states in ROSETTA structural models of K(+) channels. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(19), 7292–7297. http://doi.org/10.1073/pnas.0602350103

Yellen, G. (1998). The moving parts of voltage-gated ion channels. *Quarterly Reviews of Biophysics*, *31*(3), 239–295. http://doi.org/10.1017/S0033583598003448

Yu, F. H., & Catterall, W. A. (2004). The VGL-chanome: a protein superfamily specialized for electrical signaling and ionic homeostasis. *Science's STKE : Signal Transduction Knowledge Environment*, *2004*(253), re15. http://doi.org/10.1126/stke.2532004re15

Yus-Nájera, E., Santana-Castro, I., & Villarroel, A. (2002). The identification and characterization of a noncontinuous calmodulin-binding site in noninactivating voltage-dependent KCNQ potassium channels. *Journal of Biological Chemistry*, *277*(32), 28545–28553. http://doi.org/10.1074/jbc.M204130200

Zagotta, W. (2015). Ligand-dependent gating mechanism. *Handbook of Ion Channels*, 41–52. http://doi.org/doi:10.1201/b18027-6

Zaydman, M. A., Silva, J. R., Delaloye, K., Li, Y., Liang, H., Larsson, H. P., & Shi, J. (2013). Kv7 . 1 ion channels require a lipid to couple voltage sensing to pore opening, 1–6. http://doi.org/10.1073/pnas.1305167110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1305167110

Zhorov, B. S., & Tikhonov, D. B. (2004). Potassium, sodium, calcium and glutamate-gated channels: pore architecture and ligand action. *Journal of Neurochemistry*, *88*(4), 782–799. http://doi.org/10.1111/j.1471-4159.2004.02261.x

Zhou, Y., Morais-Cabral, J. H., Kaufman, A., & MacKinnon, R. (2001). Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. *Nature*, *414*(6859), 43–48. http://doi.org/10.1038/35102009

RATIONALE, HYPOTHESIS AND OBJECTIVES

RATIONALE

The KCNQ1/KCNE1 ion channel and its normal activity are crucial for the heart to function normally (See section 1.2). It is, therefore, important to understand the different events taking place within this channel at the atomic level. These events include the mechanisms of ion conduction and the effect of drug binding on ion permeation as well as on the structure and function of the channel. These detailed structural events are very hard to be revealed through traditional structural analysis and, therefore, modelling can be used to provide fine analysis of these mechanisms. For this reason, this Thesis aimed at building a dynamic model for the KCNQ1/KCNE1 channel complex. The model described in this work builds and improves upon the earlier models and has been validated through available experimental data.

HYPOTHESIS

There were two major hypotheses for this Thesis. First, the KCNE1 protein has an important effect on both the structure of KCNQ1 as well as on the ion permeation through the channel. Second, a strong KCNQ1 blocker can bind to a site close to the selectivity filter, allowing it to interact with specific residues at this region and also interacting with the ion influx through the channel.

OBJECTIVES

To provide a clear investigation of these two hypotheses, we established the following objectives for this Thesis:

1) To build a comprehensive homology model for the KCNQ1 protein, while taking into account the most recent experimental information.

2) To Predict the most probable interaction between the KCNQ1 and its auxiliary protein, KCNE1.

3) To study the structure of KCNQ1 and ion permeation with and without the KCNE1 protein.

4) To test a panel of I_{Ks} current blockers (consisting of Chromanol 293B and its 8 derivatives) within the model for the KCNQ1/KCNE1 ion channel and investigate their effect on ion permeation.

In short, the overall goals of this thesis are to reveal novel information that explains a fraction of the structural and functional details taking place at the atomic-level using a model that included the very minute details. I will discuss in detail my efforts toward this end in the remainder of the thesis.

CHAPTER 2: A COMPREHENSIVE STRUCTURAL MODEL FOR THE HUMAN KCNQ1/KCNE1 ION CHANNEL[1]

2.1. Introduction

Voltage-dependent Na^+, K^+ and Ca^{2+} ion channels are responsible for the generation of an action potential, which is the elementary unit of electrical signals in biology (S. Chowdhury & Chanda, 2015). The voltage-gated K^+ channels form the largest family of VGICs and are named as K_v ion channels. About 40 members of this diverse group of ion channels are identified and named K_v1 up to K_v12 with numerous subtypes to each class. Their expression sites include the heart, brain, auditory and vestibular organs, and epithelia (T J Jentsch, 2000). Amongst the currents produced by these ion channels, the I_{Ks} current plays a vital role in the repolarization phase of the cardiac action potential and thereby in the normal functioning of the heart. The IKs current is produced by the KCNQ1/KCNE1 voltage-gated ion channel.

The KCNQ1 ($K_v7.1$ or K_vLQT1) ion channel is encoded by the KCNQ1 gene and belongs to the large family of voltage-gated potassium ion channels. It was first identified in a study by Wang et al. to reveal some of the genetic causes of sudden death from cardiac arrhythmia (Q. Wang et al., 1996). The co-assembly of the KCNQ1 alpha subunit and the KCNE1 (minK) beta subunit generates the slowly activating delayed rectifier cardiac I_{Ks} current (Osteen et al., 2010). This current has a major role in controlling repolarization, and thus the duration of the cardiac action potential (Howard et al., 2007). Mutations in KCNQ1 or the KCNE1 gene that affect trafficking, assembly, or regulation lead to long QT syndrome (LQTS) also known as Romano-Ward syndrome, a hereditary cardiac condition associated with life-threatening arrhythmias, deliquium seizures and sudden death (Schwartz, Crotti, & Insolia, 2012). Null or missense mutations of $K_v7.1$ result in hearing loss and other auditory problems. Furthermore, a cardio-auditory syndrome called Jervell and Lange-Nielsen syndrome (JLNS) also results from mutations in this gene (Mousavi Nik et al., 2015). Gain-of-function mutations in $K_v7.1$ increase

[1] A version of this Chapter has been submitted to The Journal of Molecular Modeling as *Jalily Hasani H, Ahmed M, Ganesan A, Barakat K. Ion Permeation Studies in the Human KCNQ1/KCNE1 Ion Channel.*

current flow and lead to shortening of the cardiac action potential; as seen in a number of cardiac rhythm disorders such as Short QT syndrome (SQTS) and atrial fibrillation (Wulff et al., 2009).

Despite of the large number of studies, the molecular mechanisms by which these channels mediate ion transport remains a significant unsolved problem. In addition, there has been a number of important structural information that have to be constantly included in the models of proteins. This study describes the efforts taken towards constructing a comprehensive and accurate homology model for the KCNQ1 protein. Several state-of-the-art computational techniques were applied to build this sophisticated 3-dimensional model while accommodating and fitting the model to all available experimental data in the literature. Following the construction of KCNQ1 protein, we applied protein-protein docking algorithms to assemble the KCNQ1/KCNE1 complex. Throughout this process, we used extensive molecular dynamics (MD) simulations to refine the KCNQ1 homology model, relax the KCNQ1: KCNE1 complex and understand how they interact in the membrane environment at physiological conditions.

2.2. Methods

Homology Modelling Methods and Protocol

The template for the modeling step was chosen based on the results from several sequence alignment and threading tools; namely LOMETS (S. Wu & Zhang, 2007), BLAST (Stephen F Altschul et al., 1990), HHpred (Soding et al., 2005) and HHblits (Remmert, Biegert, Hauser, & Soding, 2012). The process of modeling the KCNQ1 protein, we obtained ten top models using I-TASSER standalone tool (Yang Zhang, 2008), with acceptable scores (average C-Score: -0.1, average TM-Score: 0.7). The final best model was picked based on validations from Protein Quality Predictor (ProQ) (Bjorn Wallner & Elofsson, 2003), ProSA (Sippl, 1993; Wiederstein & Sippl, 2007), Verify3D (Eisenberg, Luthy, & Bowie, 1997; Luthy, Bowie, & Eisenberg, 1992), which also had the highest percentage of residues in the favoured regions of Ramachandran plot (95.6% of residues in the favourable regions, 4% in the allowed region and 0.4% of residues in the disallowed outlier regions). Ramachandran plot is a fundamental tool in analysing the structure of proteins. It serves as a validation tool for visualizing the allowed

conformations of proteins through plotting the torsion angles of the amino acids in the protein structure (Hollingsworth & Karplus, 2010).

MD Simulation: System Preparation in the membrane[2]

When membrane proteins are studied, it is crucial to have a proper inclusion of a membrane to ensure accuracy because these proteins are highly sensitive to their surrounding environment. Similarly, it was essential to include a membrane system for the different systems in this study, i.e. KCNQ1 alone, KCNE1 alone, KCNQ1/KCNE1 complexes. The arrangement of the protein in the membrane was built using the Membrane Builder of CHARMM GUI (*http://www.charmm gui.org*). The tetrameric protein was embedded in a bilayer of Palmitoyloleoylphosphatidylcholine (POPC) and Phosphatidylinositol 4,5-bisphosphate (PIP_2) in the ratio of 10:1 respectively. The system was further hydrated with 20 Å (TIP3P water model) on upper and lower leaflets. An ionic concentration of 150 mM KCl solution was maintained in the system, both in the upper and lower regions and neutralized with counter ions. Protein, lipids and ionic parameters were assigned using the CHARMM36 force field (Brooks et al., 1983). NAMD package (Phillips et al., 2005), version 2.10 and 4,096 processors on the Blue Gene\Q supercomputer were employed for running the Molecular Dynamics (MD) simulations.

MD Simulation: Parameters and Protocol

Prior to running the full dynamics, each system was subjected to energy minimizations in two stages. In the first minimization round of 50,000 steps, the protein and the lipid heads were fixed whereas the lipid tails, water and ions were free to relax. This step was essential to remove any existing steric clashes as a result of the improper packing of the membrane around the protein. In the second stage of minimization, a constraint of 100 kcal/mol was placed on the entire system and energy minimization was performed. This constraint was gradually removed during four more rounds of minimizations. Each minimization stage was of 50,000 steps and the constraints on the protein backbone were reduced to 50, 10, 5 and finally 1 kcal/mol constraining force. The systems were then

[2] The concepts of MD simulations and its enhanced variants are provided in Appendix A.

energy minimized by for 5,000 steps. This was followed by heating the system to 310 K for 10 ns while retaining the 1 kcal/mol backbone restraint. Equilibration phase was next, consisting of NVT and NPT steps for 250 ps, each.

Simulations were performed with an integration time step of 2 fs, under periodic boundary conditions. The Langevin dynamics were adopted for temperature (310 K) and pressure control (1 bar). Bonded interactions were computed every one timestep, short-range non-bonded interactions every two timesteps, and long-range electrostatic interactions every four time steps. A cutoff of 12 Å was used for van der Waals and short-range electrostatic interactions; with a switching function starting at 10 Å for van der Waals interactions to ensure a smooth cutoff. The simulations were performed under periodic boundary conditions; long-range electrostatic interactions were calculated by using the particle-mesh Ewald (PME) method. The unit cells were large enough that adjacent copies of the protein were never close enough for short-range interactions to apply. The trajectory frames were written to file every 200 ps.

REMD Simulation: System Preparation[3]

The S1-S2 helices were excised from our homology model of KCNQ1 ion channel. This model was built using the I-TASSER package (J. Yang et al., 2014) employing the $K_V1.2$-$K_V2.1$ paddle chimera channel (PDB ID: 2R9R) as the template. The two helices and the loop attaching them, consist of 82 residues spanning from residue number 116 to 197, based on the sequence of the human KCNQ1 protein in the UniProt database (UniProt ID: P51787). The simulation system was prepared using the CHARMM-GUI (E. L. Wu et al., 2014) membrane builder package. The helices were embedded in a lipid bilayer of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC). The lipid layers were further packed with TIP3P water molecules and an ionic concentration of 150 mM KCl solution, both in the upper and lower regions and neutralized with counter ions. Protein, lipids and ions parameters were assigned using the CHARMM36 force field.

[3] Theory of REMD simulations is described in Appendix A.

REMD Simulation: Parameters and Protocol

MPI version of NAMD package (Phillips et al., 2005), version 2.10, was employed for running the REMD simulations. 8,192 processors were adopted on the Blue Gene/Q Compute Canada supercomputer facility. The whole system was minimized and equilibrated prior to REMD simulation to obtain a well-equilibrated starting system. For the REMD, we used 128 replicas distributed linearly in the temperature range from 280 to 440 K. The highest temperature (i.e. 440 K) at which the lipid bilayer system retains its integrity was obtained by performing 10 short classical MD simulations of 10 ns each at different temperatures and finally selecting 440 K as the maximum temperature to be used in the REMD. Several optimization runs were carried out to decide on appropriate simulation parameters. Because of the large size of the system (~35,000 atoms), the parameters defined for the REMD simulation were set to allow 25,000 steps (250 ps) before an attempt to swap temperatures was made. This ensured that the system at each replica equilibrates and explores all the possible conformations in the space at each respective temperature. The final system was run using very weak harmonic restraints on the lipid heads of the bilayer.

The system was simulated for 25 ns, allowing ~1000 swaps between different temperatures. Therefore, the cumulative simulation time of all replicas added up to 3.2 μs. The structures were saved every 10 ps (i.e. every 1000 frames). The average replica exchange ratios were 25.5%. The built in sortreplicas program of NAMD was used for sorting the simulation frames based on target temperature. The final sorted trajectories of the physiologic temperature range, i.e. 299-311 K were used for further analysis.

Protein-Protein Docking Method and Protocol[4]

We used the HADDOCK (High Ambiguity Driven protein-protein DOCKing) tool (Dominguez et al., 2003) version 2.2, to assemble the KCNQ1 tetramer with KCNE1 proteins in a 4:2 stoichiometry. The docking was done in a data-driven manner using experimental data from the literature. The restrains were imposed as a set of ambiguous interaction restrains (AIRs) in HADDOCK. The complete docking protocol and strategy is explained under section 2.3.7, in the current chapter.

[4] For a detailed description of protein-protein docking, please refer to Appendix C.

Binding Free Energy

As an additional measure for choosing the best complexes formed between KCNQ1 and KCNE1 proteins, we calculated the binding free energies between the proteins using the molecular mechanics Generalized-Born surface area (MM-GBSA) method as implemented in AMBER12 (Case et al., 2015). The method combines molecular mechanics with continuum solvation models. The total free energy is the sum of average molecular mechanical gas-phase (E_{MM}), solvation free energies (G_{solv}) and entropy contributions ($-TS_{solute}$) of the binding reaction:

$$G = E_{MM} + G_{solv} - TS_{solute} \qquad \textbf{Eqn. 1}$$

The molecular mechanical (E_{MM}) energy of each snapshot in the MD trajectory was calculated using the SANDER module of AMBER12. The solvation free energy (G_{solv}) was estimated as the sum of electrostatic solvation free energy, calculated by the finite-difference solution of the Poisson–Boltzmann equation in the Adaptive Poisson–Boltzmann Solver (APBS) and non-polar solvation energy, calculated from solvent-accessible surface area (SASA) algorithm:

$$\Delta G^0 = \Delta G_{gas}^{KCNQ1-KCNE1} + \Delta G_{solv}^{KCNQ1-KCNE1} - \{ \Delta G_{solv}^{KCNE1} + \Delta G_{solv}^{KCNQ1} \} \qquad \textbf{Eqn. 2}$$

The other parameters used include a dielectric constant for the protein–protein complex of 1, a dielectric constant for the water of 80, an ionic concentration of 0.15 M and a surface tension of 0.005 with a surface offset of zero, to estimate the non-polar contribution of the solvation energy. The snapshots were sampled every 1 ns of each MD trajectory.

Analysis & Visualizations

Analysis of protein-protein interactions were performed using VMD (Humphrey, Dalke, & Schulten, 1996). Visualizations were done using both Chimera suite (Pettersen et al., 2004) and VMD (Humphrey, Dalke, & Schulten, 1996). Plots were generated using

Gnuplot and GraphPad Prism version 6.0 [GraphPad Software, La Jolla California USA, www.graphpad.com].

2.3. Results and Discussion

2.3.1. Homology Modeling of the KCNQ1 Protein

A prerequisite for this Thesis was to build a robust model for the KCNQ1 protein. There is currently no available experimental structure for human KCNQ1 protein. This is a common problem for almost all human ion channels and it is mainly due to the fact that these are large membrane proteins, which are not amenable to experimental structure determination techniques. As such, one of the best approaches to address this problem is the use of *in silico* homology modeling tools, which rely on the sequence similarity of the target protein to related proteins of known structure (See Appendix B). In this thesis, we adopted a combination of cutting-edge structure prediction tools to build the initial model for KCNQ1 protein, as will be discussed in the following sections.

i) Modeling Approach

To begin the modeling process, the sequence of the KCNQ1 protein was retrieved from the UniProt database (UniProt ID: P51787). In order to identify the best available template for the modeling step, tools such as LOMETS (S. Wu & Zhang, 2007), BLAST (Stephen F Altschul et al., 1990), HHpred (Soding et al., 2005) and HHblits (Remmert, Biegert, Hauser, & Soding, 2012) were used. Table 2.1 shows the results from the template search to identify the best templates with the relevant scores and comparative parameters. Eventually the 2R9R (Long et al., 2007) PDB entry was chosen as the template to be used for the homology modeling because of a higher sequence similarity and other comparative values such as E-value and GMQE (Global Model Quality Estimation) (Schwede, Kopp, Guex, & Peitsch, 2003). The 2R9R structure belongs to a $K_V1.2$-$K_V2.1$ paddle chimera channel (Long et al., 2007) in an open conformation and has a 34.5% sequence similarity to the human KCNQ1. Figure 2.1, shows the sequence alignment of 2R9R to the human KCNQ1 (H-KCNQ1) sequence.

Table 2.1. The final result of template search and comparison between the top two templates.

Crystal Structure (PDB ID)	Sequence Similarity	GMQE (Global Model Quality Estimation)	HH-Pred Score	E-Value	Z-Score
2R9R	34.5%	0.67	169.1	1.3e-18	81.300
2A79	32.8%	0.50	167.7	1.08e-18	61.900

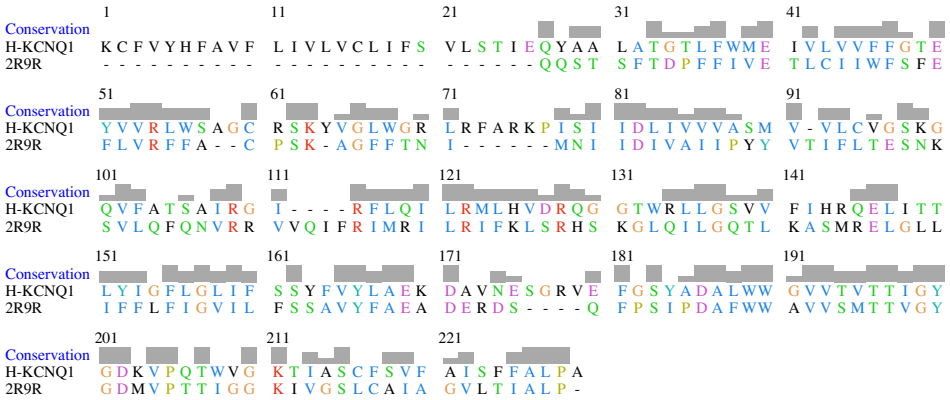


Figure 2.1. Sequence alignment between human KCNQ1 (H-KCNQ1) and K_V1.2-K_V2.1 paddle chimera channel (PDB ID: 2R9R). The sequence conservation pattern is displayed with bars. Regions of low sequence similarity or the gapped regions belong mainly to the linkers and loops in the structure.

We employed the I-TASSER standalone tool (Yang Zhang, 2008) to initiate the modeling process. I-TASSER has been consistently ranked as the top tool for protein structure prediction in the 7th to 12th community-wide blind CASP experiments (D. Xu, Zhang, Roy, & Zhang, 2011; Yang Zhang, 2007, 2009, 2014). Figure 2.2 shows the overall workflow of the I-TASSER tool. I-TASSER uses a combination approach for protein structure modeling using *ab initio*, threading and homology modeling. The template (PDB ID: 2R9R) was manually specified to I-TASSER. The modeling starts with the construction of fragments for the aligned regions from the template. The unaligned regions are predicted via *ab initio* modeling. As shown in the alignment between H-KCNQ1 and 2R9R there are a few gaps, which mainly belong to the loops and the linkers connecting the different segments. Predicating these unaligned gap

regions using the *ab initio* technique was an additional measure to increase the model quality of the whole protein.

Following the construction of the fragments forming the KCNQ1 protein, we assembled them through Replica exchange Monte Carlo simulations. These simulations were performed at different temperatures and were clustered by SPICKER (Yang Zhang & Skolnick, 2004). The clustered conformations represent the models of the lowest free-energy state in the Monte Carlo simulations. In the next step, the models were refined through a second round of minimization simulations to remove any existing local clashes between the different residues. During this second round of simulations, spatial restraints identified by TM-align (Yang Zhang & Skolnick, 2005) from the PDB library, were added to prevent large structural deviations in the overall KCNQ1 protein. The final model was built by ModRefiner (D. Xu & Zhang, 2011) based on the lowest energy states as obtained from the second round of simulation trajectories. Initially, ModeRefiner constructs the model from the C-alpha traces and then adds the side-chain atoms from a rotamer library. The full atomic conformations were then refined by energy minimization and a combination of physics- and knowledge-based force fields methods. I-TASSER provides the C-Score and TM-Score for the final models. In this modeling process, we obtained 10 models from I-TASSER with acceptable score (average C-Score: -0.1, average TM-Score: 0.7).

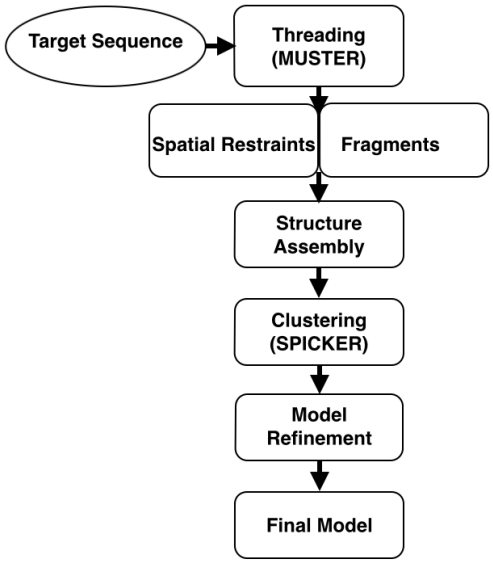


Figure 2.2. Flowchart of the I-TASSER working protocol.
Adapted from Figure 1 of (Roy, Kucukural, & Zhang, 2010)

ii) Model Refinement

Despite the high level of accuracy taken into account at each step of modeling by I-TASSER, there could still be imprecisions in the predicted models. Therefore, we did not rely solely on the final results from I-TASSER, however, we performed several other post-modeling assessments on the final top 5 models. This was important to rule out any inaccuracies and to select the best and final model to move forward. All top five structures were first evaluated for their stereochemistry using the PROCHECK tool (Laskowski et al., 1993). Furthermore, they were validated using various tools such as Protein Quality Predictor (ProQ) (Bjorn Wallner & Elofsson, 2003), ProSA (Sippl, 1993; Wiederstein & Sippl, 2007), Verify3D (Eisenberg, Luthy, & Bowie, 1997; Luthy, Bowie, & Eisenberg, 1992). A consensus among the different tools allowed us to choose the final model provided by I-TASSER. The superimposed structures of the top 5 models are shown in Figure 2.3.a. The final best model is shown in Figure 2.3.b along with its Ramachandran plot (Figure 2.3.c) which placed only 0.4% of residues in the disallowed outlier regions, 4% in the allowed region and 95.6% of residues in the favourable regions of the Ramachandran plot.

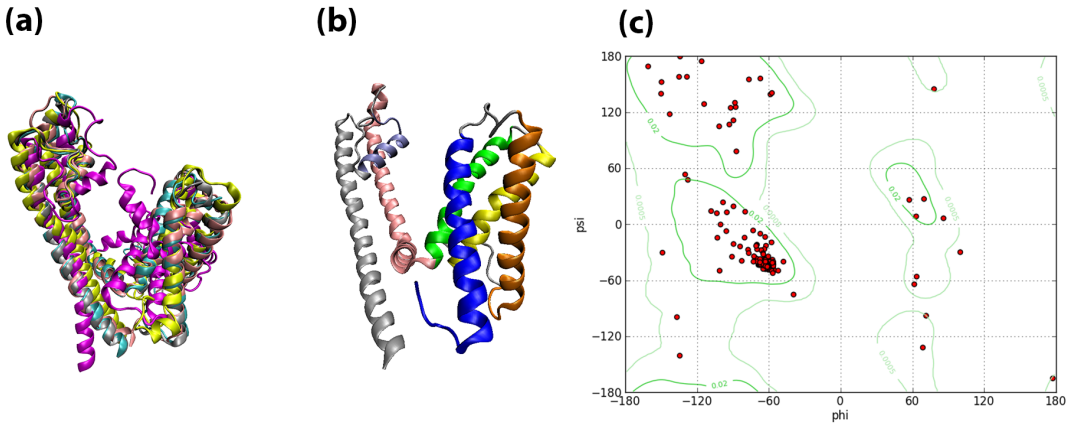


Figure 2.3. (a) Superimposed structures of the top 5 models. (b) The final model chosen from the top 5 models of I-TASSER modeling based on the post-modeling assessment. The 6 helices are coloured differently (S1: blue, S2: orange, S3: yellow, S4: green, S5: pink, S5-S6 linker: iceblue, S6: silver). (c) Ramachandran plot of the final model.

2.3.2. Fitting the Model to Experimental Data

Any implementation of modeling techniques, even the most accurate ones are associated with assumptions that can lead to ambiguities in the final results. Therefore, it is crucial to continually update the models by accommodating as much reliable experimental data as possible throughout the modelling process. For this reason, we performed an extensive literature review to identify any available structural information related to the KCNQ1. This was done to ensure that the most recent and updated structural information from the literature is included in our model.

Through this investigation, we identified two NMR studies that were focused on the structural investigation of the KCNQ1 protein. The first study by Gayen et al. (Gayen et al., 2015) comprised of the secondary structure investigation of the S4-S5 Linker using solution NMR techniques. Based on their findings, the S4-S5 linker of KCNQ1 adopts a helical structure. The S4-S5 linker is located between the VSD and the pore domain and hence plays an important role in determining the channel gating. Precise structural features of such central components in the ion channels are of utmost importance to increase the accuracy of the predicted models and thereby the consequent studies. Therefore, we investigated the secondary structure of this linker in our final model. We observed that it complies with the structural information provided by the NMR spectroscopy studies such that the S4-S5 linker possessed a helical structure (see Figure 2.4). This observation was made after completion of the modeling and without any previous structural bias, which validates the accuracy of our modeling approach. Figure 2.4 shows the structure of the S4-S5 linker in KCNQ1 along with the sequence of this linker.

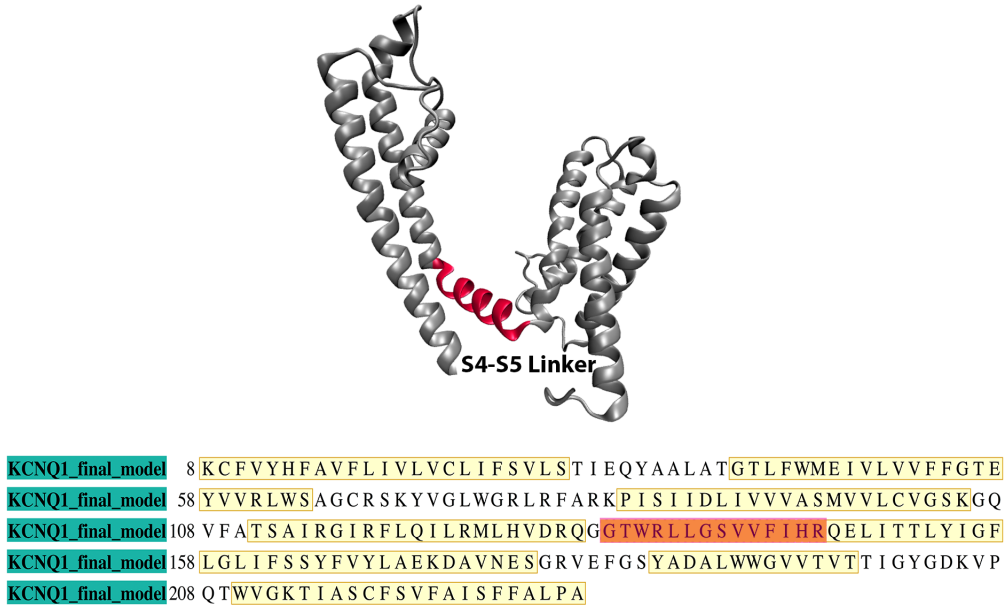


Figure 2.4. The structure and sequence of the S4-S5 linker (coloured in red) of KCNQ1. The linker adopts a helical structure in the final model. The sequence of the linker is highlighted in red. The light-yellow highlights indicate the helices (S1 to S6).

The second study we found was also an NMR investigation by Peng et al. (D. Peng et al., 2014) that has also reported novel structural information about the KCNQ1 protein. In this paper, the authors found that the two helices; S1 and S2, which are located in the voltage sensing domain (VSD), acquire a secondary structure which is significantly different from the subtype ion channels and several homologous members of the superfamily of voltage gated ion channels. Upon structural investigations on our model, there was no compliance between the data reported in the NMR study and our final model. Thus, for this reason we decided to explore the secondary structure of the two helices and the possibility of forming a helical structure as reported by Peng et al.

We began this investigation by carrying out classical MD simulations of the two helices (S1 and S2) alone for a large MD simulation (~200 ns). We hypothesized that if the helical state is energetically favourable, it may be adopted by the protein during this molecular dynamics simulation. We tried the simulations in two different settings: in water and in membrane bilayer (Figure 2.5). However, neither of the simulations resulted in an elongation of the helices.

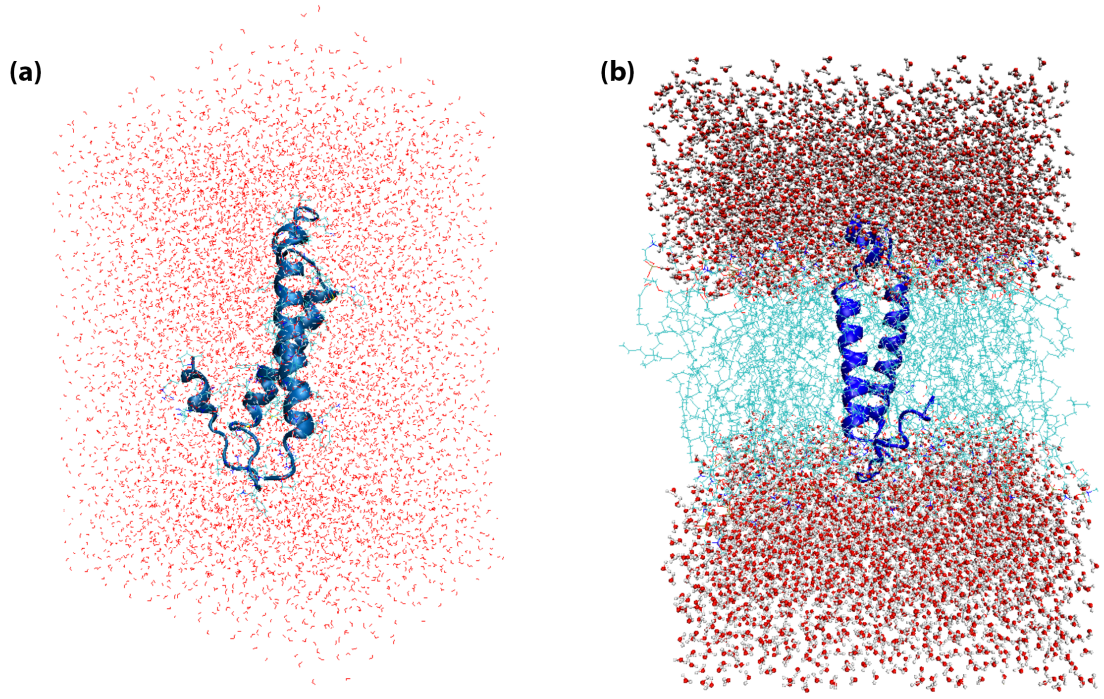


Figure 2.5. The classical MD simulation setup for the S1-S2 classical MD simulation. (a) in water, (b) in membrane lipid bilayer and water.

Following this, we tried a loop modeling approach, called the SuperLooper prediction server (P. W. Hildebrand et al., 2009) which is optimized for modeling of loops in globular and membrane proteins. It employs a database of almost half a billion segments of water-soluble proteins to find the right template for modeling the loops in the query. However, this tool also was not able to predict a helical structure for the two helices.

Eventually and after several attempts of trying to impose the structural conformation for the two helices we decided to use an enhanced MD sampling technique to explore the possible secondary structure properties of these two helices. Our hypothesis in this regard was that in order to explore the precise structure of S1-S2 helices in KCNQ1 protein we needed to overcome the large energy barriers that may trap the system in a single energy state. Therefore, an enhanced sampling technique, namely Replica Exchange Molecular Dynamics (REMD) can enable the sampling of a wider phase space. This facilitates the observation of the most prominent secondary structure

that these helices obtain by overcoming the energy barriers. The details of the REMD approach and the corresponding results are discussed in the next section.

i) Replica Exchange Molecular Dynamics (REMD) On S1 And S2.

The S1-S2 helices were excised from our top homology model of KCNQ1 ion channel. The two helices and the loop attaching them consisted of 82 residues spanning from residue number 116 to 197. The REMD simulated system was prepared using the CHARMM-GUI (E. L. Wu et al., 2014) membrane builder package. The helices were embedded in a lipid bilayer of 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and Phosphatidylinositol 4,5-bisphosphate (PIP_2). The lipid layers were further packed with TIP3P water molecules and an ionic concentration of 150 mM KCl solution, both in the upper and lower regions and were neutralized using counter ions. Protein, lipids and ions parameters were assigned using the CHARMM36 force field.

The MPI version of the NAMD package (Phillips et al., 2005), version 2.10, was employed to run the REMD simulations. 8,192 processors were used on the Blue Gene/Q Compute Canada supercomputer facility to run the final simulation. The whole system was minimized and equilibrated prior to REMD simulation to obtain a well-equilibrated starting system. For the REMD, we used 128 replicas distributed linearly in the temperature range from 280 to 440 K. The highest temperature (i.e. 440 K) at which the lipid bilayer system retains its integrity was obtained by performing 10 short classical MD simulations of 10 ns each at different temperatures and finally selecting 440 K as the maximum temperature to be used in the REMD. Several optimization runs were carried out to decide on the appropriate simulation parameters. Because of the large size of the system (~35,000 atoms), the parameters defined for the REMD simulation were set to allow 25,000 steps (250 ps) before an attempt to swap temperatures. This ensured that the system at each replica equilibrated at the respective temperature and explored all the possible conformations in the space before any swap occurred. The final system was run using very weak harmonic restraints on the lipid heads of the bilayer to ensure membrane integrity.

The system was simulated for 25 ns, allowing ~1000 swaps between different temperatures. Therefore, the cumulative simulation time of all the 128 replicas (each one

being 25 ns long) added up to 3.2 µs. The structures were saved every 10 ps (i.e. every 1000 frames). The average replica exchange ratios were 25.5%. The built-in sortreplicas utility in NAMD was used to sort the simulation frames based on their target temperatures. The final sorted trajectories at the physiological temperature range, i.e. 299-311 K were used for further analysis as discussed below.

ii) Replica Exchange Molecular Dynamics Analysis

During a REMD simulation, it is essential to ensure that each replica has visited a wide range of temperatures during the course of the simulation. This ensures that efficient sampling has taken place by overcoming the potential energy barriers at higher temperatures. Figure 2.6 shows the temperature evolution of the first 20 replicas during the whole trajectory. It is clear that each replica has visited several temperatures from the beginning to end and has done exchanges of temperature with many other replicas.

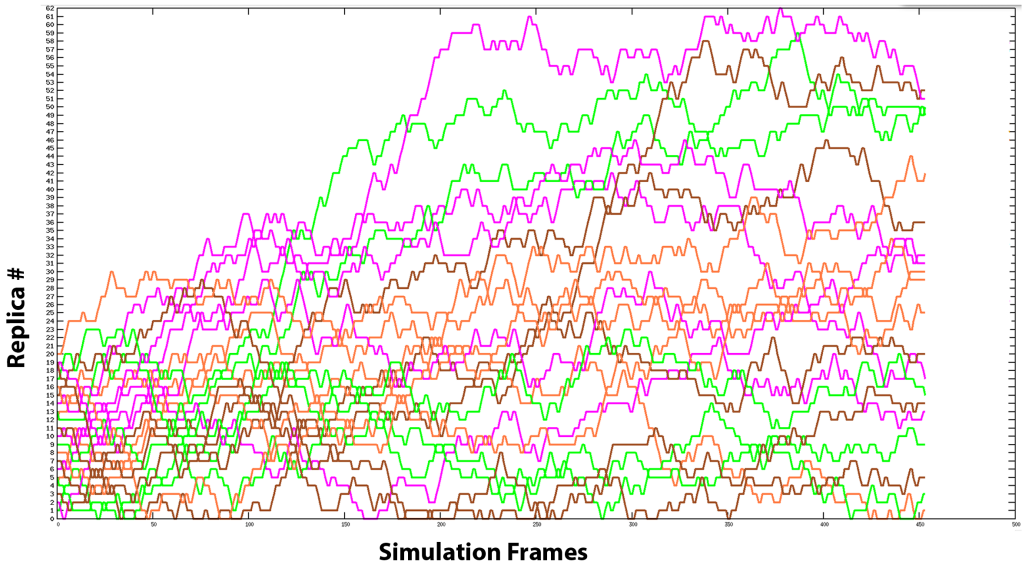


Figure 2.6. Temperature evolution of the first 20 replicas (#1-20). The plot shows the exchanges with higher or lower temperatures.

Figure 2.7.a illustrates the potential energy distributions of first 20 replicas showing a suitable overlap and corresponds to a high rate of exchange. Figure 2.7.b shows the potential energy distribution of all the 128 replicas. Both the graphs give enough evidence that the necessary criteria for enough sampling are met and that the present replica-exchange simulation was indeed performed properly and effectively. The curves

for neighbouring replicas overlap considerably, ensuring sufficient exchange between replicas. The downhill decrease in potential energy with the replicas at high temperatures (Figure 2.7.b) corresponds to the variety of folded conformations appearing at high temperatures, which is in well agreement with experimental observations (Bakk, Høye, & Hansen, 2001).

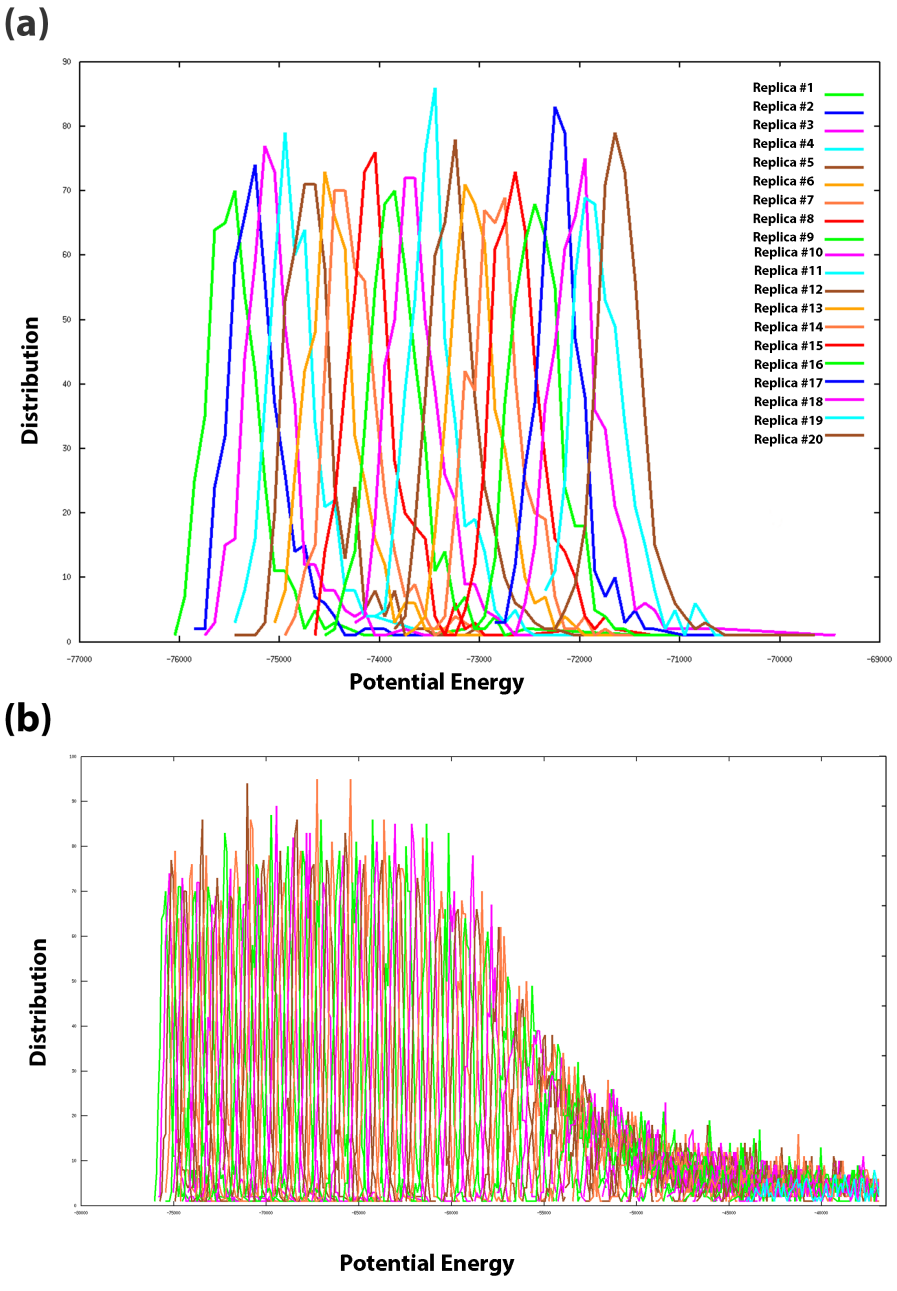


Figure 2.7. Potential energy histogram, (a) First 20 replicas, (b) all 128 replicas.

Another important parameter is the relative mean square deviation (RMSD) of the protein under study. RMSD is a representation of the stability of a protein in the simulation environment. Likewise, in REMD the protein at each temperature should energetically become stable. To check this parameter, we plotted the RMSDs of the protein at a number of temperatures after the sorting process. See Figure 2.8, for the RMSD graph of 5 replicas (~every 30 replicas) i.e. replica # 1, 30, 60, 90, 128. The RMSDs are stable and this stability indicates that the proteins are not experiencing huge structural drifts. This also indicates that the cumulative duration of simulating the replicas has been appropriate and RMSD graph converges at the respective temperature. However, it is clear from the image that Replica #128 (450 K) has the highest RMSD up to 8 Å which is due to the high temperature and unfolding and adoption of numerous conformations, resulting in high fluctuations.

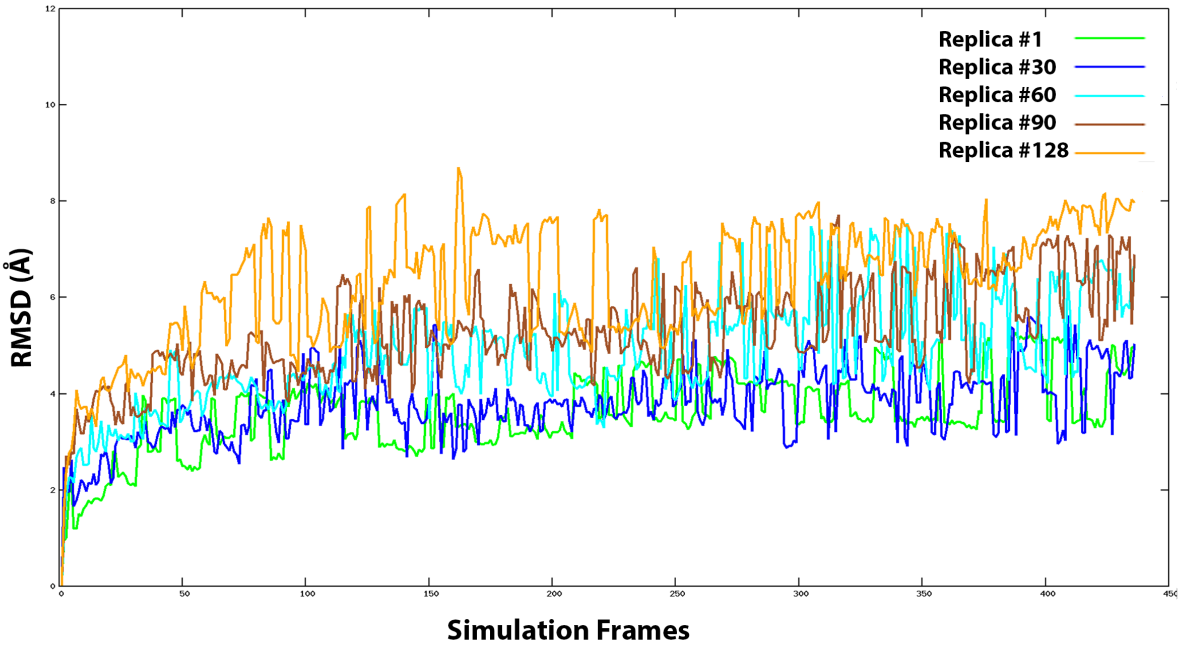


Figure 2.8. RMSD of the protein (S1-S2 segments) backbone during the REMD simulation.

1. **Extracted all the frames of the REMD trajectories**
2. **(50,000 frames)**

2. Secondary structure analysis of all frames:

Step i: Secondary structure assignment with STRIDE:

>frame999 TTCTTTGGGHHHHHHHHHHHHHHHTTTTTTTCCCCC
>frame1000 TTTCCCHHHHHHHHHHHHHHHHHHHCCCTTTTCCCC
>frame1001 TTTCCCHHHHHHHHHHHHHHHHHHHCCCTTTTCCCC
>frame1002 TTTCTTTTTHHHHHHHHHHHHHHHHCCCTTTTCCCC

Step ii: Count the number of helix assignments for each frame:

>frame999	15
>frame1000	9
>frame1001	10
>frame1002	11

Step iii: Sorting based on the length of helices in each frame

For S1: **82 conformations**: 4 residues increase in length
For S2: **11 conformations**: 3 residue increase in length

3. Further evaluation of top structures:

Step i: Visual Check

S1: 14 structures
S2: 6 structures

Step ii: Ranking

Rosetta scoring function
TM scoring (Moderefiner)
PROCHECK and Ramachandran plots

4) The two helices were first attached together and then attached to the rest of the segment

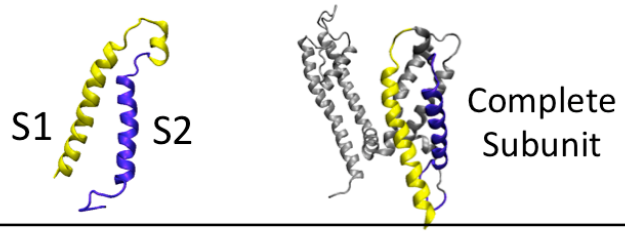


Figure 2.9. The complete workflow of analyzing the results of REMD. The steps show the process of how the S1-S2 conformations and the final structures were chosen to be incorporated into the model for KCNQ1.

To extract the dominant conformations of S1 and S2 from the REMD simulation trajectory, we adopted a scheme shown in Figure 2.9. An in-house script was built which would extract all the frames of the complete REMD trajectories. Next, the secondary structure assignment of the protein conformation in each replica is calculated. Secondary structure analysis was performed using the implementation of the STRIDE stand-alone program (Frishman & Argos, 1995). The conformations were then ranked based on the length of the helices. For the S1 helix, 82 conformations showed an increase of 4 residues in the length of the helix. These were visualized to check the integrity and their hydrogen bonding. Finally, 14 structures were considered acceptable for further analysis. The same procedure was performed for the S2 helix. The top ranked structures, which corresponded to 3 residues increase in length of the helix, were obtained. This gave rise to 11 conformations for S2 and after the visualization check, only 6 of them were retained.

The final structures for the two helices were chosen based on the secondary structure evaluation as ranked by their TM Score from the ModRefiner server (D. Xu & Zhang, 2011), Rosetta scoring function for membrane proteins (Alford et al., 2015) and PROCHECK tool (Laskowski et al., 1993). The Ramachandran plot of the final S1-S2 model is shown in Figure 2.10.a. The final model was refined and minimized to remove any existing clashes. Using the Chimera (Pettersen et al., 2004) build structure tool, the two helices were attached to the rest of the KCNQ1 voltage-sensing domain. The final S1-S2 structure is shown in Figure 2.10.b, superimposed on the model from I-TASSER. As can be seen from the figure there is a slight increase in the length of the two helices.

Overall, the results from our REMD structural investigation indicates that there is a chance of observing the change in the length of the two helices by overcoming the energy barriers and exploring all the possible conformational states of this protein segment. However, it has been frequently discussed that the environment used in the NMR studies i.e. micelles affects the behaviour of proteins (Bond & Sansom, 2003). Although we do not reject the claim of the paper by Peng et al. (D. Peng et al., 2014) regarding the difference in the KCNQ1 VSD secondary structure, but we suspect that their observations could have been biased by the micellar conditions used in NMR spectroscopy. Further discussion on assembling the tetrameric channel is discussed in the next section.

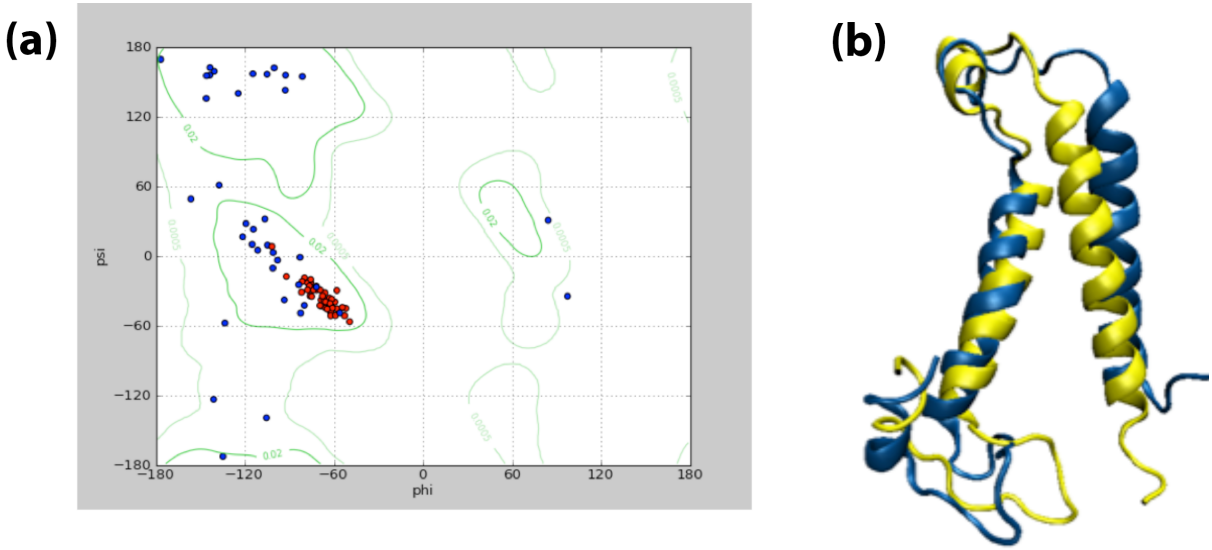


Figure 2.10. The final structure of the S1-S2 helices. (a) Ramachandran plot of the final S1-S2 model. Red dots on the plot are the S1 helix residues and blue dots are the S2 helix residues. (b) superimposed structures of the S1-S2 from I-TASSER (blue) and from REMD structural investigation (yellow).

2.3.3. The KCNQ1 Protein Assembly

The final refined KCNQ1 subunit is shown in Figure 2.11.a. The complete tetrameric KCNQ1 channel was assembled using the Chimera software. Four copies of the subunit were superimposed on the $K_V1.2$-$K_V2.1$ paddle chimera channel (PDB ID: 2R9R) to obtain the tetrameric arrangement shown in Figure 2.11.b (top view) and 7.11.c (lateral view).

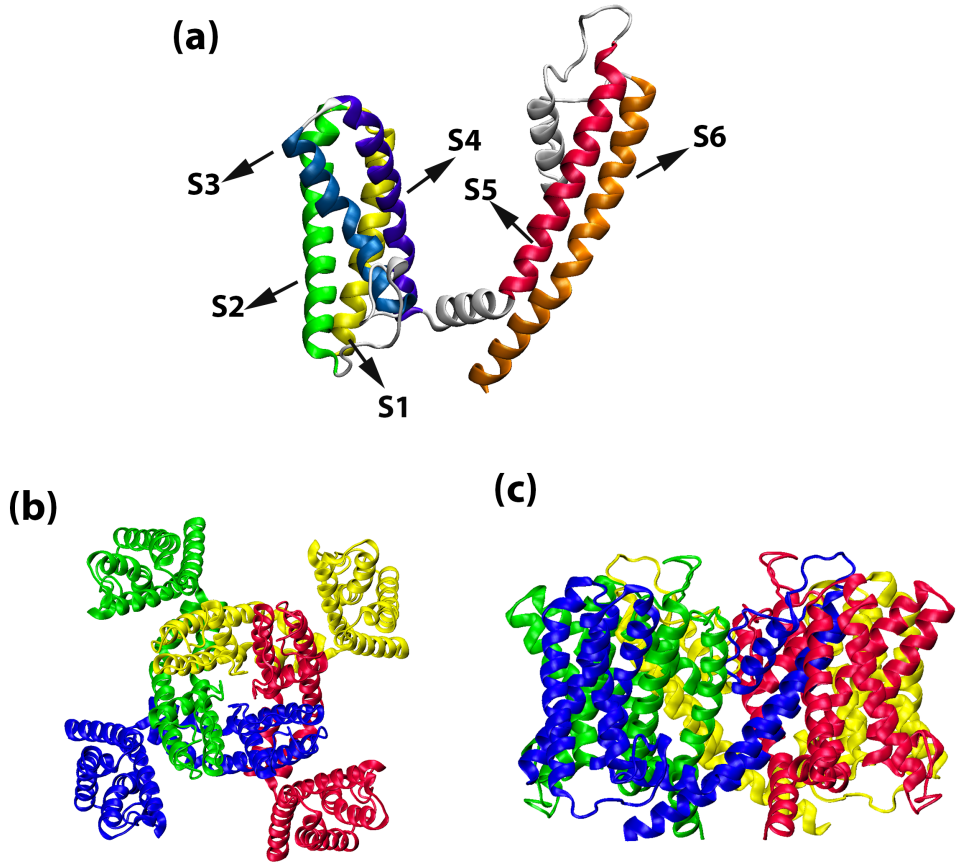


Figure 2.11. KCNQ1 single subunit and assembled structure. (a) A single subunit of KCNQ1 model, coloured to show the 6 segments (S1-S6) with their loops and linkers. (b) Top view of the assembled tetrameric KCNQ1 channel in open state. (c) Lateral view of the assembled tetrameric KCNQ1 channel in open state.

2.3.4. KCNQ1 Refinement Using Classical Molecular Dynamics Simulations

MD simulations in this study was done for two reasons. First, models generated through homology modeling might not represent the lowest energy conformation for a protein. MD simulations allow the relaxation of the protein and the removal of structural clashes that can otherwise lead to further imprecisions in the desired results. Secondly, for a complete KCNQ1/KCNE1 complex generation, an ensemble of KCNQ1 structures was required. This conformational ensemble can be obtained from MD trajectories (as will be discussed in the next section). And finally, MD simulations allowed us to confer several findings related to ion conduction as a measure of model validation, which will be discussed in the next chapter. The MD simulation system was prepared as per the

protocol discussed in the Methods section. The simulation system of KCNQ1 is shown in Figure 2.12. The system consisted of a lipid bilayer composed of POPC and PIP_2 lipid molecules in the ratio of 1:10, water and KCl ionic environment (150 mM). The KCNQ1 protein was simulated for ~800 ns.

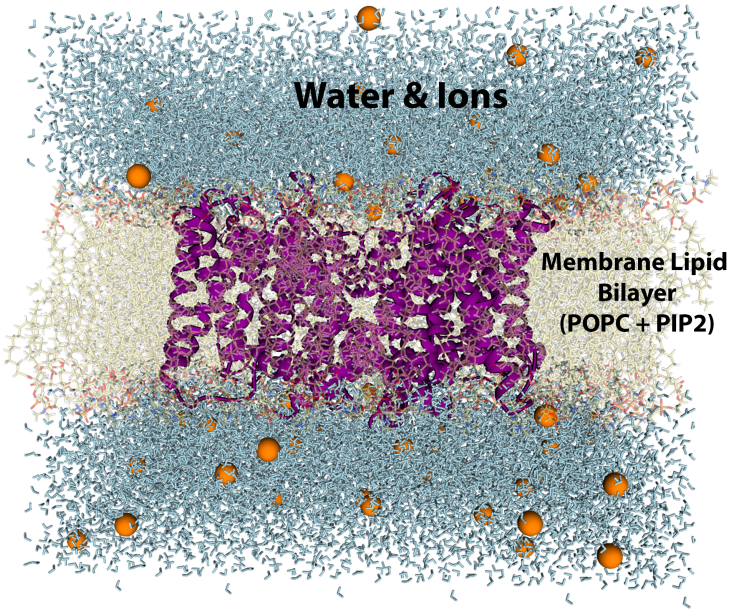


Figure 2.12. The simulation system of KCNQ1, in lipid (POPC + PIP_2) bilayer, water and ions.

i) MD Analysis

Upon completion of the simulation, the RMSD and RMSF graphs were plotted (see Figure 2.13.a and 2.13.b respectively). The RMSD graph confirms that the system has reached equilibrium, which is an indication of structural and energetic relaxation of the system. The RMSF graph shows the fluctuations of the four subunits in the structure of the KCNQ1 protein over the simulation time. The main fluctuations in the structure belong to the several loops and linkers that connect the helical segments of each subunit. These flexible regions with frequent motions include residue number 60 to 80 which correspond to the S2-S3 linker, residues 97-110 corresponding to S3-S4 linker, residues 170-180 corresponding to the loop in between S5 and the P-loop and finally, the C-terminal region of the subunits from residue 230 to 247. It is clear from Figure 2.13.b, that although the monomeric subunits forming the KCNQ1 protein are structurally the

same, they exhibited different relative motions. For example, subunit 3 showed the most fluctuations in the S2-S3 linker region and subunit 4 had the most fluctuating part in the C-terminal region. This observation confirms that MD simulation is indeed required to capture the wide range of motional behaviour of the protein over time, rather than considering a protein as a static object. Moreover, it is this differential motion that confers functional differences to the different parts of the protein.

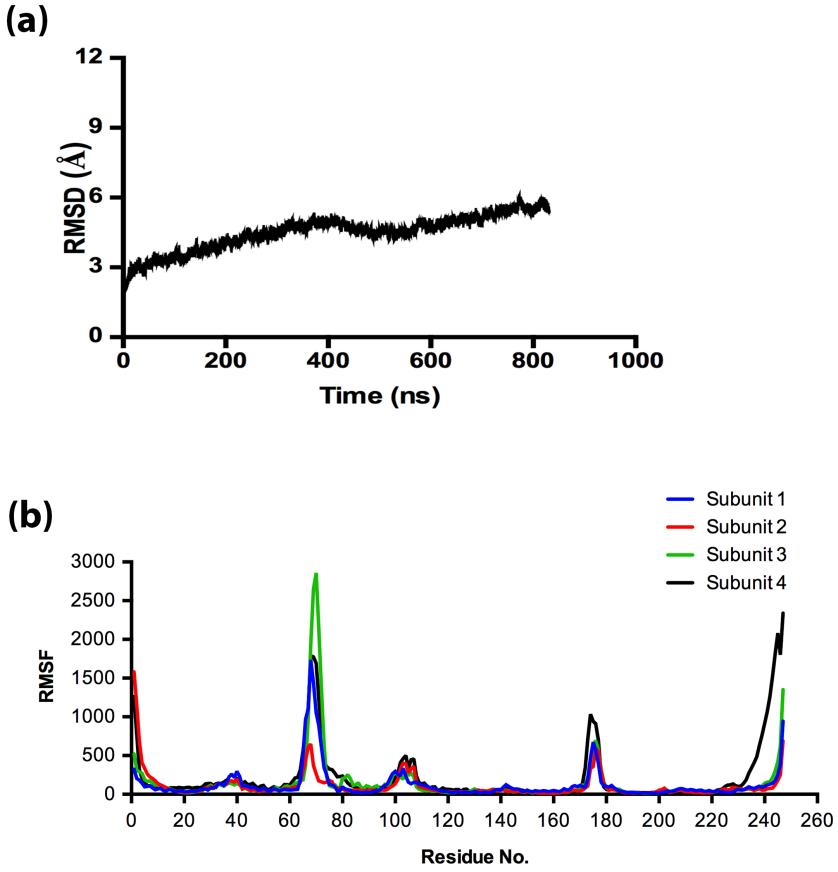


Figure 2.13. (a) C-alpha backbone RMSD of KCNQ1, and (b) C-alpha backbone RMSF of KCNQ1, during the long MD simulation of ~800 ns.

ii) KCNQ1 Model Validation

To validate the refined KCNQ1 protein, we compared it to several pieces of information from the literature. In 2015, Gayen et al. (Gayen et al., 2015) found that the W248 residue located in the S4-S5 linker of KCNQ1 subunit is involved in protein-lipid interactions. This residue corresponds to W133, W380, W627, W874 in the four subunits of our assembled KCNQ1 model. The MD trajectories were analyzed to check if this

interaction has taken place in our model and it was found that this residue in all the four subunits interacts with the lipid membrane, which further validates our model in two ways. Figure 2.14.a shows the overall lipid packing of KCNQ1 whereas Figure 2.14.b shows the specific interaction of the W248 with a PIP_2 molecule. Firstly, for these residues to interact with the membrane, they need to have the correct orientation. Secondly, this finding also confirms that the simulation system has been built in the right way such that the membrane wrapping allows the interaction to occur.

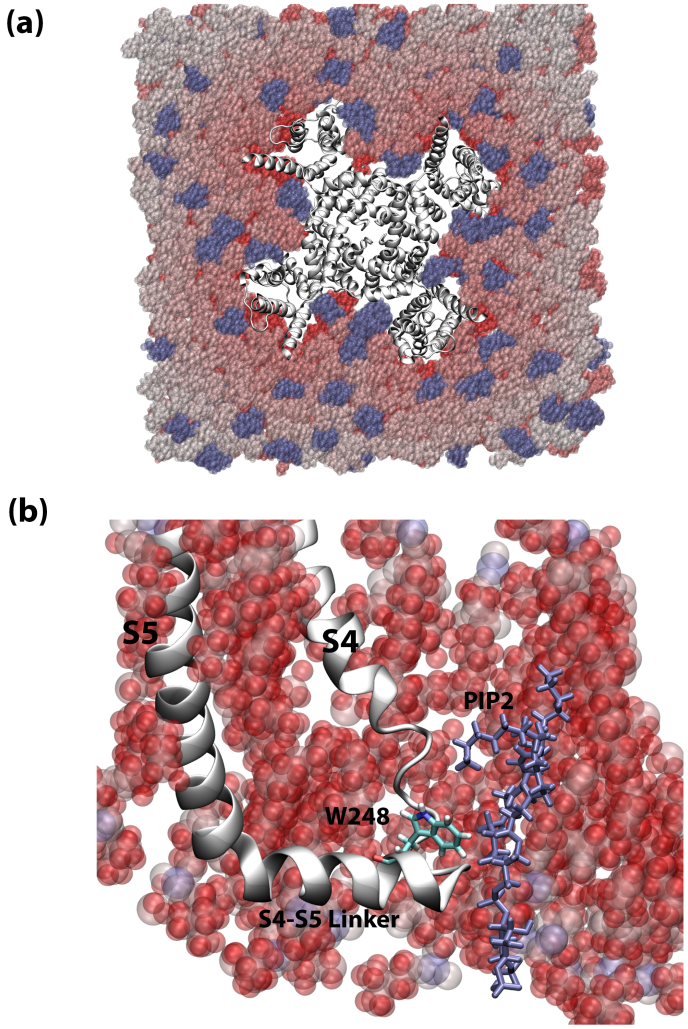


Figure 2.14. KCNQ1 protein-lipid interactions. (a) Overall packing of the KCNQ1 protein (shown in cartoon) by POPC (shown in pinkish red spheres), and PIP_2 lipid molecules (shown in blue spheres). (b) Interaction of W311 residue, located on the S-S5 linker, with a PIP_2 molecule.

In addition, the orientation of another group of residues on the S4-S5 linker (L250, L251, V254, V255 and H258) are the same as reported in the same study (Gayen et al., 2015). R259 also was suspected to have an affinity toward and facing the membrane bilayer, because of its positively charged side chain. This was also in agreement with our model. These interactions are confirmed to play an important role in the channel gating (Gayen et al., 2015).

Another validation was the unbiased events of ion conduction that took place during the MD simulation of KCNQ1. As reported in the literature, potassium ion channels including the KCNQ1 channel have well-defined binding sites for potassium in their selectivity filter (Shian Liu et al., 2015; A. N. Thompson et al., 2009; J. Thompson & Begenisich, 2001). Figure 2.15 shows the binding sites of K^+ ion in the KcsA ion channel. Although there is a difference in the sequence of KcsA ion channel's selectivity filter compared to KCNQ1, but the four binding sites remain the same in the two proteins. These binding sites are formed by the carbonyl atoms of the residues in the selectivity filter motif (TIGYG) of KCNQ1. These backbone carbonyl groups stabilize the ions through an interaction with their dipoles and form the four binding sites. However, the entrance of the potassium ions into the pore and consequently their passage through the selectivity filter requires them to overcome several high energy barriers and constrictions formed by the residues lining this pathway (as will be discussed in the subsequent sections) (Aqvist & Luzhkov, 2000).

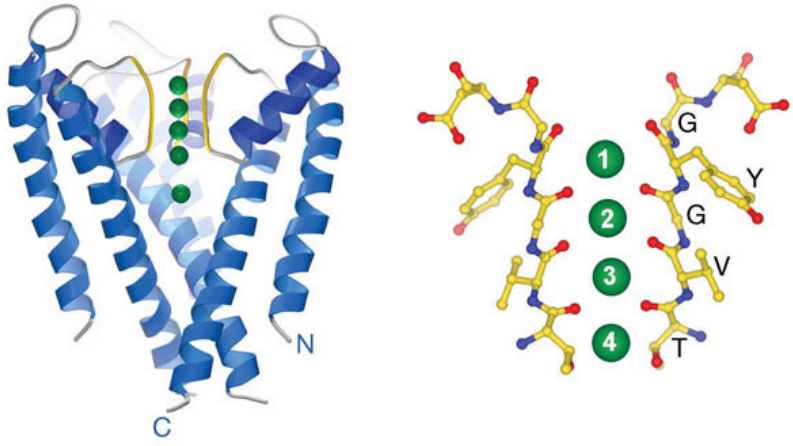


Figure 2.15. Binding sites for K^+ ions in potassium ion channels (KcsA ion channel). Reprinted by permission from Macmillan Publishers Ltd: [Nature](Morais-Cabral, Zhou, & MacKinnon, 2001), copyright (2001)

An accurate model of the KCNQ1 channel protein should possess the ability to allow the process of ion conduction to happen. This phenomenon was confirmed in our model. During the course of the 800 ns MD simulation, potassium ions entered the pore of the channel without any external bias (such as any pulling force or electric potential) and occupied the selectivity filter binding sites within the channel. The ions entered the pore at different timestamps of the simulation. Figure 2.16 illustrates the pathway that one of such potassium ions took to pass through the pore domain and to ultimately reach the selectivity filter, in a time frame of 40 ns of the MD trajectory. Several other ions also entered the pore domain and occupied the binding sites in the selectivity filter later on during the simulation.

These findings confirm that our model includes all the required components for a functional channel at physiological level. Interestingly, the same events occurred during the MD simulations of the KCNQ1/KCNE1 complex (see below), which again confirms the fact that the systems are valid and close to the real physiological environment.

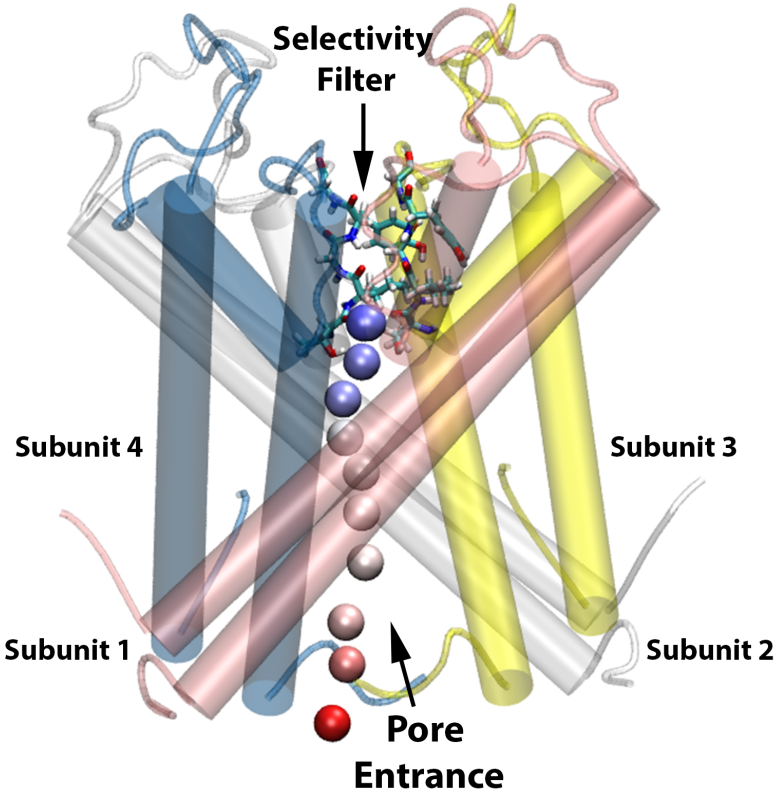


Figure 2.16. The pathway of a K^+ ion passing through the KCNQ1 channel during MD simulation. The ion is coloured by time-step, from red to blue as time progresses. The four KCNQ1 subunits are coloured in pink, silver, yellow and blue. The selectivity filter residues are depicted with bonds.

2.3.5. Preparation of the KCNE1: Pre-Assembly to KCNQ1

As discussed in the introductory chapters, a fully functional KCNQ1 channel has to be associated with its auxiliary protein, KCNE1. Therefore, to ensure that our model is comprehensive, we decided to include this protein in the final model. However, prior to that there were certain steps needed to be taken. Fortunately, there was an NMR structure (PDB ID 2K21) available for the human KCNE1 protein. Consequently, we did not need to use homology or *ab initio* modeling techniques for this protein. Nevertheless, the NMR structure of KCNE1 had several structural issues due to the experimental conditions that were used during the NMR spectroscopy imaging. As shown in Figure 2.17.a the transmembrane (TMD) region of KCNE1 possessed a conformational bent within the N- and C- terminals, folding them back into the membrane. These structural problems were due to the micellar phase that was used in the structure determination experiments.

Therefore, one step toward preparing the KCNE1protein for binding to the KCNQ1 channel was to adjust these such structural irregularities, making the KCNE1 protein in a linear form. Therefore, we adjusted several dihedral angles to correct this bent in the TMD, the N and C terminals. The residues involved in this correction were E43, H73, S74 and D76. After the dihedral angle adjustment, the structure obtained a close to linear conformation as shown in Figure 2.17.b However, it was not yet ready to be used in the next steps, as there could be structural clashes associated with this adjustment. Also, to increase the sampling of the protein conformations, we needed to generate several conformations for KCNE1. Therefore, MD simulation was performed for ~250 ns with the inclusion of lipid bilayer, water and ions (see Figure 2.18) and same parameters as discussed in the previous section for KCNQ1 protein.

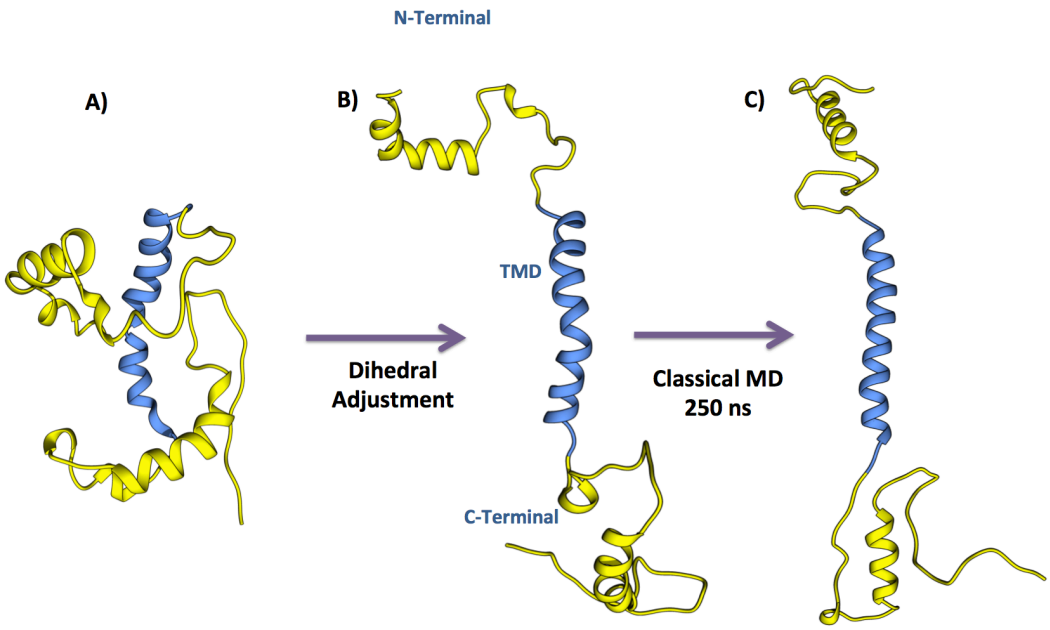


Figure 2.17. KCNE1 (PDB ID: 2K21) 3D structure. (a) KCNE1 NMR structure before refinement, (b) after dihedral angle adjustments and (c) after MD simulation.

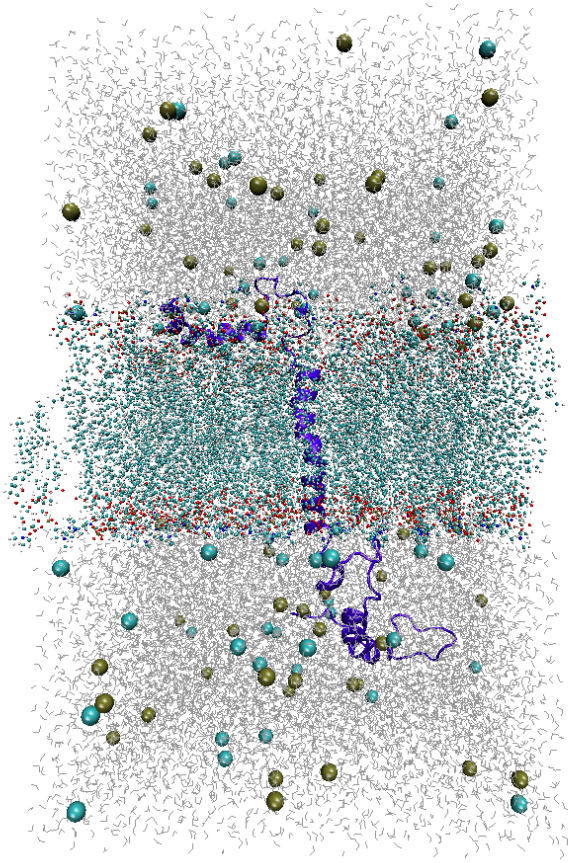


Figure 2.18. The simulation system of KCNE1. The system consisted of lipid bilayer membrane (POPC + PIP_2), water and ions (K^+, Cl^-).

i) MD Simulations Analysis for KCNE1 Refinement

After the completion of the MD simulation, the RMSD (Figure 2.19) and RMSF (Figure 2.20) graphs for the protein were plotted to ensure the stability of the protein during simulation and to infer information regarding the motion of the protein, respectively. The RMSD graph of KCNE1 in Figure 2.19, suggests that the structure has reached an acceptable equilibrium after x ns, fluctuating at around 7 Å. This value is indicative of the high flexibility of KCNE1, which also explains its versatile functions and interaction with several types of proteins. The same observation is revealed from the RMSF graph shown in Figure 2.20. The N-terminal of the protein denoted as NT in the graph, and the transmembrane helix (TMD) shows minimal structural fluctuations, whereas the C-terminal domain (CT) shows the highest motional variability, due to the long loop attachment that gives this protein region additional freedom to move.

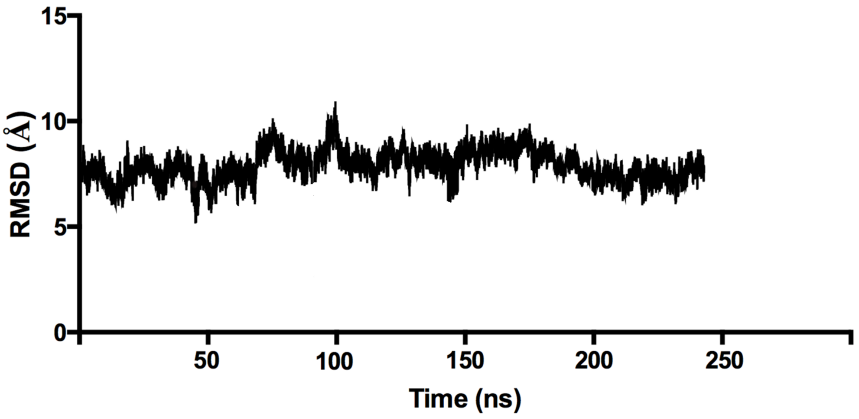
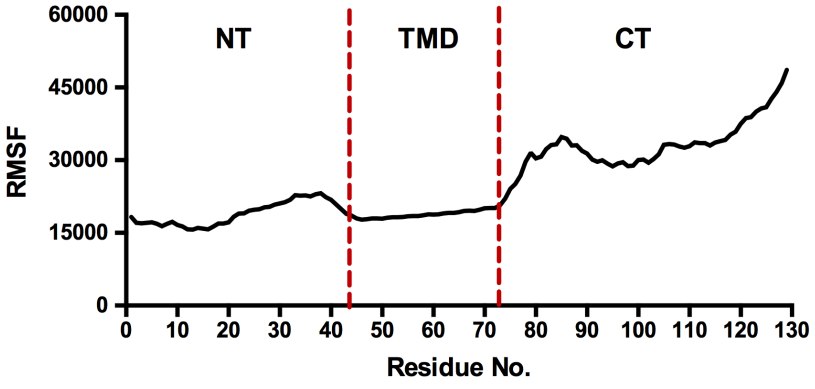


Figure 2.19. RMSD graph of KCNE1 during the MD simulation (~250 ns).

Figure 2.20. RMSF graph of the complete KCNE1 protein during the MD simulation (~250 ns). The plot is divided into 3 regions: NT (N-terminal domain), TMD (transmembrane domain) and CT (C-terminal domain).



2.3.6. Clustering Analysis of KCNQ1 and KCNE1

The MD simulations of the two proteins, i.e. KNQ1 and KCNE1 resulted in quite long trajectories that contained massive amounts of structural data. One way to mine this date for dominant protein conformations is to use clustering techniques. Clustering a trajectory is an efficient way of grouping the information in a meaningful way and helps in obtaining dominant structural representatives from the whole trajectory. These structures are considered as the most favourable conformations that were adopted by the protein during the trajectory.

In this work, we adopted the Average-Linkage algorithm using a code in PTRAJ program of AMBER (Roe & Cheatham III, 2013). We ran the average-linkage algorithm for a number of clusters ranging from 5 to 100. Structures were extracted at 8 ps intervals over the entire simulation time (800 ns). In this algorithm, cluster-to-cluster distance is defined as the average of all distances between individual points of the two clusters. Clustering quality is determined through the calculation of a number of clustering metrics including the Davies-Bouldin index (DBI) (Davies & Bouldin, 1979) and the "elbow criterion" (Shao, Tanner, Thompson, & Cheatham III, 2007). These metrics help in identifying the optimal number of clusters to be extracted and their population size. In order to remove the extra noise from the data as a result of rotations and translations, all the non-hydrogen heavy atoms were fitted to the minimized initial structure.

Next, the RMSD is used to cluster the residues at the binding interface residues between the KCNQ1 and the KCNE1 proteins (see Figure 2.21). These residues were clustered into groups of similar conformations using the atom-positional RMSD, as the similarity criterion. In each cluster, the structure that has the minimum RMSD (also called the cluster centroid) was chosen as the cluster representative. In the case of KCNQ1, 25 such dominant conformations were obtained from the clustering analyses and alignment of the DBI and SSR/SST parameters (see Figure 2.22.a) to be used in the subsequent protein-protein docking simulations. These 25 conformations represent more than 95% of the structure variability during the MD trajectory. Figure 2.22.b shows the superimposed structures of the 25 cluster representatives.

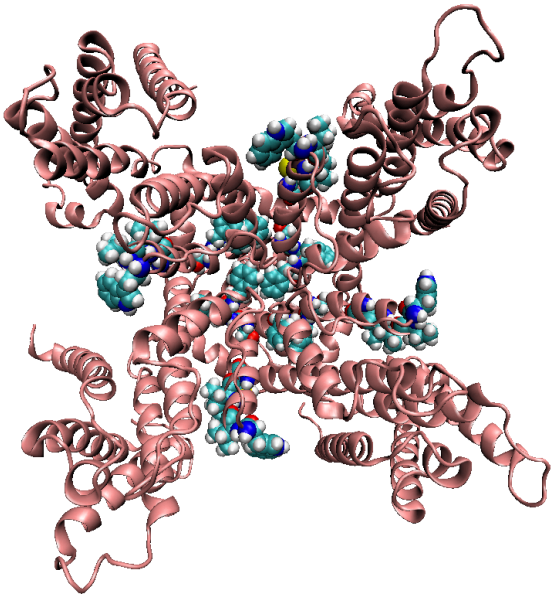


Figure 2.21. The KCNE1 binding interface residues on the KCNQ1 structure. These residues were used for clustering of the KCNQ1 MD trajectories as well as guiding the protein-protein docking simulations.

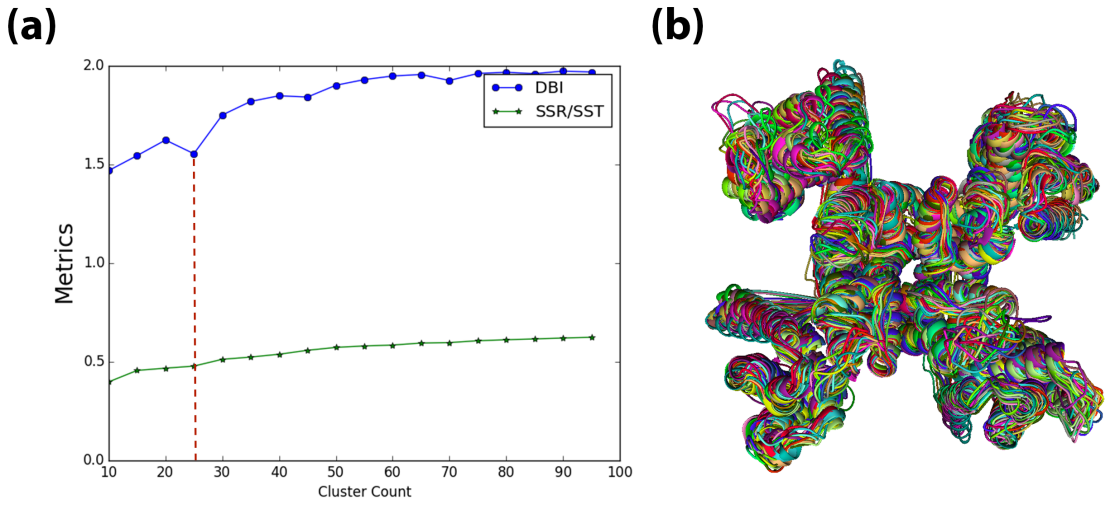


Figure 2.22. Clustering analysis of KCNQ1 protein. (a) The clustering plot for KCNQ1. (b) The 25 superimposed conformations of KCNQ1 from the clustering analysis, representing more than 95% of the structure variability.

Similarly, to obtain representative structures for the KCNE1 protein, the same clustering algorithm and protocol were applied. For KCNE1, only the TMD of the protein was clustered to eliminate the noise arising from the structural variability of the N and C terminals. The plot in Figure 2.23.a shows the results, wherein 30 conformers of KCNE1

were chosen (shown in Figure 2.23.b), for the subsequent docking studies. Furthermore, the C-terminal portion of the KCNE1 interaction was excluded from the docking studies. This was mainly because the KCNQ1 homology model does not include the cytoplasmic domains that interact with the c-terminal of KCNE1 (J. Chen et al., 2009; Howard et al., 2007; Q. Xu & Minor, 2009).

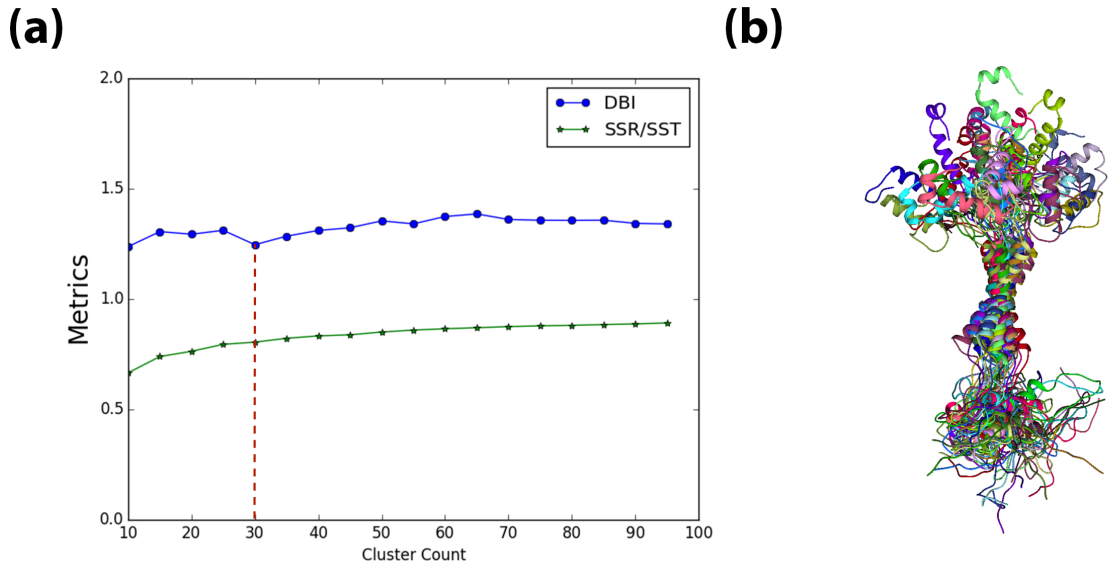


Figure 2.23. Clustering analysis of KCNE1 protein. (a) The clustering plot for KCNE1. (b) The 25 superimposed conformations of KCNE1 from the clustering analysis, representing more than 95% of the structure variability.

2.3.7. Assembly of The KCNQ1:KCNE1 Complex: Protein-Protein Docking

To obtain a physiologically relevant complex for KCNQ1/KCNE1, we implemented a data-driven protein-protein docking simulation. The HADDOCK (High Ambiguity Driven protein-protein DOCKing) tool (Dominguez et al., 2003) was used to perform the docking of KCNQ1 and KCNE1. HADDOCK is known to be one of the best information driven approaches to model biomolecular complexes. It is capable of incorporating input information related to the interface region between two interacting molecules. It represents the best tool for this study for this task, allowing us to drive the complex formation through fitting experimental data into our docking study.

In performing this task, we also tried to amend some of the well-known shortcomings in protein-protein docking simulations, such as ignoring protein flexibility

during docking (Bonvin, 2006), poor scoring methods and failure to cover the huge search space spanning the protein surfaces (Huang, 2014; Moal, Torchala, Bates, & Fernández-Recio, 2013). This involved two main stages throughout the docking workflow. First, we restricted the docking search space to the interface binding site of KCNQ1/KCNE1 as shown in Figure 2.21. The HADDOCK docking algorithm allows for restricting the search space to a limited region or group of residues on the surface of the proteins. Secondly, we provided HADDOCK with experimental restraints to probe the interaction between KCNQ1 and KCNE1, as will be described in the following section. HADDOCK also incorporates flexibility in the docking simulation, which is a significant improvement over other existing protein-protein docking algorithms.

It is well documented in the literature that two molecules of KCNE1 are associated with one KCNQ1 tetrameric protein (Nakajo, Ulbrich, Kubo, & Isacoff, 2010; Wrobel, Tapken, & Seebohm, 2012). The two KCNE1 proteins interact with KCNQ1 by binding to two opposite clefts in the structure. As shown in Figure 2.24, we named these possible sites as A, B C and D. As such, for a complete KCNQ1/KCNE1 complex, two KCNE1 proteins should occupy either the A/C or the B/D sites. Since the KCNQ1 ion channel is perfectly symmetrical, the two sites A/C and B/D are similar in residue composition. However, as we did not want to miss any trivial detail and also to increase the sampling space, we decided to perform two sets of dockings, named here as Set 1 and Set 2. The first set (Set 1) included the docking of two KCNE1 proteins at the A/C site and the second docking set (Set 2) encompassed the B/D site.

Therefore, the docking involved a 4:2 stoichiometry i.e. the KCNQ1 tetramer docked to two KCNE1 proteins. Furthermore, to guide the protein-protein docking sampling, we included several experimental interaction restraints adopted from the literature (Boulet, Labro, Raes, & Snyders, 2007; D. Y. Chung et al., 2009; Gofman, Shats, Attali, Haliloglu, & Ben-Tal, 2012; Kang et al., 2008; Kasimova et al., 2015; Panaghie, Tai, & Abbott, 2006; Strutz-Seebohm et al., 2011; Tapper & George, 2001; Y. Wang et al., 2012; Y. H. Wang et al., 2011; X. Xu et al., 2008). These restraints mainly came from disulfide mapping studies, site-directed mutagenesis analysis, combined with patch clamp electrophysiology techniques. HADDOCK uses the experimental information as ambiguous interaction restraints (AIRs) that will force the interfaces to

come together without imposing a particular orientation. The several steps that were implemented in HADDOCK to obtain the complexes formed by KCNQ1 and KCNE1 proteins are depicted in Figure 2.25 and explained in the following paragraphs.

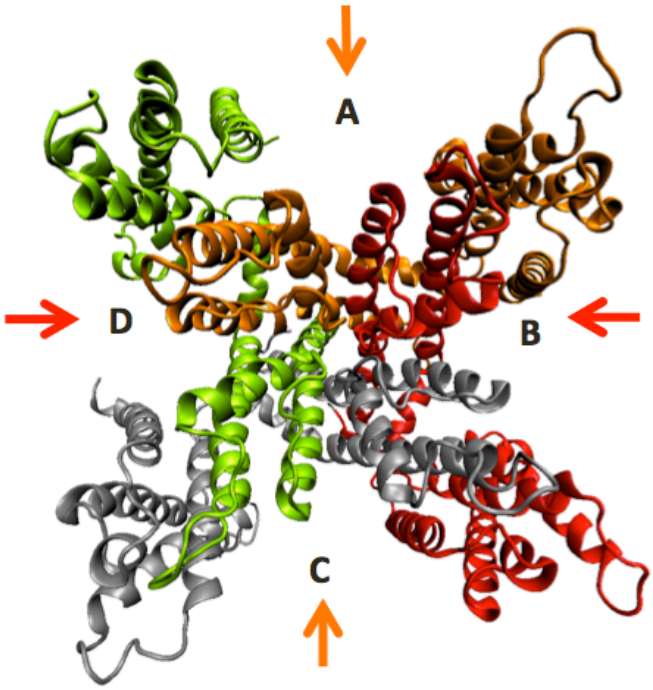


Figure 2.24. The KCNQ1 structure, showing the possible binding clefts for KCNE1. Two KCNE1 proteins would occupy either the A and C or the B and D clefts.

i) Docking Strategy

The 25 dominant conformations of KCNQ1 and 30 conformations of KCNE1 were fed into HADDOCK (Dominguez et al., 2003) to perform the 4:2 (KCNQ1 tetramer docked to 2 KCNE1 proteins) docking simulations and to probe the interaction between the two protein. HADDOCK 2.2 (Dominguez et al., 2003) is one of the very few software that takes protein flexibility into account on the fly. A typical protein-protein docking simulation with HADDOCK involves three main stages, (1) a rigid body docking stage, (2) a semi-flexible refinement stage and (3) a water refinement stage where the produced complexes are subjected to short MD simulations in water to improve their structural complementarities and remove the existing steric clashes. In general, protein-protein complexes produced by HADDOCK are of acceptable structural quality, requiring

minimal minimization refinement. To guide the docking simulation, we extracted all known "hot-spot" residues of KCNQ1 and KCNE1 interface from the literature (Boulet et al., 2007; D. Y. Chung et al., 2009; Gofman et al., 2012; Kang et al., 2008; Kasimova et al., 2015; Panaghie et al., 2006; Strutz-Seebohm et al., 2011; Tapper & George, 2001; Y. Wang et al., 2012; Y. H. Wang et al., 2011; X. Xu et al., 2008). We used these information as a set of ambiguous interaction restrains (AIRs) that are randomly excluded at each docking trial, guaranteeing a higher likelihood for converging to the best possible docking solution. This is different from other methods for imposing restraints on data driven protein-protein docking available in other docking packages.

In our docking simulations, 30,000 docking solutions were generated, out of which only the best 3,000 complexes (1,500 for each set) were retained. The 3,000 complexes after the initial rigid body docking were subject to semi-flexible and flexible docking. The result from this step was 500 complexes (250 from each set). The water refinement stage of HADDOCK was skipped for this docking simulation. The reason was because both KCNQ1 and KCNE1 are membrane proteins, which require hydrophobic environment. As such we did not want to stress the proteins structure and conformation by exposing them to the hydrophilic water environment. Next, the 3 step MD-based refinement for interface packing was implemented. This refinement stage resulted in 400 complexes in total, which were preserved for further analysis. These binding modes were rigorously visualized and filtered through a sophisticated protocol to identify the optimal binding conformation between KCNQ1 and KCNE1 proteins. These filtering steps are discussed below. The visualizations and distance calculations for the filtering of the poses were performed using TCL scripts in VMD (Humphrey, Dalke, & Schulten, 1996), and UCSF-Chimera (Pettersen et al., 2004).

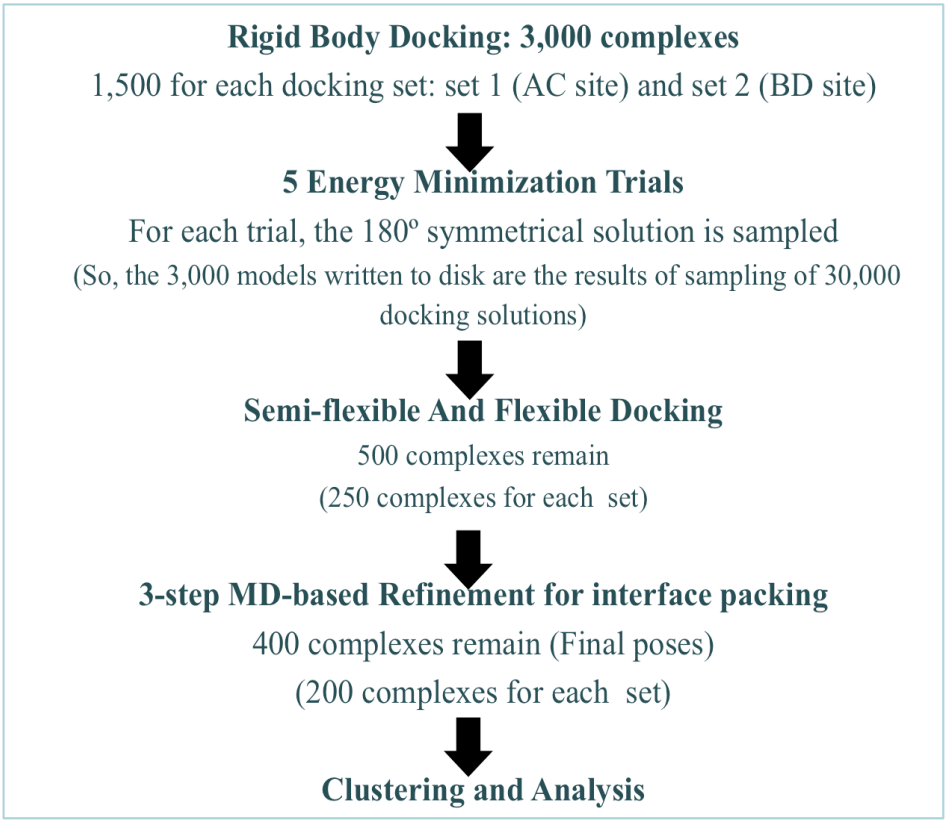


Figure 2.25. The protein-protein (KCNQ1-KCNE1) docking protocol in HADDOCK.

i) Analysis and Ranking of the Poses

The final 400 complexes (200 poses from each set) were retained for further analysis. These binding modes were rigorously filtered and visualized through a sophisticated protocol (See Figure 2.26) to identify the optimal binding conformation between KCNQ1 and KCNE1 proteins. These filtering steps were necessary to reduce the number of complexes to a manageable number for molecular dynamics (MD) simulations. The first criterion for filtering the poses was the angle of deflection and orientation of the KCNE1 protein towards the KCNQ1 interface. This was implemented by keeping an angular restrain of not less than 40 degrees between the planes of the proteins. This was done to remove the loosely packed complexes resulting in a significant elimination of poses. In the second phase of filtering only those complexes, which had the key interaction between KCNQ1 W208 and KCNE1 L45, were retained.

Similarly, the third filter involved the residue pair interaction between KCNQ1 C216 and KCNE1 G55. The fourth restraint-based filter was based upon the KCNQ1 R32

and KCNE1 K41 interaction. Finally, the 66 complexes after the fourth stage were visualized to identify the ones in which KCNE1 had a proximal distance to the S1 helix of KCNQ1 protein. After the several filtering stages, only 17 complexes out of the 400 met all the experimental criteria. These included 10 complexes from Set 1 and 7 complexes from Set 2, which were then rescored using the ZRANK scoring function (B. Pierce & Weng, 2007) (see Table 2.2).

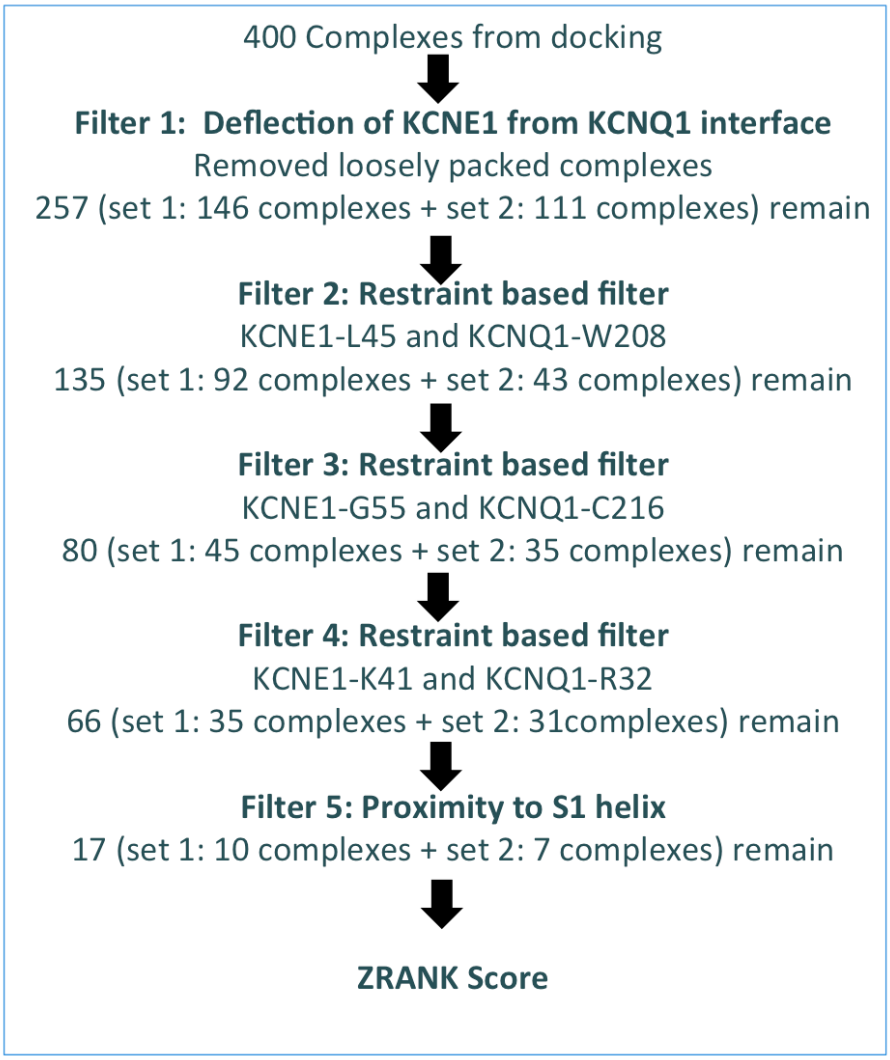


Figure 2.26. The filtering process of the top 400 complexes from the docking.

Table 2.2. Ranking of the final 17 complexes by ZRANK.

*Scoring was done twice for each complex (KCNE-1 and KCNE1-2).

**Set 1: docking at A/C site, Set 2: docking at B/D site with reference to Figure 2.24.

Complex #	ZRANK score: KCNE1-1*	ZRANK score: KCNE1-2*	Average ZRANK Score
SET 1**			
154	-130.976	-143.556	-137.266
117	-114.825	-157.732	-136.2785
118	-151.721	-115.758	-133.7395
93	-130.789	-132.23	-131.5095
79	-108.245	-133.513	-120.879
133	-100.555	-133.627	-117.091
59	-91.0198	-140.073	-115.5464
163	-105.005	-118.004	-111.5045
46	-115.778	-82.8347	-99.30635
139	-93.7249	-102.063	-97.89395
SET 2**			
51	-90.8565	-212.027	-151.44175
103	-136.779	-150.875	-143.827
112	-166.731	-82.83	-124.7805
157	-104.211	-125.349	-114.78
15	-103.757	-119.458	-111.6075
31	-98.7309	-105.247	-101.98895
187	-85.8292	-83.8901	-84.85965

Finally, the top 4 complexes (top 2 from each set) based on ZRANK scoring, were prepared for a short MD simulation and free energy calculations (MMGBSA) in AMBER. The results of the final scoring are shown in Table 2.3. The free energy calculation for each complex was done individually for each KCNE1 protein. Finally, the

average binding affinity was considered to select the top 2 docking poses, namely Complex #154 and #117.

Table 2.3. Ranking of the top 4 complexes by MMGBSA.

*Free energy calculation was done twice for each complex (KCNE-1 and KCNE1-2).

Complex #	Binding Affinity (kcal/mol): KCNE1-1*	Binding Affinity (kcal/mol): KCNE1-2*	Average Binding Affinity (kcal/mol)
117 (Set 1)	-32.29	-30.55	-31.42
154 (Set 1)	-31.92	-22.52	-27.22
51 (Set 2)	-30.67	-20.0	-25.34
103 (Set 2)	-26.74	-19.20	-22.97

2.3.8. Refinement of the Complexes: MD Simulation of KCNQ1/KCNE1

The final two complexes, i.e. complexes 154 and 117 were selected for subsequent studies. As a post-processing stage to the docked protein complexes, long classical MD simulations were conducted. The same protocol as before (see Methodology), with the inclusion of lipid bilayer membrane composed of POPC and PIP_2 lipids in the ratio of 10:1, water and ions were used. The two complexes are shown in Figure 2.27. The MD simulation for each complex was run for ~240 ns, during which the two proteins were allowed to interact with each other in a dynamic environment. Subsequent to the MD simulations, the two complexes were analyzed closely to infer more information related to the interaction of the two proteins. This analysis is described in the following section.

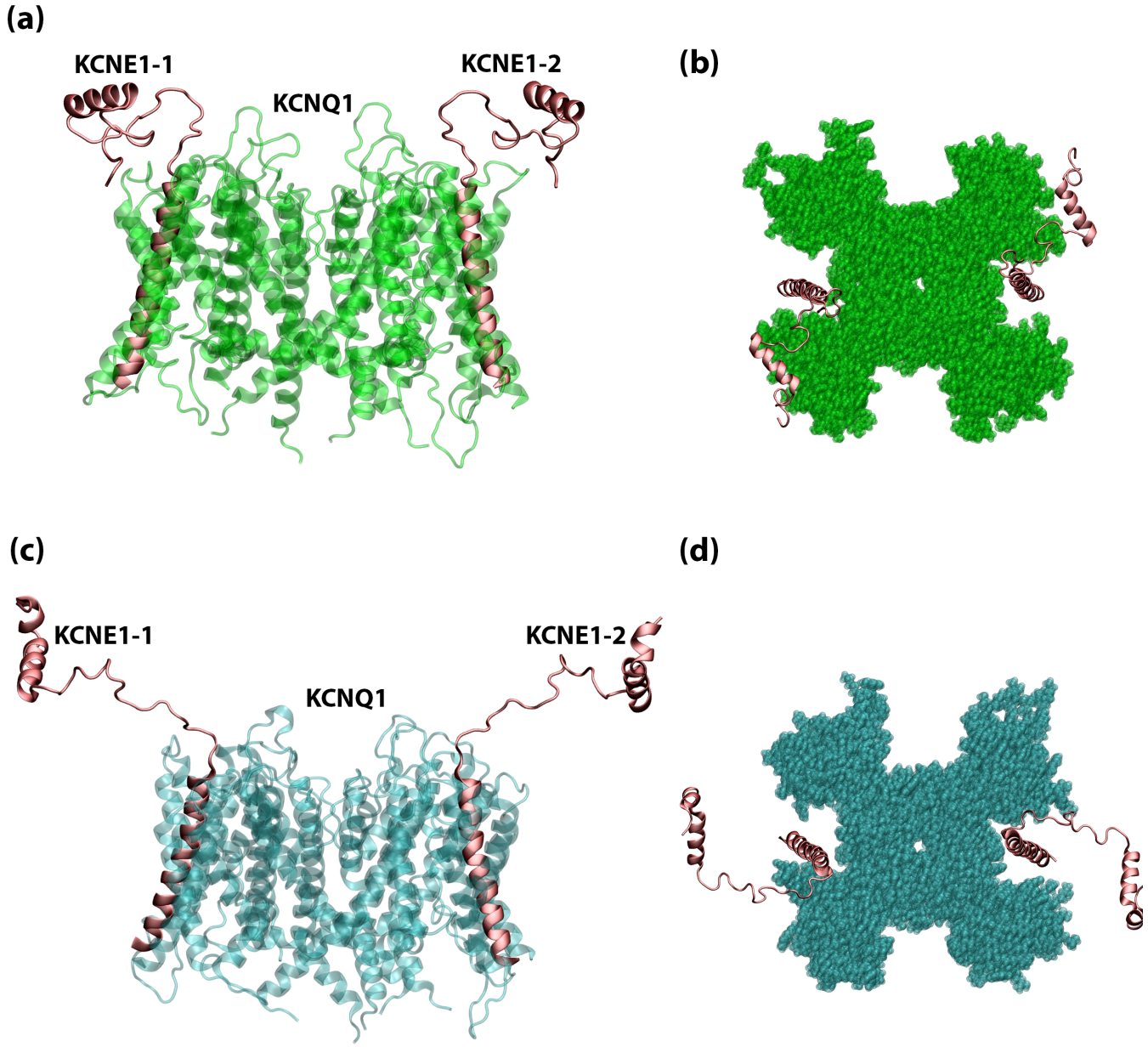


Figure 2.27. The structure of the final two complexes. Complex #154; (a) lateral view, (b) top view and Complex #117; (c) lateral view, (d) top view.

2.3.9. Analysis of KCNQ1/KCNE1 Interactions

To understand the interaction between the two proteins, we performed energetic evaluation as well as structural analysis. Energetic evaluation was done using extensive MM-GBSA calculations along with a breakdown of the total binding energy by both pair-wise and per residue decomposition analysis. Calculation of the binding energy allowed us to confirm that there is no energetically unfavourable interaction happening between the two proteins. Furthermore, we were able to understand the contributing energies, e.g. Vdw, electrostatic interactions using the decomposition analysis. The analysis was done for the both the final two complexes as well as for both the bound KCNE1 proteins (KCNE1-1, KCNE1-2) on the two opposite sides of the KNQ1 channel.

Structural analysis provided insights into the specific binding residues and hot spots at the protein-protein interface. The contacts made during the dynamic interaction of the two proteins over the 240 ns of the simulation were compared to the experimental restraints, from the cross-linking and mutagenesis studies (Boulet et al., 2007; D. Y. Chung et al., 2009; Gofman et al., 2012; Kang et al., 2008; Kasimova et al., 2015; Panaghie et al., 2006; Strutz-Seebohm et al., 2011; Tapper & George, 2001; Y. Wang et al., 2012; Y. H. Wang et al., 2011; X. Xu et al., 2008). The H-bond occupancies were measured with a distance cut-off of 3.5 and a cut-off angle of 60 degrees for every 3 frames of the MD trajectory. Tables B.1-B.4 in Appendix D show the H-bonds formed between the KCNE1-1 and KCNE1-2 residues and KCNQ1 for the two complexes. Only the residues having an occupancy of 10% or greater were retained. The hydrophobic interaction analysis was performed similarly and Tables D.5-D.8 in Appendix D, show these results. The overall number of hydrogen bonds formed during the MD simulation are plotted in Figure 2.28. As shown in the graphs, complex #154 possessed a larger number of H-bonds ranging between 30-100 compared to those observed in complex #117, namely in the range of 10-40 bonds.

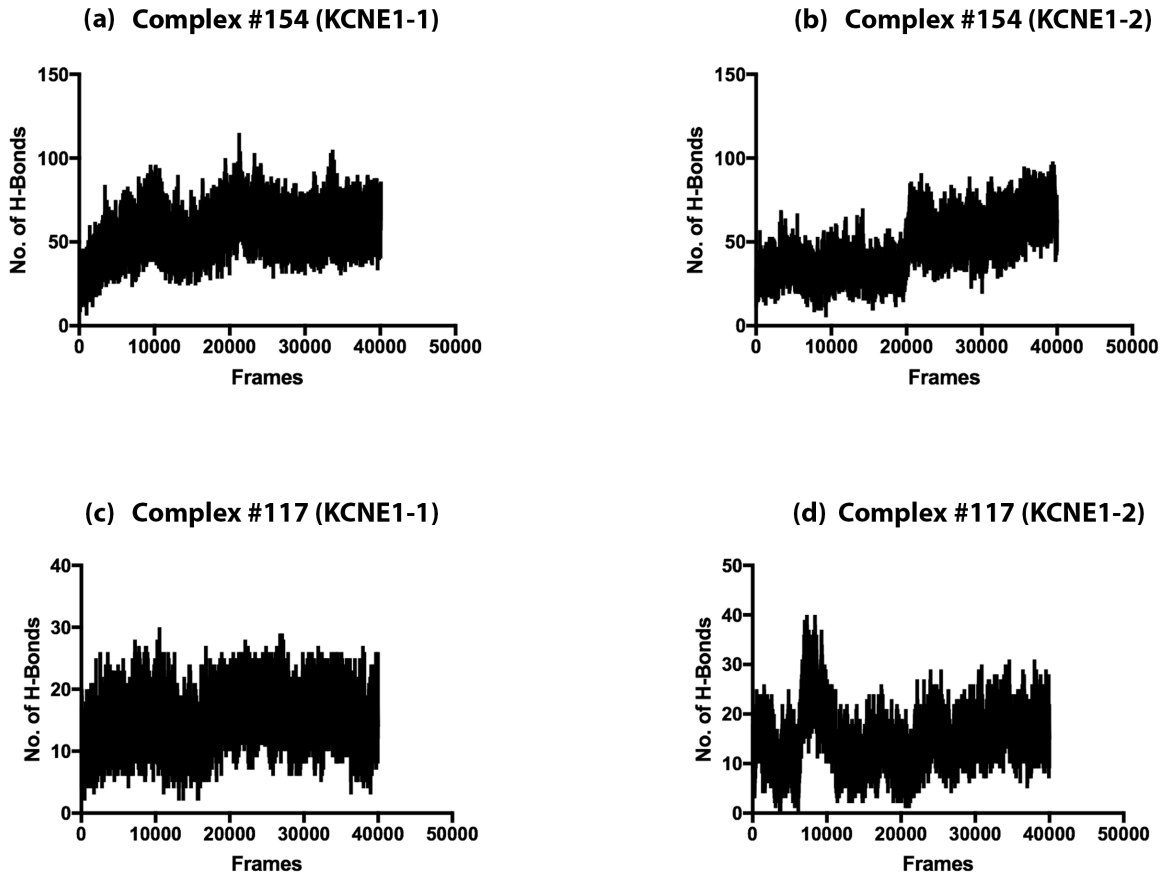


Figure 2.28. Hydrogen bond formation between KCNQ1 and KCNE1 in the two complexes.
Number of hydrogen bonds formed during the MD simulation for (a) KCNE1-1 and (b) KCNE1-2 in complex #154, (c) KCNE1-1 and (d) KCNE1-2 in complex #117.

The data obtained from the MM-GBSA binding free energy calculation as well as energy decomposition analysis together with the structural analysis confirmed that the KCNE1 proteins interacted favourably with KCNQ1 in the two complexes. The key interactions were present in both the complexes. However, there was one evident difference between complex #154 and #117. This difference mainly lies in the interactions of the N-terminal of the KCNE1 proteins. In Complex #154, the N-terminal tails are folded back on top of the KCNQ1 TMD region, which significantly increases the number of contacts. The N-terminal residues made extensive contacts with the S1-S2 linker, S3 segments, S3-S4 linker and S5-S6 linkers of the adjacent subunits. These interactions are absent in complex #117, as the tails are far from the KCNQ1 TMD

surface and are extended outwards. Figure 2.27 clearly shows this difference in the starting structures.

Figure 2.29 also shows the interactions of KCNE1 NT regions with KCNQ1. These contacts are mainly made with the S1-S2 linker, S3-S4 linker and S5-S6 linker. Of utmost importance are the hydrophobic contacts made between PHE12, LEU13, LEU16, SER34 of KCNE1 and SER217, ALA149, LEU142 and LEU156 of KCNQ1. The trend of NT interaction was similar with the KCNE1-1 and KCNE-2.

The cytoplasmic interactions on the other hand, involved the residues from S1, S5, S5-P linker, P-loop and a limited number of contacts with S6. These S1 interactions are shown in Figure 2.30.a, P-loop and S5 interactions in Figure 2.30.b. The latter are significant interactions as they may affect the pore conformation and thereby the ion conduction.

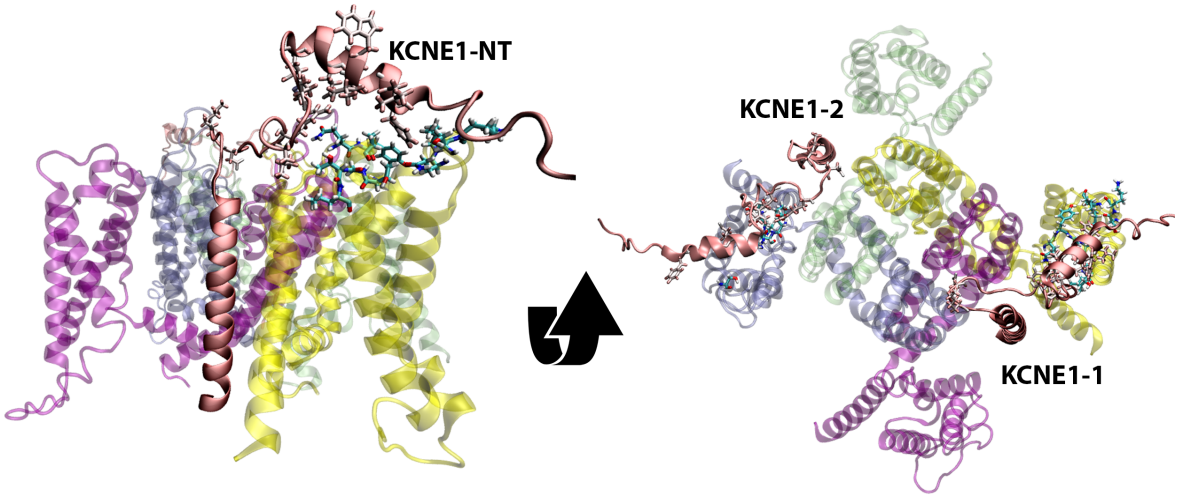


Figure 2.29. Interaction of KCNE1 NT (complex #154) with KCNQ1.

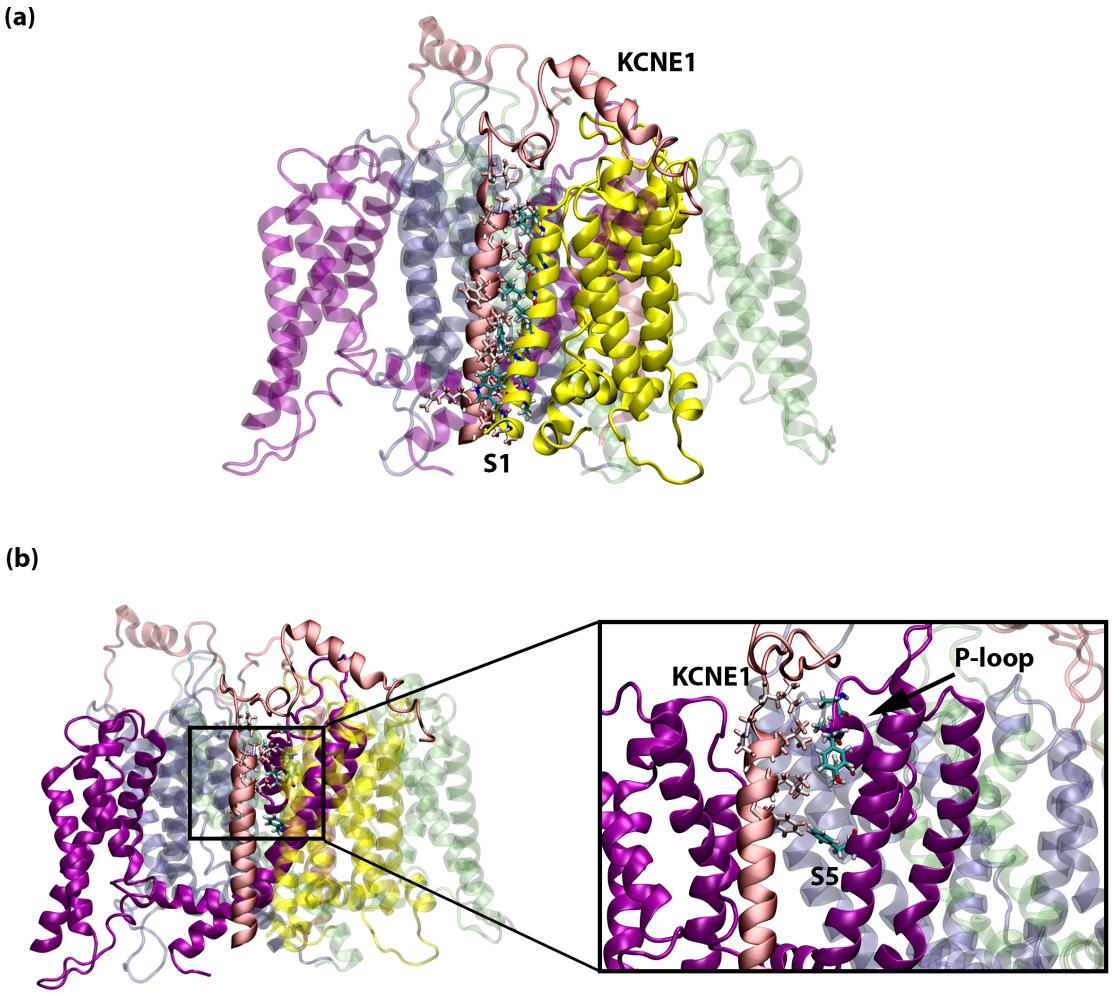


Figure 2.30. Interaction of KCNE1 TMD with KCNQ1. (a) contacts made with S1, (b) contacts made with P-loop.

To understand more about the dynamics of the two proteins in complex, the RMSF graphs of the proteins were plotted. Figure 2.31.a and 7.31.b shows the RMSF of KCNE1-1 and KCNE1-2 proteins in complex #154, while 7.31.c and 7.31.d belong to KCNE1-1 and KCNE1-2 proteins in complex #117. It is clear from the graphs that the NT of KCNE1 proteins in complex #154 had minimal fluctuations while those of KCNE1's in complex #117, were highly flexible and fluctuated significantly during the MD simulation. The structural stability in complex #154 stems from the fact that the NT of KCNE1 is engaged in interactions with several regions of KCNQ1, while these interactions are absent in complex #117 and are free to move. This agrees with the interaction analysis discussed above regarding the main difference between the two

complexes. As for the TMD, it is very interesting to see almost zero fluctuations taking place during the simulation. KCNE1 possessed high flexibility and backbone motion when it was simulated alone in lipid bilayer (section 7.5). The interaction with KCNQ1 has stabilized the TMD backbone and all the fluctuations have disappeared.

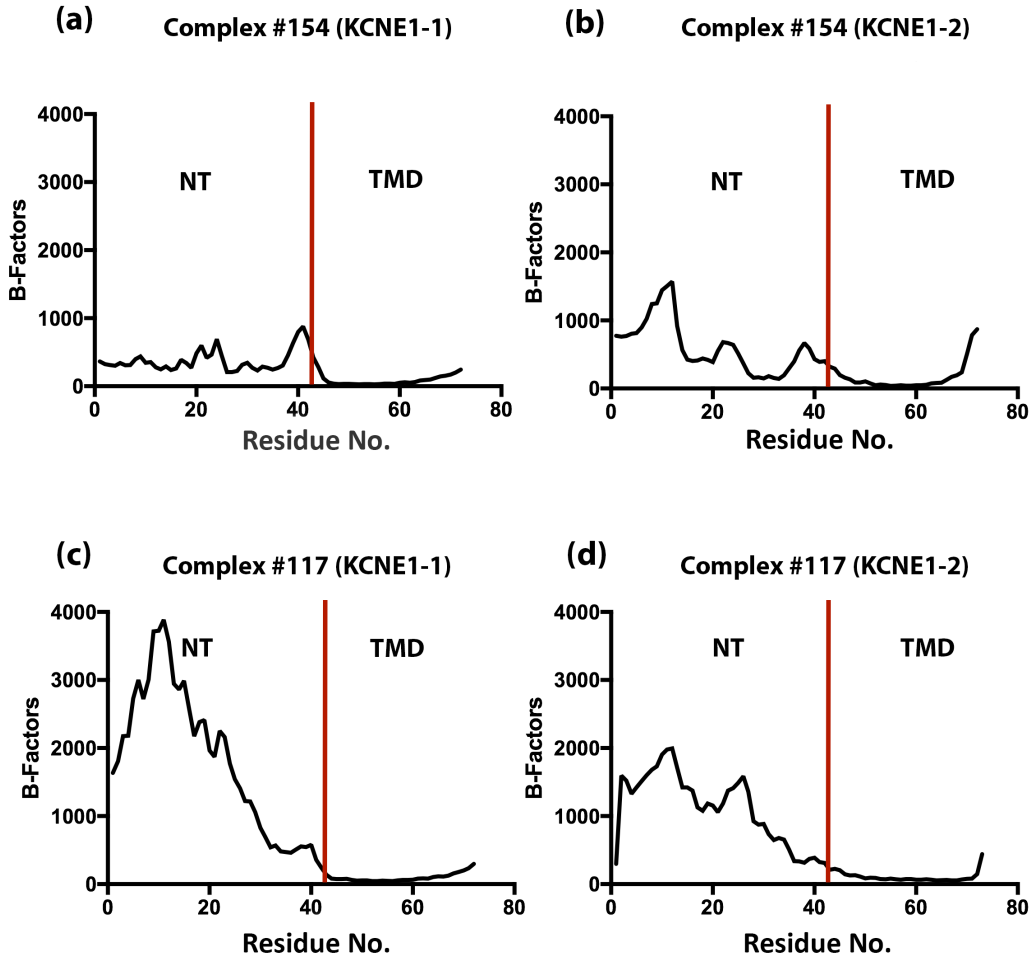


Figure 2.31. C-alpha backbone RMSF graphs of KCNE1 in the two complexes. (a) KCNE1-1 in complex #154, (b) KCNE1-2 in complex #154, (c) KCNE1-1 in complex #117 and (d) KCNE1-2 in complex #117.

Similarly, the RMSF graphs of the KCNQ1 protein in complex (#154) are shown in Figure 2.32. To have a better idea of how the presence of KCNE1 affected the structure of KCNQ1, this plot was combined with the RMSF of KCNQ1 when it was simulated alone in Figure 2.33. It is clear from the graphs that the motion of the KCNQ1 subunit structures has been significantly reduced. This graph confirms that the presence of KCNE1 has limited the structural fluctuations of the KCNQ1 structure; indicative of

the effective interaction that is taking place between the two proteins. Also, the RMSD graphs of KCNQ1 alone, KCNQ1 in complex #154 and in complex #117, were plotted together (see Figure 2.34). The structural variability of KCNQ1 is modestly affected by the presence of KCNE1 and resulted in a more stable RMSD graphs. The backbone RMSDs of the two complexes during the MD simulations were similar. Further discussion on the structural effects KCNE1 imposed on KCNQ1 during the MD simulation will be discussed in the following chapters.

Overall, from our analysis, we hypothesize that two modes of interaction between KCNQ1 and KCNE1 can exist. First possibility is that the N-terminals of KCNE1 will be highly participating in the interaction and are similar to complex #154. The second possibility involves an extended configure ration of the KCNE1 N-terminals as in complex #117, such that they are oriented in outward direction rather than folded back into the membrane. However, for the further applications and studies on this ion channel, we chose complex #154 to proceed with. This choice was mainly made because complex #154, not only satisfied all key interactions that must take place between KCNQ1 and KCNE1, but also this complex possessed more favourable interactions (H-bond and hydrophobic). This was in a better agreement with the available experimental information, and also because the KCNE1 in this complex imposed a profound structural effect on the KCNQ1 protein during the MD simulation in this complex.

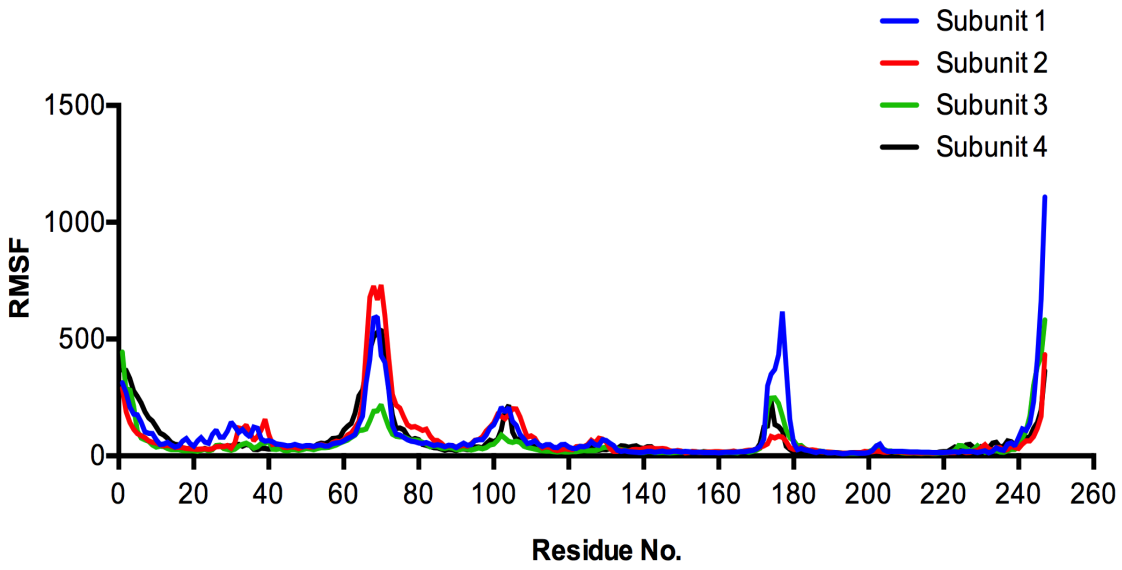


Figure 2.32. C-alpha backbone RMSF of four KCNQ1 subunits in complex # 154.

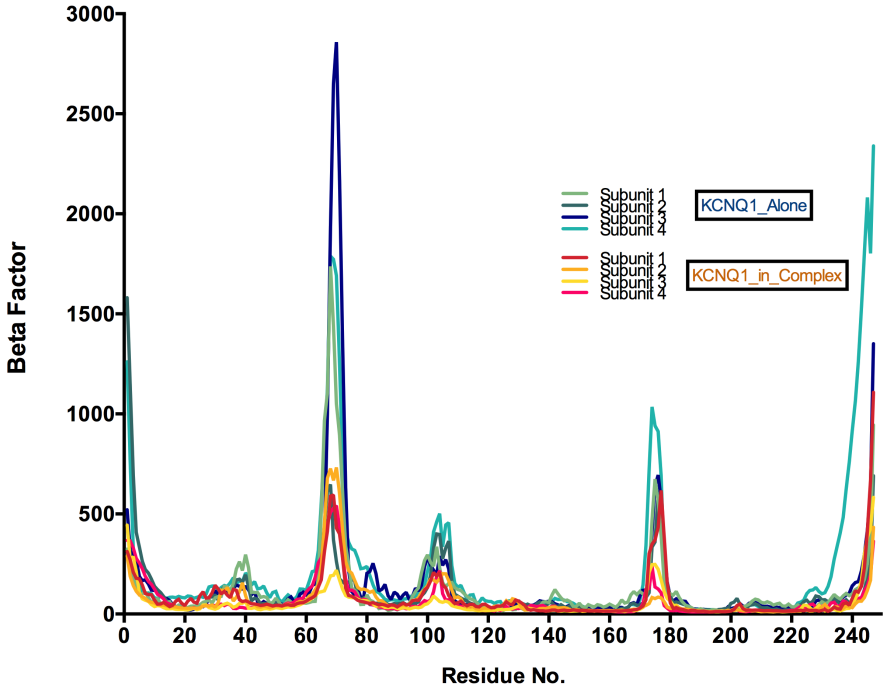


Figure 2.33. RMSF fluctuations of the four KCNQ1 subunits, when in complex with KCNE1 (shown in in red, orange, yellow, pink) compared to when simulated alone in absence of KCNE1 protein (shown in shades of blue).

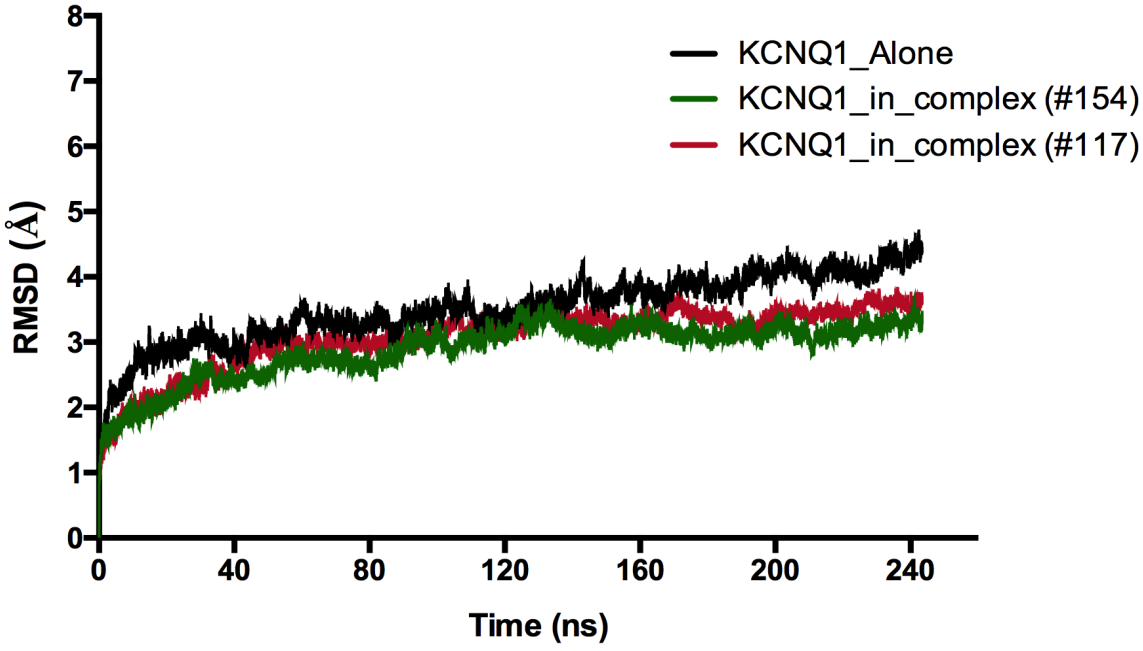


Figure 2.34. RMSD graphs of KCNQ1 alone, KCNQ1 in complex #154 and in complex #117.

2.4. Conclusion

Using a combination of comparative modeling and enhanced MD simulations, we built robust model for the KCNQ1 protein, which accommodated the most recent experimental findings. The inclusion of experimental data was done using an enhanced MD simulation; REMD to impose structural details that were reported in the literature. The biological complex that was formed between KCNQ1 and KCNE1in this study was modelled in order to mimic the physiological open state of this ion channel in complex with the KCNE1 auxiliary protein. With the help of MD simulations, we were able to capture the dynamical behaviour of the two proteins, alone and in complex. Comparison and analysis of these MD simulations highlighted several findings related to their effect on each other and on their interaction. We concluded that KCNE1 can interact with KCNQ1 in two possible conformations and in general the presence of KCNE1 proteins significantly minimized the level of structural fluctuations for KCNQ1 protein.

Amongst the two complexes that were chosen for deeper analysis, we ended up selecting complex #154 in which the N-terminal of KCNE1 proteins form more interactions with the KCNQ1 channel. This was based on an extensive analysis of the interactions of the two proteins as well as based on the analysis of their MD simulations. In general, the association of kCNQ1 with KCNE1 significantly reduced the magnitude of the VSD fluctuations. In this final complex, the two proteins satisfied all experimental constraints and exhibited several favourable contacts that had been retained throughout the MD. The current model of KCNQ1/KCNE1 protein complex can aid in future studies related to function, structure, effect of small molecule blockers and cardiotoxicity studies of drug molecules.

2.5. References

Alford, R. F., Koehler Leman, J., Weitzner, B. D., Duran, A. M., Tilley, D. C., Elazar, A., & Gray, J. J. (2015). An Integrated Framework Advancing Membrane Protein Modeling and Design. *PLoS Computational Biology*, *11*(9), e1004398. http://doi.org/10.1371/journal.pcbi.1004398

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. http://doi.org/http://dx.doi.org/10.1016/S0022-2836(05)80360-2

Aqvist, J., & Luzhkov, V. (2000). Ion permeation mechanism of the potassium channel. *Nature*, *404*(6780), 881–884. http://doi.org/10.1038/35009114

Bakk, A., Høye, J. S., & Hansen, A. (2001). Heat Capacity of Protein Folding. *Biophysical Journal*, *81*(2), 710–714. http://doi.org/http://dx.doi.org/10.1016/S0006-3495(01)75735-9

Bond, P. J., & Sansom, M. S. P. (2003). Membrane protein dynamics versus environment: simulations of OmpA in a micelle and in a bilayer. *Journal of Molecular Biology*, *329*(5), 1035–1053.

Boulet, I. R., Labro, A. J., Raes, A. L., & Snyders, D. J. (2007). Role of the S6 C-terminus in KCNQ1 channel gating. *The Journal of Physiology*, *585*(Pt 2), 325–337. http://doi.org/10.1113/jphysiol.2007.145813

Chen, J., Zheng, R., Melman, Y. F., & McDonald, T. V. (2009). Functional interactions between KCNE1 C-terminus and the KCNQ1 channel. *PLoS ONE*, *4*(4), 1–9. http://doi.org/10.1371/journal.pone.0005143

Chowdhury, S., & Chanda, B. (2015). Basic mechanisms of voltage sensing. *Handbook of Ion Channels*, 25–40. http://doi.org/doi:10.1201/b18027-5

Chung, D. Y., Chan, P. J., Bankston, J. R., Yang, L., Liu, G., Marx, S. O., … Kass, R. S. (2009). Location of KCNE1 relative to KCNQ1 in the IKS potassium channel by disulfide cross-linking of substituted cysteines. *Proceedings of the National Academy of Sciences* , *106*(3), 743–748. Retrieved from http://www.pnas.org/content/106/3/743.abstract

Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *1*(2), 224–227.

Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–1737. http://doi.org/10.1021/ja026939x

Eisenberg, D., Luthy, R., & Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in Enzymology*, *277*, 396–404.

Frishman, D., & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function and Genetics*. http://doi.org/10.1002/prot.340230412

Gayen, S., Li, Q., & Kang, C. (2015). Structural analysis of the S4–S5 linker of the human KCNQ1 potassium channel. *Biochemical and Biophysical Research Communications*, *456*(1), 410–414. http://doi.org/10.1016/j.bbrc.2014.11.097

Gofman, Y., Shats, S., Attali, B., Haliloglu, T., & Ben-Tal, N. (2012). How does KCNE1 regulate the Kv7.1 potassium channel? Model-structure, mutations, and dynamics of the Kv7.1-KCNE1 complex. *Structure*, *20*(8), 1343–1352. http://doi.org/10.1016/j.str.2012.05.016

Hildebrand, P. W., Goede, A., Bauer, R. A., Gruening, B., Ismer, J., Michalsky, E., & Preissner, R. (2009). SuperLooper--a prediction server for the modeling of loops in globular and membrane proteins. *Nucleic Acids Research*, *37*(Web Server issue), W571-4. http://doi.org/10.1093/nar/gkp338

Howard, R. J., Clark, K. a., Holton, J. M., & Minor, D. L. (2007). Structural Insight into KCNQ (Kv7) Channel Assembly and Channelopathy. *Neuron*, *53*(5), 663–675. http://doi.org/10.1016/j.neuron.2007.02.010

Jentsch, T. J. (2000). Neuronal KCNQ potassium channels: physiology and role in disease. *Nature Reviews. Neuroscience*, *1*(1), 21–30. http://doi.org/10.1038/35036198

Kang, C., Tian, C., Sönnichsen, F. D., Smith, J. A., Meiler, J., George, A. L., … Sanders, C. R. (2008). Structure of KCNE1 and Implications for How It Modulates the KCNQ1 Potassium Channel. *Biochemistry*, *47*(31), 7999–8006. http://doi.org/10.1021/bi800875q

Kasimova, M. a, Zaydman, M. a, Cui, J., & Tarek, M. (2015). PIP2-dependent coupling is prominent in Kv7.1 due to weakened interactions between S4-S5 and S6, 2–10. http://doi.org/10.1038/srep07474

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(C), 283–291. http://doi.org/10.1107/S0021889892009944

Liu, S., Focke, P. J., Matulef, K., Bian, X., Moënne-Loccoz, P., Valiyaveetil, F. I., & Lockless, S. W. (2015). Ion-binding properties of a K^+ channel selectivity filter in different conformations. *Proceedings of the National Academy of Sciences*, *112*(49), 15096–15100. http://doi.org/10.1073/pnas.1510526112

Long, S. B., Tao, X., Campbell, E. B., & MacKinnon, R. (2007). Atomic structure of a voltage-dependent K+ channel in a lipid membrane-like environment. *Nature*, *450*(7168), 376–382. http://doi.org/10.1038/nature06265

Luthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, *356*(6364), 83–85. http://doi.org/10.1038/356083a0

Morais-Cabral, J. H., Zhou, Y., & MacKinnon, R. (2001). Energetic optimization of ion conduction rate by the K+ selectivity filter. *Nature*, *414*(6859), 37–42. Retrieved from http://dx.doi.org/10.1038/35102000

Mousavi Nik, A., Gharaie, S., & Jeong Kim, H. (2015). Cellular mechanisms of mutations in Kv7.1: auditory functions in Jervell and Lange-Nielsen syndrome vs. Romano–Ward syndrome. *Frontiers in Cellular Neuroscience*, *9*, 32. http://doi.org/10.3389/fncel.2015.00032

Nakajo, K., Ulbrich, M. H., Kubo, Y., & Isacoff, E. Y. (2010). Stoichiometry of the KCNQ1 - KCNE1 ion channel complex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(44), 18862–18867. http://doi.org/10.1073/pnas.1010354107

Osteen, J. D., Sampson, K. J., & Kass, R. S. (2010). The cardiac IKs channel, complex indeed. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(44), 18751–18752. http://doi.org/10.1073/pnas.1014150107

Panaghie, G., Tai, K.-K., & Abbott, G. W. (2006). Interaction of KCNE subunits with the KCNQ1 K+ channel pore. *The Journal of Physiology*, *570*(Pt 3), 455–467. http://doi.org/10.1113/jphysiol.2005.100644

Peng, D., Kim, J., Kroncke, B. M., Law, C. L., Xia, Y., Droege, K. D., … Sanders, C. R. (2014). Purification and Structural Study of the Voltage-Sensor Domain of the Human KCNQ1 Potassium Ion Channel. *Biochemistry*, *53*(12), 2032–2042. http://doi.org/10.1021/bi500102w

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. http://doi.org/10.1002/jcc.20084

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., … Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802. http://doi.org/10.1002/jcc.20289

Pierce, B., & Weng, Z. (2007). ZRANK: Reranking Protein Docking Predictions With an Optimized Energy Function. *Proteins*, *67*, 1078–1086.

Remmert, M., Biegert, A., Hauser, A., & Soding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Meth*, *9*(2), 173–175. Retrieved from http://dx.doi.org/10.1038/nmeth.1818

Roe, D. R., & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. *J Chem Theory Com*, *9*(7), 3084–3095. http://doi.org/10.1021/ct400341p

Shao, J. ., Tanner, S. W. ., Thompson, N. ., & Cheatham III, T. E. (2007). Clusteing molecular dynamic trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.*, *3*, 2312–2334. http://doi.org/10.1021/ct700119m

Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins*, *17*(4), 355–362. http://doi.org/10.1002/prot.340170404

Soding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, *33*(Web Server issue), W244-8. http://doi.org/10.1093/nar/gki408

Strutz-Seebohm, N., Pusch, M., Wolf, S., Stoll, R., Tapken, D., Gerwert, K., … Seebohm, G. (2011). Structural basis of slow activation gating in the cardiac I Ks channel complex. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *27*(5), 443–452. http://doi.org/10.1159/000329965

Tapper, A. R., & George, A. L. (2001). Location and Orientation of minK within the IKs Potassium Channel Complex. *Journal of Biological Chemistry*, *276*(41), 38249–38254. http://doi.org/10.1074/jbc.M103956200

Thompson, A. N., Kim, I., Panosian, T. D., Iverson, T. M., Allen, T. W., & Nimigean, C. M. (2009). Mechanism of potassium channel selectivity revealed by Na(+) and Li(+) binding sites within the KcsA pore. *Nature Structural & Molecular Biology*, *16*(12), 1317–1324. http://doi.org/10.1038/nsmb.1703

Thompson, J., & Begenisich, T. (2001). Affinity and location of an internal K+ ion binding site in shaker K channels. *The Journal of General Physiology*, *117*(5), 373–384.

Wallner, B., & Elofsson, A. (2003). Can correct protein models be identified? *Protein Science : A Publication of the Protein Society*, *12*(5), 1073–1086. http://doi.org/10.1110/ps.0236803

Wang, Q., Curran, M. E., Splawski, I., Burn, T. C., Millholland, J. M., VanRaay, T. J., … Keating, M. T. (1996). Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nature Genetics*, *12*(1), 17–23. http://doi.org/10.1038/ng0196-17

Wang, Y. H., Jiang, M., Xu, X. L., Hsu, K.-L., Zhang, M., & Tseng, G.-N. (2011). Gating-Related Molecular Motions in the Extracellular Domain of the IKs Channel: Implications for IKs Channelopathy. *The Journal of Membrane Biology*, *239*(3), 137–156. http://doi.org/10.1007/s00232-010-9333-7

Wang, Y., Zhang, M., Xu, Y., Jiang, M., Zankov, D. P., Cui, M., & Tseng, G.-N. (2012). Probing the structural basis for differential KCNQ1 modulation by KCNE1 and KCNE2. *The Journal of General Physiology*, *140*(6), 653–69. http://doi.org/10.1085/jgp.201210847

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server issue), W407-10. http://doi.org/10.1093/nar/gkm290

Wrobel, E., Tapken, D., & Seebohm, G. (2012). The KCNE tango - how KCNE1 interacts with Kv7.1. *Frontiers in Pharmacology*, *3 AUG*(August), 1–14. http://doi.org/10.3389/fphar.2012.00142

Wu, E. L., Cheng, X., Jo, S., Rui, H., Song, K. C., Dávila-Contreras, E. M., … Im, W. (2014). CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *Journal of Computational Chemistry*, 1997–2004. http://doi.org/10.1002/jcc.23702

Wu, S., & Zhang, Y. (2007). LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Research*, *35*(10), 3375–3382. http://doi.org/10.1093/nar/gkm251

Wulff, H., Castle, N. A., & Pardo, L. A. (2009). Voltage-gated Potassium Channels as Therapeutic Drug Targets. *Nature Reviews. Drug Discovery*, *8*(12), 982–1001. JOUR. http://doi.org/10.1038/nrd2983

Xu, D., Zhang, J., Roy, A., & Zhang, Y. (2011). Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins*, *79 Suppl 1*, 147–160. http://doi.org/10.1002/prot.23111

Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, *101*(10), 2525–2534. http://doi.org/10.1016/j.bpj.2011.10.024

Xu, Q., & Minor, D. L. (2009). Crystal structure of a trimeric form of the KV7.1 (KCNQ1) A-domain tail coiled-coil reveals structural plasticity and context dependent changes in a putative coiled-coil trimerization motif. *Protein Science*, *18*(10), 2100–2114. http://doi.org/10.1002/pro.224

Xu, X., Jiang, M., Hsu, K.-L., Zhang, M., & Tseng, G.-N. (2008). KCNQ1 and KCNE1 in the IKs channel complex make state-dependent contacts in their extracellular domains. *The Journal of General Physiology*, *131*(6), 589–603. http://doi.org/10.1085/jgp.200809976

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2014). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, *12*(1), 7–8. http://doi.org/10.1038/nmeth.3213

Zhang, Y. (2007). Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins*, *69 Suppl 8*, 108–17. http://doi.org/10.1002/prot.21702

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, *9*(1), 40. http://doi.org/10.1186/1471-2105-9-40

Zhang, Y. (2009). I-TASSER: fully automated protein structure prediction in CASP8. *Proteins*, *77 Suppl 9*, 100–113. http://doi.org/10.1002/prot.22588

Zhang, Y. (2014). Interplay of I-TASSER and QUARK for template-based and ab initio protein structure prediction in CASP10. *Proteins*, *82 Suppl 2*, 175–187. http://doi.org/10.1002/prot.24341

Zhang, Y., & Skolnick, J. (2004). SPICKER: a clustering approach to identify near-native protein folds.

Journal of Computational Chemistry, *25*(6), 865–871. http://doi.org/10.1002/jcc.20011
Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, *33*(7), 2302–2309. http://doi.org/10.1093/nar/gki524

CHAPTER 3: ION PERMEATION STUDIES IN KCNQ1/KCNE1 ION CHANNEL[5]

3.1. Introduction

The cardiac KCNQ1/KCNE1 ion channel is a voltage-gated potassium channel, that has a unique ability to selectively allow potassium ions to pass through the cellular membrane. The maintenance of this normal ion flux gives the KCNQ1/KCNE1 ion channel, its unique role in controlling the duration of the repolarization phase of the cardiac action potential. The selective behavior of the KCNQ1/KCNE1 channel results from the intricate structural properties its pore domain, which forms the path of the ions' passage (Doyle et al., 1998; Gary Yellen, 2002). The KCNQ1 pore domain has two openings; one in the intracellular part of the protein and the other is on the extracellular region. This study represents the first application of our generated KCNQ1/ KCNE1 model and summarizes our findings on how potassium ions pass through the channel. Furthermore, it is known from the literature that the association of KCNE1 with KCNQ1 slows the activation of KCNQ1 by 5- to 10-folds (Geoffrey W. Abbott, 2014; Melman et al., 2004; Seebohm, 2013). Another profound effect of the KCNE1 protein on KCNQ1 is the paradoxical slowing of the gating associated with KCNQ1 channel inactivation, which otherwise takes place in a fast manner (Tristani-Firouzi & Sanguinetti, 1998). In this chapter, we also present our findings related to the mechanisms by which the presence of KCNE1 protein affects the ion permeation through KCNQ1 channel. For this purpose, we used an advanced MD simulation technique, called Steered Molecular Dynamics (SMD), to address this specific research question and investigate how potassium ions cross the channel and what is the role of KCNE1 in this process.

Furthermore, many drugs offer cardiotoxic effects secondary to blocking the voltage-gated ion channels in the heart, which represents in the form of Torsade de Points (TdP) arrhythmia, acquired LQTS and eventually sudden cardiac death (Kannankeril, Roden, & Darbar, 2010; Yap & Camm, 2003). Of utmost importance in this regard are

[5] A version of this Chapter has been submitted to The Journal of Molecular Modeling as *Jalily Hasani H, Ahmed M, Ganesan A, Barakat K. Ion Permeation Studies in the Human KCNQ1/KCNE1 Ion Channel.*

the HERG, KCNQ1/KCNE1, SCN5A (Nav1.5), Cav1.2 and Kir2.1 that are implicated in the lethal cardiotoxicity of drugs (Anwar-Mohamed et al., 2014; Paulussen et al., 2004; Witchel & Hancox, 2000). Therefore in this realm, it is important to get insights into the mechanisms by which small molecule drugs block these ion channels. *In silico* approaches have been successfully employed to maximize the chances of an accurate assessment of whether a drug is capable of blocking ion channels (Khaled Barakat, 2015; Mirams et al., 2011), and thereby minimizing the risk in an area where prolonged QT can result in death. In this study, we have made attempts to reveal some of the basic mechanisms by which blockers of the KCNQ1/KCNE1 ion channel affect the K+ ion permeation. To address this research question, we employed a structural model of KCNQ1/KCNE1 constructed in our earlier study and used a combination of small molecule docking calculations and SMD simulations. To validate our KCNQ1 model and to confirm our simulations' protocol, we investigated the channel's interactions with a well-known KCNQ1 blocker, namely, Chromanol 293B and its derivatives. While docking simulations placed the tested compounds in their favourable locations within the channel, SMD simulations were used to pull potassium ions through the selectivity filter of the channel in the presence of the compounds. This sophisticated simulation approach allowed us to understand the atomistic effects of the docked ligands on the ion passage. Furthermore, the results make us believe that the model is capable of predicting other off-target blockage, in the future and aid in future studies of drug-induced cardiotoxicity. To our knowledge, this is a unique and novel aspect of this work and has not been studied before in such structural details.

3.2. Methods

MD Simulation: System Preparation in the membrane[6]

When membrane proteins are studied, it is crucial to have a proper inclusion of a membrane to ensure accuracy because these proteins are highly sensitive to their surrounding environment. Similarly, it was essential to include a membrane system for the different systems in this study, i.e. KCNQ1/KCNE1 complexes and KCNQ1/KCNE1

[6] The concepts of MD simulations and its enhanced variants are provided in Appendix A.

docked to ligands. The arrangement of the protein in the membrane was built using the Membrane Builder of CHARMM GUI (*http://www.charmm gui.org*). The tetrameric protein was embedded in a bilayer of Palmitoyloleoylphosphatidylcholine (POPC) and Phosphatidylinositol 4,5-bisphosphate (PIP_2) in the ratio of 10:1 respectively. The system was further hydrated with 20 Å (TIP3P water model) on upper and lower leaflets. An ionic concentration of 150 mM KCl solution was maintained in the system, both in the upper and lower regions and neutralized with counter ions. Protein, lipids and ionic parameters were assigned using the CHARMM36 force field. NAMD package (Phillips et al., 2005), version 2.10 and 4,096 processors on the Blue Gene\Q supercomputer were employed for running the Molecular Dynamics (MD) simulations.

Classical MD Simulation: Parameters and Protocol

Prior to performing SMD simulations, each system was prepared by running a short classical MD simulation. The protein systems along with lipid and water were subjected to energy minimizations in two stages. Prior to running the full dynamics, each system was subjected to energy minimizations in two stages. In the first minimization round of 50,000 steps, the protein and the lipid heads were fixed whereas the lipid tails, water and ions were free to relax. This step was essential to remove any existing steric clashes as a result of the improper packing of the membrane around the protein. In the second stage of minimization, a constraint of 100 kcal/mol was placed on the entire system and energy minimization was performed. This constraint was gradually removed during four more rounds of minimizations. Each minimization stage was of 50,000 steps and the constraints on the protein backbone were reduced to 50, 10, 5 and finally 1 kcal/mol constraining force. The systems were then energy minimized by for 5,000 steps. This was followed by heating the system to 310 K for 10 ns while retaining the 1 kcal/mol backbone restraint. Equilibration phase was next, consisting of NVT and NPT steps for 250 ps, each.

Simulations were performed with an integration time step of 2 fs, under periodic boundary conditions. The Langevin dynamics were adopted for temperature (310 K) and pressure control (1 bar). Bonded interactions were computed every one timestep, short-

range non-bonded interactions every two timesteps, and long-range electrostatic interactions every four time steps. A cutoff of 12 Å was used for van der Waals and short-range electrostatic interactions; with a switching function starting at 10 Å for van der Waals interactions to ensure a smooth cutoff. The simulations were performed under periodic boundary conditions; long-range electrostatic interactions were calculated by using the particle-mesh Ewald (PME) method. The unit cells were large enough that adjacent copies of the protein were never close enough for short-range interactions to apply. The trajectory frames were written to file every 200 ps.

Steered Molecular Dynamics Simulation[7]

The starting systems for SMD simulations were obtained from our previous simulation systems in the presence and absence of KCNE1 protein. The snapshots to be used for SMD were taken from the equilibrated phase of the MD simulation. SMD simulations were performed using NAMD 2.9 (Phillips et al., 2005) package, with a non-bonded cutoff of 13.5 Å. Temperature was controlled through velocity reassignment once per picosecond at 310 K. An external force of 4 kcal/mol/Å were applied on the potassium ion through the *Tcl* scripting language interface of NAMD. The ion was pulled from the intracellular pore opening up to the extracellular region outside the channel into the water phase with a constant velocity (0.25Å/ps). The force was applied along a vector normal to the channel axis pointing from the axis to the initial position of the atom. During the simulation, Cα atoms of the Asp186 residue located on the S5/P-loop linker in the four subunits were constrained along the Z-direction with a force of 1 kcal/mole. This was done to prevent any structural drifts in the protein and its location in the membrane, while the ion was being pulled. With this setting the center of mass movement was limited to 0.5 Å. The direction and magnitude of all applied forces were held constant during the simulations. The protein experienced no appreciable drift in the plane of the membrane, so the applied forces may be considered to be radial at all points in the simulations.

[7] Theory of SMD simulations is described in Appendix A.

Small Molecule Docking [8]

Molecular docking calculations were carried out using the most recent version of smina (Koes, Baumgartner, & Camacho, 2013), a version of AutoDock Vina which offers a better control over the docking and scoring parameters (Trott & Olson, 2010). The protein structures were prepared using the protein preparation wizard in the Schrodinger software package (Madhavi Sastry, Adzhigirey, Day, Annabhimoju, & Sherman, 2013). The protonation states were assigned at the pH of 7. Protein structures were then saved as PDB files and converted to PDBQT format using the AutoDock Tools (Morris et al., 2009) to be used as inputs for smina.

Ligand structures were prepared using the ligprep ("Schrödinger Release 2017-2: LigPrep, Schrödinger, LLC, New York, NY, 2017.," 2017) module of Schrodinger and saved as mol2 files. The ligand protonation states and tautomeric states were assigned at neutral pH. The geometry of the ligands were optimized through the OPLS2005 force field (Banks et al., 2005). The docking search space was confined to a 20*20*20 Å box around the ligand-binding site with an exhaustiveness search parameter of 20 (default is 8). The binding site residues of Chromanol 293B were obtained from a mutational study by Lerche et al. (Lerche et al., 2007) who had confirmed the binding mode and residues responsible for interaction with Chromanol 293B. These residues included Thr197, Phe225 and Ile222 from the four subunits.

Analysis & Visualizations

Analysis of simulation trajectories was performed using VMD (Humphrey, Dalke, & Schulten, 1996) and Chimera suite (Pettersen et al., 2004). Pore radius profiles were calculated using HOLE program (O S Smart, Goodfellow, & Wallace, 1993; Oliver S. Smart, Neduvelil, Wang, Wallace, & Sansom, 1996). Plots were generated using Gnuplot and GraphPad Prism version 6.0 [GraphPad Software, La Jolla California USA, www.graphpad.com]. Protein-Ligand interaction maps were generated with the Maestro suite ("Schrödinger Release 2017-2: MS Jaguar, Schrödinger, LLC," 2017).

[8] For a basic description of docking, please refer to Appendix C.

3.3. Results and Discussion

3.3.1. Steered Molecular Dynamics (SMD) Simulations

A brief theory of Steered Molecular Dynamics (SMD) simulations and its variants are provided in Appendix A. In this study, a single K^+ ion was pulled from the intracellular region at the bottom of the pore, to the extracellular region. The length of this pathway was ~34 Å long (i.e. starting from the intracellular channel entrance to the extracellular loops above the selectivity filter). The protocol of the SMD simulations is explained in details in the Methodology section.

We performed two sets of SMD simulations; the first system consisted of the KCNQ1 protein alone and the second set was focused on the KCNQ1/KCNE1 complex. The SMD simulations started from equilibrated structures of the protein obtained from our previous simulations (discussed in Chapter 2). For each set, we performed ten SMD simulations, repeated for 4 ns each. These repeats were performed to ensure that the results we obtained were consistent and reproducible. Each run started from the same equilibrated configuration but was sampling a different trajectory. The potassium ion was pulled with a force of 4 kcal/mol/Å and with a velocity of 0.025 Å/ps. The choice of the force and velocity parameters was based on a previous study from our group, in which we benchmarked the different forces and velocities to find the optimum combination for ion channels (Ahmed M, Jalily Hasani H, Ganesan A, Houghton M, 2017). Our findings from these simulations for the two sets are discussed in the following sections.

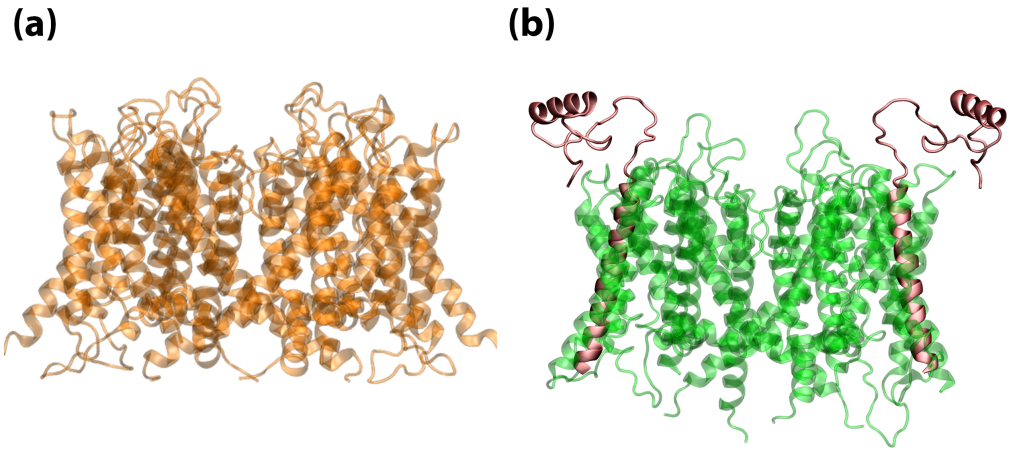


Figure 3.1. The starting systems for SMD simulation. (a) KCNQ1 protein alone, (b) KCNQ1/KCNE1 protein complex

i) Ion Permeation in KCNQ1 Protein Alone

For the KCNQ1 protein system, the force profiles of the ten repeats are shown in Figure 3.2.a The force profiles shown in the figure illustrates the different energy barriers a K^+ ion has to pass across while moving along the pulling pathway. For a clearer view of such barriers, Figure 3.2.b shows one single force profile as a representative case. As seen from the plot, there are four well-defined peaks close to one another. These peaks represent the energy barriers on the way of the ion. These peaks are also shown as B1, B2, B3 and B4 in the figure. The trajectories from the SMD simulations were carefully analyzed to spot the events that take place while the potassium ion passes across these large energy barriers. Figure 3.2.c shows the trajectory frame corresponding to the force profile. The potassium ion has no obstruction while entering the pore from the intracellular region. The first barrier (B1) on its way is the first set of electronegative carbonyl atoms formed by the selectivity filter residues. The THR197 residues from the four subunits form the B1 obstruction, which imposes a highly electronegative effect on the K^+ ion and thereby acts as an energy barrier. Next, the K^+ ion comes across the latter two barriers (B2 and B3), which are about 200 pN higher in energy than the first peak. The reason for this increased energy is because although the potassium ion has left the first energy barrier, B1, it is still under the influence of the Thr197 residues when it reaches B2. Therefore, the amount of force experienced by the K^+ ion becomes a synergistic energy barrier. The B2 obstruction is formed by the Ile198 residues from the four subunits. The same synergistic effect applies when the potassium ion is captured by the Gly198 at the third energy barrier, B3. Finally, after the potassium ion reaches the fourth energy barrier, B4 (formed by Tyr200 and Gly201), the energy profile collapses to 200 pN, which is of the same intensity as the first barrier, B1. This is again because the potassium ion has no more barriers to cross on its way out of the channel. After crossing the B4 energy barrier, the ion enters the extracellular region, which imposes no obstruction to its movement, and thus the plot returns to the equilibrium condition, similar to the initial entry phase.

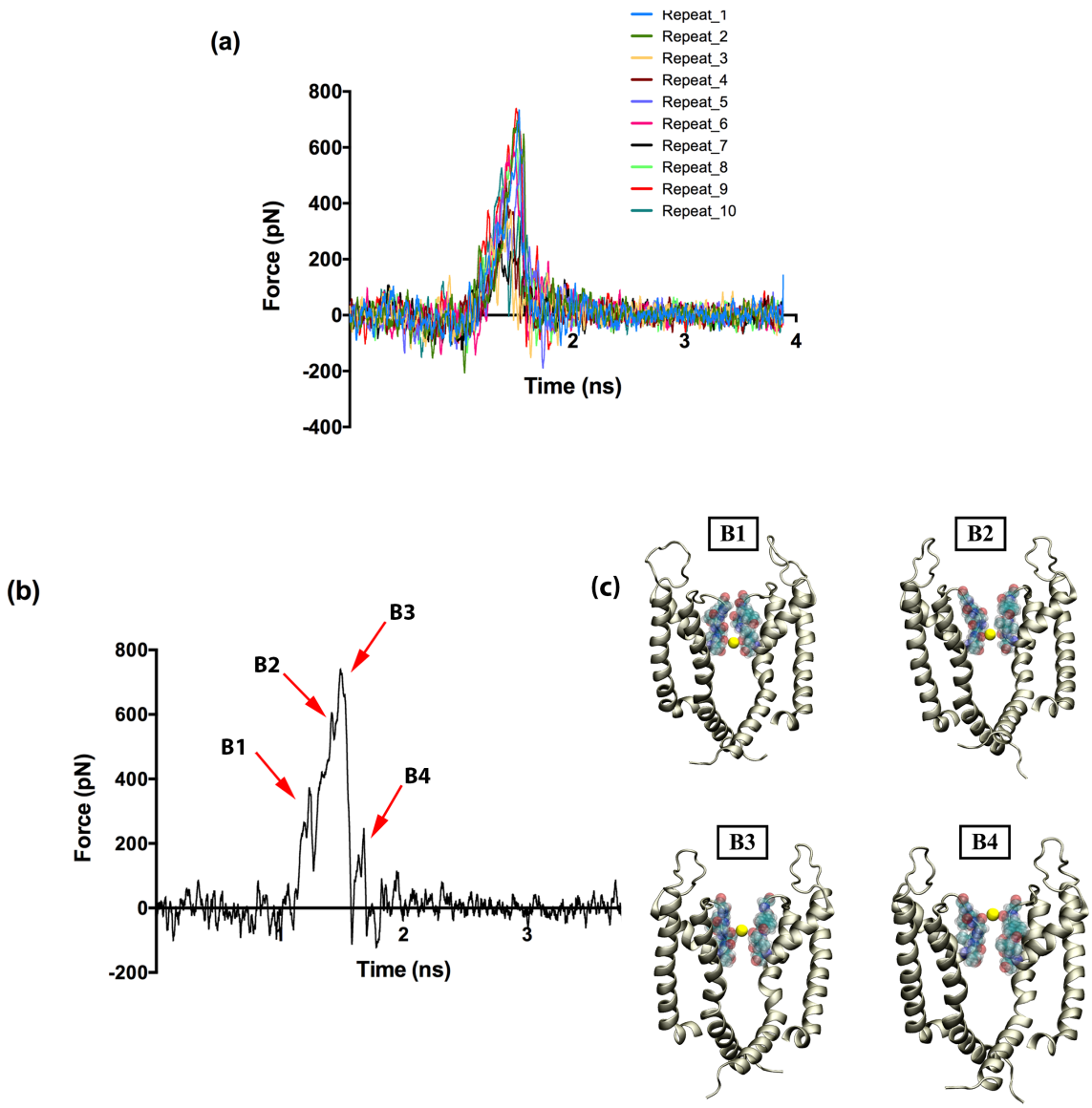


Figure 3.2. The force profile of K^+ ion pulled through the KCNQ1 protein alone. (a) the ten repeats of SMD simulation, (b) one single force profile to show the different barriers marked as B1, B2, B3 and B4. (c) The barriers formed by the selectivity filter residues (Thr-Ile-Gly-Tyr-Gly). The potassium ion is shown in yellow, the S5, S6 and P-loop of two subunits are shown in cartoon. The KCNQ1 VSD and the other two subunits are not shown for clarity.

Furthermore, we analyzed the presence and movement of water molecules during the SMD simulation. As shown in Figure 3.3, the entrance pore is widely open for water molecules to enter. However, as they approach the selectivity filter, the number of water

molecules that can be accommodated is significantly reduced. This explains why the selectivity filter acts as a sieve and does not allow the passage of other ions of different size, density and charge to pass the pore.

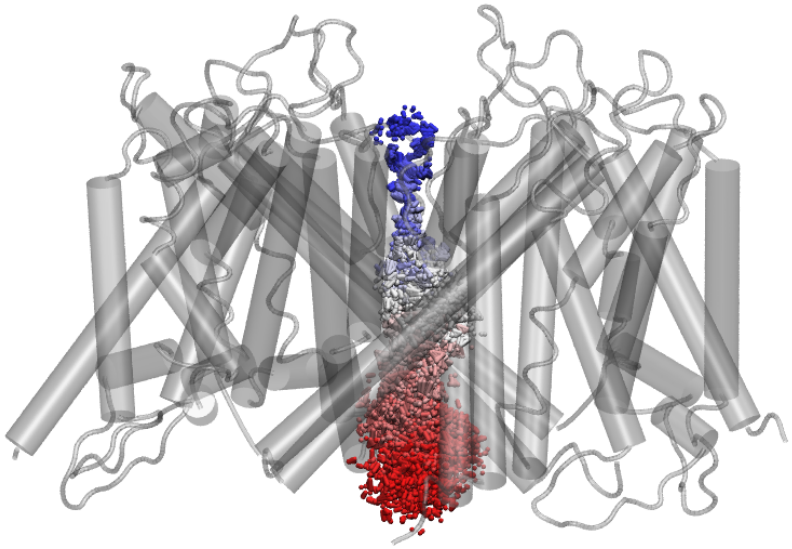


Figure 3.3. Movement of water through the pore of the KCNQ1 channel. The water molecules are colored from red to blue based on the simulation timestep. The protein is shown in cartoon presentation.

These findings are consistent with previous studies from the literature, related to the selectivity filter properties and the binding sites of potassium through the channel (Ceccarini, Masetti, Cavalli, & Recanatini, 2012; Shian Liu et al., 2015; Lockless, Zhou, & MacKinnon, 2007; Morais-Cabral et al., 2001). Our model was perfectly capable of predicting this phenomenon. The four reported binding sites for potassium were identified clearly from our SMD simulations. We were also able to quantify their energy barriers along the pathway of the ion. The selectivity filter residues motif TIGYG (Thr-Ile-Gly-Tyr-Gly) formed four binding sites for the potassium ion. For an ion to successfully pass through the filter, it has to overcome each of the carbonyl atom cages created by these residues, one after the other. However, based on our findings these binding sites do not impose equal obstacles on the way of the ion.

The force needed for a potassium ion to migrate from the first binding site formed by Thr residues (B1) is lower than the one formed by the Ile and Gly residues (B2 and B3) at least by half. That might explain the reason as to why a second potassium ion is

usually needed to push the first one ahead to the next binding site by repulsive electrostatic forces. However, entering the first binding site is rather an easy task for a single potassium ion to be achieved on its own. The last binding site that is formed by the Tyr and Gly residues (B4) is also of lower intensity, enabling the exit of the potassium ion from the selectivity filter and its complete release into the extracellular region.

ii) Ion Permeation in KCNQ1/KCNE1 Complex

To understand the effect of KCNE1 on the ion permeation process, we performed the same SMD simulations on the complex. We used the exact same parameters and simulation conditions for the equilibrated KCNQ1/KCNE1 complex systems. Figure 3.4.a shows the superimposed force profiles obtained from the ten SMD repeats for KCNQ1/KCNE1 system. As an overall and interesting observation from these plots, there are two additional peaks that appear just before the normal four peaks described earlier for the KCNQ1-alone system. To illustrate this better, Figure 3.4.b shows a representative force profile for the potassium ion pulled through the KCNQ1/KCNE1 channel complex. These new two peaks are marked as BA1 and BA2 on the plot and appear early on before passing the selectivity filter. These two peaks were completely absent in the force profile of K^+ ion for the KCNQ1 protein alone. BA1 consists of several smaller peaks that are assembled together to form a continuous energy barrier.

To understand these events better, we analyzed the trajectories resulting from the SMD simulations. The potassium ion spends a significant amount of time trying to overcome the BA1 barrier (~ 0.4 ns). The peaks in BA1 marked by red arrow in the plot were formed by a group of residues, which acted as a cage, trapping the ion. These residues include Ser234, Ala723, Gly724, Ser728, Ala970, Gly971 and Ser975 located on the lower portion of the S6 helices (see Figure 3.4.c). In the previous system of KCNQ1 alone, these residues were facing away from the pore and did not form any barrier on the way of the ion. Their inclination towards the center of the pore was a result of the shift in the location of the S6 helices with respect to the center of the pore. This shift is a direct effect of the presence of KCNE1 and its interaction with the pore domain of the channel.

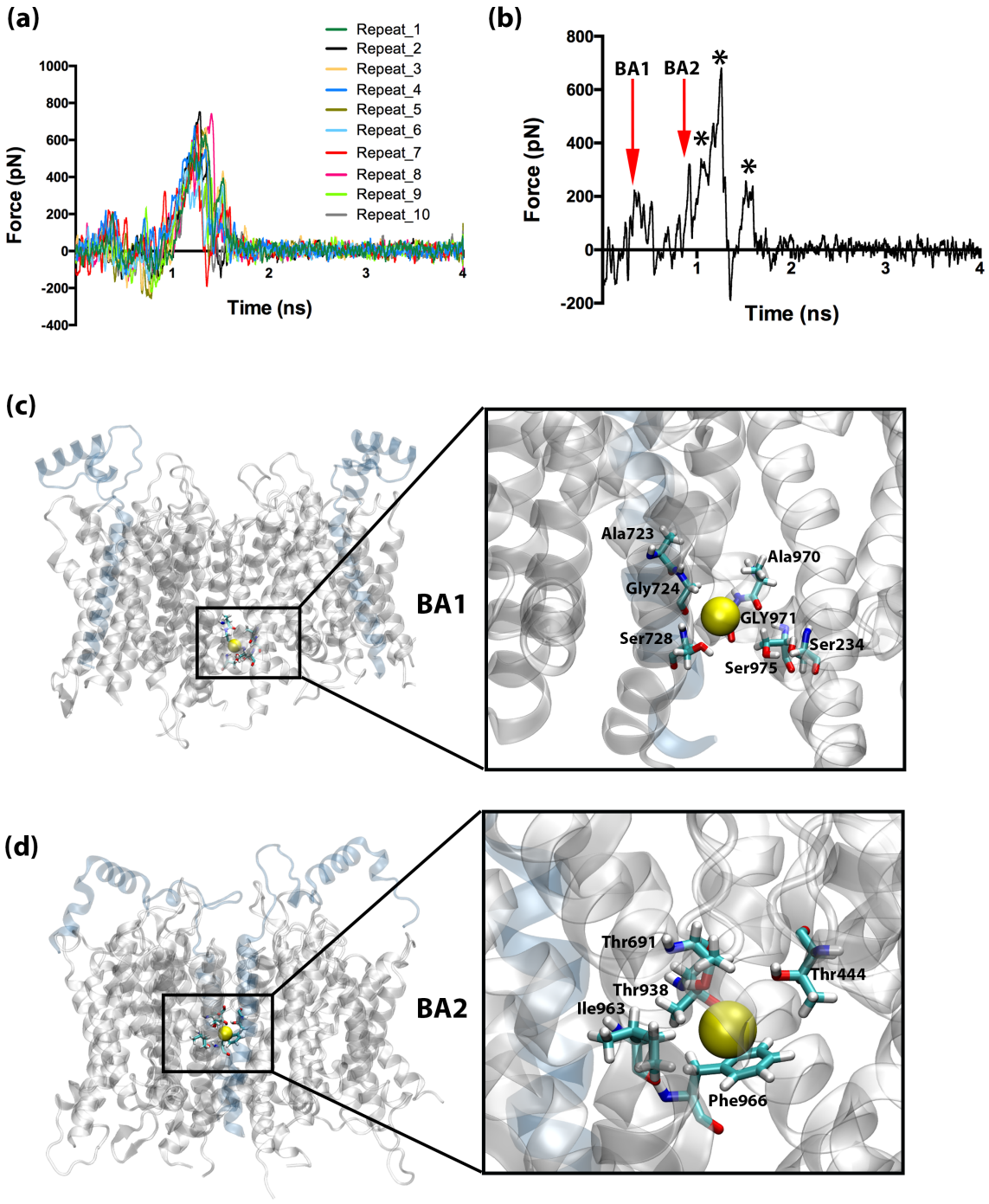


Figure 3.4. The force profile of K^+ ion pulled through the KCNQ1/KCNE1 protein complex. (a) all ten repeats of SMD simulation, (b) one single force profile to show the different barriers marked as BA1, BA2 and asterisks, (c) Close up view of the BA1 barrier, (d) Close up view of the BA2 barrier. The potassium ion is shown in yellow sphere, the protein structure is shown in cartoon presentation and the residues are depicted by bonds.

Figure 3.5 shows the difference between the conformation of KCNQ1 alone (in yellow) and KCNQ1/KCNE1 complex (in purple). There is a clear shrink in the opening of the pore by the shift of the S6 helices towards the center of the pore. This finding is in agreement with a previous study by Hoshi et al. (Hoshi & Armstrong, 2013) who put forth that the elimination of the slow inactivation in potassium channels, might be linked to structural rearrangement in the pore and/or selectivity filter region constriction.

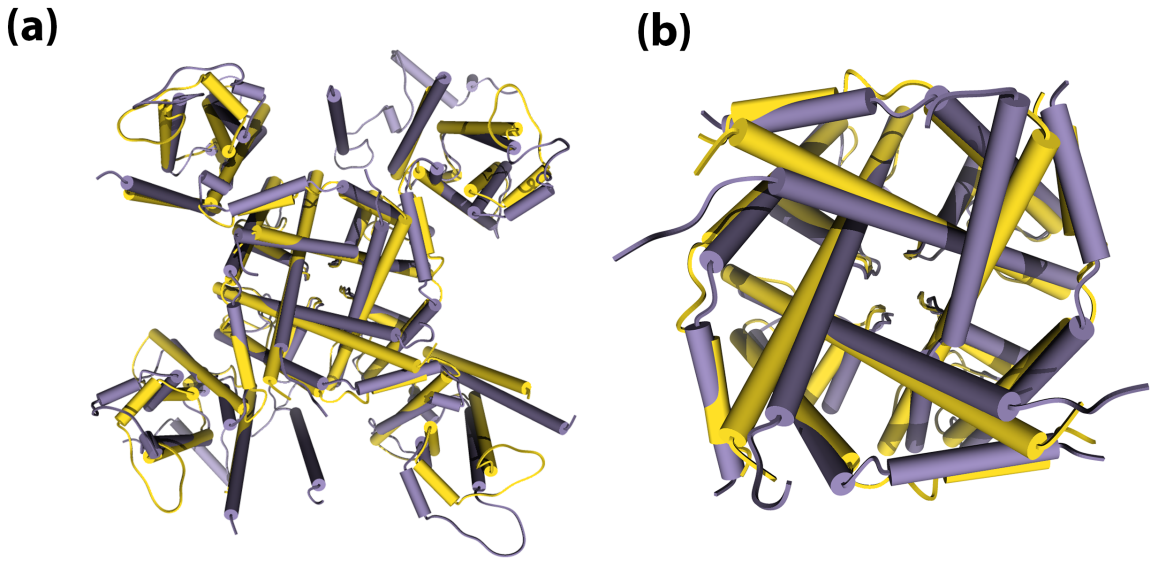


Figure 3.5. Superimposition of the KCNQ1 alone (yellow) and KCNQ1/KCNE1 complex (purple). (a) rear view of the channel, (b) zoomed view to show the shift in the S6 helices.

The BA2 barrier is located right below the selectivity filter and is formed mainly by Thr444, Thr691, Thr938, Ile963 and Phe966 residues. These residues are also located on the S6 helix and the S5-P-loop linker. Of specific interest is the Phe966 residue, which forms a cation-π interaction and stacks with the potassium ion while the other residues (Thr444, Thr691, Thr938, Ile963) prevent it from escaping. After crossing these two energy barriers, the potassium ion will face a similar force profile to the one it passed through in KCNQ1 protein alone system. The other three important peaks are marked by asterisks on the plot in Figure 3.2.b. These peaks correspond to the selectivity filter binding sites discussed earlier for the KCNQ1 alone (B1, B2, B3 and B4)

Overall, the findings from the SMD simulations and force profiles of potassium ion indicate that the KCNE1 induces its main effect on the pore opening and on the

region below the selectivity. It induces an inward shift to the S6 helices, while not substantially provoking any conformational changes in the selectivity filter.

3.3.2. Pore Dimension Analysis of Lone KCNQ1 and KCNQ1/KCNE1 Complex

We also analysed the pore dimensions of the KCNQ1 ion channel with and without KCNE1, using the HOLE program (O S Smart et al., 1993; Oliver S. Smart et al., 1996). The HOLE program adopts a Monte Carlo simulated annealing approach to find the best path for a sphere (of variable radius) to squeeze through the channel. HOLE has been successfully used in complementing the analysis of ion channels in several studies (Fowler & Sansom, 2013; Kutteh, Vandenberg, & Kuyucak, 2007; Linder, de Groot, & Stary-Weinzinger, 2013; Sula et al., 2017). The main objective of this analysis was to support our findings related to the identified force profiles of K^+ ion in the two systems; i.e. KCNQ1 alone, and KCNQ1 in complex with KCNE1.

As shown in Figure 3.6.a, the pore dimensions in the KCNQ1 protein alone have a wide opening in the bottom, continuing upward to the selectivity filter. This indicated a completely open pore without any constriction on the way for a potassium ion to pass through. The same topology has been retained after the long MD simulations were run (see Chapter 2). This also confirms that the pore constriction did not occur as a result of an artefact during the MD simulation. The only difference between the KCNQ1 pore dimension before and after MD was the loss of the perfect symmetry, which existed in the original homology model. This is an absolutely normal phenomenon since no single biological entity, which possesses dynamicity, would remain 100% symmetric over time.

The pore dimensions shown in Figure 3.6.b, which belong to KCNQ1 in complex with KCNE1, has significantly contracted at the bottom opening. The pore has become even narrower in the proximity of the selectivity filter (shown in green color). This constriction is absent in the lone KCNQ1 protein an dappears close to the selectivity filter. This structural narrowing in the overall topology of the pore domain is a direct result of the presence of KCNE1 and its interaction with KCNQ1 protein.

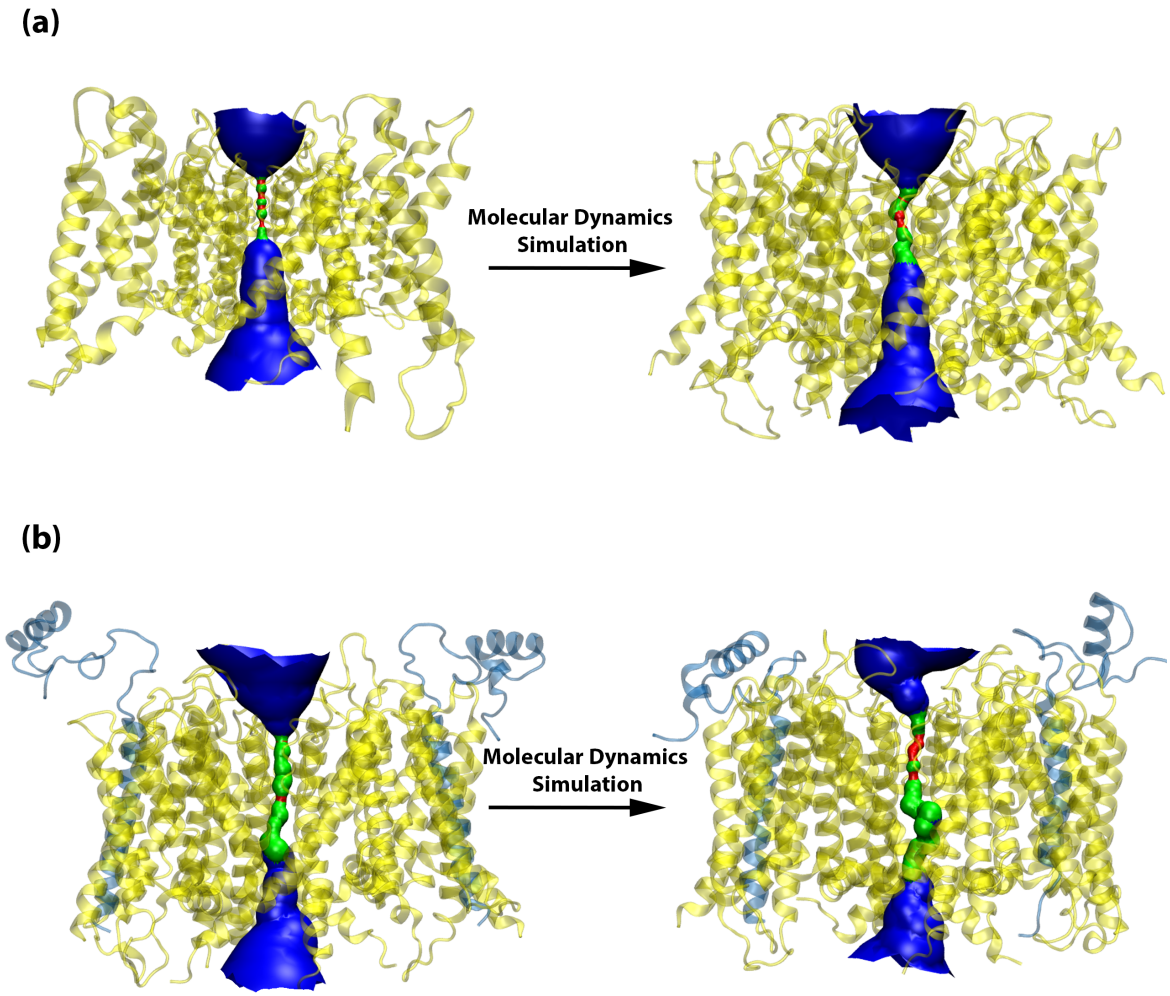


Figure 3.6. The dimensions of the pore shown in surface representation. Color code is explained below*. for the two model states: (a) KCNQ1 without KCNE1 before and after MD, (b) KCNQ1 in complex with KCNE1 before and after MD. *Pore colour code: Red is where the pore radius is too tight for a water molecule. Green where there is room for a single water molecule. Blue is where the radius is double the minimum for a single water molecule.

Along the same line, we plotted the radius of the pore in the two systems (see Figure 3.7). The red color line shows the radius of the pore domain in KCNQ1 system alone. The pore opening has a radius of around 6 Å, which reduces to less than 2 Å, while approaching the selectivity filter. However, the radius of the pore in the KCNQ1/KCNE1 system (green color) shows a completely different pattern. The constrictions of the pore begin early on at the opening of the pore domain, below the selectivity filter. The radius at this

region is around 4 Å, which increases to 7 Å marking the pore cavity. The radius of the pore is again reduced to ~0.4 Å and continues to be narrow along the selectivity filter.

With these findings, we were able to infer more insights about the effect of KCNE1 on the KCNQ1 channel. From our analysis, the KCNQ1 pore domain has gained completely different characteristics once associated with KCNE1 protein. At some points of the pore, the radius was reduced to 1 Å. This is despite of the fact that the radius of a potassium ion is around 2.8 Å, indicating that the potassium ion conduction would be reduced due to the additional energy imposed by the KCNE1 protein.

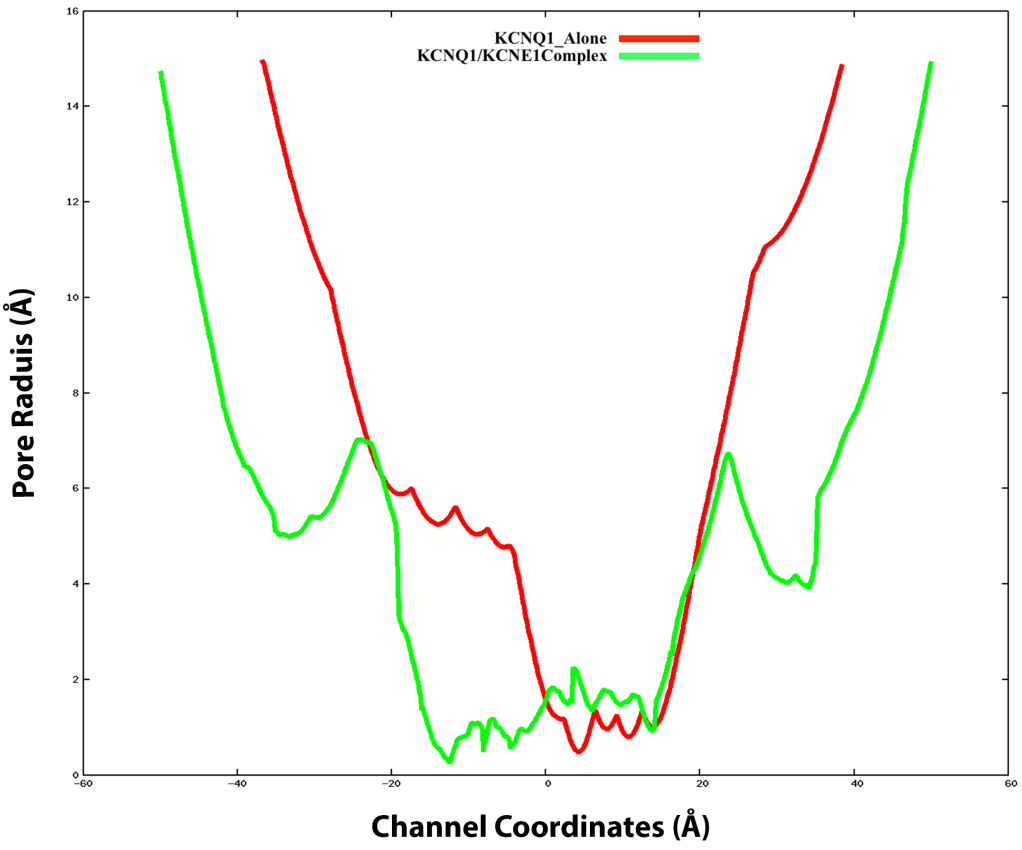


Figure 3.7. Pore radius plot of the KCNQ1 alone (red) and KCNQ1/KCNE1 complex (green) systems. There is a continuous constriction from the pore opening up to the selectivity filter in the KCNQ1/KCNE1 system as compared to the KCNQ1 pore which has a wide opening.

3.3.3. Small Molecule Docking

To make a short list of a set of small molecule KCNQ1 blockers, we filtered the ChEMBL database (Bento et al., 2014) for known blockers of the I_{KS} current. As mentioned in Chapter 1.2, this is the main cardiac current generated by the KCNQ1/KCNE1 complex. In this regard, we were looking for a panel of compounds ranked by two main criteria. First, to ensure consistency of the data and to minimize experimental data variability, we looked for compounds whose activities were determined by the same experimental assay. Secondly, we needed compounds possessing a broad range of activity to study their differential effects on ion permeation. By imposing these two criteria, we selected Chromanol 293B and its derivatives. Chromanol 293B was discovered in 1996 (Bleich et al., 1997; Busch et al., 1996; Suessbrich et al., 1996) as a selective blocker of the I_{KS} current. In 2001, Gerlach et al. (Gerlach et al., 2001) synthesized several derivatives of this lead compound to enhance its activity as antiarrhythmic drugs. These compounds and their structures are listed in Table 3.1. This set of compounds met our selection criteria. The data came from the same study, and they had a broad range of activity (IC_{50}: 50 nm to 58,000 nm). The blocking activities of the compunds were determined by initially expressing the KCNQ1/KCNE1 channel complex in Xenopus oocytes and subjecting them to two-microelectrode voltage clamp electrophysiological experiemnts. And lastly, Chromanol 293B is found to have an affinity for the open state of the KCNQ1 channel (Lerche et al., 2007) and was therefore the best choice for our model, which is also built in its open state.

Chromanol 293B along with its 8 derivatives were docked against 15 KCNQ1/KCNE1 protein conformations. This represented a total of 135 independent docking simulations. Docking every small molecule to an ensemble of protein conformations ensures the accommodation of the protein flexibility during the docking workflow. This is important to address any conformational dynamics that can lead to a better docking pose for the tested compound within the binding site of the protein. All protein conformations were generated through the clustering of the long MD simulation trajectories (as discussed in Chapter 2). The 15 representative conformations of the

KCNQ1/KCNE1 complex protein are shown in Figure 3.8. All docking simulations were performed using the smina docking tool (Koes et al., 2013), a version of AutoDock Vina which offers a better control over the docking and scoring parameters (Trott & Olson, 2010). See the Methods section for parameters and details of the docking calculations.

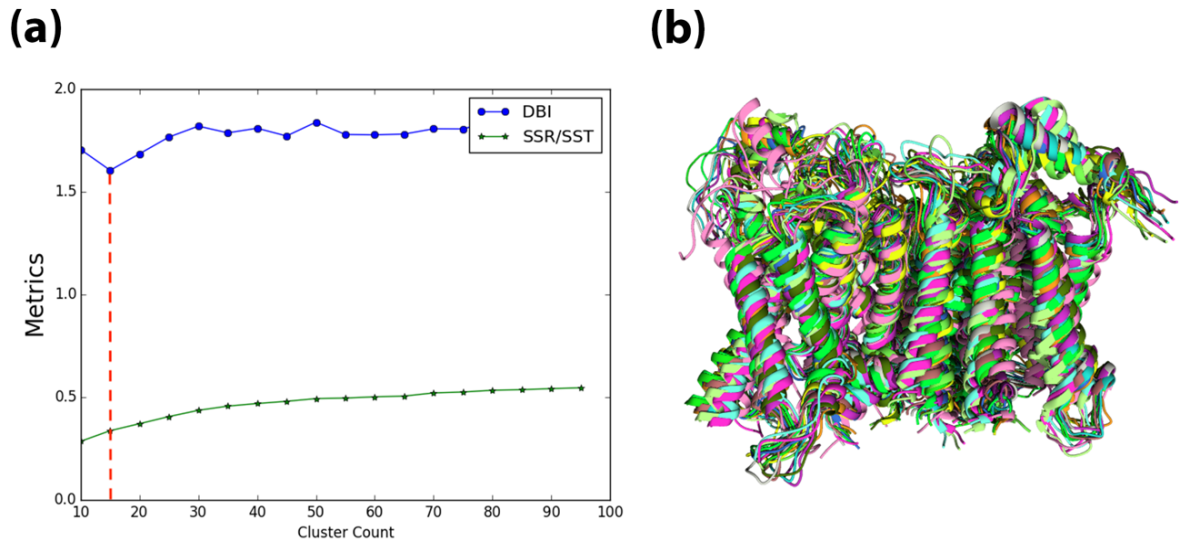
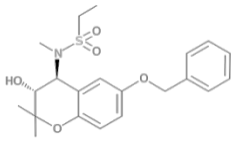
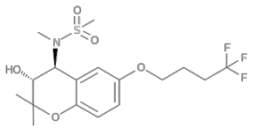
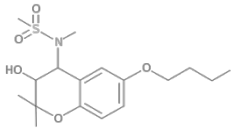
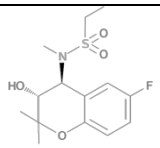
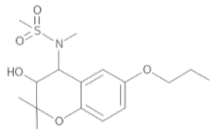
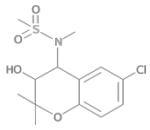
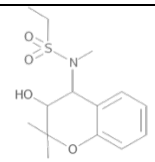
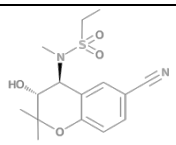
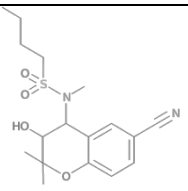


Figure 3.8. Clustering analysis of the KCNQ1/KCNE1 channel complex from the MD simulation. (a) The clustering plot of DBI and SSR/SST parameters, (b) The 15 cluster representative conformations of the KCNQ1/KCNE1 complex protein.

Table 3.1. The panel of compounds chosen for docking, consisting of Chromanol 293B and its 8 derivatives.

Compound (#. CHEMBL_ID)	Structure
1. CHEMBL124454	
2. CHEMBL298475	
3. CHEMBL125307	
4. CHEMBL340025	
5. CHEMBL124810	
6. CHEMBL338171	
7. CHEMBL125259	
8. CHEMBL434045	
9. CHEMBL330993	

3.3.4. Analysis of Drug-Protein Poses

For each of the tested compounds, 20 different poses were obtained and ranked by AutoDock Vina scoring function (Trott & Olson, 2010). The best poses for all the ligands determined by Vina were then re-ranked with two other machine learning scoring functions; the NNScore 2.0 (Durrant & McCammon, 2011) and the RF-Score-VS (Wojcikowski, Ballester, & Siedlecki, 2017). NNScore 2.0 is a neural-network based scoring function, devised to aid the computational identification of small-molecule ligands by providing a single pK_d (binding affinity) value. The RF-Score-VS (Wojcikowski et al., 2017), is another machine learning scoring function which has shown significant improvement in the performance of virtual screening studies (Wojcikowski et al., 2017). Machine-learning scoring functions trained on protein-ligand complexes have shown great promise in small tailored studies as compared to conventional scoring functions such as Vina scoring function, alone (Ain, Aleksandrova, Roessler, & Ballester, 2015; Khamis, Gomaa, & Ahmed, 2015; Wojcikowski et al., 2017).

The final docking score considered for each ligand was the average of the results from three scoring functions, explained above, i.e. AutoDock Vina, NNScore 2.0 and RF-Score-VS. The poses were then closely visualized for their proper filling of the designated binding site (see Figure 3.9). The binding site of the ligand was selected based on the study by Lerche et al. (Lerche et al., 2007) who had investigated the binding mode of Chromanol 293B within the KCNQ1 pore, using a single point mutational approach. The binding site is located right below the selectivity filter and is formed by three residues from each subunit namely “Thr197, Ile222 and Phe225”, “Thr444, Ile469 and Phe472”, “Thr716, Ile719 and Phe691”, “Thr938, Ile963 and Phe966”. In the same study, it was found that the single point mutations at these specific positions intensely reduced current inhibition and had the strongest effects on blocking activity of Chromanol 293B (Lerche et al., 2007).

Table 3.2 lists the in vitro biological activity (IC_{50} and pIC_{50} values) of the ligands compared with their docking scores as obtained from our docking calculations. The first

three compounds have an IC_{50} ranging from 50 to 250 μM, considered as the most potent compounds in this class. Compounds #3-7 with IC_{50} of 700 up to 1,100 represent blockers of average inhibition and finally the last two compounds (#8 and #9) are considered as weak blockers with 5,000 and 58,000 μM activity. Next, a Pearson product-moment correlation coefficient was computed to assess the relationship between the scores calculated from the docking poses of the ligands against the KCNQ1/KCNE1 protein complexes, and the pIC_{50}s of the compounds. Interestingly, there was a strong, positive correlation ($r_{pearson} = 0.7$) between the two variables, indicating that our model has been successful in discriminating blockers of variable activity. A scatter plot in Figure 3.10 summarizes the results.

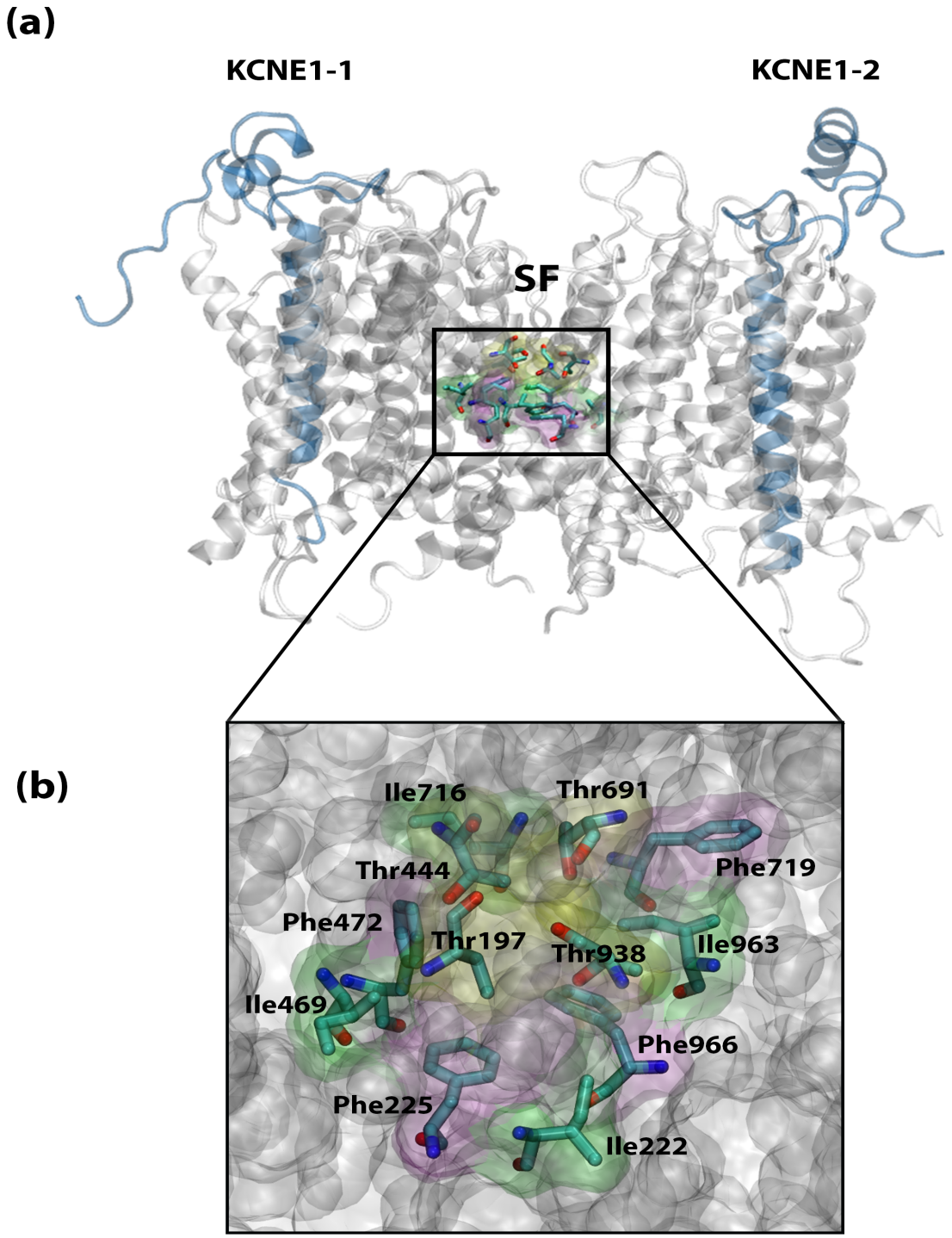


Figure 3.9. The binding site residues of the Chromanol 293B and its derivatives shown in the structure of KCNQ1 (grey color cartoon) in complex with KCNE1 proteins (blue color cartoon). (a) The binding cavity shown with bonds in licorice presentation is located right below the selectivity filter (SF) of the channel. (b) The residues with their names are shown on the lower close-up of the binding site.

Table 3.2. Ranking of the ligands by their pIC_{50}s compared with their IC_{50} values and docking scores. The docking score in the table is the average from 3 scoring functions: AutoDock Vina, NNScore 2.0 and RF-Score-VS. * Ligands used for ion permeation studies.

Compound #	CHEMBL _ID	Docking Score	IC_{50} (nM)	pIC_{50}
1 *	CHEMBL124454	7.216	50	7.3
2 *	CHEMBL298475	7.144	120	6.9
3	CHEMBL125307	7.083	250	6.6
4 *	CHEMBL340025	6.654	700	6.2
5	CHEMBL124810	7.003	900	6.1
6 *	CHEMBL338171	6.268	1100	6.0
7	CHEMBL125259	6.404	3100	5.5
8 *	CHEMBL434045	6.763	5000	5.3
9 *	CHEMBL330993	6.362	58000	4.2

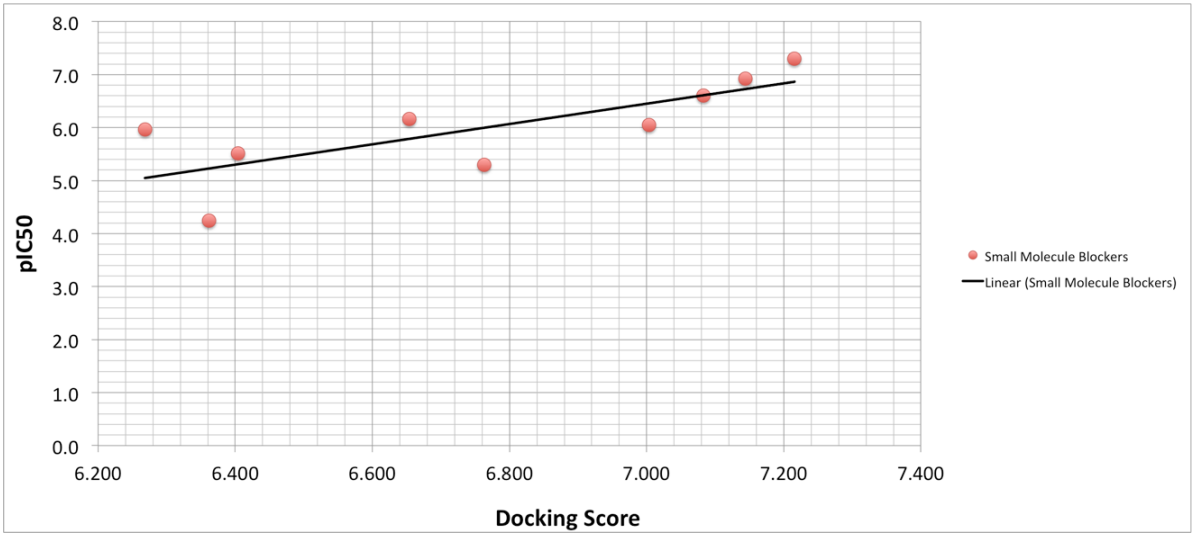


Figure 3.10. A 2D scatter plot of the compounds' docking score vs. pIC_{50}. The linear line shows the positive correlation between the two variables ($r_{pearson}$ = 0.7).

To study the interactions of the ligand and their mode of binding within the pore of the channel, we chose representatives from the high activity, average activity and low activity ligands. Figure 3.11 shows the interaction map of ligand #2 (IC_{50} = 120 μM) as well as its placement within the binding site while interacting with the protein residues. Ligand #2 possesses a trifluoro-butoxy substitution at the 6-position on the aromatic ring. This bulky side chain interacts with the residues in the centre of the pore, underneath the selectivity filter. The interactions of Ligand #2 with the binding site residues include Phe225, Phe472, Ile963, and Phe966, Thr716 and Phe691. The sulfonyl group makes contacts with Pro475, while the trifluoro group interacts with Phe472, Ala715 and Ile963. This mode of binding is consistent with the general mode of interaction of chromanols within the KCNQ1/KCNE1 ion channel (Lerche et al., 2007). Furthermore, the protrusion of the lengthy side chain of the molecule towards the centre of the pore can explain its high blocking activity. This extension is also responsible for the interaction of the ligand with potassium ions as will be discussed in the next section of ion permeation.

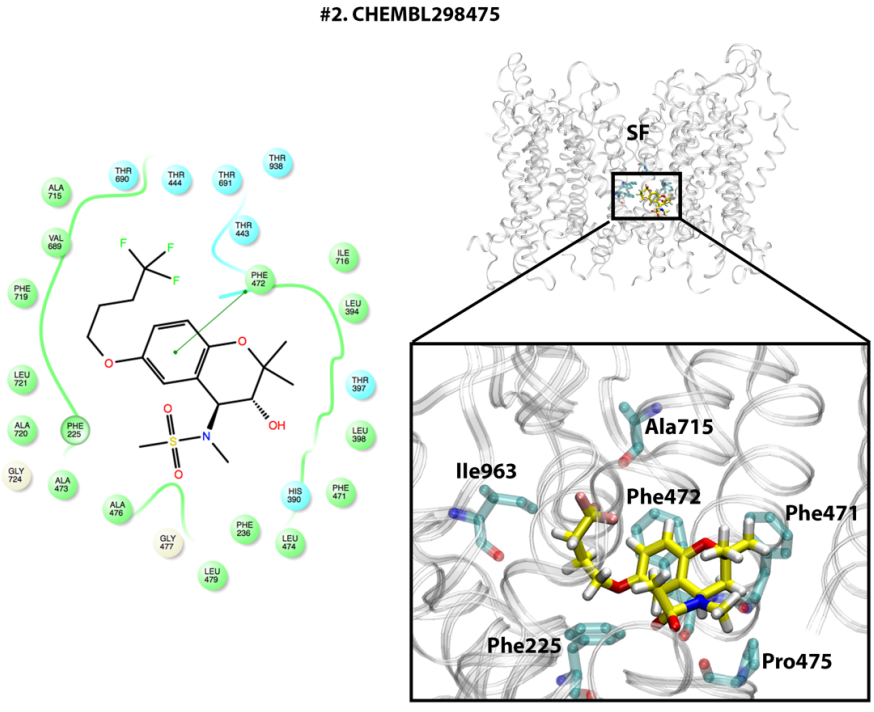


Figure 3.11. The interaction map and binding mode of Ligand #2 (CHEMBL298475). The ligand is docked right below the selectivity filter (SF), in the binding site pocket. The trifluoro side chain is extended towards the centre of the pore.

We also investigated the binding mode of Ligand #8 and #9 as representatives of the weak blockers (See Figure 3.12 and Figure 3.13). This investigation helped us understand the main reason behind the large difference in their IC_{50}s compared to ligand #1-3. Both ligands can slightly fit within the binding site pocket, albeit with an inclination away from the pore and the selectivity filter. Ligand #8, for example, makes minimal contacts with the Thr691, Thr444, Thr938 residues, which are all located right below the selectivity filter. However, the ethyl group that is attached to the sulfonyl residue of the ligand has a tendency to interact with residues that are not central to the axis of the pore, e.g Gly477. As seen in Figure 3.12 and 3.13, these two ligands have cyanide substitutions at the 6-position of the aromatic ring, which is substantially shorter and less bulkier than those of the strong blockers (ligand #1, #2, #3).

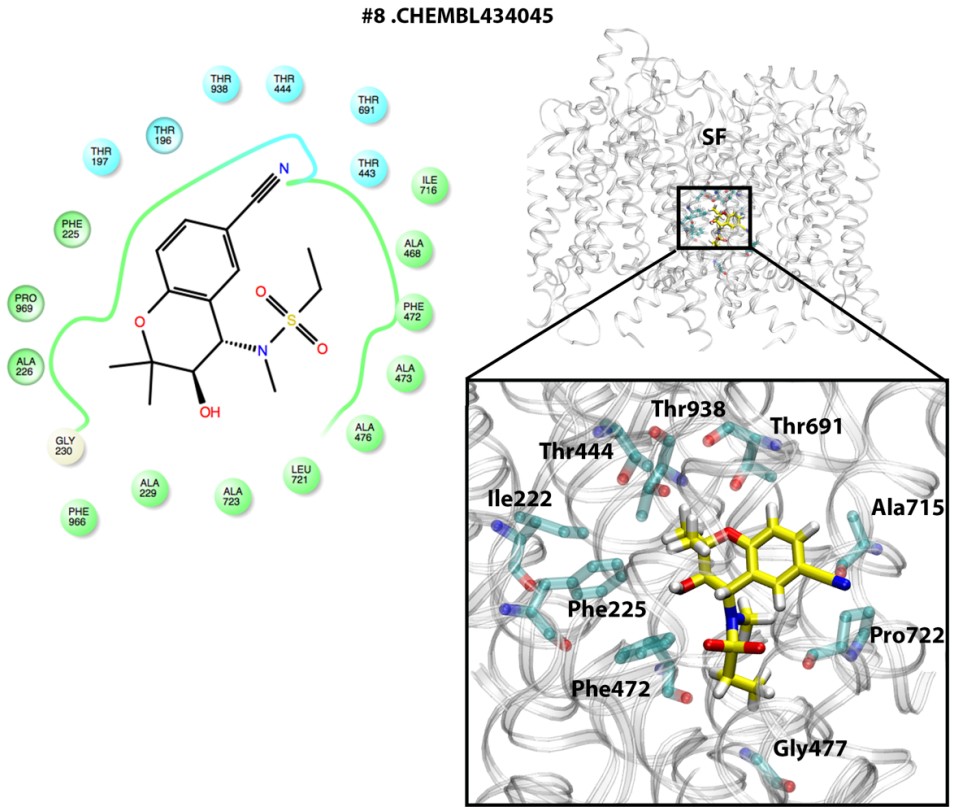


Figure 3.12. The interaction map and binding mode of Ligand #8 (CHEMBL434045). The ligand interacts with the binding site residues with an inclination, away from the pore. The cyanide group substitution is also extended in opposite direction of the pore.

Also, it is evident that increasing the size of the sulfonyl residue has a direct effect on the potency of the compounds, as is the case with #9, possessing a butyl substitution extending away from the sulfonyl group. The latter effect can be explained by the fact that the butyl entity interacts with residues on the periphery of the pore and thereby pulls the ligand away from its binding site. This hinders the physical presence of the drug molecule in the pore. Furthermore, contrary to ligand #2, the 6-position substitution in #8 and #9, i.e. the cyanide group is facing away from the pore and extends towards the opposite direction. This reduces the blockage of the ion passage and can also be a second reason behind their reduced potency.

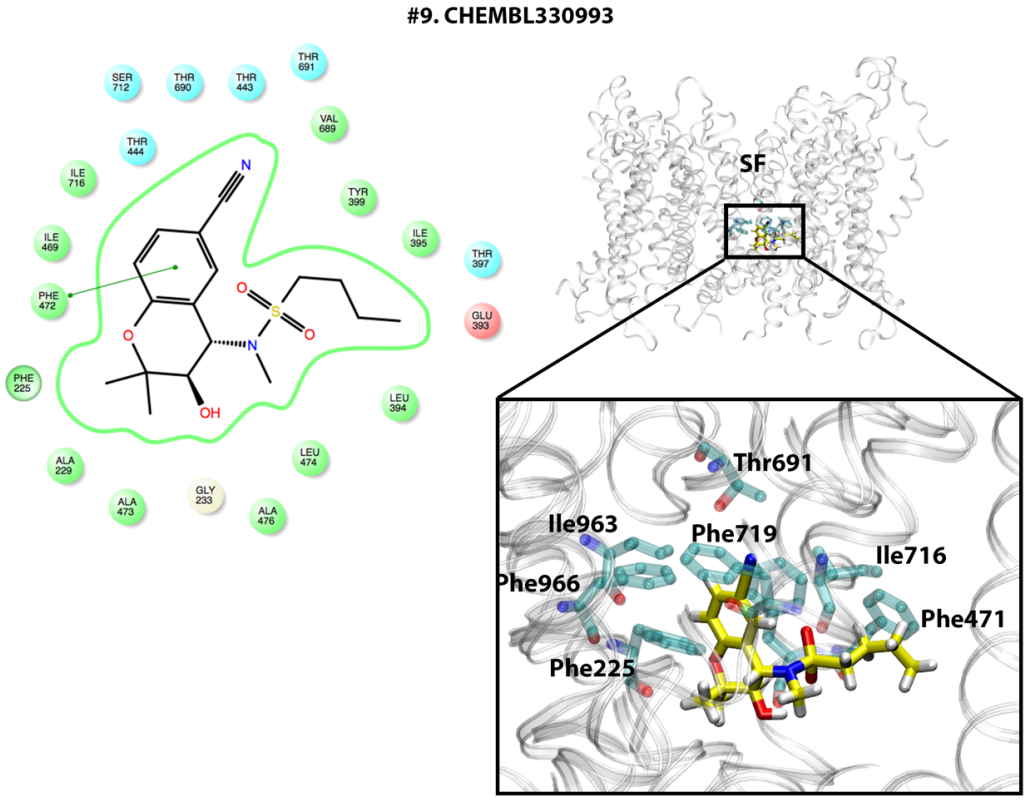


Figure 3.13. The interaction map and binding mode of Ligand #9 (CHEMBL330993). The butyl group attached to the sulfonyl residue pulls the ligand away from the pore. Similarly, the cyanide group substitution is facing away from the pore, explaining its lower potency.

In addition, the interaction of one of the median potency blockers, i.e. ligand #4, which has an IC_{50} of 700 µM, was investigated (See Figure 3.14). This compound also possesses a short substitution (a single fluorine group) at the 6-position of the benzene

ring, similar to #8 and #9. However, compared to #8 and #9, it is more inclined towards the pore. This is clear from the contact it makes with Thr691, which is placed at the mouth of the selectivity filter. This interaction is completely absent in #8 and #9 as their cyanide group substitution is facing away from this residue. The ligand also makes contacts with Thr938, Ile46 Phe225, Phe472 and Ile716 which are all amongst the binding site residues indicating that it is central to the binding site cavity, right below the selectivity filter. However, the fluorine substitution does not occupy as much space as the side chain of Ligand #1 and #2 and therefore, its lower potency compared to the strong blockers may be justified in this way.

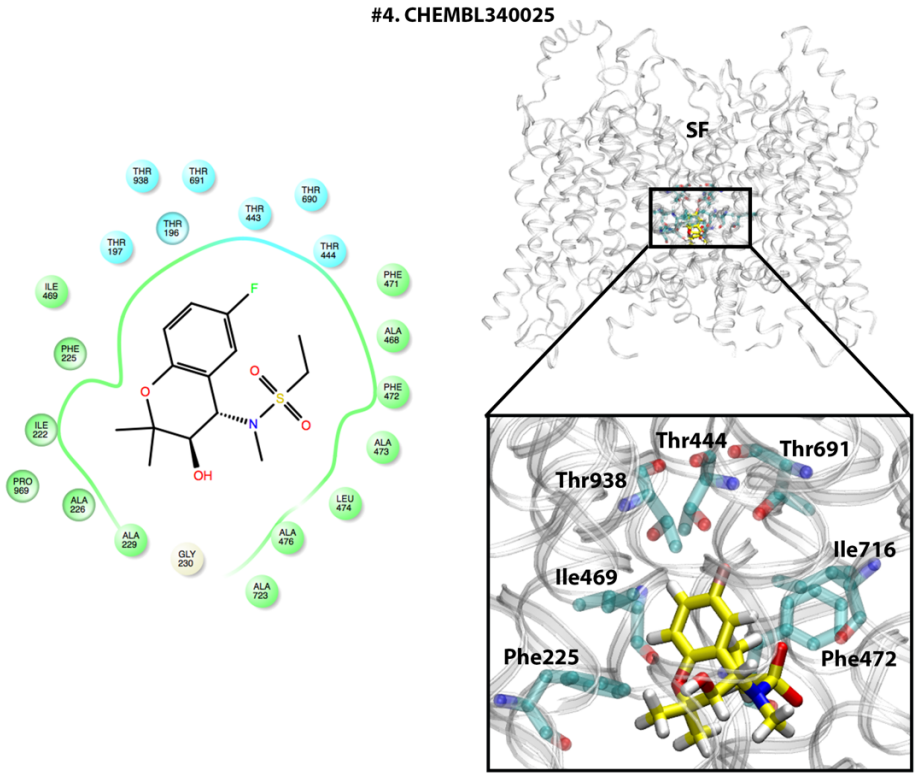


Figure 3.14. The interaction map and binding mode of Ligand #4 (CHEMBL340025). The fluorine substitution is facing away from the pore.

Overall, the results from the docking simulations enabled us to confirm the structure activity relationship of the Chromanol blockers and the reason behind their differential activity. The substitutions at the 6-position on the aromatic ring is the first determinant of potency. This substitution, depending on its size and direction of

extension towards the central axis of the pore, can have differential effect on the conduction of the potassium ions. Secondly, the sulfonyl residue substitution also affects the potency of the compounds. This effect may be produced because the substitutions at this position can have an affinity to interact with the residues on the periphery of the pore. This interaction draws the drug molecule further away from the pore and thereby reducing their ability to produce physical blockage. This effect was clear in ligand #9, which has a butyl group extending from the sulfonyl entity. The encouraging success of the model in predicting the activity and blockage capacity indicated that the model is capable of predicting the correct binding mode and the interaction of the ligands with the channel. Thereby, this adds one more validation measure for our model with regards to its ability for predicting the off-target interactions of other drugs. Also, given the acceptable results we obtained from the docking studies, we decided to take our research question to the next stage, i.e. testing the effect of drugs on the potassium ion permeation. The results from this study are presented below.

3.3.5. Effect of Blockers on Ion Permeation: SMD Simulations

Our objective was to understand how the binding of the drug within the channel would affect the passage of ions. To carry out this study, we decided to focus on 6 compounds, marked with asterisks in Table 3.2. The compounds were selected such that they represent each category of the compounds; i.e. strong, median and weak blockers. In this context, compounds #1 and #2 represented the group of strong blockers with activity ranging from 50 nM to 120 nM respectively. Compounds #4 and #6 represented the ligands with median inhibition activity, ranging from 700 nM to 1,100 nM respectively. Finally, compounds #8 and #9 were selected as representatives of the weak blockers with IC_{50} ranging from 5,000 nM to 58,000 nM, respectively. The complex poses with the bound blocker was obtained from the docking simulations described above and were minimized, heated and equilibrated with the same parameters and in the same simulation setup used in the other MD simulations as described in the previous chapters. This post-docking step was done to relax the protein-ligand complex. Next, we performed SMD simulations, wherein a single potassium ion was pulled through the intracellular pore

towards the exterior of the channel. The superimposed force profiles of the three SMD repeats, for each of the ligands are shown in Figure 3.15.

We compared the force profiles of the tested compounds to the ones we obtained from studying ion passage through the KCNQ1/KCNE1 complex (see previous sections of this Chapter). Interestingly, the energy barrier formed by the selectivity filter in the presence of the compounds remained intact and similar to the ligand-free channel. This indicated that the ligands induced their blocking activity on the pore region rather than on the selectivity filter. However, we noticed that the energy barrier caused by the presence of the KCNE1 protein was synergistic with the energy barrier imposed by the ligands. This is shown in the energy profile as a larger peak of higher intensity and duration. This synergistic effect appeared in the form of a peak with an intensity of approximately 600 pN.

This was much higher than the peaks we observed while studying the KCNQ1/KCNE1 systems, which were in the range of 200 pN. This can also be attributed to the fact that the residues forming the energy barrier in the KCNQ1/KCNE1 systems coincide well with that of the binding site residues of the ligand (see Figure 3.9). These residues that formed the energy barrier as a result of the presence of KCNE1 included Phe966, Thr444, Thr691, Thr938, Ile963 and Thr444 which were all also included in the ligand binding site, as well. This coincidence explains the synergistic peak that is observed in the graphs of Figure 3.15.

The SMD simulations were able to clearly illustrate the effect of the blocking compounds on the energy barriers for the ions in the channel pore. However, we did not observe a significant difference between the force profiles of the two strong blockers (#1 and #2) and those for the medium blockers (#4 and #6). We hypothesize that this is happening because these two groups of molecules can block the ion permeation with almost the same degree. Meaning that, their presence in the pore can block the passage of the potassium ions equally, despite that their IC_{50}s were different from one another. Future work and optimization is required to increase the sensitivity of the SMD simulation to show differential force profiles for the channel blockers. Application of Adapting Biasing Force (ABF) simulations and Potential of Mean Force (PMF) along with SMD simulations can be one solution to make this improvement.

As for the weak blockers, the SMD simulations and the resulting force profiles were perfectly consistent with their lower biological activity, i.e. blocking the channel pore with a lower potency. As can be seen in Figure 3.15, the force profiles of Ligands #8 and #9 do not show the peak that was observed for the previous four ligands. There are a few low intensity peaks of about 200 pN, which is similar to that of the apo KCNQ1/KCNE1 systems. This indicates that the ligands have very minimal or no effect on the passage of the potassium ions.

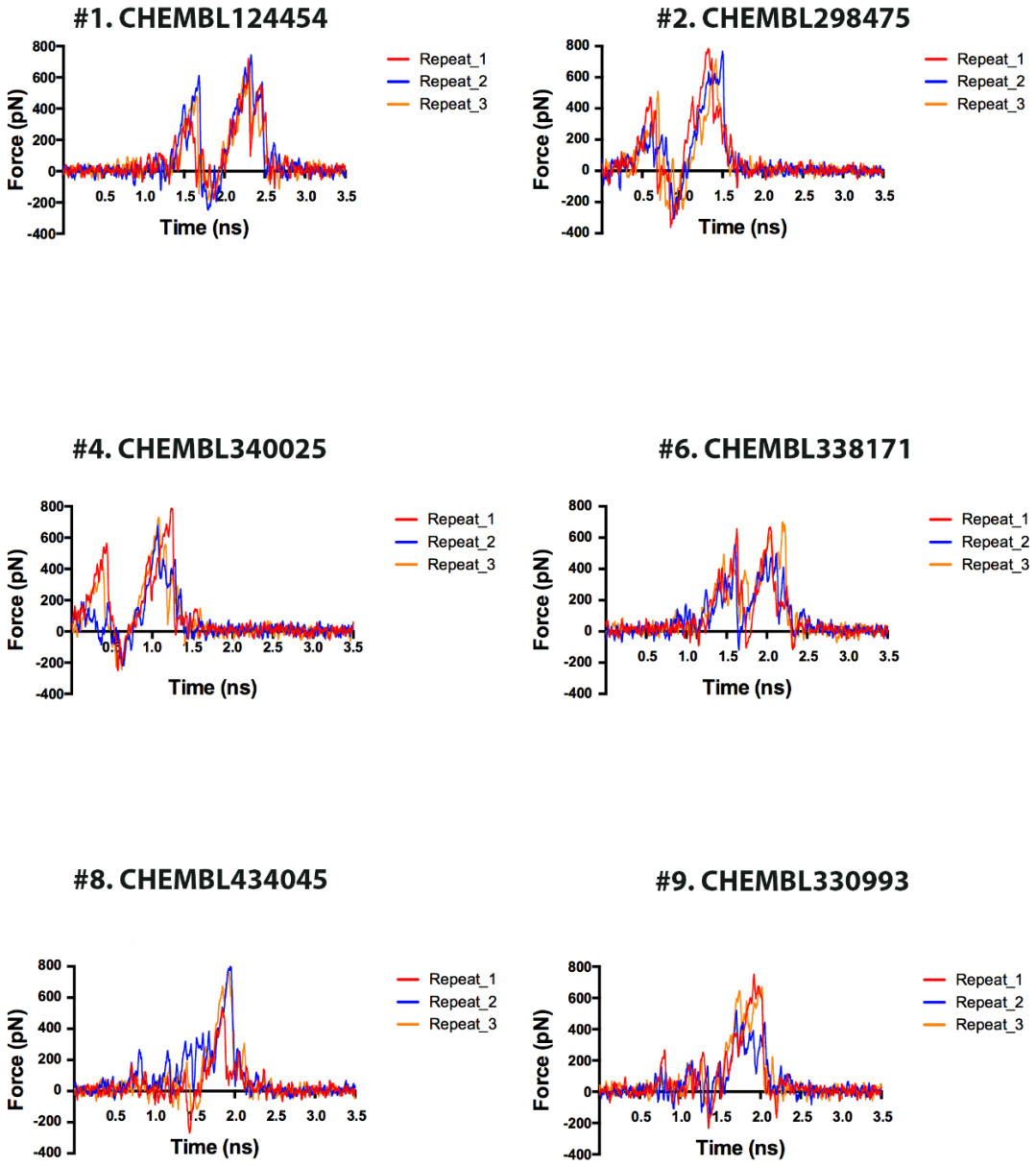


Figure 3.15. The force profiles of the three SMD repeats for each ligand (#1, #2, #4, #6, #8 and #9).

Next, we investigated the ion permeation path with the three groups of ligands, i.e. strong, median and weak blockers. Figure 3.16.a shows the involvement of ligand #1 (strong blocker) during the ion permeation process. As can be seen from the figure, the extended side chain of the drug molecule is directly in contact with the ion and interferes with its passage. This is consistent with the SAR information discussed in the previous section. Figure 3.16.b belongs to the involvement of Ligand #4 (median blocker). Although the fluoro substitution is extended towards the pore and forms a barrier for the ion, but it is not as long as the side chain of Ligand #1 to interfere much with the permeation. And finally Figure 3.16.c belongs to the weak ligand representative, i.e. ligand #8. This ligand similar to #9 has imposed the least energy barrier on the potassium ion. As can been seen from the figure, the cyano group substitution of the aromatic ring is completely drifted from the pathway of the potassium ion. Therefore, the only minimal barrier that it forms on the way of the ion are of the carbonyl groups which form comparatively weaker interactions with the ion and do not appear as huge peaks on the force profile plot of the ion (See Figure 3.15).

To explore the interactions that had taken place between the potassium ion and the ligands, we analyzed the SMD trajectories in light of the obtained force profiles. We observed three main modes of ligand-ion interactions (See Figure 3.17). Our first observation was related to ligand #1, the potassium ion was involved in an electrostatic interaction with the sulfonyl group in the ligand. This is a very strong type of interaction that acted as an energy barrier during the pulling of the ligand. Ligand #8 on the other hand, had a different mode of interaction with the potassium ion, involving a cation-π interaction with the aromatic benzene rings in the ligand structure. Ligands #2 and #9 also possessed the same types of the interactions, similar to ligand #1 and #16, respectively. Finally, we observed a third type of electrostatic interaction between the potassium ion and ligand #4. This interaction engaged the hydroxyl group on the aromatic ring of the ligand. However, this latter interaction was weaker in nature than the former types. Other ligands are not shown because more or less they had the same types of interactions with the K^+ ion, although not consistent with respect to their potency of action.

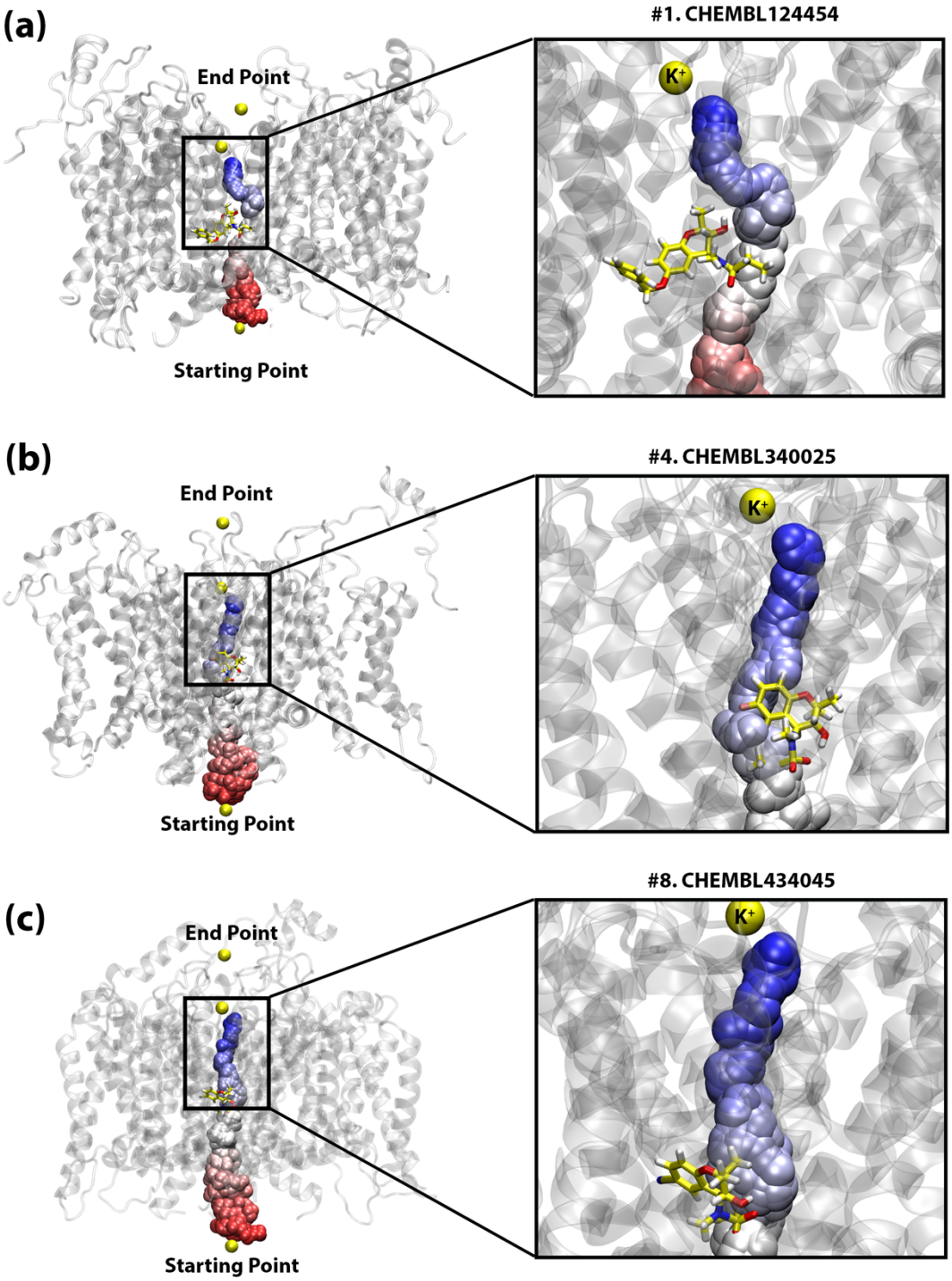


Figure 3.16. The involvement of the ligands with the potassium ion permeation pathway shown from the starting point (red) up to the end point (blue) of pulling. (a) Ligand #1 (strong blocker). (b) Ligand #4 (median blocker). (c) Ligand #8 (weak blocker).

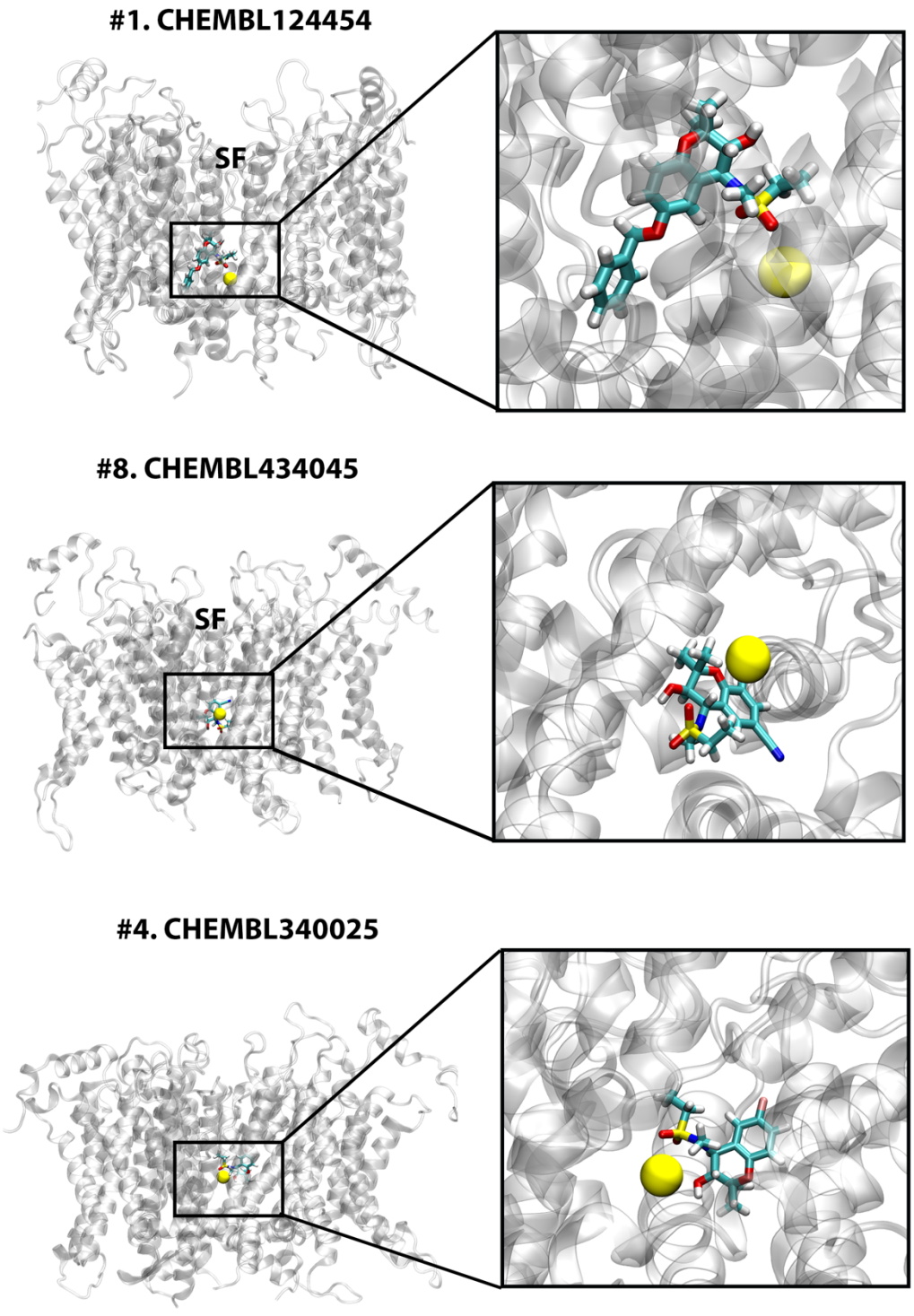


Figure 3.17. The different modes of interaction between the ligands (#1, #8, #4) and the potassium ion (yellow sphere). Ligand #1 possesses an electrostatic interaction of sulfonyl group with the K^+ ion. Ligand #16 involves cation-π interactions. Ligand #11 interacts with the K^+ ion by electrostatic forces through the phenolic OH group.

3.3.6. Pore Dimension Analysis of Drug-Protein Systems

Similar to the analysis we performed in the lone KCNQ1 and KCNQ1/KCNE1 systems, we also investigated the effects of the ligands on the dimensions of the pore. We performed pore dimension analysis using the HOLE program (O S Smart et al., 1993; Oliver S. Smart et al., 1996). Figure 3.18 shows the results for the 6 ligands and their effects on the pore size. Correlating with observation regarding the force profiles (see section 9.3), we also observed similar effects for the strong (#1 and #2) and the medium (#4 and #6) blockers on the pore dimensions. The two groups had roughly the same effect. On the other hand, the weak blockers (#16 and #17) induced a slightly lesser constriction in the pore radius. The pore radius profiles of the 6 systems (see Figure 3.19) clearly shows that the effects of the ligands appear only within the pore region. For all compounds, the topology of the selectivity filter remained almost the same as was observed with the lone KCNQ1 and the KCNQ1/KCNE1 systems.

Based on the abovementioned observations and analysis, it is clear that although strong and median activity blockers despite of having a broad range of KCNQ1 inhibitory activity, their effects on the ion permeation remains approximately the same. However, the very weak blockers are exempt from this observation, where they have a lesser effect on the ion permeation as well as on the pore dimensions. These in depth structural analyses suggest that small molecules can effectively block the KCNQ1 ion channel as long as their inhibitory activity are within the strong to medium range and as long as they interact similarly to key residues within the pore. Therefore, assessment of drugs for cardiotoxicity should involve close examination and investigations on the channel residues and ion interactions with the ligands.

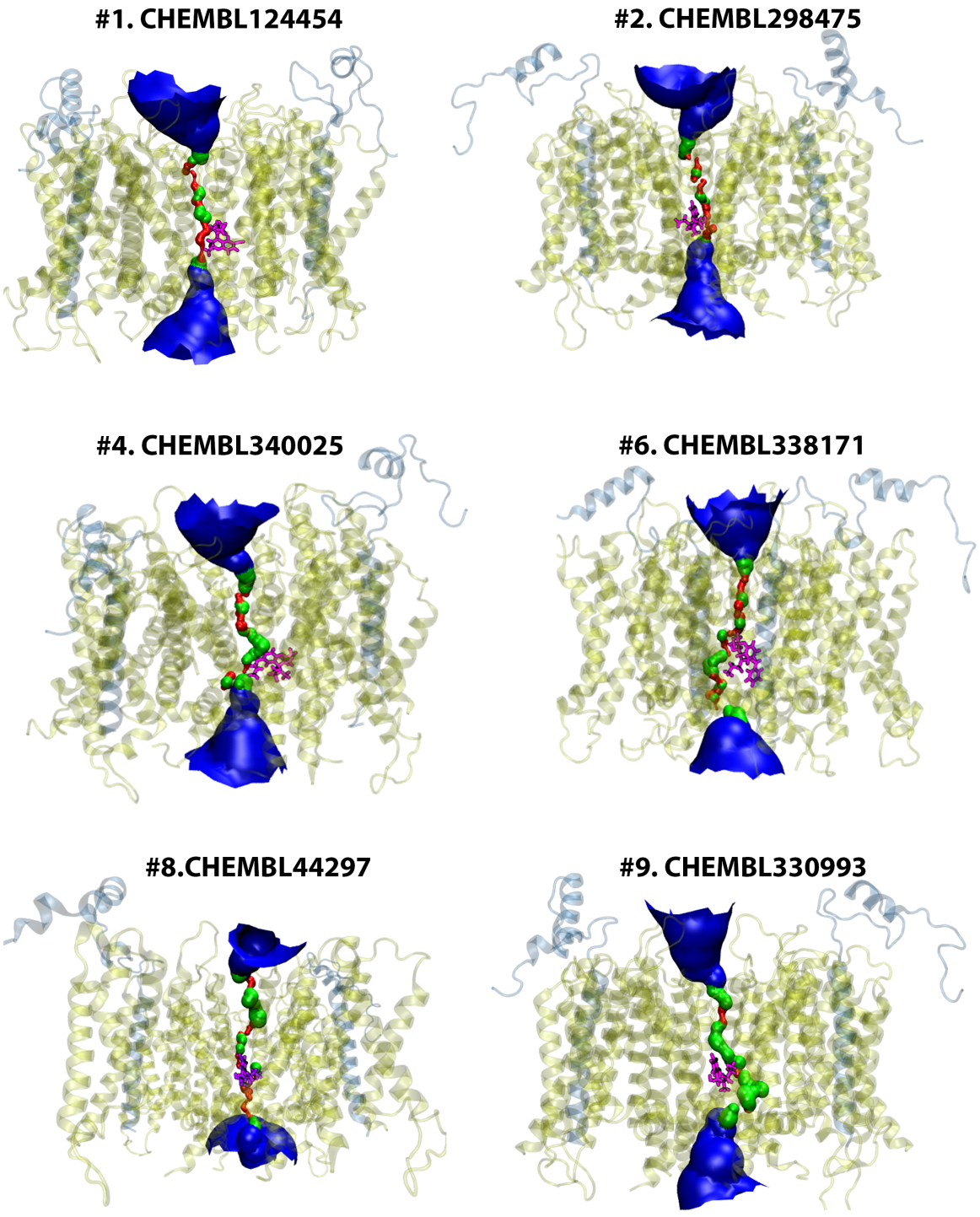


Figure 3.18. The dimensions of the pore (shown in surface representation) in KCNQ1/KCNE1 systems with the 6 docked ligands (shown in purple color). *Pore colour code: Red is where the pore radius is too tight for a water molecule. Green where there is room for a single water molecule. Blue is where the radius is double the minimum for a single water molecule.

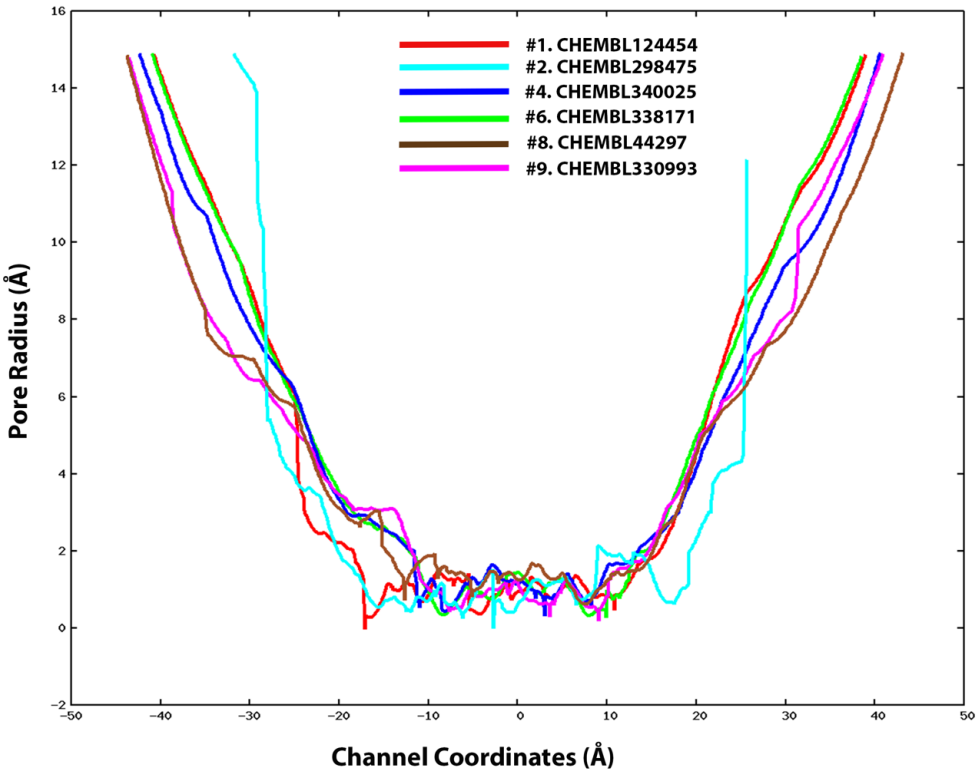


Figure 3.19. Pore radius vs. channel coordinates for the selected 6 ligands in the KCNQ1/KCNE1 channel complex. Ligands #8 and #9 (brown and pink lines) show the least constriction of pore radius, while the other four ligands impose more or less similar effect on the pore radius.

3.4. Conclusion

In this study, we presented the SMD simulations of potassium permeation across the KCNQ1 channel in a fully hydrated membrane system in the absence and presence of the KCNE1 protein. The main goal was to elucidate the dynamics of the residues lining the pore domain as a potassium ion passes through the channel. It is quite encouraging that our SMD simulations provided a direct visualization of the key ion permeation events through the KCNQ1 potassium ion channel. Furthermore, our findings provided insights about the effect of the KCNE1 protein on the K^+ ion permeation mode. The findings reported in this chapter, also serve as an additional validation for our model. In addition, the information provided regarding the effect of KCNE1 in terms of ion permeation, are novel and have not been studied before. A future direction to this study would be to test

the passage of other ion types, such as chloride and sodium and characterize the differential permeation that would exist between these ions.

We also performed small molecule docking simulations to obtain valid poses for a panel of known KCNQ1/KCNE1 blockers (Chromanol 293B and its derivatives) in the pore of the channel. To validate our docking results, we compared the scores form the docking scoring functions to the *in vitro* pIC_{50}s of the compounds under investigation. Interestingly, there was a strong and positive correlation of 0.7 between the experimental and predicted values. Furthermore, the binding mode of the ligands complied very well with the available experimental data regarding the binding site residues. From the binding and interaction mode of the ligands, we inferred their differential structure activity relationship. This high degree of agreement between results, motivated us to move to the next step and examine the effect of these drugs on the permeation of K^+ ions in the KCNQ1/KCNE1 channels. We used SMD simulations and pore dimension analysis to address this research question. Our findings illustrate how blockers of KCNQ1/KCNE1 channel perturb the normal passage of potassium ion with different modes of interaction with the ion in the pore. An interesting finding was that the effect of Chromanol 293B and its derivatives appears mainly within the channel pore and the presence of the compounds did not affect the topology of the selectivity filter. Overall, the findings in this chapter served as another level of validation for our model with regards to its ability for predicting the activity of drug molecule channel blockers. That being said, we believe the model is validated enough by different means and a future step would involve testing of other drug molecules to identify the off-target interactions with KCNQ1/KCNE1 ion channel.

3.5. References

Abbott, G. W. (2014). Biology of the KCNQ1 Potassium Channel. *New Journal of Science*, *2014*, 1–26. http://doi.org/10.1155/2014/237431

Ahmed M, Jalily Hasani H, Ganesan A, Houghton M, B. K. (2017). Modeling the human Nav1.5 sodium channel: structural and mechanistic insights of ion permeation and drug blockade. *Drug Design, Development and Therapy*, Manuscript accepted for publication.

Ain, Q. U., Aleksandrova, A., Roessler, F. D., & Ballester, P. J. (2015). Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdisciplinary Reviews. Computational Molecular Science*, *5*(6), 405–424. http://doi.org/10.1002/wcms.1225

Anwar-Mohamed, A., Barakat, K. H., Bhat, R., Noskov, S. Y., Tyrrell, D. L., Tuszynski, J. a., & Houghton, M. (2014). A human ether-á-go-go-related (hERG) ion channel atomistic model generated by long supercomputer molecular dynamics simulations and its use in predicting drug cardiotoxicity. *Toxicology Letters*, *230*(3), 382–392. http://doi.org/10.1016/j.toxlet.2014.08.007

Banks, J. L., Beard, H. S., Cao, Y., Cho, A. E., Damm, W., Farid, R., … Levy, R. M. (2005). Integrated Modeling Program, Applied Chemical Theory (IMPACT). *Journal of Computational Chemistry*, *26*(16), 1752–1780. http://doi.org/10.1002/jcc.20292

Barakat, K. (2015). Modelling Off-target Interactions (I): Cardiotoxicity. *Journal of Pharmaceutical Care & Health Systems*, *2*(3), 10–11. http://doi.org/10.4172/2376-0419.1000e131

Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., … Overington, J. P. (2014). The ChEMBL bioactivity database: an update. *Nucleic Acids Research*, *42*(D1), D1083–D1090. Retrieved from http://dx.doi.org/10.1093/nar/gkt1031

Bleich, M., Briel, M., Busch, A. E., Lang, H. J., Gerlach, U., Gogelein, H., … Kunzelmann, K. (1997). KVLQT channels are inhibited by the K+ channel blocker 293B. *Pflugers Archiv : European Journal of Physiology*, *434*(4), 499–501.

Busch, A. E., Suessbrich, H., Waldegger, S., Sailer, E., Greger, R., Lang, H.-J., … Maylie, J. G. (1996). Inhibition of IKs in guinea pig cardiac myocytes and guinea pig IsK channels by the chromanol 293B. *Pflügers Archiv - European Journal of Physiology*, *432*(6), 1094–1096. http://doi.org/10.1007/s004240050240

Ceccarini, L., Masetti, M., Cavalli, A., & Recanatini, M. (2012). Ion Conduction through the hERG Potassium Channel. *PLOS ONE*, *7*(11), e49017. Retrieved from https://doi.org/10.1371/journal.pone.0049017

Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., … MacKinnon, R. (1998). The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science*, *280*(5360), 69–77. http://doi.org/10.1126/science.280.5360.69

Durrant, J. D., & McCammon, J. A. (2011). NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *Journal of Chemical Information and Modeling*, *51*(11), 2897–2903. http://doi.org/10.1021/ci2003889

Fowler, P. W., & Sansom, M. S. P. (2013). The pore of voltage-gated potassium ion channels is strained when closed. *Nature Communications*, *4*(May), 1872. http://doi.org/10.1038/ncomms2858

Gerlach, U., Brendel, J., Lang, H.-J., Paulus, E. F., Weidmann, K., Brüggemann, A., … Greger, R. (2001). Synthesis and Activity of Novel and Selective I Ks -Channel Blockers. *Journal of Medicinal Chemistry*, *44*(23), 3831–3837. http://doi.org/10.1021/jm0109255

Hoshi, T., & Armstrong, C. M. (2013). C-type inactivation of voltage-gated K+ channels: Pore constriction or dilation? *The Journal of General Physiology*, *141*(2), 151–160. http://doi.org/10.1085/jgp.201210888

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38. http://doi.org/10.1016/0263-7855(96)00018-5

Kass, R. S. (2005). The channelopathies: novel insights into molecular and genetic mechanisms of human disease. The Journal of Clinical Investigation, 115(8), 1986–1989. http://doi.org/10.1172/JCI26011

Kannankeril, P., Roden, D. M., & Darbar, D. (2010). Drug-Induced Long QT Syndrome. *Pharmacological Reviews*, *62*(4), 760–781. http://doi.org/10.1124/pr.110.003723

Khamis, M. A., Gomaa, W., & Ahmed, W. F. (2015). Machine learning in computational docking. *Artificial Intelligence in Medicine*, *63*(3), 135–152. http://doi.org/10.1016/j.artmed.2015.02.002

Koes, D. R., Baumgartner, M. P., & Camacho, C. J. (2013). Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *Journal of Chemical Information and Modeling*, *53*(8), 1893–1904. http://doi.org/10.1021/ci300604z

Kutteh, R., Vandenberg, J. I., & Kuyucak, S. (2007). Molecular dynamics and continuum electrostatics studies of inactivation in the HERG potassium channel. *Journal of Physical Chemistry B*, *111*(5), 1090–1098. http://doi.org/10.1021/jp066294d

Lerche, C., Bruhova, I., Lerche, H., Steinmeyer, K., Wei, A. D., Strutz-Seebohm, N., … Seebohm, G. (2007). Chromanol 293B binding in KCNQ1 (Kv7.1) channels involves electrostatic interactions with a potassium ion in the selectivity filter. *Molecular Pharmacology*, *71*, 1503–1511. http://doi.org/10.1124/mol.106.031682

Linder, T., de Groot, B. L., & Stary-Weinzinger, A. (2013). Probing the Energy Landscape of Activation Gating of the Bacterial Potassium Channel KcsA. *PLOS Computational Biology*, *9*(5), e1003058. Retrieved from https://doi.org/10.1371/journal.pcbi.1003058

Liu, S., Focke, P. J., Matulef, K., Bian, X., Moënne-Loccoz, P., Valiyaveetil, F. I., & Lockless, S. W. (2015). Ion-binding properties of a K^+ channel selectivity filter in different conformations. *Proceedings of the National Academy of Sciences*, *112*(49), 15096–15100. http://doi.org/10.1073/pnas.1510526112

Lockless, S. W., Zhou, M., & MacKinnon, R. (2007). Structural and Thermodynamic Properties of Selective Ion Binding in a K+ Channel. *PLOS Biology*, *5*(5), e121. Retrieved from https://doi.org/10.1371/journal.pbio.0050121

Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, *27*(3), 221–234. http://doi.org/10.1007/s10822-013-9644-8

Melman, Y. F., Um, S. Y., Krumerman, A., Kagan, A., & McDonald, T. V. (2004). KCNE1 binds to the KCNQ1 pore to regulate potassium channel activity. *Neuron*, *42*(6), 927–937. http://doi.org/10.1016/j.neuron.2004.06.001

Mirams, G. R., Cui, Y., Sher, A., Fink, M., Cooper, J., Heath, B. M., … Noble, D. (2011). Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovascular Research*, *91*(1), 53–61. http://doi.org/10.1093/cvr/cvr044

Morais-Cabral, J. H., Zhou, Y., & MacKinnon, R. (2001). Energetic optimization of ion conduction rate by the K+ selectivity filter. *Nature*, *414*(6859), 37–42. Retrieved from http://dx.doi.org/10.1038/35102000

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *Journal of Computational Chemistry*, *30*(16), 2785–2791. http://doi.org/10.1002/jcc.21256

Paulussen, A. D. C., Gilissen, R. A. H. J., Armstrong, M., Doevendans, P. A., Verhasselt, P., Smeets, H. J. M., … Aerssens, J. (2004). Genetic variations of KCNQ1, KCNH2, SCN5A, KCNE1, and KCNE2 in drug-induced long QT syndrome patients. *Journal of Molecular Medicine*, *82*(3), 182–188. http://doi.org/10.1007/s00109-003-0522-z

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. http://doi.org/10.1002/jcc.20084

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., … Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802. http://doi.org/10.1002/jcc.20289

Schrödinger Release 2017-2: LigPrep, Schrödinger, LLC, New York, NY, 2017. (2017). *New York, NY*.

Schrödinger Release 2017-2: MS Jaguar, Schrödinger, LLC. (2017). *New York, NY*.

Seebohm, G. (2013). A complex partnership: KCNQ1 and KCNE1. *Biophysical Journal*, *105*(11), 2437–2438. http://doi.org/10.1016/j.bpj.2013.10.022

Smart, O. S., Goodfellow, J. M., & Wallace, B. A. (1993). The pore dimensions of gramicidin A. *Biophysical Journal*, *65*(6), 2455–2460. http://doi.org/10.1016/S0006-3495(93)81293-1

Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A., & Sansom, M. S. P. (1996). HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics*, *14*(6), 354–360. http://doi.org/10.1016/S0263-7855(97)00009-X

Suessbrich, H., Bleich, M., Ecke, D., Rizzo, M., Waldegger, S., Lang, F., … Busch, A. E. (1996). Specific blockade of slowly activating I(sK) channels by chromanols - Impact on the role of I(sK) channels in

epithelial. *FEBS Letters*, *396*(2–3), 271–275. http://doi.org/10.1016/0014-5793(96)01113-1

Sula, A., Booker, J., Ng, L. C. T., Naylor, C. E., DeCaen, P. G., & Wallace, B. A. (2017). The complete structure of an activated open sodium channel. *Nature Communications*, *8*, 14205. http://doi.org/10.1038/ncomms14205

Tristani-Firouzi, M., & Sanguinetti, M. C. (1998). Voltage-dependent inactivation of the human K+ channel KvLQT1 is eliminated by association with minimal K+ channel (minK) subunits. *The Journal of Physiology*, *510*(1), 37–45. http://doi.org/10.1111/j.1469-7793.1998.037bz.x

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. http://doi.org/10.1002/jcc.21334

Witchel, H. J., & Hancox, J. C. (2000). Familial and acquired long qt syndrome and the cardiac rapid delayed rectifier potassium current. *Clinical and Experimental Pharmacology & Physiology*, *27*(10), 753–766.

Wojcikowski, M., Ballester, P. J., & Siedlecki, P. (2017). Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*, *7*, 46710. http://doi.org/10.1038/srep46710

Yap, Y. G., & Camm, A. J. (2003). Drug induced QT prolongation and torsades de pointes. *Heart*, *89*(11), 1363–1372. http://doi.org/10.1136/heart.89.11.1363

Yellen, G. (2002). The voltage-gated potassium channels and their relatives. *Nature*, *419*(6902), 35–42. http://doi.org/http://dx.doi.org/10.1038/nature00978

CHAPTER 4. GENERAL DISCUSSION

The human KCNQ1 (Kv7.1 or KvLQT1) ion channel belongs to the large family of voltage gated potassium ion channels. The KCNQ1 protein is expressed in the heart, muscles, pancreas, kidney, brain and inner ear. The cardiac KCNQ1 ion channel is one of the most significant proteins in the heart, mainly due to the crucial role it plays in regulating the heart's function through the generation of the native I_{Ks}. The I_{Ks} current controls the cardiac repolarization, and thus the duration of the cardiac action potential. The importance of the I_{Ks} current is evident from the number of pathophysiological conditions that can result from mutations affecting the biophysical and regulatory properties of the KCNQ1/KCNE1 channel. These include long QT syndrome (LQTS), deliquium seizures and sudden death. Null or missense mutations in Kv7.1 can also result in hearing loss and other auditory problems. Furthermore, a cardio-auditory syndrome called Jervell and Lange-Nielsen syndrome (JLNS) also results from mutations in this gene (Mousavi Nik, Gharaie, & Jeong Kim, 2015). Gain-of-function mutations in Kv7.1 increase the current flow and can lead to shortening of the cardiac action potential; as seen in a number of cardiac rhythm disorders such as Short QT syndrome (SQTS) and atrial fibrillation (Wulff, Castle, & Pardo, 2009).

Despite of the large number of studies, the molecular mechanisms and structural details associated with the KCNQ1 ion channel remain a significant unsolved problem. The literature illustrates many earlier attempts to understand the structure and function of this channel, on both experimental and computational avenues. We identified the pitfalls and limitations of the modeling studies in order to address them in our model. These limitations in the earlier els included, but were not limited to, the absence of KCNE1 protein as a part of the model (Kasimova, Zaydman, Cui, & Tarek, 2015), failure in inclusion of all the necessary physiologic components such as the PIP2 lipids (Xu et al., 2013), and lastly the models are considered outdated because of the continuous generation of structural data related to KCNQ1 structure, as will be described below.

The last few years witnessed the accumulation of very valuable structural information that motivated us to create a more robust model compared to those described in the literature (Kang et al., 2008; Kasimova et al., 2015; Smith, Vanoye, Jr, Meiler, & Sanders, 2007; Van Horn, Vanoye, & Sanders, 2011; Xu et al., 2013). Of utmost

importance are the NMR studies that reported minute structural details in the KCNQ1 channel (I.D. et al., 2015; Peng et al., 2014), specifically, the details related to the structure of the VSD. That is, helices' x and y in the VSD should be more elongated by z turns compared to earlier models. This is of utmost importance as this can significantly affect the passage of ions through the channel. In our new model, we included these structural components and refined them using REMD simulations to predict their most accurate conformational state as described in Chapters 2 and 3.

This thesis investigated the structural and functional properties of the KCNQ1 (KV7.1) ion channel in complex with its accessory beta subunit; KCNE1 protein and in its most relevant physiological conditions. Towards this goal, we combined several advanced computational tools and techniques while incorporating the latest experimental data to construct a comprehensive model for the KCNQ1/KCNE1. The Thesis was presented in the following way. The subsections under Chapter 1 provided the essential background on ion channels in general and on the KCNQ1/KCNE1 channel in particular. These sections also summarized the aims and research questions of this Thesis. That is, the generation of a comprehensive model for KCNQ1 channel associated with its beta subunit KCNE1 proteins (Chapter 2), the investigation of their interaction and effect on each another and finally the studying of ion permeation in the presence and absence of KCNE1 and small molecule channel blockers (Chapter 3). Appendix A provided a complete introduction to the concepts of MD simulations and its enhanced variants, which were also extensively used in this Thesis. Appendices B and C presented an over view of two important techniques that formed the basis for the *in silico* methods used in this Thesis, namely, homology modeling (Appendix B) and protein-protein docking (Appendix C). Along the same line,

Chapter 2 described the construction of the structural model for KCNQ1 channel, while fitting the model to all available experimental data in the literature. Following the KCNQ1 model building, we used replica exchange MD (REMD) simulations to explore the conformational space of the KCNQ1 voltage sensing domain (VSD). REMD is one of the main techniques to investigate transitions in proteins secondary structure and the folding of peptides from first principles. We took advantage of the capability of the REMD method to explore the possible secondary structure properties of the S1 and S2

helices in the KCNQ1 VSD. As described above, these two helices acquire a secondary structure, which is significantly different from the subtype ion channels and several homologous members of the superfamily of voltage gated ion channels. Therefore, including this important piece of information in our model was an essential task especially due to the critical location of these helices in the KCNQ1 VSD. To our knowledge, the inclusion of these S1-S2 structural details forms a novel part of this Thesis. Following the identification of the most relevant conformations of the VSD, the whole KCNQ1 model was refined using a substantially long classical MD simulation for 240 ns. This long MD simulation was done for a couple of reasons. First, MD simulations allowed the relaxation of the protein and the removal of structural clashes that can otherwise lead to further imprecisions in the desired results. Secondly, for a complete KCNQ1/KCNE1 complex generation, an ensemble of KCNQ1 structures was required. This conformational ensemble was obtained from the MD trajectories, which represented a wide variety of structural representatives for the KCNQ1 ion channel. The final refined KCNQ1 model was validated using known experimental data and complied very well in several aspects including the exposure of certain residues to the lipid environment and secondary structure of important protein segments (Gayen, Li, & Kang, 2015). The exposure of these residues located on the S4-S5 linker of KCNQ1, to the lipid environment in our model adds to the validation of our modeling approach in two ways. Firstly, for these residues to interact with the membrane, they need to have the correct orientation. Secondly, this finding also confirms that the simulation system has been built in the right way such that the membrane wrapping allows the interaction to occur.

To construct a physiologically relevant KCNQ1 ion channel, we had to carry on a very important step, namely assembling a complex between the KCNQ1 channel and its auxiliary protein KCNE1 protein. Despite the complexity of this step there were several pieces of structural information that guided us to move forward with this difficult step. These experimental information related to the contacts made between KCNQ1 and KCNE1 came from cross-linking and mutagenesis studies (Boulet, Labro, Raes, & Snyders, 2007; Chung et al., 2009; Gofman, Shats, Attali, Haliloglu, & Ben-Tal, 2012; Kang et al., 2008; Kasimova, Zaydman, Cui, & Tarek, 2015; Panaghie, Tai, & Abbott, 2006; Strutz-Seebohm et al., 2011; Tapper & George, 2001; Y. Wang et al., 2012; Y. H.

Wang et al., 2011; X. Xu, Jiang, Hsu, Zhang, & Tseng, 2008a). Based on this information, we adopted data-driven protein-protein docking simulations to assemble the KCNQ1:KCNE1 complex in a 4:2 stoichiometry. We analyzed the interaction between the two proteins after running MD simulations as a measure of structural refinement. This analysis led to interesting conclusions that were consistent with the reported information in the literature. Firstly, all the essential contacts necessary for the two proteins to associate together were retained during MD, and in some cases even made stronger. Secondly, we observed two main modes of interaction between KCNQ1 tetramer and the two KCNE1 proteins. In the first mode, the N-terminals of the KCNE1 were extended in outward directions, whereas in the second mode of interaction the KCNE1 N-terminals folded back on top of the intracellular regions of KCNQ1 structure. In general, the association of KCNQ1 with KCNE1 significantly reduced the magnitude of the VSD fluctuations. This latter finding is directly consistent with several of the previous experimental data (Kang et al., 2008; X. Xu, Jiang, Hsu, Zhang, & Tseng, 2008b; Y. Xu et al., 2013), which adds another level of validation to our model.

The consistency of the data from the model with experimental data and its success in various validation steps, motivated us to carry out further studies on the KCNQ1/KCNE1 ion channel complex in terms of ion permeation mechanisms and interactions with small molecules. We divided these applications into two main categories, which form the core of the research paper presented in Chapter 3. The first application investigated ion permeation mechanisms through the KCNQ1 ion channel and the effect of the KCNE1 accessory protein on ion permeation. Using SMD simulations, we derived several findings on the differential pore characteristics of the channel in the absence and presence of KCNE1 protein. In the absence of the KCNE1 protein, the potassium ion came across the selectivity filter residues of KCNQ1 as a natural sieving process. For an ion to successfully pass through the filter, it has to overcome each of the carbonyl atom cages created by these residues, one after the other. However, based on our findings these binding sites do not impose equal obstacles on the way of the ion. The force needed for a potassium ion to migrate from the first binding site formed by Thr residues in the TIGYG motif is lower than the one formed by the Ile and Gly residues at least by half. That might explain the reason as to why a second potassium ion is usually

needed to push the first one ahead to the next binding site by repulsive electrostatic forces. However, entering the first binding site is rather an easy task for a single potassium ion to be achieved on its own. The last binding site that is formed by the Tyr and Gly residues (B4) is also of lower intensity, enabling the exit of the potassium ion from the selectivity filter and its complete release into the extracellular region.

On the other hand, in the presence of KCNE1 protein, we observed a completely different force profile characterises for the ion. This differential characteristic was imposed by the KCNE1 proteins through formation of two main energetic barriers inside the pore region before the ion approaches the barriers formed by the selectivity filter. These two barriers were formed each by a group of residues, one of which engaged the potassium ion in a cation-π interaction and the other trapped the ion by strong electronegative interactions. The pore in the lone KCNQ1 channel possessed a wider radius (~ 6 Å) whereas this was considerably constricted (~1.5 Å) when the channel was associated with the KCNE1 proteins. This effect was due to the structural changes and orientation of S5 and S5-S6 linker residues in the pore region while the characteristics of the selectivity filter remained completely intact with a radius of ~1.8 Å in both the systems. We were also able to link this information to the existing data regarding the effect of KCNE1 on KCNQ1 function. It is known from the literature that the association of KCNQ1 with KCNE1 proteins, imposes a slow activation by affecting the movement of the VSD (Osteen et al., 2010).This effect has also been linked to a collapse in the pore region with rearrangement of residues in the pore (Hoshi & Armstrong, 2013). Both these experimental evidences are in accordance with our findings regarding the changes in the KCNQ1 protein, when it is in complex with KCNE1. Our modeling study clearly demonstrated the specific structural changes and residues which can play an important role in slowing of the channel and changing its ion transport mechanisms. To our knowledge, this study represents the first attempt to use in silico techniques to understand ion conduction through the KCNQ1 channel and investigate the structural effect of KCNE1 as a direct modifier of ion permeation process.

The second category of applications described in Chapter 3 focused on the mechanism of channel blockage by small molecules and their effects on ion permeation. Through this study, we investigated the binding of a set of known I_{KS} current blockers,

using small molecule docking simulations. The compounds consisted of Chromanol 293 B and its derivatives (9 compounds in total), which came from the same study, ensuring data consistency and they had a broad range of activity (IC_{50}: 50 nM to 58,000 nM), which was ideal for validating our model and testing its competence in predicting the differential blocking effect of drug molecules. Interestingly, the results from our docking simulations correlated very well with the experimental and *in vitro* activities of the compounds. The Pearson product-moment correlation coefficient was computed to assess the relationship between the scores calculated from the docking poses of the ligands against the KCNQ1/KCNE1 protein complexes, and the pIC_{50}s of the compounds, revealing a strong, positive correlation ($r_{pearson} = 0.7$). The model was therefore capable of discriminating blockers of differential activity from each other. For example, the most potent blocker (#1) having an IC_{50} of 50 nM ($pIC_{50} = 7.3$) was also detected by our scoring functions to be the strongest blocker. The trend remains the same even for the weak blockers such as ligand #9, which has an IC_{50} of 58,000 ($pIC_{50} = 4.2$) and a very low score of 6.3 from docking. Through this study, we were able to infer the structural components of the protein responsible for interaction with the blockers and derived remarkable structure activity relationship data. These interactions occurred in a well-defined pocket right below the selectivity filter, as was reported before in previous mutational studies in the literature. The contacts made by the blockers were both with the VSD domain (S5, S6, S5-P-loop linkers) as well as the selectivity filter residues. The substitutions at the 6-position on the aromatic ring is the first determinant of potency. This substitution, depending on its size and direction of extension towards the central axis of the pore, can have differential effect on the conduction of the potassium ions. Secondly, the sulfonyl residue substitution also affects the potency of the compounds. This effect may be produced because the substitutions at this position can have an affinity to interact with the residues on the periphery of the pore. This interaction draws the drug molecule further away from the pore and thereby reducing their ability to produce physical blockage.

Following the placement of the blockers within their suitable modes of binding within the channel, we examined the effect of these drugs on the permeation of potassium ions in the KCNQ1/KCNE1 channels, using SMD simulations. The results revealed the

energy barriers formed by the small molecule blockers and the synergistic blockage of the pore residues secondary to the presence of the drug molecules. We hypothesize that the energy barrier formed by KCNE1 and the ligands overlap as they both appear at the same structural location in the pore region of the channel. This explains the higher intensity of the peak in the force profile of the ion in the presence of the drug molecules as compared with the KCNQ1/KCNE1 systems without drug. The potassium ions passing through the KCNQ1/KCNE1 channel in complex with different drug molecules, revealed three possible modes of interactions. These included cation-π interactions and two variants of electrostatic interactions involving sulfonyl and hydroxyl groups of the ligands.

Further analysis of the pore topology allowed us to understand the effect of the blockers on the physical characteristics of the channel pore. The strong and median activity blockers despite of having a broad range of KCNQ1 inhibitory activity, imposed the same effects on the pore topology. However, the very weak blockers are exempt from this observation, where they have a lesser effect on the pore dimensions. These in depth structural analyses suggest that small molecules can effectively block the KCNQ1 ion channel as long as their inhibitory activity are within the strong to medium range and as long as they interact similarly to key residues within the pore. Therefore, assessment of drugs for cardiotoxicity should involve close examination and investigations on the channel residues and ion interactions with the ligands. The results combined from the two parts of this research project that is presented in Chapter 2 and 3, shed light on the pore topology and the effect of different components on the ion permeation process in the KCNQ1/KCNE1 channel. These data are significant from both drug design point of view as well as cardiotoxicity studies of the KCNQ1/KCNE1 channel. Designing pharmacological KCNQ1/KCNE1 channel openers as novel anti-arrhythmic drugs has been an ongoing research for years (Barbuti et al., 2011; Jow et al., 2006; Yu et al., 2013). The structural details from our studies has revealed details related to the pore and effect of different intrinsic as well as extraneous components which can disrupt the normal permeation of potassium ions. This information may open new avenues for drug design strategies to overcome the barriers that disturb the passage of ions and thereby act as channel openers in curing cardiac arrhythmias and saving patients from fatal episodes of LQTS. On the other hand, the unintentional blockage of KCNQ1/KCNE1 ion channels

is one of the main causes of drugs' cardiotoxicity. Our ion permeation studies using s Chromanol 293B drug scaffolds and their effect on ion permeation, disclosed important characteristics that make a drug capable of disrupting the ion permeation process. In addition, the success of our model in predicting the blocking activity of drugs with a high quality that competes with experimental bioactivity data, suggests a future possibility for using modeling approaches in detecting drugs of high potential for cardiotoxicity.

Overall, the findings from this Thesis are important and investigated novel aspects of the KCNQ1/KCNE1 channel complex. The model proposed in this Thesis was validated through available experimental data in the literature. However, one should look into this Thesis as a good starting point for further studies to investigate different drug scaffolds and different mechanisms by which they can affect ion permeation in KCNQ1/KCNE1 ion channels. It can be also a good starting point to apply the same techniques adopted in this Thesis to model other cardiac ion channels and combine these models to reproduce the cardiac rhythm and understand the effect of drugs on this important phenomenon.

CHAPTER 5. FUTURE PERSPECTIVES

The approach presented in this Thesis is a starting point for many subsequent structural studies and for the construction of more improved models at various conformational states. One possible direction would be the development of an inactivated/closed state of the KCNQ1/KCNE1 channel. Most recently and towards the end of this project, Sun and MacKinnon published a CryoEM structure of frog (*Xenopus laevis*) KCNQ1 structure in complex with calmodulin (CaM) (J. Sun & MacKinnon, 2017). This structure can serve as a suitable template for modeling of the decoupled state (a transition state between closed and open) of the human KCNQ1 ion channel. Developing this model would enable further studies of ion permeation to complement the findings of this thesis. These studies may also include the differential ion permeation pathways between the open state (our model) and the decoupled state. This will eventually allow testing the affinity of drug molecules for the different states of the channel.

Another expansion upon the current model would involve the inclusion of the intracellular regions (COOH terminal) of the channel. The COOH terminal of KCNQ1 is known to be involved in trafficking, tetramerization and further interactions of the channel. Therefore, including them in this model could open doors for future investigations of trafficking and a simulation of tetramerization using advanced sampling techniques such as metadynamics. Also, having the COOH terminal included in the model, allows the study of CaM association that binds to the proximal region of this domain. Mutations in the regions responsible for binding to CaM are implicated in congenital heart diseases and therefore, implicating an involvement with ion permeation. The model of KCNQ1 bound with CaM can also be a good target for ion permeation studies to understand the intricate effects on this fundamental phenomenon.

The results from our ion permeation studies make us believe that our model can be considered accurate enough for the prediction of off-target interactions of other drug molecules. However, several optimizations and the use of more sophisticated simulation techniques can enhance the limitations that our model faced. One such limitation was related to the differential effect of ligands on ion permeation, which had close but different activity. Although our approach enabled us to differentiate very strong blockers from very weak blockers, it did not show any differential results for the median potency

blockers in terms of ion permeation. This indicates that we need to increase the sensitivity of our testing approach. One possible solution to this issue is to apply more advanced computational analyses in the future. This includes the use of Potential of Mean Force (PMF) and Adaptive Biasing Force (ABF) techniques hand in hand with the SMD simulations. This should be also combined with in-depth experimental structural data (e.g. additional mutational analysis and NMR spectroscopy) to provide detailed validation of the computational data. For example, mutational analysis can be helpful to confirm the direct involvement of specific residues that we repeatedly found to have a role in the ion permeation events. In the same context of ion permeation studies, one possible expansion on the findings of this Thesis is to test the permeation of other types of ions through the KCNQ1/KCNE1 channel. This would enable the study of structural basis behind the selectivity of the KCNQ1 channel towards potassium ions.

A long term goal of this project is to follow the same approach that was developed in this Thesis to build comprehensive models for other cardiac ion channels and link them together, with the aim of constructing a dynamical model for the whole heart. Such models can enable deep studies on the collective movements of the various ions in the heart, the effect of changing concentration of one ion type on the rest of the ions and the overall cardiotoxic effect of a small molecule blocker on all the cardiac ion channels.

I end this thesis by hoping that one day, potential cardiotoxicity of drugs can become an old research story narrated to the next generations.

BIBLIOGRAPHY

Abagyan, R., Totrov, M., & Kuznetsov, D. (1994). ICM - A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, *15*(5), 488–506. http://doi.org/10.1002/jcc.540150503

Abbott, G. W. (2014). Biology of the KCNQ1 Potassium Channel. *New Journal of Science*, *2014*, 1–26. http://doi.org/10.1155/2014/237431

Abbott, G. W. (2016). KCNE1 and KCNE3: The yin and yang of voltage-gated K+ channel regulation. *Gene*, *576*(1), 1–13. http://doi.org/10.1016/j.gene.2015.09.059

Abbott, G. W., & Goldstein, S. A. (2001). Potassium channel subunits encoded by the KCNE gene family: physiology and pathophysiology of the MinK-related peptides (MiRPs). *Molecular Interventions*, *1*(2), 95–107.

Abrams, C., & Bussi, G. (2014). Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. *Entropy*, *16*(1), 163–199. http://doi.org/10.3390/e16010163

Abriel, H. (2007). Roles and regulation of the cardiac sodium channel Nav1.5: Recent insights from experimental studies. *Cardiovascular Research*, *76*(3), 381–389. http://doi.org/10.1016/j.cardiores.2007.07.019

Accardi, A., & Pusch, M. (2000). Fast and Slow Gating Relaxations in the Muscle Chloride Channel Clc-1. *The Journal of General Physiology*, *116*(3), 433–444. http://doi.org/10.1085/jgp.116.3.433

Adcock, S. A., & McCammon, J. A. (2006). Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews*, *106*(5), 1589–1615. http://doi.org/10.1021/cr040426m

Adelman, S. A., & Doll, J. D. (1976). Generalized Langevin equation approach for atom/solid-surface scattering: General formulation for classical scattering off harmonic solids. *The Journal of Chemical Physics*, *64*(6), 2375. http://doi.org/10.1063/1.432526

Ahern, C. A., Payandeh, J., Bosmans, F., & Chanda, B. (2016). The hitchhiker's guide to the voltage-gated sodium channel galaxy. *The Journal of General Physiology*, *147*(1), 1–24. http://doi.org/10.1085/jgp.201511492

Ahmed, M., & Barakat, K. (2015). Baby steps toward modelling the full human programmed Death-1 (PD-1) pathway. *Receptors & Clinical Investigation*, *1*, 1–5. http://doi.org/10.14800/rci.825

Ahmed M, Jalily Hasani H, Ganesan A, Houghton M, B. K. (2017). Modeling the human Nav1.5 sodium channel: structural and mechanistic insights of ion permeation and drug blockade. *Drug Design, Development and Therapy*, Manuscript accepted for publication.

Aidley, D. J., & Stanfield, P. R. (1996). *Ion Channels—Molecules in Action*. *Cambridge University Press*.

Ain, Q. U., Aleksandrova, A., Roessler, F. D., & Ballester, P. J. (2015). Machine-learning scoring functions to improve structure-based binding affinity prediction and virtual screening. *Wiley Interdisciplinary Reviews. Computational Molecular Science*, *5*(6), 405–424. http://doi.org/10.1002/wcms.1225

Aiyar, J., Rizzi, J. P., Gutman, G. A., & Chandy, K. G. (1996). The signature sequence of voltage-gated potassium channels projects into the external vestibule. *The Journal of Biological Chemistry*, *271*(49), 31013–31016.

Alford, R. F., Koehler Leman, J., Weitzner, B. D., Duran, A. M., Tilley, D. C., Elazar, A., & Gray, J. J. (2015). An Integrated Framework Advancing Membrane Protein Modeling and Design. *PLoS Computational Biology*, *11*(9), e1004398. http://doi.org/10.1371/journal.pcbi.1004398

Allen, M. (2004). Introduction to molecular dynamics simulation. *Computational Soft Matter: From Synthetic Polymers to Proteins*, *23*(2), 1–28. http://doi.org/10.1016/j.cplett.2006.06.020

Allison, J. R., Hertig, S., Missimer, J. H., Smith, L. J., Steinmetz, M. O., & Dolenc, J. (2012). Probing the structure and dynamics of proteins by combining molecular dynamics simulations and experimental NMR data. *Journal of Chemical Theory and Computation*, *8*(10), 3430–3444. http://doi.org/10.1021/ct300393b

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. http://doi.org/http://dx.doi.org/10.1016/S0022-2836(05)80360-2

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic*

Acids Research, *25*(17), 3389–3402.
An, W. F., Bowlby, M. R., Betty, M., Cao, J., Ling, H.-P., Mendoza, G., … Rhodes, K. J. (2000). Modulation of A-type potassium channels by a family of calcium sensors. *Nature*, *403*(6769), 553–556. Retrieved from http://dx.doi.org/10.1038/35000592
Anandakrishnan, R., Drozdetski, A., Walker, R. C., & Onufriev, A. V. (2015, March). Speed of Conformational Change: Comparing Explicit and Implicit Solvent Molecular Dynamics Simulations. *Biophysical Journal*. http://doi.org/10.1016/j.bpj.2014.12.047
Andersen, H. C. (1983). Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, *52*(1), 24–34. http://doi.org/http://dx.doi.org/10.1016/0021-9991(83)90014-1
Andrew, R. L. (1996). *Molecular modeling principles and applications*. *Longman*.
Angelo, K., Jespersen, T., Grunnet, M., Nielsen, M. S., Klaerke, D. A., & Olesen, S.-P. (2002). KCNE5 induces time- and voltage-dependent modulation of the KCNQ1 current. *Biophysical Journal*, *83*(4), 1997–2006. http://doi.org/10.1016/S0006-3495(02)73961-1
Antikainen, N. M., & Martin, S. F. (2005). Altering protein specificity: techniques and applications. *Bioorganic & Medicinal Chemistry*, *13*(8), 2701–2716. Journal Article, Research Support, U.S. Gov't, P.H.S., Review. http://doi.org/10.1016/j.bmc.2005.01.059
Anwar-Mohamed, A., Barakat, K. H., Bhat, R., Noskov, S. Y., Tyrrell, D. L., Tuszynski, J. a., & Houghton, M. (2014). A human ether-á-go-go-related (hERG) ion channel atomistic model generated by long supercomputer molecular dynamics simulations and its use in predicting drug cardiotoxicity. *Toxicology Letters*, *230*(3), 382–392. http://doi.org/10.1016/j.toxlet.2014.08.007
Aqvist, J., & Luzhkov, V. (2000). Ion permeation mechanism of the potassium channel. *Nature*, *404*(6780), 881–884. http://doi.org/10.1038/35009114
Arafat, A. S. Y., Arun, A., Ilamathi, M., Asha, J., Sivashankari, P. R., D'Souza, C. J. M., … Dhananjaya, B. L. (2014). Homology modeling, molecular dynamics and atomic level interaction study of snake venom 5' nucleotidase. *Journal of Molecular Modeling*, *20*(3), 2156. http://doi.org/10.1007/s00894-014-2156-1
Arévalo, J. C. (2015). Nedd4-2 regulation of voltage-gated ion channels: An update on structure???function relationships and the pathophysiological consequences of dysfunction. *Journal of Receptor, Ligand and Channel Research*, *8*, 53–63. http://doi.org/10.2147/JRLCR.S52534
Ashcroft, F. M. (2000). Chapter 3 – How ion channels work. In *Ion Channels and Disease* (pp. 21–41). http://doi.org/10.1016/B978-012065310-2/50004-7
Ashcroft, F. M., & Ashcroft, F. M. (2000a). Chapter 10 – Voltage-gated Cl− channels. In *Ion Channels and Disease* (pp. 185–198). http://doi.org/10.1016/B978-012065310-2/50011-4
Ashcroft, F. M., & Ashcroft, F. M. (2000b). Chapter 1 – Introduction. In *Ion Channels and Disease* (pp. 1–2). http://doi.org/10.1016/B978-012065310-2/50002-3
Ashcroft, F. M., & Ashcroft, F. M. (2000c). Chapter 2 – From gene to protein. In *Ion Channels and Disease* (pp. 3–19). http://doi.org/10.1016/B978-012065310-2/50003-5
Ashcroft, F. M., & Ashcroft, F. M. (2000d). Chapter 5 – Voltage-gated Na+ channels. In *Ion Channels and Disease* (pp. 67–96). http://doi.org/10.1016/B978-012065310-2/50006-0
Ashcroft, F. M., & Ashcroft, F. M. (2000e). Chapter 6 – Voltage-gated K+ channels. In *Ion Channels and Disease* (p. 97–II). http://doi.org/10.1016/B978-012065310-2/50007-2
Ashcroft, F. M., & Ashcroft, F. M. (2000f). Chapter 8 – Inwardly rectifying K+ channels. In *Ion Channels and Disease* (pp. 135–159). http://doi.org/10.1016/B978-012065310-2/50009-6
Aszodi, A., & Taylor, W. R. (1994). Secondary structure formation in model polypeptide chains. *Protein Eng*, *7*(5), 633–644. http://doi.org/10.1093/protein/7.5.633
Aszodi, A., & Taylor, W. R. (1996). Homology modelling by distance geometry. *Folding & Design*, *1*(5), 325–334. http://doi.org/10.1016/S1359-0278(96)00048-X
Bagal, S. K., Marron, B. E., Owen, R. M., Storer, R. I., & Swain, N. A. (2015). Voltage gated sodium channels as drug discovery targets. *Channels*, *9*(6), 360–366. http://doi.org/10.1080/19336950.2015.1079674
Bairoch, A., & Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, *28*(1), 45–48.
Bajaj, C. L., Chowdhury, R., & Siddahanavalli, V. (2011). F2dock: Fast fourier protein-protein docking. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(1), 45–58. http://doi.org/10.1109/TCBB.2009.57

Bakk, A., Høye, J. S., & Hansen, A. (2001). Heat Capacity of Protein Folding. *Biophysical Journal*, *81*(2), 710–714. http://doi.org/http://dx.doi.org/10.1016/S0006-3495(01)75735-9

Banks, J. L., Beard, H. S., Cao, Y., Cho, A. E., Damm, W., Farid, R., … Levy, R. M. (2005). Integrated Modeling Program, Applied Chemical Theory (IMPACT). *Journal of Computational Chemistry*, *26*(16), 1752–1780. http://doi.org/10.1002/jcc.20292

Barakat, K. (2015). Modelling Off-target Interactions (I): Cardiotoxicity. *Journal of Pharmaceutical Care & Health Systems*, *2*(3), 10–11. http://doi.org/10.4172/2376-0419.1000e131

Barakat, K., Gajewski, M., & Tuszynski, J. (2012). DNA Repair Inhibitors: Our Last Disposal to Improve Cancer Therapy. *Current Topics in Medicinal Chemistry*.

Barakat, K. H., Mane, J. Y., & Tuszynski, J. A. (2011). Virtual Screening : An Overview on Methods and Applications. In *L. Liu, D. Wei, Y. Li, & H. Lei (Eds.) Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications* (pp. 28–60). Hershey, PA: Medical Information Science Reference. http://doi.org/10.4018/978-1-60960-491-2.ch002

Barakat, K., Issack, B. B., Stepanova, M., & Tuszynski, J. (2011). Effects of temperature on the p53-DNA binding interactions and their dynamical behavior: comparing the wild type to the R248Q mutant. *PloS One*, *6*(11), e27651.

Barakat, K., Mane, J., Friesen, D., & Tuszynski, J. (2010). Ensemble-based virtual screening reveals dual-inhibitors for the p53-MDM2/MDMX interactions. *Journal of Molecular Graphics and Modelling*, *28*(6), 555–568. http://doi.org/10.1016/j.jmgm.2009.12.003

Barhanin, J., Lesage, F., Guillemare, E., Fink, M., Lazdunski, M., & Romey, G. (1996). K(V)LQT1 and lsK (minK) proteins associate to form the I(Ks) cardiac potassium current. *Nature*, *384*(6604), 78–80. http://doi.org/10.1038/384078a0

Barro-Soria, R., Rebolledo, S., Liin, S. I., Perez, M. E., Sampson, K. J., Kass, R. S., & Larsson, H. P. (2014). KCNE1 divides the voltage sensor movement in KCNQ1/KCNE1 channels into two steps. *Nature Communications*, *5*, 3750. http://doi.org/10.1038/ncomms4750

Bates, P. A., Kelley, L. A., MacCallum, R. M., & Sternberg, M. J. (2001). Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins*, *Suppl 5*, 39–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11835480

Bean, B. P. (2007). The action potential in mammalian central neurons. *Nat Rev Neurosci*, *8*(6), 451–465. Retrieved from http://dx.doi.org/10.1038/nrn2148

Beckstein, O., Biggin, P. C., Bond, P., Bright, J. N., Domene, C., Grottesi, A., … Sansom, M. S. P. (2003). Ion channel gating: Insights via molecular simulations. *FEBS Letters*, *555*(1), 85–90. http://doi.org/10.1016/S0014-5793(03)01151-7

Belardetti, F., & Zamponi, G. W. (2012). Calcium channels as therapeutic targets. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, *1*(4), 433–451. http://doi.org/10.1002/wmts.38

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2005). GenBank. *Nucleic Acids Research*, *33*(Database issue), D34–D38. http://doi.org/10.1093/nar/gki063

Bento, A. P., Gaulton, A., Hersey, A., Bellis, L. J., Chambers, J., Davies, M., … Overington, J. P. (2014). The ChEMBL bioactivity database: an update. *Nucleic Acids Research*, *42*(D1), D1083–D1090. Retrieved from http://dx.doi.org/10.1093/nar/gkt1031

Berendsen, H. J. C., Grigera, J. R., & Straatsma, T. P. (1987). The Missing Term in Effective Pair Potentials. *Journal of Physical Chemistry*, *91*(24), 6269–6271. http://doi.org/10.1021/j100308a038

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, a, & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(1984), 3684–3690. http://doi.org/10.1063/1.448118

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981). Interaction Models for Water in Relation to Protein Hydration. In B. Pullman (Ed.), *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13--16, 1981* (pp. 331–342). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-015-7658-1_21

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., … Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. http://doi.org/10.1093/nar/28.1.235

Bernardi, R. C., Melo, M. C. R., & Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1850*(5), 872–877. http://doi.org/http://doi.org/10.1016/j.bbagen.2014.10.019

Bers, D. M. (2002). Cardiac excitation-contraction coupling. *Nature*, *415*(6868), 198–205. Retrieved from

http://dx.doi.org/10.1038/415198a
Betts, M. J., & Sternberg, M. J. (1999). An analysis of conformational changes on protein-protein association: implications for predictive docking. *Protein Engineering*, *12*(4), 271–283. http://doi.org/10.1093/protein/12.4.271
Bezanilla, F. (2007). Voltage-Gated Ion Channels. In S.-H. Chung, O. S. Andersen, & V. Krishnamurthy (Eds.), *Biological Membrane Ion Channels: Dynamics, Structure, and Applications* (pp. 81–118). New York, NY: Springer New York. http://doi.org/10.1007/0-387-68919-2_3
Bidaud, I., Mezghrani, A., Swayne, L. A., Monteil, A., & Lory, P. (2006). Voltage-gated calcium channels in genetic diseases. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, *1763*(11), 1169–1174. http://doi.org/https://doi.org/10.1016/j.bbamcr.2006.08.049
Bleich, M., Briel, M., Busch, A. E., Lang, H. J., Gerlach, U., Gogelein, H., … Kunzelmann, K. (1997). KVLQT channels are inhibited by the K+ channel blocker 293B. *Pflugers Archiv : European Journal of Physiology*, *434*(4), 499–501.
Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E., & Thornton, J. M. (1987). Knowledge-based prediction of protein structures and the design of novel molecules. *Nature*, *326*(6111), 347–352. Retrieved from http://dx.doi.org/10.1038/326347a0
Boehmer, C., Laufer, J., Jeyaraj, S., Klaus, F., Lindner, R., Lang, F., & Palmada, M. (2008). Modulation of the voltage-gated potassium channel Kv1.5 by the SGK1 protein kinase involves inhibition of channel ubiquitination. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *22*(5–6), 591–600. http://doi.org/10.1159/000185543
Boland, L. M., & Jackson, K. A. (1999). Protein kinase C inhibits Kv1.1 potassium channel function. *The American Journal of Physiology*, *277*(1 Pt 1), C100-10.
Bond, P. J., & Sansom, M. S. P. (2003). Membrane protein dynamics versus environment: simulations of OmpA in a micelle and in a bilayer. *Journal of Molecular Biology*, *329*(5), 1035–1053.
Bonvin, A. (2013). Coming to peace with protein complexes? 5th CAPRI evaluation meeting, April 17–19th 2013 – Utrecht. *Proteins: Structure, Function, and Bioinformatics*, *81*(12), 2073–2074. http://doi.org/10.1002/prot.24431
Boulet, I. R., Labro, A. J., Raes, A. L., & Snyders, D. J. (2007). Role of the S6 C-terminus in KCNQ1 channel gating. *The Journal of Physiology*, *585*(Pt 2), 325–337. http://doi.org/10.1113/jphysiol.2007.145813
Bowie, J. U., Luthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science (New York, N.Y.)*, *253*(5016), 164–170.
Boyd, D. B., & Lipkowitz, K. B. (2001). *Reviews in Computational Chemistry* (Vol. 17). New York: Wiley-VCH.
Brackenbury, W., & Isom, L. (2011). Na+ Channel β Subunits: Overachievers of the Ion Channel Family . *Frontiers in Pharmacology* . Retrieved from http://journal.frontiersin.org/article/10.3389/fphar.2011.00053
Brannigan, J. a, & Wilkinson, A. J. (2002). Protein engineering 20 years on. *Nature Reviews. Molecular Cell Biology*, *3*(12), 964–70. http://doi.org/10.1038/nrm975
Braun, M., Ramracheya, R., Bengtsson, M., Zhang, Q., Karanauskaite, J., Partridge, C., … Rorsman, P. (2008). Voltage-gated ion channels in human pancreatic beta-cells: electrophysiological characterization and role in insulin secretion. *Diabetes*, *57*(6), 1618–1628. http://doi.org/10.2337/db07-0991
Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*(2), 187–217. http://doi.org/10.1002/jcc.540040211
Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, *41*(Web Server issue), W349-57. http://doi.org/10.1093/nar/gkt381
Bujnicki, J. M., Elofsson, A., Fischer, D., & Rychlewski, L. (2001). LiveBench-2: large-scale automated evaluation of protein structure prediction servers. *Proteins*, *Suppl 5*, 184–91. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11835496
Buraei, Z., & Yang, J. (2010). The ss subunit of voltage-gated Ca2+ channels. *Physiological Reviews*, *90*(4), 1461–1506. http://doi.org/10.1152/physrev.00057.2009
Busch, A. E., Suessbrich, H., Waldegger, S., Sailer, E., Greger, R., Lang, H.-J., … Maylie, J. G. (1996).

Inhibition of IKs in guinea pig cardiac myocytes and guinea pig IsK channels by the chromanol 293B. *Pflügers Archiv - European Journal of Physiology*, *432*(6), 1094–1096. http://doi.org/10.1007/s004240050240

Buxbaum, E. (2007). Cell skeleton. In *Fundamentals of Protein Structure and Function* (pp. 175–184). Boston, MA: Springer US. http://doi.org/10.1007/978-0-387-68480-2_11

Camacho, C. J., & Vajda, S. (2002). Protein-protein association kinetics and protein docking. *Current Opinion in Structural Biology*, *12*(1), 36–40. http://doi.org/10.1016/S0959-440X(02)00286-5

Campbell, I. D. (2002, May). Timeline: the march of structural biology. *Nature Reviews. Molecular Cell Biology*. England. http://doi.org/10.1038/nrm800

Campuzano, O., Beltran-Alvarez, P., Iglesias, A., Scornik, F., Perez, G., & Brugada, R. (2010). Genetics and cardiac channelopathies. *Genet Med*, *12*(5), 260–267. Retrieved from http://dx.doi.org/10.1097/GIM.0b013e3181d81636

Canutescu, A. A., Shelenkov, A. A., & Dunbrack, R. L. J. (2003). A graph-theory algorithm for rapid protein side-chain prediction. *Protein Science : A Publication of the Protein Society*, *12*(9), 2001–2014. http://doi.org/10.1110/ps.03154503

Carpenter, E. P., Beis, K., Cameron, A. D., & Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*, *18*(5), 581–586. http://doi.org/10.1016/j.sbi.2008.07.001

Cascio, M. (2004). Structure and function of the glycine receptor and related nicotinicoid receptors. *The Journal of Biological Chemistry*, *279*(19), 19383–19386. http://doi.org/10.1074/jbc.R300035200

Catterall, W. (2015). Voltage-Gated Sodium Channels. In *Handbook of Ion Channels* (pp. 213–226). CRC Press. http://doi.org/doi:10.1201/b18027-20

Catterall, W. a. (2012). Voltage-gated sodium channels at 60: structure, function and pathophysiology. *The Journal of Physiology*, *590*(11), 2577–89. http://doi.org/10.1113/jphysiol.2011.224204

Catterall, W. A. (2005). Voltage-Gated Calcium Channels. *Molecular Biology Intelligence Unit*, 1–373. http://doi.org/10.1016/B978-0-12-374145-5.00112-1

Cavasotto, C. N., & Phatak, S. S. (2009). Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today*, *14*(13–14), 676–83. http://doi.org/10.1016/j.drudis.2009.04.006

Ceccarini, L., Masetti, M., Cavalli, A., & Recanatini, M. (2012). Ion Conduction through the hERG Potassium Channel. *PLOS ONE*, *7*(11), e49017. Retrieved from https://doi.org/10.1371/journal.pone.0049017

Chan, P. J. (2011). Functional and Biochemical Characterization of KCNQ1/KCNE1 Subunit Interactions in the Cardiac I.

Chen, C. Y.-C., & Tou, W. leong. (2013). How to design a drug for the disordered proteins? *Drug Discovery Today*, *18*(19–20), 910–915. http://doi.org/10.1016/j.drudis.2013.04.008

Chen, J., Zheng, R., Melman, Y. F., & McDonald, T. V. (2009). Functional interactions between KCNE1 C-terminus and the KCNQ1 channel. *PLoS ONE*, *4*(4), 1–9. http://doi.org/10.1371/journal.pone.0005143

Chen, R., Li, L., & Weng, Z. (2003). ZDOCK : An Initial-Stage Protein-Docking Algorithm. *Proteins*, *52*(November 2002), 80–87.

Chen, R., & Weng, Z. (2002). Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins: Structure, Function and Genetics*, *47*(August 1999), 281–294. http://doi.org/10.1002/prot.10092

Chen, Y. C. (2015). Beware of docking! *Trends in Pharmacological Sciences*, *36*(2), 78–95. http://doi.org/10.1016/j.tips.2014.12.001

Chène, P. (2003). Inhibiting the p53–MDM2 interaction: an important target for cancer therapy. *Nat Rev Cancer*, *3*(2), 102–109. Retrieved from http://dx.doi.org/10.1038/nrc991

Cheng, T. M.-K., Blundell, T. L., & Fernandez-Recio, J. (2007). pyDock: Electrostatics and Desolvation for Effective Scoring of Rigid-Body Protein–Protein Docking. *Proteins*, *68*(2), 503–515.

Chowdhury, R., Rasheed, M., Keidel, D., Moussalem, M., Olson, A., Sanner, M., & Bajaj, C. (2013). Protein-Protein Docking with F2Dock 2.0 and GB-Rerank. *PLoS ONE*, *8*(3). http://doi.org/10.1371/journal.pone.0051307

Chowdhury, S., & Chanda, B. (2015). Basic mechanisms of voltage sensing. *Handbook of Ion Channels*, 25–40. http://doi.org/doi:10.1201/b18027-5

Chuang, G.-Y., Kozakov, D., Brenke, R., Comeau, S. R., & Vajda, S. (2008). DARS (Decoys As the Reference State) potentials for protein-protein docking. *Biophysical Journal*, *95*(9), 4217–4227.

http://doi.org/10.1529/biophysj.108.135814
Chung, D. Y., Chan, P. J., Bankston, J. R., Yang, L., Liu, G., Marx, S. O., … Kass, R. S. (2009). Location of KCNE1 relative to KCNQ1 in the IKS potassium channel by disulfide cross-linking of substituted cysteines. *Proceedings of the National Academy of Sciences* , *106*(3), 743–748. Retrieved from http://www.pnas.org/content/106/3/743.abstract
Clapham, D. E. (2007). Calcium Signaling. *Cell*, *131*(6), 1047–1058. http://doi.org/http://dx.doi.org/10.1016/j.cell.2007.11.028
Cole, J. C., Murray, C. W., Nissink, J. W. M., Taylor, R. D., & Taylor, R. (2005). Comparing protein-ligand docking programs is difficult. *Proteins: Structure, Function, and Bioinformatics*, *60*(3), 325–332. http://doi.org/10.1002/prot.20497
Coleman, J. E. (1992). Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annual Review of Biochemistry*, *61*, 897–946. http://doi.org/10.1146/annurev.biochem.61.1.897
Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science : A Publication of the Protein Society*, *2*(9), 1511–1519. http://doi.org/10.1002/pro.5560020916
Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). ClusPro: An automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, *20*(1), 45–50. http://doi.org/10.1093/bioinformatics/btg371
Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., … Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, *117*(19), 5179–5197. http://doi.org/10.1021/ja00124a002
Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*, *16*(22), 10881–10890.
Corry, B. (2015). Long Timescale Molecular Simulations for Understanding Ion Channel Function. In *Pumps, Channels, and Transporters* (pp. 411–441). John Wiley & Sons, Inc. http://doi.org/10.1002/9781119085126.ch15
Cuendet, M. A., & van Gunsteren, W. F. (2007). On the calculation of velocity-dependent properties in molecular dynamics simulations using the leapfrog integration algorithm. *The Journal of Chemical Physics*, *127*(18), 184102. http://doi.org/10.1063/1.2779878
Cui, J. (2016). Voltage-Dependent Gating: Novel Insights from KCNQ1 Channels. *Biophysical Journal*, *110*(1), 14–25. http://doi.org/10.1016/j.bpj.2015.11.023
D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, D. M. Y. and P. A. K. (2015). AMBER 12. *University of California, San Francisco*.
Dalton, J. A. R., & Jackson, R. M. (2007). An evaluation of automated homology modelling methods at low target-template sequence similarity. *Bioinformatics*, *23*(15), 1901–1908. http://doi.org/10.1093/bioinformatics/btm262
Davies, D. L., & Bouldin, D. W. (1979). A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *1*(2), 224–227.
Deane, C. M., & Blundell, T. L. (2001). CODA: a combined algorithm for predicting the structurally variable regions of protein models. *Protein Science : A Publication of the Protein Society*, *10*(3), 599–612. http://doi.org/10.1110/ps.37601
Deane, C. M., & Blundell, T. L. (2003). 27 - Protein Comparative Modelling and Drug Discovery. In C. G. B. T.-T. P. of M. C. (Second E. Wermuth (Ed.), (pp. 445–458). London: Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-012744481-9/50031-3
De Vries, S. J., Schindler, C. E. M., Chauvot de Beauchêne, I., & Zacharias, M. (2015). A Web Interface for Easy Flexible Protein-Protein Docking with ATTRACT. *Biophysical Journal*, *108*(3), 462–465. http://doi.org/10.1016/j.bpj.2014.12.015
De Vries, S. J., Van Dijk, M., & Bonvin, A. M. J. J. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*, *5*(5), 883–897. http://doi.org/10.1038/nprot.2010.32
De Vries, S., & Zacharias, M. (2013). Flexible docking and refinement with a coarse-grained protein model using ATTRACT. *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2167–2174. http://doi.org/10.1002/prot.24400

Demolombe, S., Baro, I., Pereon, Y., Bliek, J., Mohammad-Panah, R., Pollard, H., … Escande, D. (1998). A dominant negative isoform of the long QT syndrome 1 gene product. *The Journal of Biological Chemistry*, *273*(12), 6837–6843.

Dhanavade, M. J., Jalkute, C. B., Barage, S. H., & Sonawane, K. D. (2013). Homology modeling, molecular docking and MD simulation studies to investigate role of cysteine protease from Xanthomonas campestris in degradation of Aβ peptide. *Computers in Biology and Medicine*, *43*(12), 2063–2070. http://doi.org/http://dx.doi.org/10.1016/j.compbiomed.2013.09.021

Dib-Hajj, S., & Priestley, T. (2010). Chapter 2.3: Voltage-gated sodium channels. In J. N. C. Kew & C. H. Davies (Eds.), *Ion channels: from structure to function*. New York: Oxford University Press.

Dolphin, A. C. (2009). Calcium channel diversity: multiple roles of calcium channel subunits. *Current Opinion in Neurobiology*, *19*(3), 237–244. http://doi.org/http://dx.doi.org/10.1016/j.conb.2009.06.006

Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–1737. http://doi.org/10.1021/ja026939x

Doyle, D. A., Cabral, J. M., Pfuetzner, R. A., Kuo, A., Gulbis, J. M., Cohen, S. L., … MacKinnon, R. (1998). The Structure of the Potassium Channel: Molecular Basis of K+ Conduction and Selectivity. *Science*, *280*(5360), 69–77. http://doi.org/10.1126/science.280.5360.69

Drabik, P., Liwo, A., Czaplewski, C., & Ciarkowski, J. (2001). The investigation of the effects of counterions in protein dynamics simulations. *Protein Engineering, Design and Selection*, *14*(10), 747. http://doi.org/10.1093/protein/14.10.747

Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., & Shaw, D. E. (2012). Biomolecular simulation: a computational microscope for molecular biology. *Annual Review of Biophysics*, *41*, 429–452. http://doi.org/10.1146/annurev-biophys-042910-155245

Du, L. P., Li, M. Y., Tsai, K. C., You, Q. D., & Xia, L. (2005). Characterization of binding site of closed-state KCNQ1 potassium channel by homology modeling, molecular docking, and pharmacophore identification. *Biochemical and Biophysical Research Communications*, *332*(3), 677–687. http://doi.org/10.1016/j.bbrc.2005.04.165

Duhovny, D., Nussinov, R., & Wolfson, H. (2002). Efficient Unbound Docking of Rigid Molecules. *Algorithms in Bioinformatics*, 185–200. http://doi.org/10.1007/3-540-45784-4_14

Dunbrack, R. L. J., & Karplus, M. (1994). Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nature Structural Biology*, *1*(5), 334–340.

Durrant, J. D., & McCammon, J. A. (2011). NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *Journal of Chemical Information and Modeling*, *51*(11), 2897–2903. http://doi.org/10.1021/ci2003889

Dvir, M., Peretz, A., Haitin, Y., & Attali, B. (2014). Recent molecular insights from mutated IKS channels in cardiac arrhythmia. *Current Opinion in Pharmacology*, *15*, 74–82. http://doi.org/http://dx.doi.org/10.1016/j.coph.2013.12.004

Dworakowska, B., & Dolowy, K. (2000). Ion channels-related diseases. *Acta Biochimica Polonica*, *47*(3), 685–703.

Earl, D. J., & Deem, M. W. (2005). Parallel Tempering: Theory, Applications, and New Perspectives. http://doi.org/10.1039/B509983H

Eckey, K., Wrobel, E., Strutz-Seebohm, N., Pott, L., Schmitt, N., & Seebohm, G. (2014). Novel Kv7.1-Phosphatidylinositol 4,5-Bisphosphate Interaction Sites Uncovered by Charge Neutralization Scanning. *Journal of Biological Chemistry*, *289*(33), 22749–22758. http://doi.org/10.1074/jbc.M114.589796

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics (Oxford, England)*, *14*(9), 755–763.

Ehrlich, L. P., & Wade, R. C. (2001). Protein–Protein Docking. In *Reviews in Computational Chemistry* (17th ed., pp. 61–92). New York: Wiley-VCH.

Eisenberg, D., Luthy, R., & Bowie, J. U. (1997). VERIFY3D: assessment of protein models with three-dimensional profiles. *Methods in Enzymology*, *277*, 396–404.

Elcock, A. A. H., Sept, D., & McCammon, J. A. (2001). Computer simulation of protein-protein interactions. *The Journal of Physical Chemistry B*, *105*(8), 1504–1518. http://doi.org/10.1021/jp003602d

Eswar, N. (2003). Tools for comparative protein structure modeling and analysis. *Nucleic Acids Research*, *31*(13), 3375–3380. http://doi.org/10.1093/nar/gkg543

Evers, A., & Klabunde, T. (2005). Structure-based Drug Discovery Using GPCR Homology Modeling:

Successful Virtual Screening for Antagonists of the Alpha1A Adrenergic Receptor. *Journal of Medicinal Chemistry*, *48*(4), 1088–1097. http://doi.org/10.1021/jm0491804

Eyrich, V. A., Marti-Renom, M. A., Przybylski, D., Madhusudhan, M. S., Fiser, A., Pazos, F., … Rost, B. (2001). EVA: continuous automatic evaluation of protein structure prediction servers. *Bioinformatics (Oxford, England)*, *17*(12), 1242–1243.

Fallis, A. . (2013). *Fundamentals of Protein Structure and Function*. *Journal of Chemical Information and Modeling* (Vol. 53). http://doi.org/10.1017/CBO9781107415324.004

Fan, H. (2004). Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Science*, *13*(1), 211–220. http://doi.org/10.1110/ps.03381404

Feranchak, A. P. (2003). Ion Channels in Digestive Health and Disease. *Journal of Pediatric Gastroenterology and Nutrition*, *37*(3), 230–241. http://doi.org/10.1097/00005176-200309000-00006

Fernandez-Recio, J., Totrov, M., & Abagyan, R. (2002). Soft protein – protein docking in internal coordinates. *Protein Science*, *11*(2), 280–291. http://doi.org/10.1110/ps.19202.ical

Ferreira, P. M. (1995). Protein structure: By N J Darby and T E Creighton. pp 97. IRL Press, Oxford University Press. 1993. SBN 0-19-963310-X. *Biochemical Education*, *23*(1), 46. http://doi.org/10.1016/0307-4412(95)90200-7

Fiser, A., Do, R., & Sali, A. (2000). Modeling of loops in protein structures. *PRS*, *9*(9), 1753–1773.

Fiser, a, Fiser, a, Do, R. K., Do, R. K., Sali, a, & Sali, a. (2000). Modeling of loops in protein structures. *Protein Science : A Publication of the Protein Society*, *9*(9), 1753–73. http://doi.org/10.1110/ps.9.9.1753

Fisher, C. K., Huang, A., & Stultz, C. M. (2010). Modeling intrinsically disordered proteins with Bayesian statistics. *Journal of the American Chemical Society*, *132*(42), 14919–14927. http://doi.org/10.1021/ja105832g

FLORESCU, M., CINTEZA, M., & VINEREANU, D. (2013). Chemotherapy-induced Cardiotoxicity. *Mædica*, *8*(1), 59–67. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3749765/

Fowler, P. W., & Sansom, M. S. P. (2013). The pore of voltage-gated potassium ion channels is strained when closed. *Nature Communications*, *4*(May), 1872. http://doi.org/10.1038/ncomms2858

Francis-Lyon, P., & Koehl, P. (2014). Protein side-chain modeling with a protein-dependent optimized rotamer library. *Proteins: Structure, Function and Bioinformatics*, *82*(9), 2000–2017. http://doi.org/10.1002/prot.24555

Frishman, D., & Argos, P. (1995). Knowledge-based protein secondary structure assignment. *Proteins: Structure, Function and Genetics*. http://doi.org/10.1002/prot.340230412

Gaalswyk, K., & Rowley, C. N. (2016). An explicit-solvent conformation search method using open software. *PeerJ*, *4*, e2088. http://doi.org/10.7717/peerj.2088

Gabashvili, I. S., Sokolowski, B. H. A., Morton, C. C., & Giersch, A. B. S. (2007, September). Ion Channel Gene Expression in the Inner Ear. *JARO: Journal of the Association for Research in Otolaryngology*. New York. http://doi.org/10.1007/s10162-007-0082-y

Gabb, H. a, Jackson, R. M., & Sternberg, M. J. (1997). Modelling protein docking using shape complementarity, electrostatics and biochemical information. *Journal of Molecular Biology*, *272*(1), 106–120. http://doi.org/10.1006/jmbi.1997.1203

Gagneux, P. (2004). Protein Structure and Function. *Journal of Heredity*, *95*(3), 274–274. http://doi.org/10.1093/jhered/esh040

Gagnidze, K., Sachchidanand, Rozenfeld, R., Mezei, M., Zhou, M.-M., & Devi, L. A. (2008). Homology Modeling and Site-directed Mutagenesis to Identify Selective Inhibitors of Endothelin-Converting Enzyme-2. *Journal of Medicinal Chemistry*, *51*(12), 3378–3387. http://doi.org/10.1021/jm7015478

Gajda, M. J., Tuszynska, I., Kaczor, M., Bakulina, A. Y., & Bujnicki, J. M. (2010). FILTREST3D: Discrimination of structural models using restraints from experimental data. *Bioinformatics*, *26*(23), 2986–2987. http://doi.org/10.1093/bioinformatics/btq582

Galligan, J. J. (2002). Ligand-gated ion channels in the enteric nervous system. *Neurogastroenterology and Motility : The Official Journal of the European Gastrointestinal Motility Society*, *14*(6), 611–623.

Ganesan, A., Coote, M. L., & Barakat, K. (2017a). Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discovery Today*, *22*(2), 249–269. http://doi.org/http://dx.doi.org/10.1016/j.drudis.2016.11.001

Ganesan, A., Coote, M. L., & Barakat, K. (2017b). Molecular “time-machines” to unravel key biological events for drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *7*(4), e1306–n/a. http://doi.org/10.1002/wcms.1306

Gao, N., Lu, M., Echeverri, F., Laita, B., Kalabat, D., Williams, M. E., … Moyer, B. D. (2009). Voltage-gated sodium channels in taste bud cells. *BMC Neuroscience*, *10*(1), 20. http://doi.org/10.1186/1471-2202-10-20

Gayen, S., Li, Q., & Kang, C. (2015). Structural analysis of the S4–S5 linker of the human KCNQ1 potassium channel. *Biochemical and Biophysical Research Communications*, *456*(1), 410–414. http://doi.org/10.1016/j.bbrc.2014.11.097

Gerlach, U., Brendel, J., Lang, H.-J., Paulus, E. F., Weidmann, K., Brüggemann, A., … Greger, R. (2001). Synthesis and Activity of Novel and Selective I Ks -Channel Blockers. *Journal of Medicinal Chemistry*, *44*(23), 3831–3837. http://doi.org/10.1021/jm0109255

Gerstein, M. (1997). A structural census of genomes: comparing bacterial, eukaryotic, and archaeal genomes in terms of protein structure. *Journal of Molecular Biology*, *274*(4), 562–576. http://doi.org/10.1006/jmbi.1997.1412

Gez, L. S., Hagalili, Y., Shainberg, A., & Atlas, D. (2012). Voltage-driven Ca(2+) binding at the L-type Ca(2+) channel triggers cardiac excitation-contraction coupling prior to Ca(2+) influx. *Biochemistry*, *51*(48), 9658–9666. http://doi.org/10.1021/bi301124a

Ghosh, E., Kumari, P., Jaiman, D., & Shukla, A. K. (2015). Methodological advances: the unsung heroes of the GPCR structural revolution. *Nat Rev Mol Cell Biol*, *16*(2), 69–81. Retrieved from http://dx.doi.org/10.1038/nrm3933

Ghosh, S., Nunziato, D. a., & Pitt, G. S. (2006). KCNQ1 assembly and function is blocked by long-QT syndrome mutations that disrupt interaction with calmodulin. *Circulation Research*, *98*(8), 1048–1054. http://doi.org/10.1161/01.RES.0000218863.44140.f2

Ginalski, K. (2006). Comparative modeling for protein structure prediction. *Current Opinion in Structural Biology*, *16*(2), 172–177. http://doi.org/10.1016/j.sbi.2006.02.003

Ginalski, K., Elofsson, A., Fischer, D., & Rychlewski, L. (2003). 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics (Oxford, England)*, *19*(8), 1015–1018.

Gnanakaran, S., Hochstrasser, R. M., & García, A. E. (2004). Nature of structural inhomogeneities on folding a helix and their influence on spectral measurements. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(25), 9229–9234. http://doi.org/10.1073/pnas.0402933101

Godzik, A., Kolinski, A., & Skolnick, J. (1992). Topology fingerprint approach to the inverse protein folding problem. *Journal of Molecular Biology*, *227*(1), 227–238. http://doi.org/http://dx.doi.org/10.1016/0022-2836(92)90693-E

Gofman, Y., Shats, S., Attali, B., Haliloglu, T., & Ben-Tal, N. (2012). How does KCNE1 regulate the Kv7.1 potassium channel? Model-structure, mutations, and dynamics of the Kv7.1-KCNE1 complex. *Structure*, *20*(8), 1343–1352. http://doi.org/10.1016/j.str.2012.05.016

Gógl, G., Schneider, K. D., Yeh, B. J., Alam, N., Nguyen Ba, A. N., Moses, A. M., … Weiss, E. L. (2015). The Structure of an NDR/LATS Kinase–Mob Complex Reveals a Novel Kinase–Coactivator System and Substrate Docking Mechanism. *PLoS Biology*, *13*(5), e1002146. http://doi.org/10.1371/journal.pbio.1002146

Gonzalez, B. S., Noya, E. G., Vega, C., & Sese, L. M. (2010). Nuclear Quantum Effects in Water Clusters: The Role of the Molecular Flexibility. *J. Phys. Chem. B*, *114*(7), 2484–2492. http://doi.org/10.1021/jp910770y

González, M. A. (2011). Force fields and molecular dynamics simulations. *Collection SFN*, *12*, 169–200. http://doi.org/10.1051/sfn/201112009

Grant, A. O. (2009). Cardiac ion channels. *Circulation: Arrhythmia and Electrophysiology*, *2*(2), 185–194. http://doi.org/10.1161/CIRCEP.108.789081

Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. a., & Baker, D. (2003). Protein–Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. *Journal of Molecular Biology*, *331*(1), 281–299. http://doi.org/10.1016/S0022-2836(03)00670-3

Gribskov, M., McLachlan, A. D., & Eisenberg, D. (1987). Profile analysis: detection of distantly related proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(13), 4355–4358.

Grinter, S. Z., & Zou, X. (2014). Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules*, *19*(7), 10150–10176. http://doi.org/10.3390/molecules190710150

Grosdidier, S., & Fernández-Recio, J. (2012). Protein-protein docking and hot-spot prediction for drug discovery. *Current Pharmaceutical Design*, *18*(30), 4607–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22650255

Grunnet, M., Jespersen, T., Rasmussen, H. B., Ljungstrøm, T., Jorgensen, N. K., Olesen, S.-P., & Klaerke, D. A. (2002). KCNE4 is an inhibitory subunit to the KCNQ1 channel. *The Journal of Physiology*, *542*(Pt 1), 119–130. http://doi.org/10.1113/jphysiol.2002.017301

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, *18*(15), 2714–2723. http://doi.org/10.1002/elps.1150181505

Guo, F., Li, S. C., Ma, W., & Wang, L. (2013). Detecting protein conformational changes in interactions via scaling known structures. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7821 LNBI*(10), 58–74. http://doi.org/10.1007/978-3-642-37195-0_6

Gutman, G. A., Chandy, K. G., Grissmer, S., Lazdunski, M., McKinnon, D., Pardo, L. A., … Wang, X. (2005). International Union of Pharmacology. LIII. Nomenclature and molecular relationships of voltage-gated potassium channels. *Pharmacological Reviews*, *57*(4), 473–508. http://doi.org/10.1124/pr.57.4.10.1

Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., & Schwede, T. (2013). The Protein Model Portal--a comprehensive resource for protein structure and model information. *Database : The Journal of Biological Databases and Curation*, *2013*, bat031. http://doi.org/10.1093/database/bat031

Haitin, Y., & Attali, B. (2008). The C-terminus of Kv7 channels: a multifunctional module. *The Journal of Physiology*, *586*(7), 1803–1810. http://doi.org/10.1113/jphysiol.2007.149187

Haitin, Y., Wiener, R., Shaham, D., Peretz, A., Cohen, E. B.-T., Shamgar, L., … Attali, B. (2009). Intracellular domains interactions and gated motions of I(KS) potassium channel subunits. *The EMBO Journal*, *28*(14), 1994–2005. http://doi.org/10.1038/emboj.2009.157

Halling, D. B., Aracena-Parks, P., & Hamilton, S. L. (2005). Regulation of voltage-gated Ca2+ channels by calmodulin. *Science's STKE : Signal Transduction Knowledge Environment*, *2005*(315), re15. Journal Article, Review. http://doi.org/10.1126/stke.3152005re15

Han, R., Leo-Macias, A., Zerbino, D., Bastolla, U., Contreras-Moreira, B., & Ortiz, A. R. (2008). An efficient conformational sampling method for homology modeling. *Proteins*, *71*(1), 175–88. http://doi.org/10.1002/prot.21672

Hansen, S. B. (2015). Lipid agonism: The PIP2 paradigm of ligand-gated ion channels. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids*, *1851*(5), 620–628. http://doi.org/http://dx.doi.org/10.1016/j.bbalip.2015.01.011

Hardin, C., Pogorelov, T. V., & Luthey-Schulten, Z. (2002). Ab initio protein structure prediction. *Current Opinion in Structural Biology*, *12*(2), 176–181. http://doi.org/10.1016/S0959-440X(02)00306-8

Hasani, H. J., & Barakat, K. (n.d.). Protein-Protein Docking: *Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery*, 173–195. http://doi.org/10.4018/978-1-5225-0115-2.ch007

Hausdorff, S. F., Goldstein, S. A. N., Rushin, E. E., & Miller, C. (1991). Functional characterization of a minimal potassium channel expressed from a synthetic gene. *Biochemistry*, *30*(13), 3341–3346. http://doi.org/10.1021/bi00227a025

Heifetz, A., James, T., Morao, I., Bodkin, M. J., & Biggin, P. C. (2016). Guiding lead optimization with GPCR structure modeling and molecular dynamics. *Current Opinion in Pharmacology*, *30*, 14–21. http://doi.org/http://dx.doi.org/10.1016/j.coph.2016.06.004

Hernández-santoyo, A., Tenorio-barajas, A. Y., Altuzar, V., Vivanco-cid, H., & Mendoza-barrera, C. (2013). *Protein-Protein and Protein-Ligand Docking, Protein Engineering - Technology and Application*. (D. T. Ogawa, Ed.). http://doi.org/10.5772/56376

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, *18*(12), 1463–1472. http://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

Heyes, S., Pratt, W. S., Rees, E., Dahimene, S., Ferron, L., Owen, M. J., & Dolphin, A. C. (2015). Genetic disruption of voltage-gated calcium channels in psychiatric and neurological disorders. *Progress in Neurobiology*, *134*, 36–54. http://doi.org/10.1016/j.pneurobio.2015.09.002

Hildebrand, A., Remmert, M., Biegert, A., & Soding, J. (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins*, *77 Suppl 9*, 128–132. http://doi.org/10.1002/prot.22499

Hildebrand, P. W., Goede, A., Bauer, R. A., Gruening, B., Ismer, J., Michalsky, E., & Preissner, R. (2009). SuperLooper--a prediction server for the modeling of loops in globular and membrane proteins. *Nucleic Acids Research*, *37*(Web Server issue), W571-4. http://doi.org/10.1093/nar/gkp338

Hille, B. (2001). Ion Channel Excitable Membranes. *Sunderland Massachusetts USA*. http://doi.org/10.1007/3-540-29623-9_5640

Hillisch, A., Pineda, L. F., & Hilgenfeld, R. (2004). Utility of homology models in the drug discovery process. *Drug Discovery Today*, *9*(15), 659–69. http://doi.org/10.1016/S1359-6446(04)03196-4

Hodgkin, A. L., & Huxley, A. F. (1990). A quantitative description of membrane current and its application to conduction and excitation in nerve. *Bulletin of Mathematical Biology*, *52*(1–2), 25–71. http://doi.org/10.1007/BF02459568

Hofmann, F., Lacinová, L., & Klugbauer, N. (1999). Voltage-dependent calcium channels: From structure to function BT - Reviews of Physiology, Biochemistry and Pharmacology, Volume 139 (pp. 33–87). CHAP, Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/BFb0033648

Hollenhorst, M. I., Richter, K., & Fronius, M. (2011). Ion transport by pulmonary epithelia. *Journal of Biomedicine and Biotechnology*, *2011*. http://doi.org/10.1155/2011/174306

Holm, L., & Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Research* , *38*(suppl 2), W545–W549. http://doi.org/10.1093/nar/gkq366

Holm, L., & Sander, C. (1991). Database algorithm for generating protein backbone and side-chain co-ordinates from a C alpha trace application to model building and detection of co-ordinate errors. *Journal of Molecular Biology*, *218*(1), 183–194.

Hongmao, S. (2016). Chapter 4 - Homology Modeling and Ligand-Based Molecule Design. In S. B. T.-A. P. G. to R. D. D. Hongmao (Ed.), *A Practical Guide to Rational Drug Design* (pp. 109–160). Woodhead Publishing. http://doi.org/http://dx.doi.org/10.1016/B978-0-08-100098-4.00004-1

Honoré, E., Attali, B., Romey, G., Heurteaux, C., Ricard, P., Lesage, F., … Barhanin, J. (1991). Cloning, expression, pharmacology and regulation of a delayed rectifier K+ channel in mouse heart. *The EMBO Journal*, *10*(10), 2805–2811. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC452989/

Hooft, R. W., Vriend, G., Sander, C., & Abola, E. E. (1996, May). Errors in protein structures. *Nature*. ENGLAND. http://doi.org/10.1038/381272a0

Hoover. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review. A, General Physics*, *31*(3), 1695–1697.

Hoshi, T., & Armstrong, C. M. (2013). C-type inactivation of voltage-gated K+ channels: Pore constriction or dilation? *The Journal of General Physiology*, *141*(2), 151–160. http://doi.org/10.1085/jgp.201210888

Howard, R. J., Clark, K. a., Holton, J. M., & Minor, D. L. (2007). Structural Insight into KCNQ (Kv7) Channel Assembly and Channelopathy. *Neuron*, *53*(5), 663–675. http://doi.org/10.1016/j.neuron.2007.02.010

Huang, S. (2014). Search strategies and evaluation in protein – protein docking : principles , advances and challenges. *Drug Discovery Today*, *19*(8), 1081–1096. http://doi.org/10.1016/j.drudis.2014.02.005

Huang, T.-T., Hwang, J.-K., Chen, C.-H., Chu, C.-S., Lee, C.-W., & Chen, C.-C. (2015). (PS)2: protein structure prediction server version 3.0. *Nucleic Acids Research*, *454*(Ext 56921), 1–5. http://doi.org/10.1093/nar/gkv454

Hucho, F., & Weise, C. (2001). Ligand-Gated Ion Channels. *Angewandte Chemie International Edition*, *40*(17), 3100–3116. http://doi.org/10.1002/1521-3773(20010903)40:17<3100::AID-ANIE3100>3.0.CO;2-A

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38. http://doi.org/10.1016/0263-7855(96)00018-5

Ibragimova, G. T., & Wade, R. C. (1998). Importance of Explicit Salt Ions for Protein Stability in Molecular Dynamics Simulation. *Biophysical Journal*, *74*(6), 2906–2911. http://doi.org/http://dx.doi.org/10.1016/S0006-3495(98)77997-4

Isard, B. P. C. (2008). Theory and Practice in Replica-exchange Molecular Dynamics Simulation. *Biography An Interdisciplinary Quarterly*, 133. Retrieved from http://books.google.com/books/about/Theory_and_Practice_in_Replica_exchange.html?id=01OEBVckFcIC&pgis=1

Ishitani, R., Terada, T., & Shimizu, K. (2008). Refinement of comparative models of protein structure by using multicanonical molecular dynamics simulations. *Molecular Simulation*, *34*(3), 327–336.

http://doi.org/10.1080/08927020801930539
Ismail, A. M., Sharma, O. P., Kumar, M. S., Kannangai, R., & Abraham, P. (2013). Impact of rtI233V mutation in hepatitis B virus polymerase protein and adefovir efficacy: Homology modeling and molecular docking studies. *Bioinformation*, *9*(3), 121–125. http://doi.org/10.6026/97320630009121
Isom, L. L., De Jongh, K. S., & Catterall, W. A. (1994). Auxiliary subunits of voltage-gated ion channels. *Neuron*, *12*(6), 1183–1194. http://doi.org/10.1016/0896-6273(94)90436-7
Iwata, Y., Kasuya, A., & Miyamoto, S. (2002). An efficient method for reconstructing protein backbones from alpha-carbon coordinates. *Journal of Molecular Graphics & Modelling*, *21*(2), 119–128.
Jackson, R. M., Gabb, H. a, & Sternberg, M. J. (1998). Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. *Journal of Molecular Biology*, *276*(1), 265–285. http://doi.org/10.1006/jmbi.1997.1519
Jackson, W. F. (2000, January). Ion Channels and Vascular Tone. *Hypertension*.
Janin, J. (2013). The targets of CAPRI rounds 20-27. *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2075–2081. http://doi.org/10.1002/prot.24375
Janin, J., Henrick, K., Moult, J., Eyck, L. Ten, Sternberg, M. J. E., Vajda, S., … Wodak, S. J. (2003). CAPRI: A critical assessment of PRedicted interactions. *Proteins: Structure, Function and Genetics*, *52*(1), 2–9. http://doi.org/10.1002/prot.10381
Jarzynski, C. (1997). Nonequilibrium Equality for Free Energy Differences. *Physical Review Letters*, *78*(14), 2690–2693. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.78.2690
Jas, G. S., & Kuczera, K. (2004). Equilibrium Structure and Folding of a Helix-Forming Peptide: Circular Dichroism Measurements and Replica-Exchange Molecular Dynamics Simulations. *Biophysical Journal*, *87*(6), 3786–3798. JOUR. http://doi.org/10.1529/biophysj.104.045419
Jentsch, T. J. (2000). Neuronal KCNQ potassium channels: physiology and role in disease. *Nature Reviews. Neuroscience*, *1*(1), 21–30. http://doi.org/10.1038/35036198
Jentsch, T. J., Stein, V., Weinreich, F., & Zdebik, A. A. (2002). Molecular structure and physiological function of chloride channels. *Physiological Reviews*, *82*(2), 503–568. http://doi.org/10.1152/physrev.00029.2001
Jespersen, T., Grunnet, M., & Olesen, S.-P. (2005). The KCNQ1 potassium channel: from gene to physiological function. *Physiology (Bethesda, Md.)*, *20*, 408–416. http://doi.org/10.1152/physiol.00031.2005
Jiang, F., & Kim, S. H. (1991). "Soft docking": Matching of molecular surface cubes. *Journal of Molecular Biology*, *219*(1), 79–102. http://doi.org/10.1016/0022-2836(91)90859-5
Jiang, Y., Lee, A., Chen, J., Cadene, M., Chait, B. T., & MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. *Nature*, *417*(6888), 515–522. Retrieved from http://dx.doi.org/10.1038/417515a
Jiménez-García, B., Pons, C., & Fernández-Recio, J. (2013). pyDockWEB: A web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. *Bioinformatics*, *29*(13), 1698–1699. http://doi.org/10.1093/bioinformatics/btt262
Jimenez-Garcia, B., Pons, C., Svergun, D. I., Bernado, P., & Fernandez-Recio, J. (2015). pyDockSAXS: protein-protein complex structure by SAXS and computational docking. *Nucleic Acids Research*, 1–6. http://doi.org/10.1093/nar/gkv368
Johnson, M. S., & Overington, J. P. (1993). A structural basis for sequence comparisons. An evaluation of scoring methodologies. *Journal of Molecular Biology*, *233*(4), 716–738. http://doi.org/10.1006/jmbi.1993.1548
Johnson, M. S., Srinivasan, N., Sowdhamini, R., & Blundell, T. L. (1994). Knowledge-based protein modeling. *Critical Reviews in Biochemistry and Molecular Biology*, *29*(1), 1–68. http://doi.org/10.3109/10409239409086797
Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, *292*(2), 195–202. http://doi.org/10.1006/jmbi.1999.3091
Jones, D. T., Taylor, W. R., & Thornton, J. M. (1992). A new approach to protein fold recognition. *Nature*, *358*(6381), 86–89. http://doi.org/10.1038/358086a0
Jorgensen, W. L., & Madura, J. D. (1983). Solvation and conformation of methanol in water. *Journal of the American Chemical Society*, *105*(6), 1407–1413. http://doi.org/10.1021/ja00344a001
Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, *118*(45), 11225–11236. http://doi.org/10.1021/ja9621760

Jost, N., Papp, J. G., & Varro´, A. (2007). Slow Delayed Rectifier Potassium Current (I Ks) and the Repolarization Reserve. *?Ane*, *12*(36), 64–78.

Jost, N., Papp, J. G., & Varro, A. (2007). Slow delayed rectifier potassium current (IKs) and the repolarization reserve. *Annals of Noninvasive Electrocardiology : The Official Journal of the International Society for Holter and Noninvasive Electrocardiology, Inc*, *12*(1), 64–78. http://doi.org/10.1111/j.1542-474X.2007.00140.x

Kaczorowski, G. J., McManus, O. B., Priest, B. T., & Garcia, M. L. (2008). Ion Channels as Drug Targets: The Next GPCRs. *The Journal of General Physiology*, *131*(5), 399–405. http://doi.org/10.1085/jgp.200709946

Kahraman, A., Herzog, F., Leitner, A., Rosenberger, G., Aebersold, R., & Malmström, L. (2013). Cross-Link Guided Molecular Modeling with ROSETTA. *PLoS ONE*, *8*(9), e73411. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0073411

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protocols*, *7*(8), 1511–1522. Retrieved from http://dx.doi.org/10.1038/nprot.2012.085

Kang, C., Tian, C., Sönnichsen, F. D., Smith, J. A., Meiler, J., George, A. L., … Sanders, C. R. (2008). Structure of KCNE1 and Implications for How It Modulates the KCNQ1 Potassium Channel. *Biochemistry*, *47*(31), 7999–8006. http://doi.org/10.1021/bi800875q

Kann, M. G. (2007). Protein interactions and disease: Computational approaches to uncover the etiology of diseases. *Briefings in Bioinformatics*, *8*(5), 333–346. http://doi.org/10.1093/bib/bbm031

Kannan, S., & Zacharias, M. (2010). Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent. *Proteins*, *78*(13), 2809–19. http://doi.org/10.1002/prot.22796

Kannankeril, P., Roden, D. M., & Darbar, D. (2010). Drug-Induced Long QT Syndrome. *Pharmacological Reviews*, *62*(4), 760–781. http://doi.org/10.1124/pr.110.003723

Karplus, K., Barrett, C., & Hughey, R. (1998). Hidden Markov models for detecting remote protein homologies. *Bioinformatics (Oxford, England)*, *14*(10), 846–856.

Karplus, K., Karchin, R., Draper, J., Casper, J., Mandel-Gutfreund, Y., Diekhans, M., & Hughey, R. (2003). Combining local-structure, fold-recognition, and new fold methods for protein structure prediction. *Proteins*, *53 Suppl 6*, 491–496. http://doi.org/10.1002/prot.10540

Karplus, M., & McCammon, J. A. (1983). Dynamics of proteins: elements and function. *Annual Review of Biochemistry*, *52*, 263–300. http://doi.org/10.1146/annurev.bi.52.070183.001403

Kasimova, M. a, Zaydman, M. a, Cui, J., & Tarek, M. (2015). PIP2-dependent coupling is prominent in Kv7.1 due to weakened interactions between S4-S5 and S6, 2–10. http://doi.org/10.1038/srep07474

Kass, R. S. (n.d.). The channelopathies: novel insights into molecular and genetic mechanisms of human disease. *The Journal of Clinical Investigation*, *115*(8), 1986–1989. http://doi.org/10.1172/JCI26011

Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, a a, Aflalo, C., & Vakser, I. a. (1992). Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(6), 2195–2199. http://doi.org/10.1073/pnas.89.6.2195

Kelley, L. A., MacCallum, R. M., & Sternberg, M. J. E. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM1. *Journal of Molecular Biology*, *299*(2), 501–522. http://doi.org/http://dx.doi.org/10.1006/jmbi.2000.3741

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocols*, *10*(6), 845–858. Retrieved from http://dx.doi.org/10.1038/nprot.2015.053

Keskin, O., Ma, B., & Nussinov, R. (2005). Hot regions in protein-protein interactions: The organization and contribution of structurally conserved hot spot residues. *Journal of Molecular Biology*, *345*(5), 1281–1294. http://doi.org/10.1016/j.jmb.2004.10.077

Khalili-Araghi, F., Gumbart, J., Wen, P.-C., Sotomayor, M., Tajkhorshid, E., & Schulten, K. (2009). Molecular dynamics simulations of membrane channels and transporters. *Current Opinion in Structural Biology*, *19*(2), 128–137. http://doi.org/10.1016/j.sbi.2009.02.011

Khamis, M. A., Gomaa, W., & Ahmed, W. F. (2015). Machine learning in computational docking. *Artificial Intelligence in Medicine*, *63*(3), 135–152. http://doi.org/10.1016/j.artmed.2015.02.002

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., & Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*, *37*(Database issue), D387-92.

http://doi.org/10.1093/nar/gkn750
Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, *32*(Web Server issue), W526-31. http://doi.org/10.1093/nar/gkh468
Kim, D., McCoy, J., & Nimigean, C. (2015). Ion selectivity and conductance. *Handbook of Ion Channels*, 13–24. http://doi.org/doi:10.1201/b18027-4
Klibanov, A. M. (2001). Improving enzymes by using them in organic solvents. *Nature*, *409*(6817), 241–246. Retrieved from http://dx.doi.org/10.1038/35051719
Koes, D. R., Baumgartner, M. P., & Camacho, C. J. (2013). Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. *Journal of Chemical Information and Modeling*, *53*(8), 1893–1904. http://doi.org/10.1021/ci300604z
Koh, I. Y. Y., Eyrich, V. A., Marti-Renom, M. A., Przybylski, D., Madhusudhan, M. S., Eswar, N., … Rost, B. (2003). EVA: Evaluation of protein structure prediction servers. *Nucleic Acids Research*, *31*(13), 3311–3315.
Kojima, K., Konopleva, M., McQueen, T., O'Brien, S., Plunkett, W., & Andreeff, M. (2006). Mdm2 inhibitor Nutlin-3a induces p53-mediated apoptosis by transcription-dependent and transcription-independent mechanisms and may overcome Atm-mediated resistance to fludarabine in chronic lymphocytic leukemia. *Blood*, *108*(3), 993–1000. http://doi.org/10.1182/blood-2005-12-5148
Korasick, D. A., & Jez, J. M. (2016). Protein Domains: Structure, Function, and Methods. In R. A. Bradshaw & P. D. B. T.-E. of C. B. Stahl (Eds.), (pp. 91–97). Waltham: Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-0-12-394447-4.10011-2
Kortemme, T., & Baker, D. (2004). Computational design of protein-protein interactions. *Current Opinion in Chemical Biology*, *8*(1), 91–97. http://doi.org/10.1016/j.cbpa.2003.12.008
Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Vajda, S. (2013). How good is automated protein docking? *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2159–2166. http://doi.org/10.1002/prot.24403
Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., … Karplus, K. (2009). Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins: Structure, Function and Bioinformatics*, *77*(SUPPL. 9), 114–122. http://doi.org/10.1002/prot.22570
Krieger, E., Nabuurs, S. B., & Vriend, G. (2003). Homology Modeling. In *Structural Bioinformatics* (pp. 509–523). John Wiley & Sons, Inc. http://doi.org/10.1002/0471721204.ch25
Krissinel, E., & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D*, *60*(12 Part 1), 2256–2268. http://doi.org/10.1107/S0907444904026460
Krivov, G. G., Shapovalov, M. V., & Dunbrack, R. L. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins: Structure, Function and Bioinformatics*, *77*(4), 778–795. http://doi.org/10.1002/prot.22488
Kuang, Q., Purhonen, P., & Hebert, H. (2015). Structure of potassium channels. *Cellular and Molecular Life Sciences*, *72*, 3677–3693. http://doi.org/10.1007/s00018-015-1948-5
Kuo, I. Y., & Ehrlich, B. E. (2012, December). Ion Channels in Renal Disease. *Chemical Reviews*. http://doi.org/10.1021/cr3001077
Kutteh, R., Vandenberg, J. I., & Kuyucak, S. (2007). Molecular dynamics and continuum electrostatics studies of inactivation in the HERG potassium channel. *Journal of Physical Chemistry B*, *111*(5), 1090–1098. http://doi.org/10.1021/jp066294d
Lai, H. C., & Jan, L. Y. (2006). The distribution and targeting of neuronal voltage-gated ion channels. *Nat Rev Neurosci*, *7*(7), 548–562. Retrieved from http://dx.doi.org/10.1038/nrn1938
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., … Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947–2948. http://doi.org/10.1093/bioinformatics/btm404
Larsson, P., Wallner, B., Lindahl, E., & Elofsson, A. (2008). Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Science : A Publication of the Protein Society*, *17*(6), 990–1002. http://doi.org/10.1110/ps.073344908
Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(C), 283–291. http://doi.org/10.1107/S0021889892009944
Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and

PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, *8*(4), 477–486.

Laskowski, R. A., & Swaminathan, G. J. (2013). *Problems of Protein Three-Dimensional Structures*. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*. Elsevier Inc. http://doi.org/10.1016/B978-0-12-409547-2.02535-X

Launay, G., & Simonson, T. (2008). Homology modelling of protein-protein complexes: a simple method and its possibilities and limitations. *BMC Bioinformatics*, *9*, 427. http://doi.org/10.1186/1471-2105-9-427

Lehmann-Horn, F., & Jurkat-Rott, K. (1999). Voltage-Gated Ion Channels and Hereditary Disease. *Physiological Reviews*, *79*(4), 1317–1372. Retrieved from http://physrev.physiology.org/content/79/4/1317

Lemoine, D., Jiang, R., Taly, A., Chataigneau, T., Specht, A., & Grutter, T. (2012). Ligand-Gated Ion Channels: New Insights into Neurological Disorders and Ligand Recognition. *Chemical Reviews*, *112*(12), 6285–6318. http://doi.org/10.1021/cr3000829

Lensink, M. F., & Wodak, S. J. (2013). Docking, scoring, and affinity prediction in CAPRI. *Proteins: Structure, Function, and Bioinformatics*, *81*(12), 2082–2095. http://doi.org/10.1002/prot.24428

Lerche, C., Bruhova, I., Lerche, H., Steinmeyer, K., Wei, A. D., Strutz-Seebohm, N., … Seebohm, G. (2007). Chromanol 293B binding in KCNQ1 (Kv7.1) channels involves electrostatic interactions with a potassium ion in the selectivity filter. *Molecular Pharmacology*, *71*, 1503–1511. http://doi.org/10.1124/mol.106.031682

Lesk, A. M. (1997). CASP2: report on ab initio predictions. *Proteins*, *Suppl 1*, 151–166. http://doi.org/10.1002/(sici)1097-0134(1997)1+<151::aid-prot20>3.3.co;2-j

Lesk, V. I., & Sternberg, M. J. E. (2008). 3D-Garden: A system for modelling protein-protein complexes based on conformational refinement of ensembles generated with the marching cubes algorithm. *Bioinformatics*, *24*(9), 1137–1144. http://doi.org/10.1093/bioinformatics/btn093

Levitt, M. (1992). Accurate modeling of protein conformation by automatic segment matching. *Journal of Molecular Biology*, *226*(2), 507–533. http://doi.org/http://dx.doi.org/10.1016/0022-2836(92)90964-L

Li, L., Chen, R., & Weng, Z. (2003). RDOCK:Refinement of Rigid-body Protein Docking Predictions. *Proteins: Struc. Funct. Gen*, *53*(061/14), 693–707. http://doi.org/061/14

Li, M., & Wang, B. (2007). Homology modeling and examination of the effect of the D92E mutation on the H5N1 nonstructural protein NS1 effector domain. *Journal of Molecular Modeling*, *13*(12), 1237–1244. http://doi.org/10.1007/s00894-007-0245-0

Li, Y., & Zhang, Y. (2009). REMO: A new protocol to refine full atomic protein models from C-alpha traces by optimizing hydrogen-bonding networks. *Proteins: Structure, Function and Bioinformatics*, *76*(3), 665–676. http://doi.org/10.1002/prot.22380

Liang, S., & Grishin, N. V. (2002). Side-chain modeling with an optimized scoring function. *Protein Science : A Publication of the Protein Society*, *11*(2), 322–331. http://doi.org/10.1110/ps.24902

Liang, S., Liu, S., Zhang, C., & Zhou, Y. (2007). A simple reference state makes a significant improvement in near-native selections from structurally refined docking decoys. *Proteins*, *69*, 244–253.

Linder, T., de Groot, B. L., & Stary-Weinzinger, A. (2013). Probing the Energy Landscape of Activation Gating of the Bacterial Potassium Channel KcsA. *PLOS Computational Biology*, *9*(5), e1003058. Retrieved from https://doi.org/10.1371/journal.pcbi.1003058

Liu, S., Focke, P. J., Matulef, K., Bian, X., Moënne-Loccoz, P., Valiyaveetil, F. I., & Lockless, S. W. (2015). Ion-binding properties of a K^+ channel selectivity filter in different conformations. *Proceedings of the National Academy of Sciences*, *112*(49), 15096–15100. http://doi.org/10.1073/pnas.1510526112

Liu, S., & Vakser, I. a. (2011). DECK: Distance and environment-dependent, coarse-grained, knowledge-based potentials for protein-protein docking. *BMC Bioinformatics*, *12*(1), 280. http://doi.org/10.1186/1471-2105-12-280

Lockless, S. W., Zhou, M., & MacKinnon, R. (2007). Structural and Thermodynamic Properties of Selective Ion Binding in a K+ Channel. *PLOS Biology*, *5*(5), e121. Retrieved from https://doi.org/10.1371/journal.pbio.0050121

Long, S. B., Campbell, E. B., & MacKinnon, R. (2005). Crystal Structure of a Mammalian Voltage-Dependent Shaker Family K+ Channel. *Science*, *309*(5736), 897–903. http://doi.org/10.1126/science.1116269

Long, S. B., Tao, X., Campbell, E. B., & MacKinnon, R. (2007). Atomic structure of a voltage-dependent

K+ channel in a lipid membrane-like environment. *Nature*, *450*(7168), 376–382. http://doi.org/10.1038/nature06265

Lonsdale, R., Harvey, J. N., & Mulholland, A. J. (2012). A practical guide to modelling enzyme-catalysed reactions. *Chemical Society Reviews*, *41*(8), 3025–3038. http://doi.org/10.1039/c2cs15297e

Lorenzen, S., & Zhang, Y. (2007). Monte Carlo refinement of rigid-body protein docking structures with backbone displacement and side-chain optimization. *Protein Science : A Publication of the Protein Society*, *16*(12), 2716–2725. http://doi.org/10.1110/ps.072847207

Lu, H., & Skolnick, J. (2003). Application of statistical potentials to protein structure refinement from low resolution ab initio models. *Biopolymers*, *70*(4), 575–584. http://doi.org/10.1002/bip.10537

Lundstrom, K. (2007). Structural genomics and drug discovery. *Journal of Cellular and Molecular Medicine*, *11*(2), 224–238. http://doi.org/10.1111/j.1582-4934.2007.00028.x

Luthy, R., Bowie, J. U., & Eisenberg, D. (1992). Assessment of protein models with three-dimensional profiles. *Nature*, *356*(6364), 83–85. http://doi.org/10.1038/356083a0

Lvov, A., Gage, S. D., Berrios, V. M., & Kobertz, W. R. (2010). Identification of a protein-protein interaction between KCNE1 and the activation gate machinery of KCNQ1. *The Journal of General Physiology*, *135*(6), 607–618. http://doi.org/10.1085/jgp.200910386

Ma, J., & Wang, S. (2014). Chapter Five - Algorithms, Applications, and Challenges of Protein Structure Alignment. In R. D. B. T.-A. in P. C. and S. Biology (Ed.), (Vol. Volume 94, pp. 121–175). Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-0-12-800168-4.00005-6

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., … Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B*, *102*(18), 3586–3616. http://doi.org/10.1021/jp973084f

Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., & Sherman, W. (2013). Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, *27*(3), 221–234. http://doi.org/10.1007/s10822-013-9644-8

Magder, S. (1998). The Heart: Physiology, from cell to circulation. *New England Journal of Medicine*, *339*(26), 1949–1950. http://doi.org/10.1056/NEJM199812243392622

Mahoney, M. W., & Jorgensen, W. L. (2000). A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics*, *112*(20), 8910–8922. http://doi.org/10.1063/1.481505

Maier, S. K. G., Westenbroek, R. E., Schenkman, K. A., Feigl, E. O., Scheuer, T., & Catterall, W. A. (2002). An unexpected role for brain-type sodium channels in coupling of cell surface depolarization to contraction in the heart. *Proceedings of the National Academy of Sciences* , *99*(6), 4073–4078. http://doi.org/10.1073/pnas.261705699

Mandell, J. G., Roberts, V. a, Pique, M. E., Kotlovyi, V., Mitchell, J. C., Nelson, E., … Ten Eyck, L. F. (2001). Protein docking using continuum electrostatics and geometric fit. *Protein Engineering*, *14*(2), 105–113. http://doi.org/10.1093/protein/14.2.105

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., & Sali, A. (2000). Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, *29*, 291–325. http://doi.org/10.1146/annurev.biophys.29.1.291

Martinez-Martinez, P., Molenaar, P. C., Losen, M., Stevens, J., Baets, M. H. De, Szoke, A., … Rutten, B. P. F. (2013). Autoantibodies to neurotransmitter receptors and ion channels: from neuromuscular to neuropsychiatric disorders. *Frontiers in Genetics*, *4*, 181. http://doi.org/10.3389/fgene.2013.00181

Mashayak, S. Y., & Tanner, D. E. (2011). Comparing Solvent Models for Molecular Dynamics of Protein. *Time*, 1–18.

Mashiach, E., Nussinov, R., & Wolfson, H. J. (2010). FiberDock: Flexible induced-fit backbone refinement in molecular docking. *Proteins: Structure, Function and Bioinformatics*, *78*(6), 1503–1519. http://doi.org/10.1002/prot.22668

Mashiach, E., Schneidman-Duhovny, D., Andrusier, N., Nussinov, R., & Wolfson, H. J. (2008). FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Research*, *36*(Web Server issue), W229-32. http://doi.org/10.1093/nar/gkn186

Matsuzaki, Y., Uchikoga, N., Ohue, M., Shimoda, T., Sato, T., Ishida, T., & Akiyama, Y. (2013). MEGADOCK 3.0: a high-performance protein-protein interaction prediction software using hybrid parallel computing for petascale supercomputing environments. *Source Code for Biology and Medicine*, *8*(1), 18. http://doi.org/10.1186/1751-0473-8-18

Maximova, T., Moffatt, R., Ma, B., Nussinov, R., & Shehu, A. (2016). Principles and Overview of

Sampling Methods for Modeling Macromolecular Structure and Dynamics. *PLoS Computational Biology*, *12*(4), e1004619. JOUR. http://doi.org/10.1371/journal.pcbi.1004619

Mcdonough, S. I. (2013). Calcium ion channels: Challenges and successes in drug discovery. *Wiley Interdisciplinary Reviews: Membrane Transport and Signaling*, *2*(2), 85–104. http://doi.org/10.1002/wmts.71

McGuffin, L. J., Atkins, J. D., Salehe, B. R., Shuid, A. N., & Roche, D. B. (2015). IntFOLD: an integrated server for modelling protein structures and functions from amino acid sequences. *Nucleic Acids Research*, *43*(W1), W169-73. http://doi.org/10.1093/nar/gkv236

McLaughlin, S., Wang, J., Gambhir, A., & Murray, D. (2002). PIP2 and Proteins: Interactions, Organization, and Information Flow. *Annual Review of Biophysics and Biomolecular Structure*, *31*(1), 151–175. http://doi.org/10.1146/annurev.biophys.31.082901.134259

Melman, Y. F., Krummerman, A., & McDonald, T. V. (2002). KCNE regulation of KvLQT1 channels: Structure-function correlates. *Trends in Cardiovascular Medicine*, *12*(4), 182–187. http://doi.org/10.1016/S1050-1738(02)00158-5

Melman, Y. F., Um, S. Y., Krumerman, A., Kagan, A., & McDonald, T. V. (2004). KCNE1 binds to the KCNQ1 pore to regulate potassium channel activity. *Neuron*, *42*(6), 927–937. http://doi.org/10.1016/j.neuron.2004.06.001

Melo, F., Devos, D., Depiereux, E., & Feytmans, E. (1997). ANOLEA: a www server to assess protein structures. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, *5*, 187–190.

Merlino, A., Vieites, M., Gambino, D., & Laura Coitiño, E. (2014). Homology modeling of T. cruzi and L. major NADH-dependent fumarate reductases: Ligand docking, molecular dynamics validation, and insights on their binding modes. *Journal of Molecular Graphics and Modelling*, *48*, 47–59. http://doi.org/http://dx.doi.org/10.1016/j.jmgm.2013.12.001

Miceli, F., Soldovieri, M. V., Ambrosino, P., De Maria, M., Manocchio, L., Medoro, A., & Taglialatela, M. (2015). Molecular pathophysiology and pharmacology of the voltage-sensing module of neuronal ion channels . *Frontiers in Cellular Neuroscience* . Retrieved from http://journal.frontiersin.org/article/10.3389/fncel.2015.00259

Miller, C. (2000). An overview of the potassium channel family. *Genome Biology*, *1*(4), reviews0004.1-reviews0004.5. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC138870/

Mills, C. L., Beuning, P. J., & Ondrechen, M. J. (2015). Biochemical functional predictions for protein structures of unknown or uncertain function. *Computational and Structural Biotechnology Journal*, *13*, 182–191. http://doi.org/10.1016/j.csbj.2015.02.003

Minor, D. L. J., & Findeisen, F. (2010). Progress in the structural understanding of voltage-gated calcium channel (CaV) function and modulation. *Channels (Austin, Tex.)*, *4*(6), 459–474. http://doi.org/10.4161/chan.4.6.12867

Mirams, G. R., Cui, Y., Sher, A., Fink, M., Cooper, J., Heath, B. M., … Noble, D. (2011). Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovascular Research*, *91*(1), 53–61. http://doi.org/10.1093/cvr/cvr044

Mitsutake, A., Mori, Y., & Okamoto, Y. (2013). Enhanced sampling algorithms. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 924, pp. 153–195). United States. http://doi.org/10.1007/978-1-62703-017-5_7

Moal, I. H., Torchala, M., Bates, P. A., & Fernández-Recio, J. (2013). The scoring of poses in protein-protein docking: current capabilities and future directions. *BMC Bioinformatics*, *14*, 286. http://doi.org/10.1186/1471-2105-14-286

Modell, S. M., & Lehmann, M. H. (2006). The long QT syndrome family of cardiac ion channelopathies: a HuGE review. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, *8*(3), 143–155. http://doi.org/10.109701.gim.0000204468.85308.86

Moont, G., Gabb, H. a., & Sternberg, M. J. E. (1999). Use of pair potentials across protein interfaces in screening predicted docked complexes. *Proteins: Structure, Function and Genetics*, *35*(3), 364–373. http://doi.org/10.1002/(SICI)1097-0134(19990515)35:3<364::AID-PROT11>3.0.CO;2-4

Morais-Cabral, J. H., Zhou, Y., & MacKinnon, R. (2001). Energetic optimization of ion conduction rate by the K+ selectivity filter. *Nature*, *414*(6859), 37–42. Retrieved from http://dx.doi.org/10.1038/35102000

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility.

Journal of Computational Chemistry, *30*(16), 2785–2791. http://doi.org/10.1002/jcc.21256
Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., & Tramontano, A. (2014). Critical assessment of methods of protein structure prediction (CASP) - round x. *Proteins: Structure, Function and Bioinformatics*, *82*(SUPPL.2), 1–6. http://doi.org/10.1002/prot.24452
Moult, J., Pedersen, J. T., Judson, R., & Fidelis, K. (1995). A large-scale experiment to assess protein structure prediction methods. *Proteins*, *23*(3), ii–v. http://doi.org/10.1002/prot.340230303
Mousavi Nik, A., Gharaie, S., & Jeong Kim, H. (2015). Cellular mechanisms of mutations in Kv7.1: auditory functions in Jervell and Lange-Nielsen syndrome vs. Romano–Ward syndrome. *Frontiers in Cellular Neuroscience*, *9*, 32. http://doi.org/10.3389/fncel.2015.00032
Muegge, I., & Rarey, M. (2001). Small Molecule Docking and Scoring. In *Reviews in Computational Chemistry* (pp. 1–60). John Wiley & Sons, Inc. http://doi.org/10.1002/0471224413.ch1
Mulholland, A. J. (2008). Introduction. Biomolecular simulation. *Journal of the Royal Society Interface*, *5*(Suppl 3), S169–S172. http://doi.org/10.1098/rsif.2008.0385.focus
Muller, A., MacCallum, R. M., & Sternberg, M. J. (1999). Benchmarking PSI-BLAST in genome annotation. *Journal of Molecular Biology*, *293*(5), 1257–1271. http://doi.org/10.1006/jmbi.1999.3233
Murzin, A. G., Brenner, S. E., Hubbard, T., & Chothia, C. (1995). SCOP: A structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology*, *247*(4), 536–540. http://doi.org/http://dx.doi.org/10.1016/S0022-2836(05)80134-2
Nakajo, K., & Kubo, Y. (2015). KCNQ1 channel modulation by KCNE proteins via the voltage-sensing domain. *The Journal of Physiology*, *593*(12), 2617–2625. http://doi.org/10.1113/jphysiol.2014.287672
Nakajo, K., Ulbrich, M. H., Kubo, Y., & Isacoff, E. Y. (2010). Stoichiometry of the KCNQ1 - KCNE1 ion channel complex. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(44), 18862–18867. http://doi.org/10.1073/pnas.1010354107
Nayeem, A., Sitkoff, D., & Krystek, S. (2006). A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models. *Protein Science : A Publication of the Protein Society*, *15*(4), 808–824. http://doi.org/10.1110/ps.051892906
Nerbonne, J. M., & Kass, R. S. (2005). Molecular physiology of cardiac repolarization. *Physiological Reviews*, *85*(4), 1205–1253. http://doi.org/10.1152/physrev.00002.2005
Ngo, T., Kufareva, I., Coleman, J. L., Graham, R. M., Abagyan, R., & Smith, N. J. (2016). Identifying ligands at orphan GPCRs: Current status using structure-based approaches. *British Journal of Pharmacology*. http://doi.org/10.1111/bph.13452
Nillegoda, N. B., Kirstein, J., Szlachcic, A., Berynskyy, M., Stank, A., Stengel, F., … Bukau, B. (2015). Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature*, *524*(7564), 247–251. Retrieved from http://dx.doi.org/10.1038/nature14884
Norin, M., Haeffner, F., Hult, K., & Edhoim, O. (1994). Molecular Dynamics Simulations of an Enzyme Surrounded by Vacuum, Water, or a Hydrophobic Solvent. *Bipsical Journal*, *67*(August), 548–559. http://doi.org/10.1016/S0006-3495(94)80515-6
Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *Jchep*, *81*(1), 511. http://doi.org/10.1063/1.447334
Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, *302*(1), 205—217. http://doi.org/10.1006/jmbi.2000.4042
Nymeyer, H., Gnanakaran, S., & Garc??a, A. E. (2004). Atomic Simulations of Protein Folding, Using the Replica Exchange Algorithm. *Methods in Enzymology*, *383*(2000), 119–149. http://doi.org/10.1016/S0076-6879(04)83006-4
Oostenbrink, C., Villa, A., Mark, A. E., & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, *25*(13), 1656–1676. http://doi.org/10.1002/jcc.20090
Orengo, C. A., Bray, J. E., Buchan, D. W. A., Harrison, A., Lee, D., Pearl, F. M. G., … Thornton, J. M. (2002). The CATH protein family database: a resource for structural and functional annotation of genomes. *Proteomics*, *2*(1), 11–21.
Ortiz, A. R., Strauss, C. E. M., & Olmea, O. (2002). MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison. *Protein Science : A Publication of the Protein Society*, *11*(11), 2606–2621. http://doi.org/10.1110/ps.0215902
Osteen, J. D., Sampson, K. J., & Kass, R. S. (2010). The cardiac IKs channel, complex indeed. *Proceedings*

of the National Academy of Sciences of the United States of America, *107*(44), 18751–18752. http://doi.org/10.1073/pnas.1014150107

Pabo, C. O., & Sauer, R. T. (2003). Transcription Factors: Structural Families and Principles of DNA Recognition. *Annual Review of Biochemistry*. http://doi.org/10.1146/annurev.biochem.61.1.1053

Palma, P. N., Krippahl, L., Wampler, J. E., & Moura, J. J. (2000). BiGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins*, *39*(4), 372–384. http://doi.org/10.1002/(SICI)1097-0134(20000601)39:4<372::AID-PROT100>3.0.CO;2-Q [pii]

Palmer, L. G. (2007). Ion Channels in Epithelial Cells. In S.-H. Chung, O. S. Andersen, & V. Krishnamurthy (Eds.), *Biological Membrane Ion Channels: Dynamics, Structure, and Applications* (pp. 425–445). New York, NY: Springer New York. http://doi.org/10.1007/0-387-68919-2_12

Panaghie, G., Tai, K.-K., & Abbott, G. W. (2006). Interaction of KCNE subunits with the KCNQ1 K+ channel pore. *The Journal of Physiology*, *570*(Pt 3), 455–467. http://doi.org/10.1113/jphysiol.2005.100644

Park, K.-S., Yang, J.-W., Seikel, E., & Trimmer, J. S. (2008). Potassium Channel Phosphorylation in Excitable Cells: Providing Dynamic Functional Variability to a Diverse Family of Ion Channels. *Physiology*, *23*(1), 49–57. Retrieved from http://physiologyonline.physiology.org/content/23/1/49.abstract

Patel, J. S., Berteotti, A., Ronsisvalle, S., Rocchia, W., & Cavalli, A. (2014). Steered Molecular Dynamics Simulations for Studying Protein–Ligand Interaction in Cyclin-Dependent Kinase 5. *Journal of Chemical Information and Modeling*, *54*(2), 470–480. http://doi.org/10.1021/ci4003574

Paulussen, A. D. C., Gilissen, R. A. H. J., Armstrong, M., Doevendans, P. A., Verhasselt, P., Smeets, H. J. M., … Aerssens, J. (2004). Genetic variations of KCNQ1, KCNH2, SCN5A, KCNE1, and KCNE2 in drug-induced long QT syndrome patients. *Journal of Molecular Medicine*, *82*(3), 182–188. http://doi.org/10.1007/s00109-003-0522-z

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods in Enzymology*, *183*, 63–98.

Pearson, W. R. (1998). Empirical statistical estimates for sequence similarity searches. *Journal of Molecular Biology*, *276*(1), 71–84. http://doi.org/10.1006/jmbi.1997.1525

Pedotti, M., Simonelli, L., Livoti, E., & Varani, L. (2011). Computational Docking of Antibody-Antigen Complexes, Opportunities and Pitfalls Illustrated by Influenza Hemagglutinin. *International Journal of Molecular Sciences*, *12*(1), 226–251. http://doi.org/10.3390/ijms12010226

Peng, D., Kim, J., Kroncke, B. M., Law, C. L., Xia, Y., Droege, K. D., … Sanders, C. R. (2014). Purification and Structural Study of the Voltage-Sensor Domain of the Human KCNQ1 Potassium Ion Channel. *Biochemistry*, *53*(12), 2032–2042. http://doi.org/10.1021/bi500102w

Peng, J., & Xu, J. (2011). A multiple-template approach to protein threading. *Proteins: Structure, Function, and Bioinformatics*, *79*(6), 1930–1939. http://doi.org/10.1002/prot.23016

Pereon, Y., Demolombe, S., Baro, I., Drouin, E., Charpentier, F., & Escande, D. (2000). Differential expression of KvLQT1 isoforms across the human ventricular wall. *American Journal of Physiology. Heart and Circulatory Physiology*, *278*(6), H1908-15.

Peroz, D., Rodriguez, N., Choveau, F., Baró, I., Mérot, J., & Loussouarn, G. (2008). Kv7.1 (KCNQ1) properties and channelopathies. *The Journal of Physiology*, *586*(7), 1785–1789. http://doi.org/10.1113/jphysiol.2007.148254

Perry, S. R., Xu, W., Wirija, A., Lim, J., Yau, M.-K., Stoermer, M. J., … Fairlie, D. P. (2015). Three Homology Models of PAR2 Derived from Different Templates: Application to Antagonist Discovery. *Journal of Chemical Information and Modeling*, *55*(6), 1181–1191. http://doi.org/10.1021/acs.jcim.5b00087

Perutz, M. (2012). Protein Structure, 1–18. http://doi.org/10.5772/2335

Petegem, F. Van, & Minor, D. L. (2006). The structural biology of voltage-gated calcium channel function and regulation. *Biochemical Society Transactions*, *34*(Pt 5), 887–893. http://doi.org/10.1042/BST0340887

Petrey, D., Xiang, Z., Tang, C. L., Xie, L., Gimpelev, M., Mitros, T., … Honig, B. (2003). Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins*, *53 Suppl 6*, 430–435. http://doi.org/10.1002/prot.10550

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612. http://doi.org/10.1002/jcc.20084

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., … Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802. http://doi.org/10.1002/jcc.20289

Pieper, U., Eswar, N., Davis, F. P., Braberg, H., Madhusudhan, M. S., Rossi, A., … Sali, A. (2006). MODBASE: a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, *34*(Database issue), D291-5. http://doi.org/10.1093/nar/gkj059

Pierce, B. G., Wiehe, K., Hwang, H., Kim, B.-H., Vreven, T., & Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics*, *30*(12), btu097. http://doi.org/10.1093/bioinformatics/btu097

Pierce, B., & Weng, Z. (2007). ZRANK: Reranking Protein Docking Predictions With an Optimized Energy Function. *Proteins*, *67*, 1078–1086.

Pinnell, J., Turner, S., & Howell, S. (2007). Cardiac muscle physiology. *Continuing Education in Anaesthesia Critical Care & Pain*, *7*(3), 85–88. Retrieved from http://dx.doi.org/10.1093/bjaceaccp/mkm013

Pletscher-Frankild, S., Pallejà, A., Tsafou, K., Binder, J. X., & Jensen, L. J. (2015). DISEASES: Text mining and data integration of disease–gene associations. *Methods*, *74*, 83–89. http://doi.org/https://doi.org/10.1016/j.ymeth.2014.11.020

Po, S., Roberds, S., Snyders, D. J., Tamkun, M. M., & Bennett, P. B. (1993). Heteromultimeric assembly of human potassium channels. Molecular basis of a transient outward current? *Circulation Research*, *72*(6), 1326–1336. http://doi.org/10.1161/01.RES.72.6.1326

Ponder, J. W., & Case, D. A. (2003). Force fields for protein simulations. *Advances in Protein Chemistry*, *66*, 27–85. http://doi.org/10.1016/S0065-3233(03)66002-X

Pons, C., Jiménez-González, D., González-Álvarez, C., Servat, H., Cabrera-Benítez, D., Aguilar, X., & Fernández-Recio, J. (2012). Cell-Dock: High-performance protein-protein docking. *Bioinformatics*, *28*(18), 2394–2396. http://doi.org/10.1093/bioinformatics/bts454

Popov, P., Ritchie, D. W., & Grudinin, S. (2014). DockTrina: Docking triangular protein trimers. *Proteins: Structure, Function and Bioinformatics*, *82*(1), 34–44. http://doi.org/10.1002/prot.24344

Proks, P., & Lippiat, J. D. (2006). Membrane ion channels and diabetes. *Current Pharmaceutical Design*, *12*(4), 485–501. Journal Article, Review.

Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., … Lindahl, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics (Oxford, England)*, *29*(7), 845–54. http://doi.org/10.1093/bioinformatics/btt055

Pusch, M. (2007). Chloride Transporting CLC Proteins1. In S.-H. Chung, O. S. Andersen, & V. Krishnamurthy (Eds.), *Biological Membrane Ion Channels: Dynamics, Structure, and Applications* (pp. 301–333). New York, NY: Springer New York. http://doi.org/10.1007/0-387-68919-2_8

Qian, B., Ortiz, A. R., & Baker, D. (2004). Improvement of comparative model accuracy by free-energy optimization along principal components of natural structural variation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(43), 15346–15351. http://doi.org/10.1073/pnas.0404703101

Rajapaksha, H., & Petrovsky, N. (2014). <italic>In Silico</italic> Structural Homology Modelling and Docking for Assessment of Pandemic Potential of a Novel H7N9 Influenza Virus and Its Ability to Be Neutralized by Existing Anti-Hemagglutinin Antibodies. *PLoS ONE*, *9*(7), e102618. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0102618

Ramachandran, S., Kota, P., Ding, F., & Dokholyan, N. V. (2011). Automated Minimization of Steric Clashes in Protein Structures. *Proteins*, *79*(1), 261–270. http://doi.org/10.1002/prot.22879

Rappsilber, J. (2011). The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of Structural Biology*, *173*(3), 530–540. http://doi.org/http://dx.doi.org/10.1016/j.jsb.2010.10.014

Raval, A., Piana, S., Eastwood, M. P., Dror, R. O., & Shaw, D. E. (2012). Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins*, *80*(8), 2071–9. http://doi.org/10.1002/prot.24098

Ravikant, D. V. S., & Elber, R. (2010). PIE-efficient filters and coarse grained potentials for unbound protein-protein docking. *Proteins*, *78*(2), 400–419. http://doi.org/10.1002/prot.22550

Ravikant, D. V. S., & Elber, R. (2011). Energy design for protein-protein interactions. *The Journal of Chemical Physics*, *135*(6), 65102. http://doi.org/10.1063/1.3615722

Ravna, A. W., & Sylte, I. (2012). Homology modeling of transporter proteins (carriers and ion channels).

Methods in Molecular Biology (Clifton, N.J.), *857*, 281–299. http://doi.org/10.1007/978-1-61779-588-6_12

Rehm, S., Trodler, P., & Pleiss, J. (2010). Solvent-induced lid opening in lipases: A molecular dynamics study. *Protein Science : A Publication of the Protein Society*, *19*(11), 2122–2130. http://doi.org/10.1002/pro.493

Remmert, M., Biegert, A., Hauser, A., & Soding, J. (2012). HHblits: lightning-fast iterative protein sequence searching by HMM-HMM alignment. *Nat Meth*, *9*(2), 173–175. Retrieved from http://dx.doi.org/10.1038/nmeth.1818

Ritchie, D. W., & Kemp, G. J. L. (2000). Protein docking using spherical polar Fourier correlations. *Proteins: Structure, Function and Genetics*, *39*(2), 178–194. http://doi.org/10.1002/(SICI)1097-0134(20000501)39:2<178::AID-PROT8>3.0.CO;2-6

Rivas, A., & Francis, H. W. (2005). Inner ear abnormalities in a Kcnq1 (Kvlqt1) knockout mouse: a model of Jervell and Lange-Nielsen syndrome. *Otology & Neurotology : Official Publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, *26*(3), 415–424.

Robbins, J. (2001). KCNQ potassium channels : physiology , pathophysiology , and pharmacology, *90*, 1–19.

Roden, D. M. (2006). A new role for calmodulin in ion channel biology. *Circulation Research*, *98*(8), 979–981. http://doi.org/10.1161/01.RES.0000221822.22971.8c

Roden, D. M., Balser, J. R., George, A. L., & Anderson, M. E. (2002). Cardiac ion channels. *Annual Review of Physiology*, *64*, 431–75. http://doi.org/10.1146/annurev.physiol.64.083101.145105

Roden, D. M., & Yang, T. (2005, September). Protecting the heart against arrhythmias: potassium current physiology and repolarization reserve. *Circulation*. United States. http://doi.org/10.1161/CIRCULATIONAHA.105.562777

Rodrigues, J. P. G. L. M., Melquiond, A. S. J., Karaca, E., Trellet, M., Van Dijk, M., Van Zundert, G. C. P., … Bonvin, A. M. J. J. (2013). Defining the limits of homology modeling in information-driven protein docking. *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2119–2128. http://doi.org/10.1002/prot.24382

Rodriguez, R., Chinea, G., Lopez, N., Pons, T., & Vriend, G. (1998). Homology modeling, model and software evaluation: three related resources. *Bioinformatics (Oxford, England)*, *14*(6), 523–528. http://doi.org/10.1093/bioinformatics/14.6.523

Roe, D. R., & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: software for processing and analysis of molecular synamics trajectory data. *J Chem Theory Com*, *9*(7), 3084–3095. http://doi.org/10.1021/ct400341p

Rohl, C. A., Strauss, C. E. M., Chivian, D., & Baker, D. (2004). Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*, *55*(3), 656–677. http://doi.org/10.1002/prot.10629

Roux, B., Bernèche, S., Egwolf, B., Lev, B., Noskov, S. Y., Rowley, C. N., & Yu, H. (2011). Ion selectivity in channels and transporters. *The Journal of General Physiology*, *137*(5), 415–426. http://doi.org/10.1085/jgp.201010577

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, *5*(4), 725–738. http://doi.org/10.1038/nprot.2010.5

Rudy, Y. (2008). Molecular basis of cardiac action potential repolarization. *Annals of the New York Academy of Sciences*, *1123*, 113–118. http://doi.org/10.1196/annals.1420.013

Rychlewski, L., Jaroszewski, L., Li, W., & Godzik, A. (2000). Comparison of sequence profiles. Strategies for structural predictions using sequence information. *Protein Science : A Publication of the Protein Society*, *9*(2), 232–241. http://doi.org/10.1110/ps.9.2.232

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327–341. http://doi.org/http://dx.doi.org/10.1016/0021-9991(77)90098-5

Sachyani, D., Dvir, M., Strulovich, R., Tria, G., Pongs, O., Svergun, D., … Hirsch, J. a. (2013). Structural Basis of a Kv7 . 1 (KCNQ1) Potassium Channel Gating Module : Studies of the Intracellular C-terminal Domain in Complex with Calmodulin. *Structure/Folding and Design*, *1*(11), 1582–1594. http://doi.org/10.1016/j.str.2014.07.016

Šali, A., & Blundell, T. L. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology*, *234*(3), 779–815. http://doi.org/http://dx.doi.org/10.1006/jmbi.1993.1626

Sali, A., & Overington, J. P. (1994). Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Science : A Publication of the Protein Society*, *3*(9), 1582–1596. http://doi.org/10.1002/pro.5560030923

Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., & Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins*, *23*(3), 318–326. http://doi.org/10.1002/prot.340230306

Salsbury, F. R. (2010). Molecular Dynamics Simulations of Protein Dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology*, *10*(6), 738–744. http://doi.org/10.1016/j.coph.2010.09.016

Samudrala, R., & Moult, J. (1998). Determinants of side chain conformational preferences in protein structures. *Protein Engineering*, *11*(11), 991–997.

Sanbonmatsu, K. Y., & García, A. E. (2002). Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Proteins*, *46*(2), 225–34. http://doi.org/10.1002/prot.0000

Sanchez, R., & Sali, A. (1997). Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins*, *Suppl 1*, 50–58.

Sanguinetti, M. C., Curran, M. E., Zou, A., Shen, J., Spector, P. S., Atkinson, D. L., & Keating, M. T. (1996). Coassembly of K(V)LQT1 and minK (IsK) proteins to form cardiac I(Ks) potassium channel. *Nature*, *384*(6604), 80–83. http://doi.org/10.1038/384080a0

Sansom, M. S. P., Shrivastava, I. H., Bright, J. N., Tate, J., Capener, C. E., & Biggin, P. C. (2002). Potassium channels : structures , models , simulations, *1565*, 294–307.

Saxena, A., Sangwan, R. S., & Mishra, S. (2013). Fundamentals of Homology Modeling Steps and Comparison among Important Bioinformatics Tools: An Overview. *Science International*. http://doi.org/10.5567/sciint.2013.237.252

Scheraga, H. A., Khalili, M., & Liwo, A. (2007). Protein-folding dynamics: Overview of molecular simulation techniques. *Annual Review Of Physical Chemistry*, *58*, 57–83. http://doi.org/10.1146/annurev.physchem.58.032806.104614

Schleif, R. (2004). Modeling and Studying Proteins with Molecular Dynamics. In B. T.-M. in Enzymology (Ed.), *Numerical Computer Methods, Part D* (Vol. Volume 383, pp. 28–47). Academic Press. http://doi.org/http://dx.doi.org/10.1016/S0076-6879(04)83002-7

Schmitt, N., Schwarz, M., Peretz, A., Abitbol, I., Attali, B., & Pongs, O. (2000). A recessive C-terminal Jervell and Lange-Nielsen mutation of the KCNQ1 channel impairs subunit assembly. *The EMBO Journal*, *19*(3), 332–340. http://doi.org/10.1093/emboj/19.3.332

Schneider, T., Kruse, T., Wimmer, R., Wiedemann, I., Sass, V., Pag, U., … Kristensen, H.-H. (2010). Plectasin, a Fungal Defensin, Targets the Bacterial Cell Wall Precursor Lipid II. *Science* , *328*(5982), 1168–1172. http://doi.org/10.1126/science.1185723

Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*, *33*(SUPPL. 2), 363–367. http://doi.org/10.1093/nar/gki481

Schneidman-Duhovny, D., Ruth, N., & Wolfson, H. J. (2007). Automatic prediction of protein interactions with large scale motion. *Proteins*, *69*, 764–773.

Schrödinger Release 2017-2: LigPrep, Schrödinger, LLC, New York, NY, 2017. (2017). *New York, NY*.

Schrödinger Release 2017-2: MS Jaguar, Schrödinger, LLC. (2017). *New York, NY*.

Schroeder, B. C., Waldegger, S., Fehr, S., Bleich, M., Warth, R., Greger, R., & Jentsch, T. J. (2000). A constitutively open potassium channel formed by KCNQ1 and KCNE3. *Nature*, *403*(6766), 196–199. http://doi.org/10.1038/35003200

Schwartz, P. J., Crotti, L., & Insolia, R. (2012). Long-QT Syndrome. *Circulation: Arrhythmia and Electrophysiology*, *5*(4), 868 LP-877. Retrieved from http://circep.ahajournals.org/content/5/4/868.abstract

Schwede, T. (2013). Protein Modeling: What Happened to the "Protein Structure Gap"? *Structure*, *21*(9), 1531–1540. http://doi.org/http://dx.doi.org/10.1016/j.str.2013.08.007

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* , *31*(13), 3381–3385. http://doi.org/10.1093/nar/gkg520

Seebohm, G. (2013). A complex partnership: KCNQ1 and KCNE1. *Biophysical Journal*, *105*(11), 2437–2438. http://doi.org/10.1016/j.bpj.2013.10.022

Shamgar, L., Ma, L., Schmitt, N., Haitin, Y., Peretz, A., Wiener, R., … Attali, B. (2006). Calmodulin is

essential for cardiac IKS channel gating and assembly: Impaired function in long-QT mutations. *Circulation Research*, *98*(8), 1055–1063. http://doi.org/10.1161/01.RES.0000218979.40770.69

Shao, J. ., Tanner, S. W. ., Thompson, N. ., & Cheatham III, T. E. (2007). Clusteing molecular dynamic trajectories: 1. Characterizing the performance of different clustering algorithms. *J. Chem. Theory Comput.*, *3*, 2312–2334. http://doi.org/10.1021/ct700119m

Shehata, M. A., Belcik Christensen, H., Isberg, V., Sejer Pedersen, D., Bender, A., Brauner-Osborne, H., & Gloriam, D. E. (2015). Identification of the first surrogate agonists for the G protein-coupled receptor GPR132. *RSC Advances*, *5*(60), 48551–48557. http://doi.org/10.1039/C5RA04804D

Sheinerman, F. B., Norel, R., & Honig, B. (2000). Electrostatic aspects of protein-protein interactions. *Current Opinion in Structural Biology*, *10*(2), 153–159. http://doi.org/10.1016/S0959-440X(00)00065-8

Shieh, C.-C., Coghlan, M., Sullivan, J. P., & Gopalakrishnan, M. (2000). Potassium Channels: Molecular Defects, Diseases, and Therapeutic Opportunities. *Pharmacological Reviews*, *52*(4), 557 LP-594. Retrieved from http://pharmrev.aspetjournals.org/content/52/4/557.abstract

Shipston, M. J. (2011). Ion Channel Regulation by Protein Palmitoylation. *The Journal of Biological Chemistry*, *286*(11), 8709–8716. http://doi.org/10.1074/jbc.R110.210005

Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., & Sherman, W. (2010). Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation*, *6*(5), 1509–1519. http://doi.org/10.1021/ct900587b

Silva, J. R., Pan, H., Wu, D., Nekouzadeh, A., Decker, K. F., Cui, J., … Rudy, Y. (2009). A multiscale model linking ion-channel molecular dynamics and electrostatics to the cardiac action potential. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(27), 11102–11106. http://doi.org/10.1073/pnas.0904505106

Silvestrov, P., Müller, T. A., Clark, K. N., Hausinger, R. P., & Cisneros, G. A. (2014). Homology modeling, molecular dynamics, and site-directed mutagenesis study of AlkB human homolog 1 (ALKBH1). *Journal of Molecular Graphics and Modelling*, *54*, 123–130. http://doi.org/http://dx.doi.org/10.1016/j.jmgm.2014.10.013

Simms, B. A., & Zamponi, G. W. (2014). Neuronal Voltage-Gated Calcium Channels: Structure, Function, and Dysfunction. *Neuron*, *82*(1), 24–45. http://doi.org/http://dx.doi.org/10.1016/j.neuron.2014.03.016

Simon, M. I., Strathmann, M. P., & Gautam, N. (1991). Diversity of G proteins in signal transduction. *Science (New York, N.Y.)*, *252*(5007), 802–8. http://doi.org/10.1126/science.1902986

Sippl, M. J. (1993). Recognition of errors in three-dimensional structures of proteins. *Proteins*, *17*(4), 355–362. http://doi.org/10.1002/prot.340170404

Sircar, A., Kim, E. T., & Gray, J. J. (2009). RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Research*, *37*(Web Server issue), W474-9. http://doi.org/10.1093/nar/gkp387

Smart, O. S., Goodfellow, J. M., & Wallace, B. A. (1993). The pore dimensions of gramicidin A. *Biophysical Journal*, *65*(6), 2455–2460. http://doi.org/10.1016/S0006-3495(93)81293-1

Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A., & Sansom, M. S. P. (1996). HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics*, *14*(6), 354–360. http://doi.org/10.1016/S0263-7855(97)00009-X

Smith, G. R., & Sternberg, M. J. E. (n.d.). Prediction of protein – protein interactions by docking methods, 28–35.

Smith, J. a, Vanoye, C. G., Jr, a L., Meiler, J., & Sanders, C. R. (2007). Structural Models for the KCNQ1 Voltage-Gated Potassium Channel. *Biochemistry*, 14141–14152. http://doi.org/10.1021/bi701597s

Soding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, *33*(Web Server issue), W244-8. http://doi.org/10.1093/nar/gki408

Sokolova, O., Kolmakova-Partensky, L., & Grigorieff, N. (2001). Three-Dimensional Structure of a Voltage-Gated Potassium Channel at 2.5 nm Resolution. *Structure*, *9*(3), 215–220. http://doi.org/http://dx.doi.org/10.1016/S0969-2126(01)00578-0

Song, Y., DiMaio, F., Wang, R. Y.-R., Kim, D., Miles, C., Brunette, T., … Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure (London, England : 1993)*, *21*(10), 1735–1742. http://doi.org/10.1016/j.str.2013.08.005

Splawki, I., Tristani-Firouzi, M., Lehmann, M. H., Sanguinetti, M. C., & Keating, M. T. (1997). Mutations in the hminK gene cause long QT syndrome and supppress Iks function. *Nature Genetics*, *17*, 338–

340.
Splawski, I., Shen, J., Timothy, K. W., Lehmann, M. H., Priori, S., Robinson, J. L., … Keating, M. T. (2000). Spectrum of Mutations in Long-QT Syndrome Genes. *Circulation*, *102*(10), 1178 LP-1185. Retrieved from http://circ.ahajournals.org/content/102/10/1178.abstract
Splawski, I., Shen, J., Timothy, K. W., Vincent, G. M., Lehmann, M. H., & Keating, M. T. (1998). Genomic Structure of Three Long QT Syndrome Genes:KVLQT1, HERG,andKCNE1. *Genomics*, *51*(1), 86–97. http://doi.org/http://dx.doi.org/10.1006/geno.1998.5361
Stites, W. E., Meeker, A. K., & Shortle, D. (1994). Evidence for strained interactions between side-chains and the polypeptide backbone. *Journal of Molecular Biology*, *235*(1), 27–32.
Strutz-Seebohm, N., Pusch, M., Wolf, S., Stoll, R., Tapken, D., Gerwert, K., … Seebohm, G. (2011). Structural basis of slow activation gating in the cardiac I Ks channel complex. *Cellular Physiology and Biochemistry : International Journal of Experimental Cellular Physiology, Biochemistry, and Pharmacology*, *27*(5), 443–452. http://doi.org/10.1159/000329965
Suessbrich, H., Bleich, M., Ecke, D., Rizzo, M., Waldegger, S., Lang, F., … Busch, A. E. (1996). Specific blockade of slowly activating I(sK) channels by chromanols - Impact on the role of I(sK) channels in epithelial. *FEBS Letters*, *396*(2–3), 271–275. http://doi.org/10.1016/0014-5793(96)01113-1
Sugita, Y., & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, *314*(1–2), 141–151. JOUR. http://doi.org/http://dx.doi.org/10.1016/S0009-2614(99)01123-9
Suh, B.-C., & Hille, B. PIP2 is a necessary cofactor for ion channel function: how and why?, 37Annual review of biophysics 175–195 (2008). United States. http://doi.org/10.1146/annurev.biophys.37.032807.125859
Sula, A., Booker, J., Ng, L. C. T., Naylor, C. E., DeCaen, P. G., & Wallace, B. A. (2017). The complete structure of an activated open sodium channel. *Nature Communications*, *8*, 14205. http://doi.org/10.1038/ncomms14205
Sun, H. (1998). COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase ApplicationsOverview with Details on Alkane and Benzene Compounds. *The Journal of Physical Chemistry B*, *102*(38), 7338–7364. http://doi.org/10.1021/jp980939v
Sun, J., & MacKinnon, R. (2017). Cryo-EM Structure of a KCNQ1/CaM Complex Reveals Insights into Congenital Long QT Syndrome. *Cell*, *169*(6), 1042–1050.e9. http://doi.org/http://dx.doi.org/10.1016/j.cell.2017.05.019
Sutcliffe, M. J., Haneef, I., Carney, D., & Blundell, T. L. (1987). Knowledge based modelling of homologous proteins, Part I: Three-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Engineering*, *1*(5), 377–384.
Swope, W. C., Andersen, H. C., Berens, P. H., & Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, *76*(1), 637–649. http://doi.org/10.1063/1.442716
Szilagyi, A., Grimm, V., Arakaki, K., & Skolnick, J. (2005). Prediction of physical protein – protein. *Most*, *2*, 1–16. http://doi.org/10.1088/1478-3967/2/0/000
Szilagyi, A., & Zhang, Y. (2014). Template-based structure modeling of protein–protein interactions. *Current Opinion in Structural Biology*, *24*, 10–23. http://doi.org/http://dx.doi.org/10.1016/j.sbi.2013.11.005
Szklarz, G. D., & Halpert, J. R. (1997). Use of homology modeling in conjunction with site-directed mutagenesis for analysis of structure-function relationships of mammalian cytochromes P450. *Life Sciences*, *61*(26), 2507–2520.
Szklarz, G. D., Ornstein, R. L., & Halpert, J. R. (1994). Application of 3-dimensional homology modeling of cytochrome P450 2B1 for interpretation of site-directed mutagenesis results. *Journal of Biomolecular Structure & Dynamics*, *12*(1), 61–78.
Tapper, A. R., & George, A. L. (2001). Location and Orientation of minK within the IKs Potassium Channel Complex. *Journal of Biological Chemistry*, *276*(41), 38249–38254. http://doi.org/10.1074/jbc.M103956200
Tappura, K. (2001). Influence of rotational energy barriers to the conformational search of protein loops in molecular dynamics and ranking the conformations. *Proteins*, *44*(3), 167–179.
Taylor, D. W., Zhu, Y., Staals, R. H. J., Kornfeld, J. E., Shinkai, A., van der Oost, J., … Doudna, J. A. (2015). Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* ,

348(6234), 581–585. http://doi.org/10.1126/science.aaa4535
Taylor, J. S., & Burnett, R. M. (2000). DARWIN: A program for docking flexible molecules. *Proteins: Structure, Function and Genetics*, *41*(2), 173–191. http://doi.org/10.1002/1097-0134(20001101)41:2<173::AID-PROT30>3.0.CO;2-3
Taylor, W. (2004). Protein Structure Folding and Prediction. In *Compact Handbook of Computational Biology* (pp. 223–240). CRC Press. http://doi.org/doi:10.1201/9780203021415.ch6
Teichmann, S. A., Chothia, C., Church, G. M., & Park, J. (2000). Fast assignment of protein structures to sequences using the intermediate sequence library PDB-ISL. *Bioinformatics (Oxford, England)*, *16*(2), 117–124.
Terashi, G., Takeda-Shitaka, Mayuko Kazuhiko, K., Iwadate, M., Takaya, D., & Umeyama, H. (2007). The SKE-DOCK server and human teams based on a combined method of shape complementarity and free energy estimation. *Proteins*, *69*, 866–872.
Tester, D. J., & Ackerman, M. J. (2014). GENETICS OF LONG QT SYNDROME. *Methodist DeBakey Cardiovascular Journal*, *10*(1), 29–33. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4051331/
Thomas, D., Zhang, W., Wu, K., Wimmer, A.-B., Gut, B., Wendt-Nordahl, G., … Karle, C. A. (2003). Regulation of HERG potassium channel activation by protein kinase C independent of direct phosphorylation of the channel protein. *Cardiovascular Research*, *59*(1), 14–26. Retrieved from http://dx.doi.org/10.1016/S0008-6363(03)00386-9
Thompson, A. N., Kim, I., Panosian, T. D., Iverson, T. M., Allen, T. W., & Nimigean, C. M. (2009). Mechanism of potassium channel selectivity revealed by Na(+) and Li(+) binding sites within the KcsA pore. *Nature Structural & Molecular Biology*, *16*(12), 1317–1324. http://doi.org/10.1038/nsmb.1703
Thompson, J., & Begenisich, T. (2001). Affinity and location of an internal K+ ion binding site in shaker K channels. *The Journal of General Physiology*, *117*(5), 373–384.
Tian, W., & Skolnick, J. (2003). How well is enzyme function conserved as a function of pairwise sequence identity? *Journal of Molecular Biology*, *333*(4), 863–882.
Tien, J., Young, D. M., Jan, Y. N., & Jan, L. Y. (2014). Chapter 11 - Molecular Properties of Ion Channels BT - From Molecules to Networks (Third Edition) (pp. 323–348). Boston: Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-0-12-397179-1.00011-7
Tovchigrechko, A., & Vakser, I. a. (2006). GRAMM-X public web server for protein-protein docking. *Nucleic Acids Research*, *34*(WEB. SERV. ISS.), 310–314. http://doi.org/10.1093/nar/gkl206
Trezise, D., Dale, T., & Main, M. (2010). Chapter 2.3: Voltage-gated sodium channels. In J. N. C. Kew & C. H. Davies (Eds.), *Ion channels : from structure to function*. New York: Oxford University Press.
Tristani-Firouzi, M., & Sanguinetti, M. C. (1998). Voltage-dependent inactivation of the human K+ channel KvLQT1 is eliminated by association with minimal K+ channel (minK) subunits. *The Journal of Physiology*, *510*(1), 37–45. http://doi.org/10.1111/j.1469-7793.1998.037bz.x
Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. http://doi.org/10.1002/jcc.21334
Trujillo, K., Paoletta, S., Kiselev, E., & Jacobson, K. A. (2015). Molecular modeling of the human P2Y14 receptor: A template for structure-based design of selective agonist ligands. *Bioorganic & Medicinal Chemistry*, *23*(14), 4056–4064. http://doi.org/http://dx.doi.org/10.1016/j.bmc.2015.03.042
Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., … Pontén, F. (2015). Tissue-based map of the human proteome. *Science*, *347*(6220). Retrieved from http://science.sciencemag.org/content/347/6220/1260419.abstract
Uline, M. J., & Corti, D. S. (2013). Molecular dynamics at constant pressure: Allowing the system to control volume fluctuations via a "shell" particle. *Entropy*, *15*(9), 3941–3969. http://doi.org/10.3390/e15093941
Vacher, H., Mohapatra, D. P., & Trimmer, J. S. (2008). Localization and Targeting of Voltage-Dependent Ion Channels in Mammalian Central Neurons. *Physiological Reviews*, *88*(4), 1407 LP-1447. Retrieved from http://physrev.physiology.org/content/88/4/1407.abstract
Vakser, I. a. (1995). Protein docking for low-resolution structures. *Protein Engineering*, *8*(4), 371–377. http://doi.org/061/14
Vakser, I. A. (2014). Protein-Protein Docking: From Interaction to Interactome. *Biophysical Journal*, *107*(8), 1785–1793. http://doi.org/10.1016/j.bpj.2014.08.033

Vallon, V., Grahammer, F., Richter, K., Bleich, M., Lang, F., Barhanin, J., … Warth, R. (2001). Role of KCNE1-dependent K+ fluxes in mouse proximal tubule. *Journal of the American Society of Nephrology : JASN*, *12*(10), 2003–2011. Journal Article, Research Support, Non-U.S. Gov't.

van Gelder, C. W., Leusen, F. J., Leunissen, J. A., & Noordik, J. H. (1994). A molecular dynamics approach for the generation of complete protein structures from limited coordinate data. *Proteins*, *18*(2), 174–185. http://doi.org/10.1002/prot.340180209

Van Horn, W. D., Vanoye, C. G., & Sanders, C. R. (2011). Working model for the structural basis for KCNE1 modulation of the KCNQ1 potassium channel. *Current Opinion in Structural Biology*, *21*(2), 283–291. http://doi.org/10.1016/j.sbi.2011.01.001

van Vlijmen, H. W., & Karplus, M. (1997). PDB-based protein loop prediction: parameters for selection and methods for optimization. *Journal of Molecular Biology*, *267*(4), 975–1001. http://doi.org/10.1006/jmbi.1996.0857

Veerman, C. C., Verkerk, A. O., Blom, M. T., Klemens, C. A., Langendijk, P. N. J., van Ginneken, A. C. G., … Tan, H. L. (2013). Slow delayed rectifier potassium current blockade contributes importantly to drug-induced long QT syndrome. *Circulation. Arrhythmia and Electrophysiology*, *6*(5), 1002–1009. http://doi.org/10.1161/CIRCEP.113.000239

Venkatraman, V., & Ritchie, D. W. (2012). Flexible protein docking refinement using pose-dependent normal mode analysis. *Proteins: Structure, Function and Bioinformatics*, *80*(9), 2262–2274. http://doi.org/10.1002/prot.24115

Venkatraman, V., Yang, Y. D., Sael, L., & Kihara, D. (2009). Protein-protein docking using region-based 3D Zernike descriptors. *BMC Bioinformatics*, *10*, 407. http://doi.org/10.1186/1471-2105-10-407

Venselaar, H., Joosten, R. P., Vroling, B., Baakman, C. A. B., Hekkelman, M. L., Krieger, E., & Vriend, G. (2010). Homology modelling and spectroscopy, a never-ending love story. *European Biophysics Journal* , *39*(4), 551–563. http://doi.org/10.1007/s00249-009-0531-0

Verlet, L. (1967). Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, *159*(1), 98–103. http://doi.org/10.1103/PhysRev.159.98

Villoutreix, B. O., Kuenemann, M. a., Poyet, J. L., Bruzzoni-Giovanelli, H., Labbé, C., Lagorce, D., … Miteva, M. a. (2014). Drug-like protein-protein interaction modulators: Challenges and opportunities for drug discovery and chemical biology. *Molecular Informatics*, *33*(6–7), 414–437. http://doi.org/10.1002/minf.201400040

Viricel, C., Ahmed, M., & Barakat, K. (2015). Human PD-1 Binds Differently To Its Human Ligands: A Comprehensive Modelling Study. *Journal of Molecular Graphics and Modelling*, *57*, 131–142. http://doi.org/10.1016/j.jmgm.2015.01.015

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *Journal of Molecular Graphics*, *8*(1), 29,52-56.

Vyas, V., Ukawala, R., Chintha, C., & Ghate, M. (2012). Homology modeling a fast tool for drug discovery: Current perspectives. *Indian Journal of Pharmaceutical Sciences*, *74*(1), 1. http://doi.org/10.4103/0250-474X.102537

Waksman, G., & Sansom, C. (2005). Introduction : Proteomics and Protein – Protein Interactions : Biology , Chemistry , Bioinformatics , and Drug Design. In G. Waksman (Ed.), *Proteomics and Protein – Protein Interactions : Biology , Chemistry , Bioinformatics , and Drug Design* (pp. 1–18). New York: Springer.

Wallner, B., & Elofsson, A. (2003). Can correct protein models be identified? *Protein Science : A Publication of the Protein Society*, *12*(5), 1073–1086. http://doi.org/10.1110/ps.0236803

Wallner, B., & Elofsson, A. (2005). All are not equal: a benchmark of different homology modeling programs. *Protein Science : A Publication of the Protein Society*, *14*(5), 1315–27. http://doi.org/10.1110/ps.041253405

Wallqvist, A., & Mountain, R. D. (2007). Molecular Models of Water: Derivation and Description. In *Reviews in Computational Chemistry* (pp. 183–247). John Wiley & Sons, Inc. http://doi.org/10.1002/9780470125908.ch4

Wang, C. H. U., & Schueler-furman, O. R. a. (2005). Improved side-chain modeling for protein – protein docking. *Protein Science*, *14*(5), 1328–1339. http://doi.org/10.1110/ps.041222905.nent

Wang, J., Cieplak, P., Cai, Q., Hsieh, M.-J., Wang, J., Duan, Y., & Luo, R. (2012). Development of Polarizable Models for Molecular Mechanical Calculations III: Polarizable Water Models Conforming to Thole Polarization Screening Schemes. *The Journal of Physical Chemistry. B*, *116*(28), 7999–8008. http://doi.org/10.1021/jp212117d

Wang, Q., Curran, M. E., Splawski, I., Burn, T. C., Millholland, J. M., VanRaay, T. J., … Keating, M. T. (1996). Positional cloning of a novel potassium channel gene: KVLQT1 mutations cause cardiac arrhythmias. *Nature Genetics*, *12*(1), 17–23. http://doi.org/10.1038/ng0196-17

Wang, W., Flores, M. C. P., Sihn, C.-R., Kim, H. J., Zhang, Y., Doyle, K. J., … Yamoah, E. N. (2015). Identification of a key residue in K(v)7.1 potassium channel essential for sensing external potassium ions. *The Journal of General Physiology*, *145*(3), 201–212. http://doi.org/10.1085/jgp.201411280

Wang, Y. H., Jiang, M., Xu, X. L., Hsu, K.-L., Zhang, M., & Tseng, G.-N. (2011). Gating-Related Molecular Motions in the Extracellular Domain of the IKs Channel: Implications for IKs Channelopathy. *The Journal of Membrane Biology*, *239*(3), 137–156. http://doi.org/10.1007/s00232-010-9333-7

Wang, Y., Zhang, M., Xu, Y., Jiang, M., Zankov, D. P., Cui, M., & Tseng, G.-N. (2012). Probing the structural basis for differential KCNQ1 modulation by KCNE1 and KCNE2. *The Journal of General Physiology*, *140*(6), 653–69. http://doi.org/10.1085/jgp.201210847

Wass, M. N., Fuentes, G., Pons, C., Pazos, F., & Valencia, A. (2011). Towards the prediction of protein interaction partners using physical docking. *Molecular Systems Biology*, *7*(469), 469. http://doi.org/10.1038/msb.2011.3

Wen, H., & Levitan, I. B. (2002). Calmodulin is an auxiliary subunit of KCNQ2/3 potassium channels. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *22*(18), 7991–8001. http://doi.org/22/18/7991 [pii]

Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server issue), W407-10. http://doi.org/10.1093/nar/gkm290

Wiehe, K., Peterson, M. W., Pierce, B., Mintseris, J., & Weng, Z. (2008). Protein – Protein Docking : Overview and Performance Analysis. In M. Zaki & C. Bystroff (Eds.), *Protein Structure Prediction* (Second Edi, Vol. 413, pp. 283–314). Humana Press.

Wiener, R., Haitin, Y., Shamgar, L., Fernandez-Alonso, M. C., Martos, A., Chomsky-Hecht, O., … Hirsch, J. A. (2008). The KCNQ1 (Kv7.1) COOH terminus, a multitiered scaffold for subunit assembly and protein interaction. *The Journal of Biological Chemistry*, *283*(9), 5815–5830. http://doi.org/10.1074/jbc.M707541200

Wiener, R., Haitin, Y., Shamgar, L., Fernández-Alonso, M. C., Martos, A., Chomsky-Hecht, O., … Hirsch, J. a. (2008). The KCNQ1 (Kv7.1) COOH terminus, a multitiered scaffold for subunit assembly and protein interaction. *Journal of Biological Chemistry*, *283*(9), 5815–5830. http://doi.org/10.1074/jbc.M707541200

Williamson, A. R. (2000). Creating a structural genomics consortium. *Nat Struct Mol Biol*, *7*(953). http://doi.org/10.1038/80726

Witchel, H. J., & Hancox, J. C. (2000). Familial and acquired long qt syndrome and the cardiac rapid delayed rectifier potassium current. *Clinical and Experimental Pharmacology & Physiology*, *27*(10), 753–766.

Wodak, S. J., & Méndez, R. (2004). Prediction of protein-protein interactions: The CAPRI experiment, its evaluation and implications. *Current Opinion in Structural Biology*, *14*(2), 242–249. http://doi.org/10.1016/j.sbi.2004.02.003

Wojcikowski, M., Ballester, P. J., & Siedlecki, P. (2017). Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*, *7*, 46710. http://doi.org/10.1038/srep46710

Wrobel, E., Tapken, D., & Seebohm, G. (2012). The KCNE tango - how KCNE1 interacts with Kv7.1. *Frontiers in Pharmacology*, *3 AUG*(August), 1–14. http://doi.org/10.3389/fphar.2012.00142

Wu, E. L., Cheng, X., Jo, S., Rui, H., Song, K. C., Dávila-Contreras, E. M., … Im, W. (2014). CHARMM-GUI Membrane Builder toward realistic biological membrane simulations. *Journal of Computational Chemistry*, 1997–2004. http://doi.org/10.1002/jcc.23702

Wu, H., Chen, X., Cheng, J., & Qi, Y. (2016). SUMOylation and Potassium Channels: Links to Epilepsy and Sudden Death. *Advances in Protein Chemistry and Structural Biology*, *103*, 295–321. http://doi.org/10.1016/bs.apcsb.2015.11.009

Wu, S., & Zhang, Y. (2007). LOMETS: A local meta-threading-server for protein structure prediction. *Nucleic Acids Research*, *35*(10), 3375–3382. http://doi.org/10.1093/nar/gkm251

Wulff, H., Castle, N. A., & Pardo, L. A. (2009). Voltage-gated Potassium Channels as Therapeutic Drug Targets. *Nature Reviews. Drug Discovery*, *8*(12), 982–1001. JOUR. http://doi.org/10.1038/nrd2983

Xiang, Z. (2006). Advances in homology protein structure modeling. *Current Protein & Peptide Science*, *7*(3), 217–227.

Xiang, Z., & Honig, B. (2001). Extending the accuracy limits of prediction for side-chain conformations. *Journal of Molecular Biology*, *311*(2), 421–430. http://doi.org/10.1006/jmbi.2001.4865

Xiang, Z., Soto, C. S., & Honig, B. (2002). Evaluating conformational free energies: The colony energy and its application to the problem of loop prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(11), 7432–7437. http://doi.org/10.1073/pnas.102179699

Xu, D., Zhang, J., Roy, A., & Zhang, Y. (2011). Automated protein structure modeling in CASP9 by I-TASSER pipeline combined with QUARK-based ab initio folding and FG-MD-based structure refinement. *Proteins*, *79 Suppl 1*, 147–160. http://doi.org/10.1002/prot.23111

Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, *101*(10), 2525–2534. http://doi.org/10.1016/j.bpj.2011.10.024

Xu, J., & Berger, B. (2006). Fast and accurate algorithms for protein side-chain packing. *Journal of the ACM*, *53*(4), 533–557. http://doi.org/10.1145/1162349.1162350

Xu, Q., & Minor, D. L. (2009). Crystal structure of a trimeric form of the KV7.1 (KCNQ1) A-domain tail coiled-coil reveals structural plasticity and context dependent changes in a putative coiled-coil trimerization motif. *Protein Science*, *18*(10), 2100–2114. http://doi.org/10.1002/pro.224

Xu, X., Jiang, M., Hsu, K.-L., Zhang, M., & Tseng, G.-N. (2008). KCNQ1 and KCNE1 in the IKs channel complex make state-dependent contacts in their extracellular domains. *The Journal of General Physiology*, *131*(6), 589–603. http://doi.org/10.1085/jgp.200809976

Xu, Y., Wang, Y., Meng, X. Y., Zhang, M., Jiang, M., Cui, M., & Tseng, G. N. (2013a). Building KCNQ1/KCNE1 channel models and probing their interactions by molecular-dynamics simulations. *Biophysical Journal*, *105*(11), 2461–2473. http://doi.org/10.1016/j.bpj.2013.09.058

Xu, Y., Wang, Y., Meng, X., Zhang, M., Jiang, M., Cui, M., & Tseng, G. (2013b). Building KCNQ1 / KCNE1 Channel Models and Probing their Interactions by Molecular-Dynamics Simulations, *105*(December), 2461–2473. http://doi.org/10.1016/j.bpj.2013.09.058

Xu, Y., Wang, Y., Zhang, M., Jiang, M., Rosenhouse-Dantsker, A., Wassenaar, T., & Tseng, G.-N. (2015). Probing Binding Sites and Mechanisms of Action of an IKs Activator by Computations and Experiments. *Biophysical Journal*, *108*(1), 62–75. http://doi.org/10.1016/j.bpj.2014.10.059

Xun, S., Jiang, F., & Wu, Y. D. (2015). Significant refinement of protein structure models using a residue-specific force field. *Journal of Chemical Theory and Computation*, *11*(4), 1949–1956. http://doi.org/10.1021/acs.jctc.5b00029

Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2014). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, *12*(1), 7–8. http://doi.org/10.1038/nmeth.3213

Yang, T., Smith, J. a, Leake, B. F., Sanders, C. R., Meiler, J., & Roden, D. M. (2013). An allosteric mechanism for drug block of the human cardiac potassium channel KCNQ1. *Molecular Pharmacology*, *83*(2), 481–9. http://doi.org/10.1124/mol.112.081513

Yang, W.-P., Levesque, P. C., Little, W. a, Conder, M. L., Shalaby, F. Y., & Blanar, M. a. (1997). KvLQT1, a voltage-gated potassium channel responsible for human cardiac arrhythmias. *Proceedings of the National Academy of Sciences*, *94*(April), 4017–4021. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=20560&tool=pmcentrez&rendertype=abstract

Yang, Y., Faraggi, E., Zhao, H., & Zhou, Y. (2011). Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics (Oxford, England)*, *27*(15), 2076–2082. http://doi.org/10.1093/bioinformatics/btr350

Yang, Y., & Sigworth, F. J. (1998). Single-Channel Properties of I(Ks) Potassium Channels . *The Journal of General Physiology*, *112*(6), 665–678. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2229447/

Yap, Y. G., & Camm, A. J. (2003). Drug induced QT prolongation and torsades de pointes. *Heart*, *89*(11), 1363–1372. http://doi.org/10.1136/heart.89.11.1363

Yarov-Yarovoy, V., Baker, D., & Catterall, W. a. (2006). Voltage sensor conformations in the open and closed states in ROSETTA structural models of K(+) channels. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(19), 7292–7297. http://doi.org/10.1073/pnas.0602350103

Yellen, G. (1998). The moving parts of voltage-gated ion channels. *Quarterly Reviews of Biophysics*, *31*(3), 239–295. http://doi.org/10.1017/S0033583598003448

Yellen, G. (2002). The voltage-gated potassium channels and their relatives. *Nature*, *419*(6902), 35–42. http://doi.org/http://dx.doi.org/10.1038/nature00978

Yona, G., & Levitt, M. (2002). Within the twilight zone: a sensitive profile-profile comparison tool based on information theory. *Journal of Molecular Biology*, *315*(5), 1257–1275. http://doi.org/10.1006/jmbi.2001.5293

Yonath, A. (2011). X-ray crystallography at the heart of life science. *Current Opinion in Structural Biology*, *21*(5), 622–626. http://doi.org/http://dx.doi.org/10.1016/j.sbi.2011.07.005

Yu, F. H., & Catterall, W. A. (2004). The VGL-chanome: a protein superfamily specialized for electrical signaling and ionic homeostasis. *Science's STKE : Signal Transduction Knowledge Environment*, *2004*(253), re15. http://doi.org/10.1126/stke.2532004re15

Yus-Nájera, E., Santana-Castro, I., & Villarroel, A. (2002). The identification and characterization of a noncontinuous calmodulin-binding site in noninactivating voltage-dependent KCNQ potassium channels. *Journal of Biological Chemistry*, *277*(32), 28545–28553. http://doi.org/10.1074/jbc.M204130200

Zacharias, M. (2003). Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Science : A Publication of the Protein Society*, *12*(6), 1271–1282. http://doi.org/10.1110/ps.0239303

Zagotta, W. (2015). Ligand-dependent gating mechanism. *Handbook of Ion Channels*, 41–52. http://doi.org/doi:10.1201/b18027-6

Zaydman, M. A., Silva, J. R., Delaloye, K., Li, Y., Liang, H., Larsson, H. P., & Shi, J. (2013). Kv7 . 1 ion channels require a lipid to couple voltage sensing to pore opening, 1–6. http://doi.org/10.1073/pnas.1305167110/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1305167110

Zhang, C., & Ma, J. (2008). Comparison of sampling efficiency between simulated tempering and replica exchange. *The Journal of Chemical Physics*, *129*(13), 134112. http://doi.org/10.1063/1.2988339

Zhang, X., & Cheng, X. (n.d.). Structure of Protein, *34*, 978–981.

Zhang, Y. (2007). Template-based modeling and free modeling by I-TASSER in CASP7. *Proteins*, *69 Suppl 8*, 108–17. http://doi.org/10.1002/prot.21702

Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, *9*(1), 40. http://doi.org/10.1186/1471-2105-9-40

Zhang, Y. (2009). I-TASSER: fully automated protein structure prediction in CASP8. *Proteins*, *77 Suppl 9*, 100–113. http://doi.org/10.1002/prot.22588

Zhang, Y. (2014). Interplay of I-TASSER and QUARK for template-based and ab initio protein structure prediction in CASP10. *Proteins*, *82 Suppl 2*, 175–187. http://doi.org/10.1002/prot.24341

Zhang, Y., & Skolnick, J. (2004). SPICKER: a clustering approach to identify near-native protein folds. *Journal of Computational Chemistry*, *25*(6), 865–871. http://doi.org/10.1002/jcc.20011

Zhang, Y., & Skolnick, J. (2005). TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Research*, *33*(7), 2302–2309. http://doi.org/10.1093/nar/gki524

Zhang, Y., Zheng, N., Hao, P., Cao, Y., & Zhong, Y. (2005). A molecular docking model of SARS-CoV S1 protein in complex with its receptor, human ACE2. *Computational Biology and Chemistry*, *29*, 254–257. http://doi.org/10.1016/j.compbiolchem.2005.04.008

Zhang, Z., Schindler, C. E. M., Lange, O. F., & Zacharias, M. (2015). Application of Enhanced Sampling Monte Carlo Methods for High-Resolution Protein-Protein Docking in Rosetta. *Plos One*, *10*(6), e0125941. http://doi.org/10.1371/journal.pone.0125941

Zhorov, B. S., & Tikhonov, D. B. (2004). Potassium, sodium, calcium and glutamate-gated channels: pore architecture and ligand action. *Journal of Neurochemistry*, *88*(4), 782–799. http://doi.org/10.1111/j.1471-4159.2004.02261.x

Zhou, Y., Morais-Cabral, J. H., Kaufman, A., & MacKinnon, R. (2001). Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 A resolution. *Nature*, *414*(6859), 43–48. http://doi.org/10.1038/35102009

Zhu, J., Fan, H., Periole, X., Honig, B., & Mark, A. E. (2008). Refining Homology Models by Combining Replica-Exchange Molecular Dynamics and Statistical Potentials. *Proteins*, *72*(4), 1171–1188. JOUR. http://doi.org/10.1002/prot.22005

Zhu, K., Day, T., Warshaviak, D., Murrett, C., Friesner, R., & Pearlman, D. (2014). Antibody structure

determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins*, *82*(8), 1646–55. http://doi.org/10.1002/prot.24551

Zorko, M. (2009). Protein Folding. In *Introduction to Peptides and Proteins* (pp. 101–122). CRC Press. http://doi.org/doi:10.1201/b15106-11

APPENDIX A: MOLECULAR DYNAMIC SIMULATIONS

A.1. Introduction

Different techniques of Molecular Dynamic Simulations, both classical and advanced ones are employed in this Thesis. The applications of these MD simulations are discussed in Chapter 2 and 3 and therefore the current chapter serves as an introduction to these techniques. Understanding the structure and function of biological molecules (e.g. proteins, nucleic acids) has been an inspiring element for many researchers. Structural biology techniques such as X-ray crystallography, Nuclear Magnetic Resonance (NMR) spectroscopy and Cryo-electron microscopy (cryo-EM) have been quite successful in providing high-resolution atomic structures for these molecules. These structures have made a tremendous contribution to our understanding of molecular biology (Campbell, 2002). However, in many cases these are just static structures adopting an average conformation in time and space. These static images represent ensemble average properties and can only depict a minor illustration to what real biological systems behave in nature. Proteins, for example, are highly dynamic molecular entities and possess what is usually referred to as "jiggling and wiggling" at the atomic level (Mulholland, 2008). These movements are essential for the proteins to perform their vital actions. Accordingly, the dynamic conformational changes in the structure of proteins enable them to act as signalling molecules, transporters, catalysts, sensors, and mechanical effectors. In addition, it is the motion of atoms in the proteins that enables them to interact with indigenous chemicals, hormones, drugs, and other proteins (Dror, Dirks, Grossman, Xu, & Shaw, 2012). These movements can open new binding sites, close existing ones and induce significant changes to the studied molecules. Revealing all these information from a static image is not possible. In order to understand these changes, one should bring life to this static picture, allowing it to move, interact and reveal all its secrets.

A promising method to study the dynamical behaviour of macromolecules is all-atom molecular dynamics (MD) simulations (Adcock & McCammon, 2006). MD simulations enable us to zoom into the very minute dynamic motions of the atoms. Using MD simulations, it is possible to obtain information on the time evolution of

conformations of proteins and quantify the structural and biophysical properties of a system. Other fields where MD simulations are applicable include structural biochemistry, biophysics, enzymology, molecular biology, pharmaceutical chemistry, drug design and biotechnology (Mulholland, 2008).

The aim of this reiew is to discuss the key attributes of MD as a contributing technique in the study of biomolecular systems and proteins. First, the basic theory of classical MD simulation will be discussed, followed by various important aspects. Also, a brief description of enhanced MD simulations with an emphasis on Replica Exchange Molecular Dynamics (REMD) technique, is provided under this Appendix.

A.2. Classical MD Simulation

In classical MD simulation, the precise position of each atom at any instant in time for a single protein molecule are followed and iteratively calculated. In this regard, atoms are treated as solid spheres connected together via springs (see Figure A.1). The motions of the atoms are then computed based on Newton's equations of motion (see equation 1) in an iterative manner over varying time-scales.

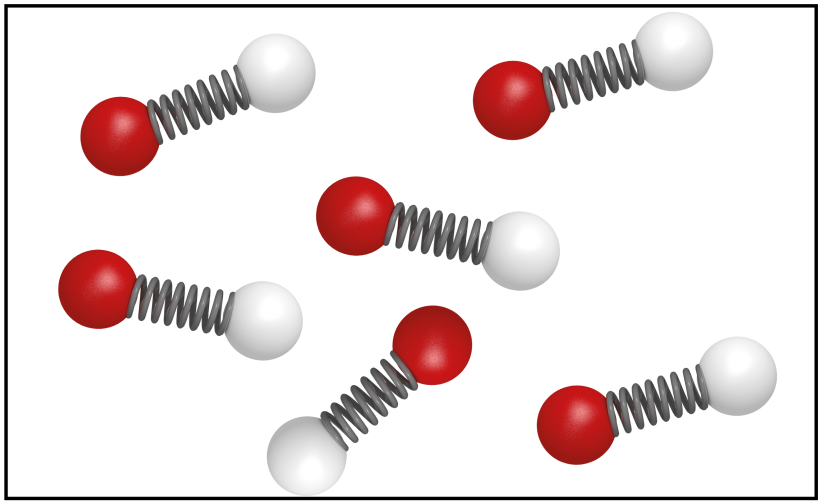


Figure A.1. Spring-atom model of water molecules.

There are three primary requirements to perform an MD simulation. First, one needs a set of initial conditions including the positions and velocities of each particle in the system. The positions can either be obtained from an experimental structure, e.g. X-ray crystal structure or assigned randomly in the case of a disordered system. The

velocities are specified randomly using a Maxwellian distribution centered on the desired temperature and then they are adjusted in order to zero the angular momentum and the center of mass velocity of the total system (González, 2011). The second requirement is a force field; which is a set of parameterized terms and energy functions and also defines the forces acting between the particles. These parameters are obtained either from experimental and/or quantum mechanical studies of small molecules or fragments, that are considered to be transferrable to larger systems (González, 2011). The force field defines the bonded (bonds, angles, and dihedrals) and non-bonded (van der Waals potentials and Coulomb potentials) energies between all the atoms in the system. Choosing the suitable type of energy function to describe the intermolecular and intra-molecular interactions is an essential prerequisite for a successful MD simulation (Ponder & Case, 2003). Some of the top notch force fields developed for biological systems and proteins are CHARMM (MacKerell et al., 1998), AMBER (Cornell et al., 1995) and GROMOS (Oostenbrink, Villa, Mark, & van Gunsteren, 2004).

Finally, one needs to specify the surrounding space of the system. As simulation can run in vacuum (although this will not be physiologically relevant) or in solvent (usually water). Having all the essential components, next we need to solve the classical equation of motion. For a simple atomic system:

$$m_i \frac{d^2 r_i}{dt^2} = F_i \qquad \textbf{Eqn. A.1}$$

Where m_i is the mass of atom i, at position r_i and at time t. F_i is the net force acting on the ith atom. The force (F_i) can then be computed as:

$$F_i = -\frac{\delta}{\delta r_i} U(r_1, r_2, ..., r_N) \qquad \textbf{Eqn. A.2}$$

Where $U(r_1, r_2, r_3, \ldots r_N)$ is the potential energy depending on the coordinates of the N particles, and specified by the force field. The equations can be solved at each step of the MD simulation for all the atoms in the system. However, it is impractical to numerically solve the equations, for the thousands of degrees of freedom that may exist in an MD

simulation. Therefore, a time integrator is used to advance over small time steps. Numerical integration algorithms include the Verlet integrator (Verlet, 1967), velocity Verlet integrator (Swope, Andersen, Berens, & Wilson, 1982), and leapfrog integrator (Cuendet & van Gunsteren, 2007).

In classical MD, the most commonly used time integration algorithm is the Verlet algorithm. The basic form of the Verlet integration algorithm is shown below in Eqn. A.3, where the positions are **r(*t*)**, one forward and one backward in time, **v** is the velocity, **a** is the acceleration and **b** is the third derivatives of **r** with respect to ***t***. The Taylor expansions of +**Δ***t,* -**Δ***t,* the terms in **Δ***t* , **Δ***t*^{3} etc. are summed, cancelled and then we obtain what is shown in Eqn. A.4:

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + (1/2)\mathbf{a}(t)\Delta t^2 + (1/6)\mathbf{b}(t)\Delta t^3 + O(\Delta t^4) \quad \textbf{Eqn. A.3}$$

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)\Delta t^2 + O(\Delta t^4) \quad \textbf{Eqn. A.4}$$

Since we are integrating Newton's equations of motion, **a(*t*)** is just the force divided by the mass, and the force is in turn a function of the positions **r(*t*)**

$$\mathbf{a}(t) = -(1/m)\nabla V(\mathbf{r}(t)) \quad \textbf{Eqn. A.5}$$

The velocities are not directly calculated in the algorithm, but they can be obtained as:

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t} \quad \textbf{Eqn. A.6}$$

A.3. Running a MD Simulation

To run a classical MD simulation, several basic steps need to be followed. Similarly, many decisions are taken by the user, based on the purpose of simulation and the type of system under study. Figure A.2, shows a general workflow for a MD simulation. The basic MD parameters, over which the user usually has a choice, are described here.

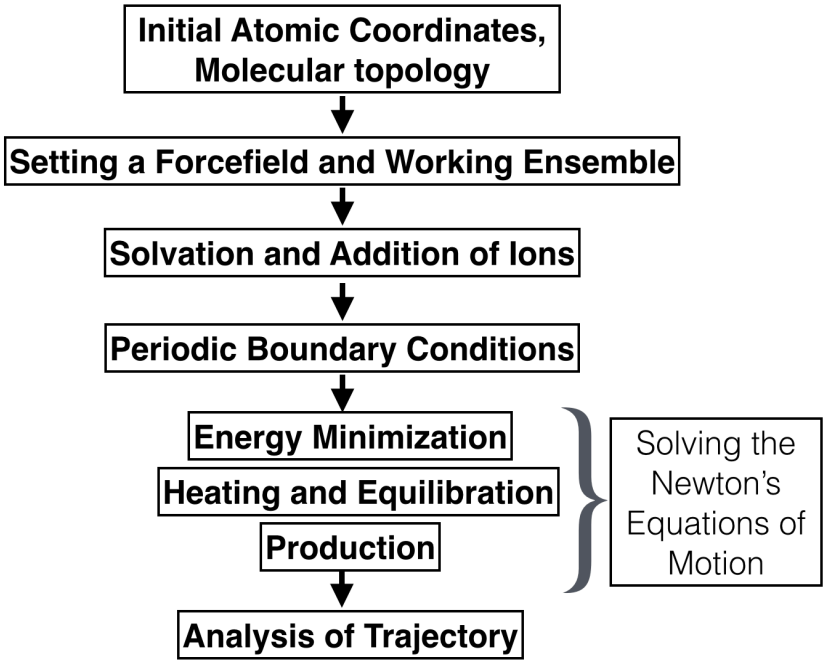


Figure A.2. Workflow of a classical MD simulation.

1. Choosing a force field and the MD tool

As described earlier, a force field is a set of mathematical expressions and parameters, which describe the relation between the energy of a system on the atomic coordinates. It consists of an analytical form of the interatomic potential energy, U (r_1, r_2, ... , r_N) and a set of parameters entering into this form. These parameters include information related to bond stretching, angle bending, dihedral and improper angles, intramolecular and intermolecular interactions such as electrostatic and Van der Waals energies, etc. (González, 2011). Force field parameters are obtained either from *ab initio* or semi-empirical quantum mechanical calculations or by fitting to experimental data such as neutron, X-ray and electron diffraction, NMR, infrared, Raman and neutron spectroscopy, etc.

An ideal force field should be simple to facilitate a quick evaluation but also detailed to reproduce reliable data. There are a large number of force fields available. They mainly differ in their level of complexity and are optimized for specific kinds of systems. Some of the popular force fields include CHARMM (MacKerell et al., 1998), AMBER (Cornell et al., 1995) and GROMOS (Oostenbrink et al., 2004), OPLS(Jorgensen, Maxwell, & Tirado-Rives, 1996), and COMPASS (H. Sun, 1998). The

first three force fields are optimized for biomolecules, the latter two were originally developed to simulate condensed matter.

Similar to force fields, there are various tools available to perform the MD simulation. Some of the most commonly used MD tools for simulating biological systems are NAMD (Phillips et al., 2005), CHARMM (Brooks et al., 1983), AMBER (D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, 2015), GROMACS (Pronk et al., 2013), and Desmond (Shivakumar et al., 2010). The different MD programs possess several common basic features, but do vary in their capabilities, algorithms and applicability to different problems. The choice of the MD program also depends on the level of expertise and computational knowledge of the user. For example, CHARMM (Brooks et al., 1983) demands high skills of a complex scripting language, but instead rewards the user with the possibility of performing a wide variety of professional simulations, analyses and manipulations to suit the purpose of the study. NAMD, on the other hand, is a user-friendly tool with a simpler scripting language that can be equally handled by the inexperienced users (Salsbury, 2010).

GROMACS and AMBER are similar to NAMD in their scope and complexity. GROMACS is well known for its several tools that can be used in a variety of trajectory analysis. Amongst all, NAMD has been the most frequently used package for long all atom MD simulations for large proteins and protein complexes. Some examples include

2. Input Information and Files

Each MD tool requires a number of input information; the most basic of them are an atomic coordinates file, parameter file and topology files. The coordinate file contains the positional information of each atom relative the others in space. It may be a PDB coordinate file or modifications of the same. The parameter and topology files contain information related to the different residues and their connectivity. These files are the least required components for the program to perform further operations on the protein

such as energy determination, minimization, residual interactions, structure modification, or the running of dynamics simulations (Schleif, 2004).

3. Working Ensemble (NVE, NVT or NPT)

One of the objectives of simulating biomolecular systems is to mimic the experimental and/or physiological conditions that a protein experiences in an experiment or in the human body (Uline & Corti, 2013). These conditions that can be adjusted accordingly, include temperature, pressure, number and type of solvent molecules and ionic concentrations, etc. Some of these parameters are specified with the help of statistical mechanics' ensembles that act as thermostats and barostats during an MD simulation. An ensemble is a set of all the possible systems that have an identical macroscopic or thermodynamic state but a different microscopic state, such as temperature, pressure and volume (Scheraga, Khalili, & Liwo, 2007).

The most basic ensemble is the microcanonical ensemble or the NVE ensemble, in which the number of particles N, the volume of the simulation cell V, and the total energy of the system E, are kept constant. However, this ensemble is associated with unrealistic energy drifts and since we generally prefer to compare the simulation results with experiment, we need to control the temperature and pressure (Uline & Corti, 2013). The other types of ensembles are the canonical type (NVT) and the isothermal-isobaric ensemble (NPT). In the latter two types temperature and pressure are controlled using several types of thermostats and barostats, namely the Langevin (Adelman & Doll, 1976), Berendsen (Berendsen, Postma, van Gunsteren, DiNola, & Haak, 1984), and Nose-Hoover (Hoover, 1985; Nosé, 1984).

4. Solvent and Ions

As mentioned above, the user has a choice of either simulating the system under study, in vacuum or choosing a type of liquid to immerse the protein. Many of the protein-related phenomena such as folding and thermodynamic stability are dependent on the surrounding environment. Since water constitutes the environment in which the majority of proteins interact, it is usually the solvent of choice (Mashayak & Tanner, 2011). There are several types of water models that mimic the specific nature of water molecules in

MD. These models are obtained from quantum mechanics, molecular mechanics, experimental data and/or a combination of these methods. The difference between the various water models are mainly the geometrical properties, polarization characteristics and interaction points (Wallqvist & Mountain, 2007).

The water component used during an MD simulation can either be explicit or implicit (see Figure A.3). In the former, the water molecules are individually present in the core of the solvent. Each water molecule can interact with the protein or other molecules in the system independently.

Table A.1. Some of the existing water models used in MD simulations. Source: (Gonzalez, Noya, Vega, & Sese, 2010)(J. Wang et al., 2012)

Water Models	Types
SPC	• Flexible (SPC/F) • Rigid (SPC/E)
TIP3P	• Original TIP3P • Modified CHARMM version
TIP4P	• Flexible (q-TIP4P/F) • Rigid (TIP4P/2005 & TIP4P/Ew)
TIP5P	• Rigid
POL3	• POL3-ET • POL3-LT

The most common explicit water models are SPC (Berendsen, Postma, van Gunsteren, & Hermans, 1981), SPC/E (Berendsen, Grigera, & Straatsma, 1987) available in GROMOS force field, TIP3P (Jorgensen & Madura, 1983) in AMBER and CHARMM, TIP4P (Jorgensen & Madura, 1983) and TIP5P (Mahoney & Jorgensen, 2000) in OPLS force field. In the implicit model, water is represented as a continuous medium. It is the method of choice wherein the properties and distribution of individual water molecules in the solvent-solute interface is not of interest. Implicit solvent models are not as computationally demanding as the explicit mode of solvent representation and result in faster simulations. Some common models in this category are Solvent accessible surface area (SAS) models, Poisson-Boltzmann and Generalized Born models (Anandakrishnan, Drozdetski, Walker, & Onufriev, 2015).

Although water is the solvent of choice for proteins, the user has a number of other choices over the type of solvent. Organic solvents for example are such alternatives used for specific purposes, such as mimicking a highly hydrophobic environment which can affect protein folding, structure and properties (Norin, Haeffner, Hult, & Edhoim, 1994). Enzymes for example, exhibit a number of unique properties such as higher catalytic power, increased stability, enantiomeric selectivity, etc., and have found numerous potential applications. Some of the commonly used organic solvents for this purpose are hexane, ethanol, acetonitrile, diisopropyl ether, etc. Almost all packages and force fields enable the use of organic solvents in an MD simulation (Klibanov, 2001; Rehm, Trodler, & Pleiss, 2010).

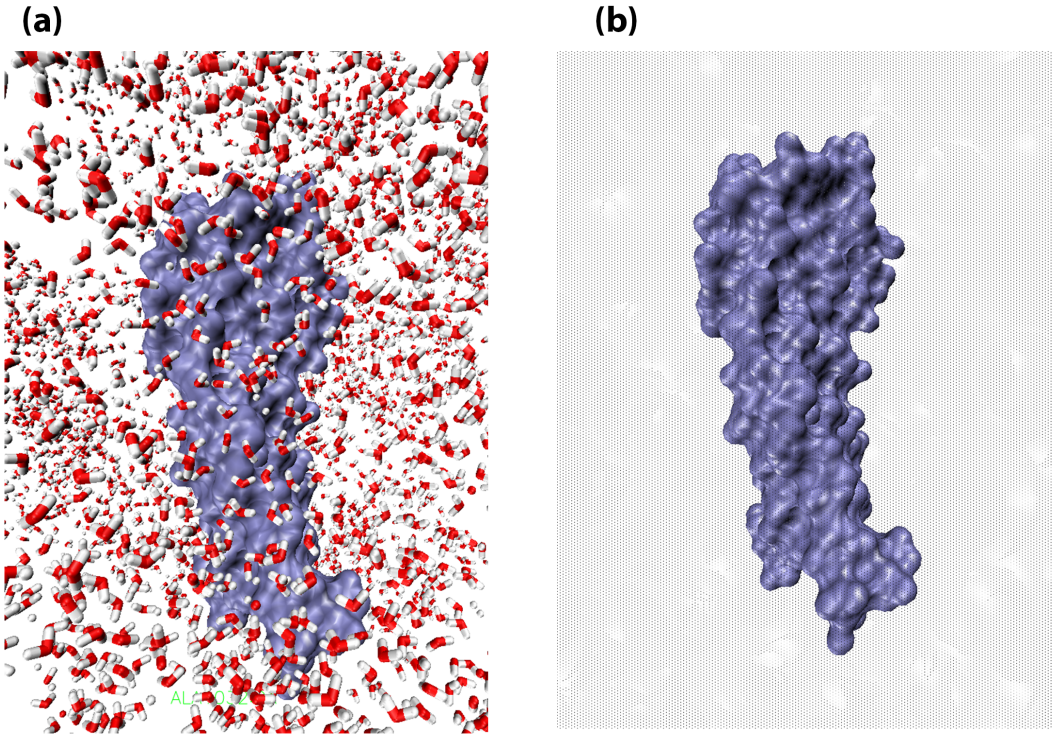


Figure A.3. Explicit (a) and implicit (b) water models. In the explicit water model the water is physically represented by atoms and bonds with full atomic details. In the implicit models, water is included as an effective (averaged) interaction with no atomic detail.

Another important step in setting up a system for MD simulation is to neutralize the existing charge of proteins and bring the net charge of the system down to zero. This

is mainly done by the addition of sodium and chloride to neutralize the excess negative and positive charge, respectively. After neutralization, it is then critical to add the right amount of counter ions to maintain the right ionic concentration, i.e. the one existing in physiological environments. Both of these steps are necessary for the stability of the protein structure during the MD simulation (Ibragimova & Wade, 1998). However, there are also some controversies discussed in (Drabik, Liwo, Czaplewski, & Ciarkowski, 2001)

5. Periodic Boundary Conditions

Boundary conditions are defined during an MD simulation to confine the movement of the molecules in the system. When proteins are simulated in a bulk of solid and/or liquid, the number of atoms in the entire system significantly increases. To make the calculations easier and reduce the computational demand, periodic boundary conditions (PBCs) are used (Allen, 2004).

PBCs resemble a box in which the atoms are enclosed. This box has several copies that are replicated to infinity by translation in all the three Cartesian directions. The images of each particle from the simulated box move accordingly. Such that when a water molecule, exits to the right side of the box it reappears on the left. The box should be sufficiently large so that molecules would not interact with their copies. This depends on the non-bonded type of interactions, which should be summed over all neighbours in the resulting infinite periodic system. In this way surface artifacts are avoided. Another advantage of using a PBC is that macroscopic properties can be simulated and thereby calculated from lesser number of particles (Scheraga et al., 2007).

6. Energy Minimization

Energy minimization is done before initiating MD, to provide a reasonable starting structure in terms of geometry and solvent orientation. The ideal outcome from an energy minimization is an optimized arrangement of the atoms that possess a local minimum or less frequently a global minimum. This is a stable point or a minimum on the potential energy surface (PES) where the net force on each atom is close to zero (Adcock & McCammon, 2006).

In case of proteins, known constraints from experiments can be used during the minimization phase. In this way, one can guide the system towards the right energy minimum. For proteins, a minimized structure would be the closest to its corresponding natural state and therefore the conformational search during the following MD simulation will be more meaningful (Lonsdale, Harvey, & Mulholland, 2012). In addition, there are always a number of atoms, which are too close to each other, causing steric clashes in the system. These structural issues may lead to huge drifts in the energy of the system at later stages. Constrained type of equilibration ensures that these unfavourable events are removed before starting the production run (Ramachandran, Kota, Ding, & Dokholyan, 2011).

7. Equilibration

Before initiating an MD production, equilibrium has to be achieved for the system. An equilibrated structure is not necessarily the lowest energy configuration but it is a favourable state for the system with the given amount of energy. This step is done for a couple of reasons. First, a configuration from the minimization is at zero kelvin temperature. Thus, it is necessary to adjust the temperature of the system and to raise it to the desired range (González, 2011). For that, the velocities are assigned at a low temperature first and then a few steps of dynamics are performed. The temperature is gradually increased during several more iterations, until the desired temperature is reached. During the first phase of equilibration, the "isothermal-isochoric" NVT working ensemble (described in the previous sections) is used. It is continued until the temperature of the system attains a plateau at the desired range of temperature. Following this, the "isothermal-isobaric" NPT ensemble is employed to achieve equilibrium density consistent with the desired pressure and temperature (Uline & Corti, 2013).

It is critical to extend the equilibration for a sufficient duration of time. To ensure an equilibrated state is achieved, one needs to follow and monitor the different system parameters such as pressure, density and the energy components. RMSD is another parameter, which can define when the system has reached an equilibrium state (M. Karplus & McCammon, 1983). More or less, all the mentioned parameters should fluctuate around some average value without showing any huge drift.

8. Production

After implementing the two equilibration phases, the system is now well equilibrated with respect to the desired temperature and pressure. The MD simulation is ready to record the positions and momenta of the atoms as a function of time. This is where the integration of Newton's equations of motion is applied to simulate atomic movements, vibrations and interactions based on the force field. During MD simulation, quite different molecular conformations of the system are sampled. This sampling, if sufficient, would lead to the observation of important properties or events during an MD simulation (Andrew, 1996).

The output of a production run is a trajectory file that records the positions and relevant information about the system at each time step. It is important to decide upon the right time step for an MD simulation. The time step being too small may improve the accuracy of the numerical solution of equations, however this might be at the expense of more computational resources and time to complete the required length of simulation. Conversely, a very large time step might lead to energy fluctuations and instabilities during the course of the simulation.

Ideally, we also need to calculate the non- bonded interactions for each time step during the simulation. Increasing the time step is an alternative for advancing simulation performance. However, this is an erroneous solution, since inaccuracies of bond vibrations begin even at a time-step of one femtosecond. However, the bond vibrations and dynamics of covalent bonds involving hydrogen atoms are not of interest per se, and can be completely removed using a bond constraint algorithm. These include SHAKE (Ryckaert, Ciccotti, & Berendsen, 1977), LINCS (Hess, Bekker, Berendsen, & Fraaije, 1997), and RATTLE (Andersen, 1983). Introducing constraints is another way of extending the time step to 2 fs, and fixed-length bonds represent better approximations of the quantum mechanical grounds state compared with the harmonic springs. As such, for biological systems and proteins, a time-step of 1.5 up to 2 fs is shown to be suitable during an MD simulation to observe the required phenomena (González, 2011).

Another important choice is the right amount of time to run the simulation. The length of the production run is significantly depended on the desired outcome as well as the available computational resources. However, lack of sufficient sampling in MD simulation has always been a challenge in observing important characteristics of biomolecular systems. The latter topic is elaborated more in the following sections covering enhanced techniques of MD.

9. Analysis

Once the system is simulated for a suitable amount of time, different types of analysis can be done to obtain information about the system and the evolution of different events during the simulation. The collected trajectory data of the protein at every time step is analyzed for finding the proteins motions, diffusion coefficient, Root Mean Square Deviation (RMSD) calculation, Radius of Gyration, Radial distribution, interaction energies, or any other relevant quantifiable property from the simulation. The type of analysis depends on the biological question that is being asked (Ganesan, Coote, & Barakat, 2017a).

Figure A.4 shows some of the most common and primary analysis of MD simulations. Figure A.4.a displays the RMSD graph for a protein. RMSD is simply the distance between the different conformations that a protein attains with respect to their initial structure, during a MD simulation. In the example shown the protein fluctuated at about 2 Angstroms, which is within the acceptable range of RMSD (2-5 A). Typically, a protein should structurally remain stable or stabilize over time and have an almost constant value of RMSD during a simulation. A non-converging RMSD graph indicates that there are still energetic and structural problems in the system that have led to huge structural drifts. This problem can usually be removed by elongating the equilibration time.

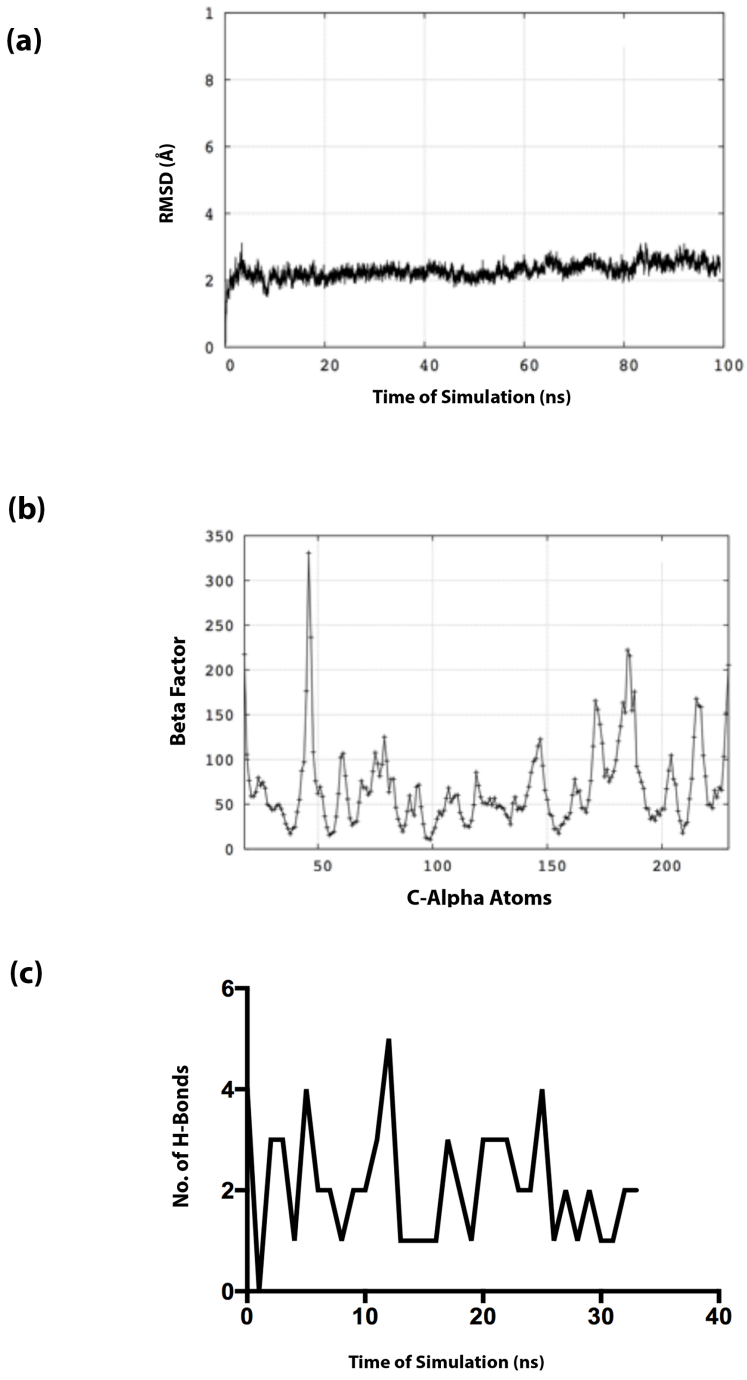


Figure A.4. Examples of graphs for MD trajectory analysis. (a) RMSD graph, (b) RMSF graph (c) Hydrogen bonding analysis.

Figure A.4 (b) is an example of a Relative Mean Square Fluctuations (RMSF) graph. It is simply a measure of the average atomic mobility of the backbone (usually C-alpha) atoms in the protein with respect to its initial coordinates. The RMSF values usually correlate with the secondary structure of the protein. Alpha helices and beta sheets are structurally more stable whereas loops have a higher rate of motions. RMSF graphs can also be used as a measure of the interaction of proteins with either another protein or drug molecules. When a protein has lesser mobility, it could imply that it is engaged in some sort of atomic interaction rather than being free to move. However, a more precise and detailed type of analysis to understand the interaction of proteins during a MD simulation is by analyzing hydrogen bond, salt-bridge and hydrophobic interactions individually. Figure A.4 (c) shows the frequency of hydrogen bond formation of one amino acid residue during the simulation time.

A.4. Enhanced MD simulations

Insufficient sampling has often been a limitation for the applicability of MD to different problems. This limitation is due to rough energy landscapes, with many local minima separated by high-energy barriers, which govern the biomolecular motion. The sampling problem mainly refers to the fact that MD simulation does not cover all the relevant conformational states. As a result, the dynamics and function of the molecule of interest cannot be meaningfully characterized (Bernardi, Melo, & Schulten, 2015). For example, the huge conformational changes or transitions of proteins, which are substantial to their function, may not be completely captured by classical MD simulations. Another example is transport through membrane proteins, channels and transporters that have to undergo large conformational changes to allow the influx or efflux of substances (Khalili-Araghi et al., 2009). These intricate and time-consuming processes are beyond the ability of straightforward MD simulations and enhanced sampling algorithms are needed.

Several methods have been developed to address the sampling problem and to explore events that occur on timescales, inaccessible to classical all-atom MD simulation (Mitsutake, Mori, & Okamoto, 2013). These include metadynamics, replica exchange molecular dynamics (REMD), random acceleration molecular dynamics (RAMD), steered molecular dynamics (SMD), umbrella sampling and adaptive bias force steering

(ABFS) (Abrams & Bussi, 2014). Here we present a brief overview of the REMD technique, which is one of the tools employed in this thesis.

A.4.1. Replica Exchange Molecular Dynamics (REMD)

REMD also known as parallel tempering is a popular enhancement of the classical MD approach. The REMD algorithm evolves numerous independent copies of a system, each using a separate process. The copies, or replicas are identical, except for their temperature. The replicas periodically exchange temperatures with their neighbours, through a Monte Carlo (MC) move that maintains detailed balance. By iteratively exchanging target temperatures with neighbouring replica systems, a given replica is capable of sampling a wide range of temperatures, enhancing conformational samplings (Sugita & Okamoto, 1999). Figure A.5 displays a general scheme for a REMD simulation of four replicas simulated at temperatures T1, T2, T3, and T4.

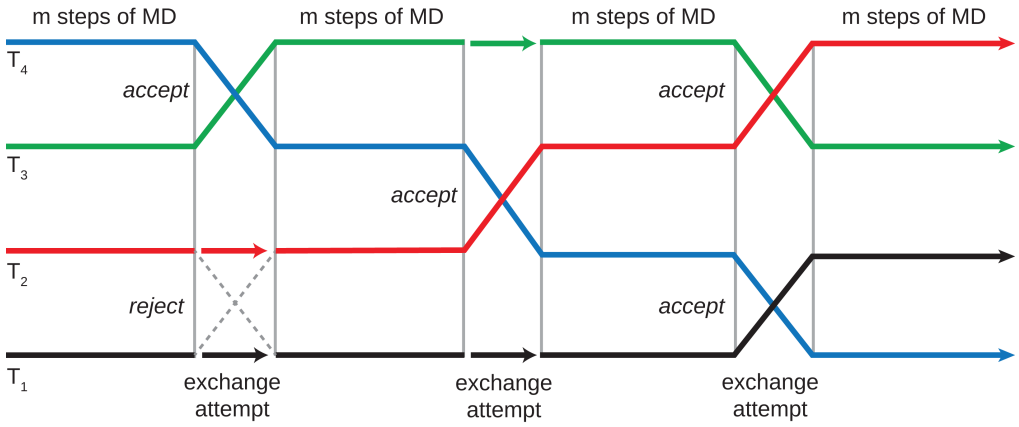


Figure A.5. A general scheme of REMD of four replicas simulated at temperatures T1, T2, T3, and T4. Source: (Gaalswyk & Rowley, 2016) **(DOI: 10.7717/peerj.2088/fig-2)**

A comparison of this algorithm with constant temperature classical MD applied to peptides at room temperature has shown that this algorithm decreases the sampling time by factors of 20 or more (Sanbonmatsu & García, 2002). One of the main applications of REMD is the investigation of transitions in proteins secondary structure and the *ab initio* folding of peptides from first principles (Earl & Deem, 2005). For example, increases in

temperature provided by the replica exchange technique enable transitions from non-helical to helical structures which is often not possible due to kinetic trapping (Gnanakaran, Hochstrasser, & García, 2004; Jas & Kuczera, 2004; Nymeyer, Gnanakaran, & Garc??a, 2004; Sanbonmatsu & García, 2002).

Replica-exchange molecular dynamics arises by applying the parallel tempering method to molecular dynamics (MD) simulation. As described earlier, the REMD algorithm simulates multiple replicas in parallel at a sequence of increasing temperatures ($T_0, T_1, ...T_n$), whose coordinates are represented by (q_1, q_2,...., q_n) and alternatingly attempt to exchange simulations between temperatures. Such that, every L steps, two chains j and k are randomly chosen (or j random and k = j + 1), and the main basis of whether an exchange is accepted or rejected is given by the Metropolis ratio below:

$$\min\left\{1, \frac{\pi_{T_k}(\mathbf{q}_j, \mathbf{p}_j)\pi_{T_j}(\mathbf{q}_k, \mathbf{p}_k)}{\pi_{T_j}(\mathbf{q}_j, \mathbf{p}_j)\pi_{T_k}(\mathbf{q}_k, \mathbf{p}_k)}\right\} \qquad \textbf{Eqn. A.7}$$

The distribution $\pi_{Ti}(q_j, p_j)$ is the Boltzmann distribution for replica *j* at temperature *Ti* is given below:

$$\pi_T(\mathbf{x}) = \frac{e^{-E(\mathbf{x})/(k_B T)}}{\int_{(\mathbf{q},\mathbf{p})} e^{-E(\mathbf{x})/(k_B T)}} \qquad \textbf{Eqn. A.8}$$

The resulting REMD algorithm is a stochastic dynamical system on X = R2dn. The added stochastic element enables the crossing of large energy barriers and escaping a local minimum by making conformations accessed at higher temperatures available to those at lower temperatures, thus enhancing sampling capability (Maximova, Moffatt, Ma, Nussinov, & Shehu, 2016). This has given REMD the power for simulations of complex molecules such as polymers, peptides and proteins (Isard, 2008). In addition, the remarkable ability to parallelize REMD and the evidence that parallel tempering simulations equilibrate dramatically faster, adds to their efficiency (C. Zhang & Ma, 2008).

A.4.2. Steered Molecular Dynamics (SMD)

Steered MD (SMD) simulation is another variation of MD categorized as an enhanced sampling technique. It is successfully applied to answer a number of questions related to association and disassociation of ligands to or from proteins, to reveal conformational changes in biomolecules (Ganesan, Coote, & Barakat, 2017b). With ion channels and membrane transporters, SMD techniques are applied to investigate the ion permeation phenomena (Khalili-Araghi et al., 2009).

In SMD simulations, a time-dependent external force is applied to an atom or group of atoms to facilitate their unbinding from the protein, which usually cannot be achieved by standard MD simulation (Patel, Berteotti, Ronsisvalle, Rocchia, & Cavalli, 2014). In particular, in SMD the transition between two states, for example the bound and unbound ones, is achieved by adding to the standard Hamiltonian, a harmonic time-dependent potential U(r,t) acting on a descriptor s(r) (e.g., the protein-ligand distance), which holds the following time dependency:

$$U(\boldsymbol{r}, t) = \frac{k}{2}\,[s(\boldsymbol{r}) - s_0(t) - vt]^2 \qquad \textbf{Eqn. A.9}$$

Where s_0 is the value of the descriptor in the initial state, $\boldsymbol{t}$ is the time and k is the constant representing the applied force of pulling. After a specified amount of time, the harmonic constraint will be centered in its final position, which is represented as:

$$s_1 = s_0 + vt_1 \qquad \textbf{Eqn. A.10}$$

During this transition, the value of the exerted force, F is calculated using the following equation:

$$F(t) = -k[s(r) - s_\lambda(t) - vt \qquad \textbf{Eqn. A.11}$$

Figure A.6 represents the force profile resulting from pulling of an ion through the KCNQ1/KCNE1 ion channel. The external work ΔW performed on the system can be calculated by integrating the power along the entire transition time:

$$\Delta W = v \int_{t0}^{t1} F(t)\, dt \qquad \textbf{Eqn. 5.12}$$

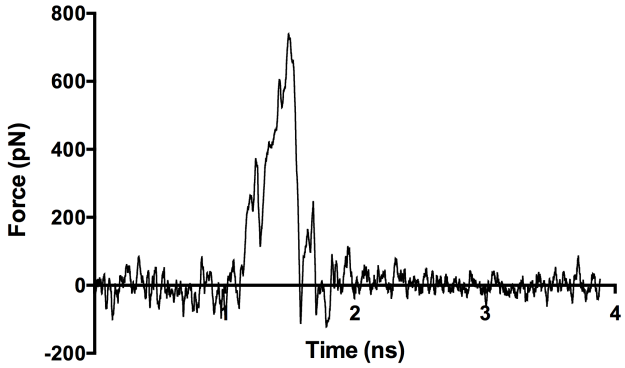


Figure A.6. The force profile resulting from the pulling of a K^+ ion through the KCNQ1/KCNE1 channel complex. The peak of the plot indicates the highest energy barrier that the ion has experienced during the time of the SMD simulation.

There are two variations of SMD, namely, constant-force SMD and constant-velocity SMD (See Figure A.7). In the former type, the atom is pulled with a constant amount of force. The changes that take place in the velocity as a result of adding this force are recorded during the simulation, which can be an indication of the obstacles and their effect on the movement of the pulled entity. The constant-velocity SMD simulation involves pulling of the selected atom or groups of atoms at a constant velocity. The force applied on the atom is then recorded, allowing one to estimate the potential of mean force (PMF) using the Jarzynski equality (Jarzynski, 1997) which averages the work over a large ensemble of simulations. The amount of force or the velocity employed in either of the variations depends on the type of property being studied as well as the stability of the system. However, benchmarking of different parameters over years have provided certain ranges for the force and velocity to be used as general starting points (Ganesan et al., 2017b).

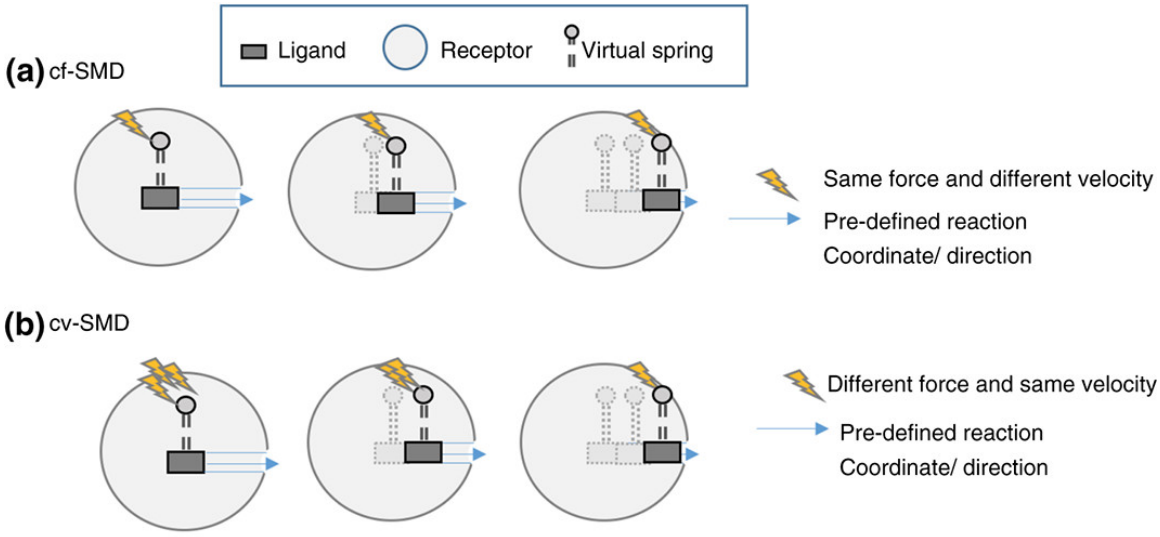


Figure A.7. Schematic diagrams describing the process of constant-force SMD (a), and constant-velocity SMD (b). In constant-force SMD a predetermined force in a pre-determined direction is applied onto the atom(s). In constant-velocity SMD, the atom is pulled in a pre-determined with a constant velocity and spring constant. The force profiles for the entire reaction coordinates are then recorded.
Adopted with permission from: (Ganesan et al., 2017b)

A.5. Summary

MD simulations are valuable tools for studying biomolecular systems and complexes. Specifically, different MD techniques have allowed the exploration of the dynamic properties, structure and inaccessible events related to proteins. The main output from a simulation consists of a phase space trajectory and it is in the hand of the user to judicially extract relevant information about the system under study.

Although the aim of a simulation is not to reproduce an experimental result, but they allow the understanding of the microscopic origin of physical properties or prediction of a certain biological behaviour that is often not accessible experimentally. Nowadays, many experimentalists rely on simulation techniques to analyze and interpret their complex experiments and/or systems at a spatial resolution.

Due to the development of specialized hardware and better parallelization algorithms all-atom simulations of proteins can now reach timescales in excess of a

millisecond. However, this is still a computationally expensive and time-consuming process for the majority of the scientific community. Currently, the advent of enhanced modifications of classical MD simulations has shielded its limitations and shortcomings. These developments, combined with the improvements to the force field models that underlie MD simulations, have allowed MD to capture atomistic detail processes such as the conformational transitions essential to protein function, the folding of proteins to their native structures, the transport of small molecules across cell membranes, and the binding of drugs to their targets. Examples of these successful approaches are REMD and SMD techniques. In this thesis, we have applied both of these enhanced techniques of MD to improve the sampling of different properties, as will be discussed in the following chapters.

A.6. References

Abrams, C., & Bussi, G. (2014). Enhanced sampling in molecular dynamics using metadynamics, replica-exchange, and temperature-acceleration. Entropy, 16(1), 163–199. http://doi.org/10.3390/e16010163

Adcock, S. A., & McCammon, J. A. (2006). Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chemical Reviews*, *106*(5), 1589–1615. http://doi.org/10.1021/cr040426m

Adelman, S. A., & Doll, J. D. (1976). Generalized Langevin equation approach for atom/solid-surface scattering: General formulation for classical scattering off harmonic solids. *The Journal of Chemical Physics*, *64*(6), 2375. http://doi.org/10.1063/1.432526

Allen, M. (2004). Introduction to molecular dynamics simulation. *Computational Soft Matter: From Synthetic Polymers to Proteins*, *23*(2), 1–28. http://doi.org/10.1016/j.cplett.2006.06.020

Anandakrishnan, R., Drozdetski, A., Walker, R. C., & Onufriev, A. V. (2015, March). Speed of Conformational Change: Comparing Explicit and Implicit Solvent Molecular Dynamics Simulations. *Biophysical Journal*. http://doi.org/10.1016/j.bpj.2014.12.047

Andersen, H. C. (1983). Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, *52*(1), 24–34. http://doi.org/http://dx.doi.org/10.1016/0021-9991(83)90014-1

Andrew, R. L. (1996). *Molecular modeling principles and applications*. *Longman*.

Berendsen, H. J. C., Grigera, J. R., & Straatsma, T. P. (1987). The Missing Term in Effective Pair Potentials. *Journal of Physical Chemistry*, *91*(24), 6269–6271. http://doi.org/10.1021/j100308a038

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, a, & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, *81*(1984), 3684–3690. http://doi.org/10.1063/1.448118

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981). Interaction Models for Water in Relation to Protein Hydration. In B. Pullman (Ed.), *Intermolecular Forces: Proceedings of the Fourteenth Jerusalem Symposium on Quantum Chemistry and Biochemistry Held in Jerusalem, Israel, April 13--16, 1981* (pp. 331–342). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-94-015-7658-1_21

Bernardi, R. C., Melo, M. C. R., & Schulten, K. (2015). Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochimica et Biophysica Acta (BBA) - General Subjects*, *1850*(5), 872–877. http://doi.org/http://doi.org/10.1016/j.bbagen.2014.10.019

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*(2), 187–217. http://doi.org/10.1002/jcc.540040211

Campbell, I. D. (2002, May). Timeline: the march of structural biology. *Nature Reviews. Molecular Cell Biology*. England. http://doi.org/10.1038/nrm800

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., … Kollman, P. A. (1995). A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, *117*(19), 5179–5197. http://doi.org/10.1021/ja00124a002

Cuendet, M. A., & van Gunsteren, W. F. (2007). On the calculation of velocity-dependent properties in molecular dynamics simulations using the leapfrog integration algorithm. *The Journal of Chemical Physics*, *127*(18), 184102. http://doi.org/10.1063/1.2779878

D.A. Case, J.T. Berryman, R.M. Betz, D.S. Cerutti, T.E. Cheatham, III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K.M. Merz, D. M. Y. and P. A. K. (2015). AMBER 12. *University of California, San Francisco*.

Drabik, P., Liwo, A., Czaplewski, C., & Ciarkowski, J. (2001). The investigation of the effects of counterions in protein dynamics simulations. *Protein Engineering, Design and Selection*, *14*(10), 747. http://doi.org/10.1093/protein/14.10.747

Dror, R. O., Dirks, R. M., Grossman, J. P., Xu, H., & Shaw, D. E. (2012). Biomolecular simulation: a computational microscope for molecular biology. *Annual Review of Biophysics*, *41*, 429–452. http://doi.org/10.1146/annurev-biophys-042910-155245

Earl, D. J., & Deem, M. W. (2005). Parallel Tempering: Theory, Applications, and New Perspectives. http://doi.org/10.1039/B509983H

Gaalswyk, K., & Rowley, C. N. (2016). An explicit-solvent conformation search method using open software. *PeerJ*, *4*, e2088. http://doi.org/10.7717/peerj.2088

Ganesan, A., Coote, M. L., & Barakat, K. (2017a). Molecular dynamics-driven drug discovery: leaping forward with confidence. *Drug Discovery Today*, *22*(2), 249–269. http://doi.org/http://dx.doi.org/10.1016/j.drudis.2016.11.001

Ganesan, A., Coote, M. L., & Barakat, K. (2017b). Molecular "time-machines" to unravel key biological events for drug design. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, *7*(4), e1306–n/a. http://doi.org/10.1002/wcms.1306

Gnanakaran, S., Hochstrasser, R. M., & García, A. E. (2004). Nature of structural inhomogeneities on folding a helix and their influence on spectral measurements. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(25), 9229–9234. http://doi.org/10.1073/pnas.0402933101

Gonzalez, B. S., Noya, E. G., Vega, C., & Sese, L. M. (2010). Nuclear Quantum Effects in Water Clusters: The Role of the Molecular Flexibility. *J. Phys. Chem. B*, *114*(7), 2484–2492. http://doi.org/10.1021/jp910770y

González, M. A. (2011). Force fields and molecular dynamics simulations. *Collection SFN*, *12*, 169–200. http://doi.org/10.1051/sfn/201112009

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, *18*(12), 1463–1472. http://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H

Hoover. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review. A, General Physics*, *31*(3), 1695–1697.

Ibragimova, G. T., & Wade, R. C. (1998). Importance of Explicit Salt Ions for Protein Stability in Molecular Dynamics Simulation. *Biophysical Journal*, *74*(6), 2906–2911. http://doi.org/http://dx.doi.org/10.1016/S0006-3495(98)77997-4

Isard, B. P. C. (2008). Theory and Practice in Replica-exchange Molecular Dynamics Simulation. *Biography An Interdisciplinary Quarterly*, 133. Retrieved from http://books.google.com/books/about/Theory_and_Practice_in_Replica_exchange.html?id=01OEBVckFcIC&pgis=1

Jarzynski, C. (1997). Nonequilibrium Equality for Free Energy Differences. *Physical Review Letters*, *78*(14), 2690–2693. Retrieved from https://link.aps.org/doi/10.1103/PhysRevLett.78.2690

Jas, G. S., & Kuczera, K. (2004). Equilibrium Structure and Folding of a Helix-Forming Peptide: Circular Dichroism Measurements and Replica-Exchange Molecular Dynamics Simulations. *Biophysical Journal*, *87*(6), 3786–3798. JOUR. http://doi.org/10.1529/biophysj.104.045419

Jorgensen, W. L., & Madura, J. D. (1983). Solvation and conformation of methanol in water. *Journal of the American Chemical Society*, *105*(6), 1407–1413. http://doi.org/10.1021/ja00344a001

Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, *118*(45), 11225–11236. http://doi.org/10.1021/ja9621760

Karplus, M., & McCammon, J. A. (1983). Dynamics of proteins: elements and function. *Annual Review of Biochemistry*, *52*, 263–300. http://doi.org/10.1146/annurev.bi.52.070183.001403

Khalili-Araghi, F., Gumbart, J., Wen, P.-C., Sotomayor, M., Tajkhorshid, E., & Schulten, K. (2009). Molecular dynamics simulations of membrane channels and transporters. *Current Opinion in Structural Biology*, *19*(2), 128–137. http://doi.org/10.1016/j.sbi.2009.02.011

Klibanov, A. M. (2001). Improving enzymes by using them in organic solvents. *Nature*, *409*(6817), 241–246. Retrieved from http://dx.doi.org/10.1038/35051719

Lonsdale, R., Harvey, J. N., & Mulholland, A. J. (2012). A practical guide to modelling enzyme-catalysed reactions. *Chemical Society Reviews*, *41*(8), 3025–3038. http://doi.org/10.1039/c2cs15297e

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., … Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B*, *102*(18), 3586–3616. http://doi.org/10.1021/jp973084f

Mahoney, M. W., & Jorgensen, W. L. (2000). A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics*, *112*(20), 8910–8922. http://doi.org/10.1063/1.481505

Mashayak, S. Y., & Tanner, D. E. (2011). Comparing Solvent Models for Molecular Dynamics of Protein. *Time*, 1–18.

Maximova, T., Moffatt, R., Ma, B., Nussinov, R., & Shehu, A. (2016). Principles and Overview of Sampling Methods for Modeling Macromolecular Structure and Dynamics. *PLoS Computational Biology*, *12*(4), e1004619. JOUR. http://doi.org/10.1371/journal.pcbi.1004619

Mitsutake, A., Mori, Y., & Okamoto, Y. (2013). Enhanced sampling algorithms. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 924, pp. 153–195). United States. http://doi.org/10.1007/978-1-62703-017-5_7

Mulholland, A. J. (2008). Introduction. Biomolecular simulation. *Journal of the Royal Society Interface*, *5*(Suppl 3), S169–S172. http://doi.org/10.1098/rsif.2008.0385.focus

Norin, M., Haeffner, F., Hult, K., & Edhoim, O. (1994). Molecular Dynamics Simulations of an Enzyme Surrounded by Vacuum, Water, or a Hydrophobic Solvent. *Bipsical Journal*, *67*(August), 548–559. http://doi.org/10.1016/S0006-3495(94)80515-6

Nosé, S. (1984). A unified formulation of the constant temperature molecular dynamics methods. *Jchep*, *81*(1), 511. http://doi.org/10.1063/1.447334

Nymeyer, H., Gnanakaran, S., & Garc??a, A. E. (2004). Atomic Simulations of Protein Folding, Using the Replica Exchange Algorithm. *Methods in Enzymology*, *383*(2000), 119–149. http://doi.org/10.1016/S0076-6879(04)83006-4

Oostenbrink, C., Villa, A., Mark, A. E., & van Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6. *Journal of Computational Chemistry*, *25*(13), 1656–1676. http://doi.org/10.1002/jcc.20090

Patel, J. S., Berteotti, A., Ronsisvalle, S., Rocchia, W., & Cavalli, A. (2014). Steered Molecular Dynamics Simulations for Studying Protein–Ligand Interaction in Cyclin-Dependent Kinase 5. *Journal of Chemical Information and Modeling*, *54*(2), 470–480. http://doi.org/10.1021/ci4003574

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., … Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802. http://doi.org/10.1002/jcc.20289

Ponder, J. W., & Case, D. A. (2003). Force fields for protein simulations. *Advances in Protein Chemistry*, *66*, 27–85. http://doi.org/10.1016/S0065-3233(03)66002-X

Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., … Lindahl, E. (2013). GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics (Oxford, England)*, *29*(7), 845–54. http://doi.org/10.1093/bioinformatics/btt055

Ramachandran, S., Kota, P., Ding, F., & Dokholyan, N. V. (2011). Automated Minimization of Steric Clashes in Protein Structures. *Proteins*, *79*(1), 261–270. http://doi.org/10.1002/prot.22879

Rehm, S., Trodler, P., & Pleiss, J. (2010). Solvent-induced lid opening in lipases: A molecular dynamics study. *Protein Science : A Publication of the Protein Society*, *19*(11), 2122–2130. http://doi.org/10.1002/pro.493

Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, *23*(3), 327–341. http://doi.org/http://dx.doi.org/10.1016/0021-9991(77)90098-5

Salsbury, F. R. (2010). Molecular Dynamics Simulations of Protein Dynamics and their relevance to drug discovery. *Current Opinion in Pharmacology*, *10*(6), 738–744. http://doi.org/10.1016/j.coph.2010.09.016

Sanbonmatsu, K. Y., & García, A. E. (2002). Structure of Met-enkephalin in explicit aqueous solution using replica exchange molecular dynamics. *Proteins*, *46*(2), 225–34. http://doi.org/10.1002/prot.0000

Scheraga, H. A., Khalili, M., & Liwo, A. (2007). Protein-folding dynamics: Overview of molecular simulation techniques. *Annual Review Of Physical Chemistry*, *58*, 57–83. http://doi.org/10.1146/annurev.physchem.58.032806.104614

Schleif, R. (2004). Modeling and Studying Proteins with Molecular Dynamics. In B. T.-M. in Enzymology (Ed.), *Numerical Computer Methods, Part D* (Vol. Volume 383, pp. 28–47). Academic Press. http://doi.org/http://dx.doi.org/10.1016/S0076-6879(04)83002-7

Shivakumar, D., Williams, J., Wu, Y., Damm, W., Shelley, J., & Sherman, W. (2010). Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation*, *6*(5), 1509–1519. http://doi.org/10.1021/ct900587b

Sugita, Y., & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, *314*(1–2), 141–151. JOUR. http://doi.org/http://dx.doi.org/10.1016/S0009-

2614(99)01123-9
Sun, H. (1998). COMPASS: An ab Initio Force-Field Optimized for Condensed-Phase ApplicationsOverview with Details on Alkane and Benzene Compounds. *The Journal of Physical Chemistry B*, *102*(38), 7338–7364. http://doi.org/10.1021/jp980939v
Swope, W. C., Andersen, H. C., Berens, P. H., & Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *The Journal of Chemical Physics*, *76*(1), 637–649. http://doi.org/10.1063/1.442716
Uline, M. J., & Corti, D. S. (2013). Molecular dynamics at constant pressure: Allowing the system to control volume fluctuations via a "shell" particle. *Entropy*, *15*(9), 3941–3969. http://doi.org/10.3390/e15093941
Verlet, L. (1967). Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, *159*(1), 98–103. http://doi.org/10.1103/PhysRev.159.98
Wallqvist, A., & Mountain, R. D. (2007). Molecular Models of Water: Derivation and Description. In *Reviews in Computational Chemistry* (pp. 183–247). John Wiley & Sons, Inc. http://doi.org/10.1002/9780470125908.ch4
Wang, J., Cieplak, P., Cai, Q., Hsieh, M.-J., Wang, J., Duan, Y., & Luo, R. (2012). Development of Polarizable Models for Molecular Mechanical Calculations III: Polarizable Water Models Conforming to Thole Polarization Screening Schemes. *The Journal of Physical Chemistry. B*, *116(28), 7999–8008. http://doi.org/10.1021/jp212117d*
Zhang, C., & Ma, J. (2008). Comparison of sampling efficiency between simulated tempering and replica exchange. The Journal of Chemical Physics, 129(13), 134112. http://doi.org/10.1063/1.2988339

Appendix B: HOMOLOGY MODELING: AN OVERVIEW OF FUNDAMENTALS AND TOOLS[9]

B.1. Introduction

Homology modeling is one of the main tools employed in this thesis. The comprehensive models for KCNQ1 protein were generated using a comparative modeling approach. For this reason, an overview of fundamentals and tools of homology modeling technique is given in this chapter.

Proteins play numerous roles in driving several biological processes. The structure of a protein determines its function and specifies its role(s) in the cell and in the body. These roles include enzymatic activities (Coleman, 1992), signal transduction (Simon, Strathmann, & Gautam, 1991), substance and ions' transport (Hille, 2001), gene expression (Pabo & Sauer, 2003) and also serving as structural elements in the cell (Buxbaum, 2007). Resolving the three-dimensional (3D) structure of proteins answers many intriguing biological questions and provides insights into the physiological role of the different proteins. Understanding proteins' structural differences in health and disease can help in rationally designing potent drugs that can overcome these diseases (Kann, 2007). Over the last decades, protein structure determination relied heavily on X-ray crystallography and NMR spectroscopy (Yonath, 2011). Although these techniques can provide high-resolution atomistic details of proteins' structures, they are time consuming and are expensive both in terms of labour and resources. Furthermore, they are still incapable of being successfully applied to all types of proteins, such as membrane (Carpenter et al., 2008; E. Ghosh, Kumari, Jaiman, & Shukla, 2015) and/or intrinsically disordered proteins (IDPs) (Marti-Renom et al., 2000). With these limitations, relying heavily on these experimental methods has created a huge gap between experimentally determined protein structures and sequenced genes, such that only a tiny portion of proteins with known sequences have their structures resolved and deposited in the Protein

[9] A version of this appendix has been published in: *Jalily Hasani H, Barakat K. Homology Modeling: an Overview of Fundamentals and Tools. Int Rev Model Simulations (IREMOS); Vol 10, No 2. 2017*

Data Bank (Berman et al., 2000). With this gap growing exponentially with time (Mills, Beuning, & Ondrechen, 2015), there has always been a need for more efficient techniques to work hand-in-hand with experiments and to provide a more rapid and less expensive way to determine the structural details of proteins (Hasani & Barakat, n.d.; Schwede, 2013).

For decades, protein structure prediction has puzzled the scientific community; it is an extremely important problem that is simple to define but difficult to solve. For example, although many proteins can be classified into different families that share a generalized homologous structure, a very simple amino acid sequence variation among members of the same family can determine which substrate they accept, which protein(s) they interact with and even which cell type they can be active in. Covering all this biochemical space experimentally is unfeasible. In this context, bioinformatics tools, in general, and homology modelling, in particular, have been employed to explore this space and fulfill this unmet need.

Theoretical structure prediction can be divided into two main categories, namely *ab initio* methods and homology modeling. The former's main goal is to predict the overall folding problem for a particular protein from physical chemistry principles, while the latter predicts the 3D structure of a given protein based primarily on its sequence similarity to one or more proteins of already known structures (Eswar, 2003; S. Hongmao, 2016). Despite the considerable progress accomplished to date in *ab initio* structure prediction, "comparative" or "homology" modeling methods, when applicable, provide the most reliable and accurate models of protein structure (Cavasotto & Phatak, 2009; A. M. Lesk, 1997).

This review is intended for the uninitiated reader and, therefore, it initially and briefly describes protein structure, followed by a presentation on the different steps involved in homology modeling. Introduction of the various homology-modeling tools including prediction servers and programs is given throughout the chapter. In the subsequent sections, several applications and case studies in different fields will be listed and finally, the current limitations and challenges in the face of homology modeling will be discussed.

B.2. Protein Structure

To precisely follow the concept of homology modeling, a basic understanding of protein structure and composition is required. Although a comprehensive explanation does not fit into the scope of this review, we aim to provide an overview of the relevant aspects of protein structure for the benefit of the inexperienced reader. For more detailed description of these topics, the reader is advised to consult one of the many excellent references in this area, such as (Fallis, 2013; Korasick & Jez, 2016; W. Taylor, 2004).

Figure B.1 illustrates the different levels of protein structure right from the amino acid sequence to the higher complex tertiary and quaternary structures. Each amino acid in a protein sequence is encoded by a codon, which consists of three nucleotides. Such codons are responsible for the makeup of the nucleotide sequence of the mRNA, which is translated into proteins through a set of ribosomal molecules. This marks the translation stage in the process of protein synthesis. The ribonucleoproteins translate each codon to one amino acid at a time and add it to a growing polypeptide chain. The resulting amino acid sequence is known as the primary structure of the protein (Gagneux, 2004). Amino acids are linked together via peptide bonds, which connect the amino group (NH2) of one amino acid to the carboxyl group (COOH) of the other amino acid. Each protein sequence, thus, has a free NH2 group (N-terminus) on one end and a free COOH group (C-terminus) on the other end. The amino acids differ from one another depending on the nature of their side chains or their functional groups flanking from their backbone. It is the characteristics of those side chains that make a particular amino acid possess a specific property such as being positively charged (acidic), negatively charged (basic), polar (hydrophilic) or non-polar (hydrophobic) (23).

The next level of complexity is the result of the secondary structure of proteins, consisting of the folding of the peptide backbone. The secondary structure can involve alpha helices, beta sheets, hairpins, loops, or many other structural elements. These are formed from hydrogen bonding between the NH and C=O groups of the polypeptide backbone as well as interactions and physiochemical properties of the side chains. For example, alpha helices are formed when the C=O group of one peptide bond binds to the NH group of another peptide bond, four residues further along the same polypeptide chain. This creates a right-handed twist of a typical alpha helix characterised by one full

turn every 3.6 residues. A beta sheet, on the other hand, is formed when hydrogen bonds are established between the peptide bonds of several different polypeptide chains or within a single polypeptide chain. There are two types of beta sheets; parallel in which the strands run in the same direction and antiparallel strands that run in an opposite direction. This is mainly determined by the type of amino acids predominantly present in the beta sheet region of the protein. Despite the formation of the two main secondary structures in a protein, i.e. alpha helices and beta sheets, certain amino acid combinations favour the presence of loops in the structure of the protein (Perutz, 2012).

The tertiary structure of a protein describes the 3D assembly of the polypeptide. The 3D structure of a protein is stabilized by non-covalent interactions such as ionic bonds, H-bonds, hydrophobic interactions and van der Waals forces. Importantly, covalent bonds are also formed between the SH groups of Cysteine residues to form disulphide bonds offering additional stability to the protein structure. In case of proteins possessing more than one subunit, the overall complex produced by such subunits form the quaternary structure of a protein (X. Zhang & Cheng, n.d.). Finally, the quaternary structures are produced when two tertiary domains, or more, interact together. This structural characteristic is observed in case of proteins possessing more than one subunit. Examples include multi-subunit membrane proteins such as ion channels, G-protein coupled receptors, etc.

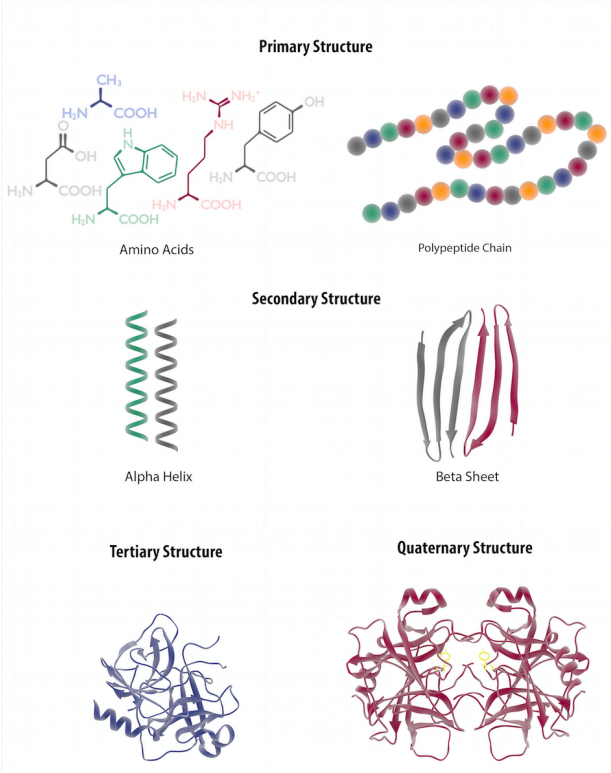


Figure B.1. Protein Structure: from primary to quaternary structure. Primary structure of a protein consists of the amino acid sequence; secondary structure consists of helices and beta sheets. The tertiary and quaternary structures are at a higher complexity level wherein the secondary structure elements and domains are assembled together.

The structure of a protein and its folding behaviour is determined by its amino acid composition. Therefore, speaking theoretically, it should be possible to use the primary structure of a protein (amino acid sequence) to predict its 3D structure (Zorko, 2009). This might sound like an easy problem, but in reality, it is still one of the main challenges in the field of protein research. Although techniques like homology modeling and *ab initio* modeling have succeeded to some extent solely based on this theory, but still are lagging behind for modeling a large number of proteins. This is in part because the laws, which govern the folding of a polypeptide into its tertiary structure, have not been fully interpreted yet. In the following sections, the different existing techniques and challenges are discussed further, to illustrate how homology modeling is linked to the above theory.

B.3. Homology Modeling

Homology modeling is based on a few related established observations; 1) Evolutionarily related sequences have similar 3D structures. As a consequence, a 3D model of a protein of interest (target) can be built from related protein(s) of known structure (template) that share statistically significant sequence similarity. 2) The structural conformation of a protein is more highly conserved than its amino acid sequence (Saxena, Sangwan, & Mishra, 2013; Vyas, Ukawala, Chintha, & Ghate, 2012). This implies that small or medium changes in sequence typically result in only small changes in the 3D structure. 3) Functional sites maintain identical structural folds and proteins of the same function and/or family have similar structure (Nayeem, Sitkoff, & Krystek, 2006).

The homology modeling procedure consists of several sequential steps, shown in Figure B.2, which are usually repeated iteratively until a satisfactory model is obtained. This multi-step process, can be summarized in the following way: 1) identification of a homologous template protein with known structure and fold assignment using sequence alignment. This step relies on the similarity between the target sequence and at least one template protein structure; 2) alignment of the target and template protein sequences; 3) identifying structurally conserved regions and predicting structurally variable regions, including insertions and missing N and C termini; 4) building a structural model for the target; 5) refinement and optimization of the predicted structure and 6) assessment and validation of the resulting model (Ginalski, 2006; Vyas et al., 2012). These sequential steps have been developed and evolved over the years and there are instances where the steps are switched back and forth to become appropriate for different modeling algorithms. Furthermore, several improvements and modifications of the general homology modeling strategy have been developed and applied to the prediction of protein structures.

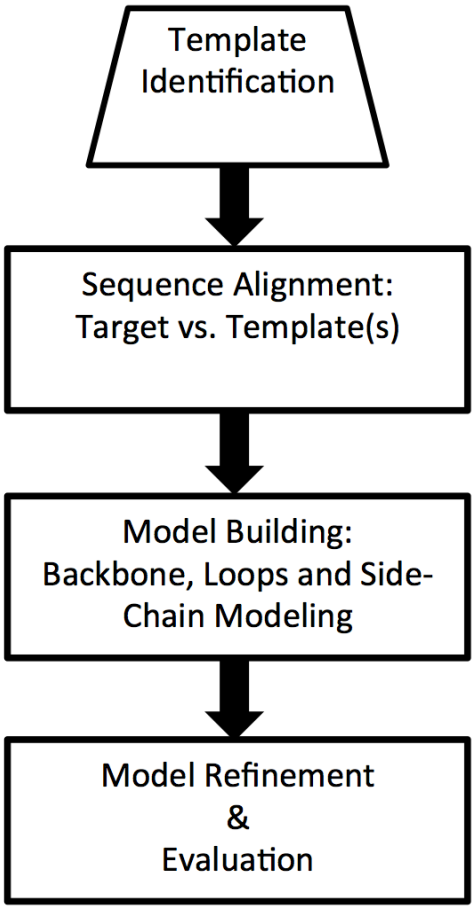


Figure B.2. The main steps in homology modeling. The process of homology modeling begins with finding a suitable template. Aligning the target and template sequence, and subsequently building the model follow this. The predicted model is exposed to refinement for improving its quality and removing errors. Evaluation is the last step for checking the accuracy and reliability of the prediction. The sequence of steps shown in the figure is sometimes switched back and forth to suit different algorithms.

The final model quality is strongly affected by any errors that could be introduced during each of these steps. A major proportion of the accuracy and quality of the resulting model also depends on the extent of similarity between the target and the template. As a general rule, models based on templates with more than 50% sequence identity are accurate and reliable enough for applications ranging from drug discovery to designing mutagenesis experiments and in vitro test assays (Hillisch, Pineda, & Hilgenfeld, 2004). Those between 30-50% similarity share at least 80% of their structures with errors located mainly in loop regions. Sequence identity of below 30% can lead to fold assignment and alignment errors, uncertainty in loop modeling and/or side-chain packing,

leading to misalignment and speculative models (a Fiser et al., 2000). However, even when the quality is low, the models can provide insights into coarse structural features, which are sometimes sufficient to identify some aspects of the underlying function and to predict the effects directing mutagenesis experiments (Marti-Renom et al., 2000). In addition, the optimal use of structural information from available templates and the precision of sequence-to-structure alignment are other significant determinants of the overall quality of the generated models (Eswar, 2003).

B.4. Detailed Steps in Homology Modeling

Fold Assignment and Template Identification

This step marks the starting point of the modeling process, wherein all protein structures related to the target sequence are identified and those with the highest sequence similarity are selected as templates. The quality of the final structures is in many parts dependent on the quality of template and, thus, this precise template selection is critical. Many protein structure databases and tools are employed to facilitate this step by using the target sequence as the query. CATH (Orengo et al., 2002), SCOP (Murzin, Brenner, Hubbard, & Chothia, 1995), DALI (Liisa Holm & Rosenström, 2010) and Protein Data Bank (Berman et al., 2000) are some examples of such databases. Based on several case studies, Saxena et al. (Saxena et al., 2013) have reported that for a given target sequence that is picked randomly from a genome, the probability of finding a related protein of known structure would vary from 20% to 70%. Some of the databases and tools that are used for template search and consequent modeling steps are listed in Table B.1.

Table B.1. Programs and online tools useful in the different steps of homology modeling.

Name	Type	URL Address/Reference
PROTEIN DATABASES		
CATH	Server	http://www.cathdb.info (Orengo et al., 2002)
Protein Data Bank (PDB)	Server	http://www.rcsb.org/pdb/ (Berman et al., 2000)
GenBank	Server	http://www.ncbi.nlm.nih.gov/genbank/ (Benson, h-Mizrachi, Lipman, Ostell, & Wheeler, 2005)
MODBASE	Server	http://modbase.compbio.ucsf.edu/modbase-cgi/ r et al., 2006)
GeneCensus	Server	http://bioinfo.mbb.yale.edu/genome/ (Gerstein, 1997)
SCOP	Server	http://scop.mrc-lmb.cam.ac.uk/scop/ (Murzin et al.,
SWISSPROT+TrEMBL	Server	http://www.ebi.ac.uk/uniprot (Bairoch & Apweiler,
TEMPLATE SEARCH		
FASTA	Server	http://www.ebi.ac.uk/Tools/sss/fasta/ (Pearson, 1990)
BLAST	Server	http://blast.ncbi.nlm.nih.gov/Blast.cgi (Stephen F ul, Gish, Miller, Myers, & Lipman, 1990)
DALI	Server	http://ekhidna.biocenter.helsinki.fi/dali_server/ (Liisa & Rosenström, 2010)
THREADER	Program	http://bioinf.cs.ucl.ac.uk (Jones, Taylor, & Thornton,
ALIGNMENT TOOLS		
BLAST	Server	http://blast.ncbi.nlm.nih.gov/Blast.cgi (Stephen F ul et al., 1990)
CLUSTAL	Server	http://www.ebi.ac.uk/Tools/msa/clustalw2/ (Larkin et 07)
MULTALIN	Server, Program	http://multalin.toulouse.inra.fr/multalin/ (Corpet,
FASTA	Server	http://www.ebi.ac.uk/Tools/sss/fasta/ (Pearson, 1990)
T-COFFEE	Server	http://www.ebi.ac.uk/Tools/msa/tcoffee/ (Notredame, ns, & Heringa, 2000)
PDBeFold	Server	http://www.ebi.ac.uk/msd-srv/ssm/ (Krissinel & ck, 2004)
MODELLING TOOLS		
ROBETTA	Server	http://www.robetta.org (D. E. Kim, Chivian, & Baker,
ICM	Program	www.molsoft.com (Abagyan, Totrov, & Kuznetsov,
MODELLER	Program	http://salilab.org/modeller/ (Šali & Blundell, 1993)
I-TASSER	Program, Server	http://zhanglab.ccmb.med.umich.edu/I-TASSER/ (Roy, Kucukural, & Zhang, 2010; Yang Zhang, 2008)
3D-JIGSAW	Server	http://bmm.crick.ac.uk/~3djigsaw/ (Bates, Kelley, allum, & Sternberg, 2001),
3D-JURY	Server	http://meta.bioinfo.pl/submit_wizard.pl (Ginalski, on, Fischer, & Rychlewski, 2003)
PSIPRED	Program, Server	http://bioinf.cs.ucl.ac.uk/psipred/ (Buchan, Minneci, t, Bryson, & Jones, 2013)(Jones, 1999)
RaptorX	Server	http://raptorx.uchicago.edu (Källberg et al., 2012)
PHYRE2	Server	http://www.sbg.bio.ic.ac.uk/phyre2 (Kelley, Mezulis, Wass, & Sternberg, 2015)

SWISS-MODEL	Server	http://swissmodel.expasy.org (Schwede, Kopp, Guex, tsch, 2003)
WHAT IF	Program	http://swift.cmbi.ru.nl/whatif/ (Vriend, 1990)
SYBYL	Program	http://tripos.com
DRAGON	Program	http://www.nimr.mrc.ac.uk/~mathbio/a-aszodi/dragon.html (Aszodi & Taylor, 1996)
SCWRL4	Program, Server	http://dunbrack.fccc.edu/scwrl4/ (Krivov, Shapovalov, & Dunbrack, 2009)
MODEL EVALUATION		
ANOLEA	Server	http://melolab.org/anolea/ (Melo, Devos, Depiereux, tmans, 1997)
AQUA		http://aquad.sourceforge.net (Laskowski, Rullmannn, rthur, Kaptein, & Thornton, 1996)
ERRAT	Server	http://services.mbi.ucla.edu/ERRAT/ (Colovos & s, 1993)
PROCHECK	Server	http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/ (Laskowski, MacArthur, Moss, & Thornton, 1993)
WHATCHECK	Server	http://swift.cmbi.ru.nl/gv/whatcheck/ (Hooft, Vriend, r, & Abola, 1996)
VERIFY3D	Program	http://services.mbi.ucla.edu/Verify_3D/ (Bowie, , & Eisenberg, 1991)
PROSA-WEB	Program	https://prosa.services.came.sbg.ac.at/prosa.php erstein & Sippl, 2007)

In general, the protein comparison methods used for template identification are classified into three main categories. The first category involves comparing the target sequence with each of the databases individually using pairwise sequence comparison. Frequently used programs in this class are FASTA (Pearson, 1998) and BLAST (22). This is a comparatively simpler and less sensitive method used to choose templates only for those targets with very high sequence similarity and close templates in PDB.

Tools in the second category, perform multiple sequence alignments using a wide array of methods like profile analysis (Gribskov, McLachlan, & Eisenberg, 1987), profile-profile comparisons (Rychlewski, Jaroszewski, Li, & Godzik, 2000; Yona & Levitt, 2002), Hidden Markov Models (Eddy, 1998; K Karplus, Barrett, & Hughey, 1998) and intermediate sequence search (Teichmann, Chothia, Church, & Park, 2000). Widely used programs in this category are PSI-BLAST (S F Altschul et al., 1997) and SAM (Kevin Karplus et al., 2003). PSI-BLAST performs the template detection by iteratively expanding the set of homologs of the query sequence such that an initial set of homologs is collected, a weighted multiple sequence alignment is made and a position specific scoring matrix is then built from the alignment (S F Altschul et al., 1997). This matrix is used to search for more homologs iteratively until no more homologs are found. Profile-profile comparison methods involve finding all the related sequences to the target to

obtain the target sequence profile. The potential templates are then found by comparing this profile with each of the sequence profiles for known structures. The multiple sequence alignment is a sensitive and fully automated method and is especially useful when the sequence identity between the target and the template is below 25% and for detecting distantly related sequence-structure relationships (Muller, MacCallum, & Sternberg, 1999). HHpred (Soding, Biegert, & Lupas, 2005) is another online tool that is based on the pairwise comparison of profile hidden Markov models (HMMs) for distant homology detection.

The third class of protein sequence comparison methods consists of threading or 3D template matching methods (Bowie et al., 1991; Godzik, Kolinski, & Skolnick, 1992; Jones et al., 1992). These methods rely on pairwise comparison of the query protein sequence and a protein of known structure. Whether or not a given target sequence adopts any one of the many known 3D folds, is predicted by an optimization of the alignment with respect to a structure-dependent scoring function, independently for each sequence-structure pair. That is, a target sequence is threaded through a library of 3D folds (J. Peng & Xu, 2011). The use of these methods is encouraged when no related sequences to the query are found and whatsoever the sensitivity of sequence profile methods, no hits are found. Examples of programs in this category are THREADER (J. Peng & Xu, 2011), 3D-PSSM (Kelley, MacCallum, & Sternberg, 2000), SPARKS-X (Yuedong Yang, Faraggi, Zhao, & Zhou, 2011) and the alignment tool employed in RaptorX server (Källberg et al., 2012).

After a list of all related (homologous) protein structures is found, it is now necessary to choose the template(s) that are appropriate for the given modeling problem. As a first measure, high sequence similarity of the template to the target implies a suitable candidate. However, several other factors need to be taken into account. One important factor to keep in mind is that templates, which are more closely related to the target in terms of protein family relationships, are better choices. Usually, a multiple sequence alignment and phylogenetic tree construction can illustrate the closeness of a target to a template as a protein subfamily (Korasick & Jez, 2016). Another selection criterion is the quality of the experimentally determined template structure. For example, in case of X-ray crystal structures, the resolution and the R-factor are indicative of the

template quality, whereas the accuracy of NMR structures is recognized by the number of restraints per residue (Charlotte M Deane & Blundell, 2003). In addition, prior biological knowledge of the target sequence is valuable information in identifying a suitable template. The template selection criteria also depends in large part on the purpose of the comparative modeling, which requires some level of expertise to obtain trustworthy results. This implies that the type of interpretations and conclusions needed to be drawn by the model at hand, have to be clearly defined.

Target-Template Alignment

The alignment between the target sequence and the template structure(s) that is built by the fold assignment methods is usually not the most optimal alignment for building the models. Thus, it is essential to apply a further step to correct the alignment of target sequence to template structure using more sophisticated methods (Ma & Wang, 2014). Because of the importance of this stage, which could ultimately lead to an erroneous model, great care should be taken.

High sequence similarities of over 40% can ensure a correct alignment, whereas lower sequence similarities (below 30%) indicate difficulties in the alignment due to the presence of more gaps and mutations (Johnson & Overington, 1993; Yona & Levitt, 2002). In difficult cases of low sequence similarity where the alignment is by no means possible, multiple structures and sequences from homologous proteins are usually used to facilitate the alignment process between the target and template. This works by creating a multiple sequence alignment profile for both the target and the template using related sequences that are easily matched. The two profiles are then aligned with inclusion of as much structural information as possible. However, it is beneficial to do a visual inspection for the final alignment to avoid unnecessary gaps in secondary structure elements of buried regions or between two residues that are far in space (Larsson, Wallner, Lindahl, & Elofsson, 2008; J. Peng & Xu, 2011). Alternative methods involve secondary structure predictions for the two profiles or 3D model evaluation instead of the alignment score (Marti-Renom et al., 2000; Sanchez & Sali, 1997).

Model Building

Once the alignment is ready, the model building can be performed. The process of building the model is categorized into three complications; 1) Backbone generation, 2) Loop modeling, and 3) Side-chain modeling, and are discussed here individually.

1. Backbone Generation

There are four methods of building the actual backbone or the framework of the target protein, depending on how the information in the known structures is transferred to the target sequence. The first method is rigid body assembly, based on the natural dissection of the protein structure into conserved core regions, variable loops that connect them and side chains (Blundell, Sibanda, Sternberg, & Thornton, 1987). COMPOSER program is an example that uses this method for generating the main building blocks of the model (Sutcliffe, Haneef, Carney, & Blundell, 1987).

The second technique involves segment matching or coordinates reconstruction through which a subset of atomic positions (usually C-alpha atoms) from the template structure are chosen as guiding positions and are used to identify and assemble short, all-atom segments that fit these guiding positions. The all-atom segments are found either by scanning all the known protein structures in a database (L Holm & Sander, 1991), or by a conformational search restrained by an energy function (Iwata, Kasuya, & Miyamoto, 2002; van Gelder, Leusen, Leunissen, & Noordik, 1994). This method can be used for main-chain, side-chain and even unaligned gapped regions. The program SegMod/ENCAD (Levitt, 1992) is one of the first segment-based methods of secondary structure construction.

Another class of methods is modeling by satisfaction of spatial restraints, in which many constraints or restraints are imposed on the structure of the target, using its alignment with template protein structures as a guide. For example, the distances between the residues in the template are considered to be similar to that of the target residues. In addition, stereo-chemical restraints on bond lengths, bond angles, dihedral angles and non-bonded atom-atom contacts, obtained from a molecular mechanic force field, are also used (Saxena et al., 2013). The model is then generated by minimizing the violations of these restraints. This approach is employed in the program Modeller (Šali & Blundell, 1993). Due to the versatility of this approach to use different types of information about

the target sequence, it is considered to be one of the most promising of all homology modeling approaches. Other types of restraints that can be used in addition to the homology-driven restraints to improve the accuracy of the model, include analyses of hydrophobicity (Aszodi & Taylor, 1994), data from NMR experiments (Allison et al., 2012), site-directed mutagenesis and cross-linking experiments (Kahraman et al., 2013; Rappsilber, 2011) and fluorescence spectroscopy (Venselaar et al., 2010).

A more recent method is called the artificial evolution model building. It was first implemented in the NEST program (Petrey et al., 2003). In this approach, the target model building in a homology modeling process should be similar to the natural process of evolving a protein that happens in multiple steps. This is possible by changing the template structure based on the alignment in several "operations". Each operation is followed by an energy minimization to compensate for the energy cost due to mutation, deletion or insertion in the template sequence. Only those operations that do not cause a significant energy penalty are retained, others are discarded. This process is repeated until the migration of the template structure to the target is accomplished (Zhexin Xiang, 2006).

2) Loop Modeling

Generally, there are functional differences between the members of the same protein family that usually result in a structural variability on the protein surface. This is observed in the form of insertions, deletions or substitution between the template and the target. Such variable regions frequently represent loop regions that connect main elements of the secondary structure in the protein fold, i.e. alpha helices and beta sheets. Loops are one of the important factors of functional specificity and extensively contribute to active and binding sites. As a result, careful considerations must be taken into account while modeling these structural elements to ensure the usefulness of comparative models in studying the interactions between the protein and its various partners; ligands, drugs or other proteins (A. Fiser, Do, & Sali, 2000).

There are two main approaches to the loop modeling step; 1) knowledge-based and, 2) energy based. The former involves a scanning of the databases (for example, PDB) of all known protein structures to find segments that can fit the endpoint regions in the protein backbone. Modeller (Sali, Potterton, Yuan, van Vlijmen, & Karplus, 1995),

Swiss-Model (Guex & Peitsch, 1997), WHAT IF (Vriend, 1990) and 3D-Jigsaw (Bates et al., 2001) utilize this approach to model the loop regions. The latter approach relies on *ab initio* fold prediction, wherein a large number of randomly chosen candidate conformations are generated and a scoring function is optimized through Monte Carlo or molecular dynamics techniques to identify the best match (Tappura, 2001). In addition, there are also methods that combine these two approaches given in (Charlotte M Deane & Blundell, 2003; van Vlijmen & Karplus, 1997). Other tools used for loop modeling are MODLOOP (A. Fiser et al., 2000), LOOPY (Zhexin Xiang, Soto, & Honig, 2002) and CODA (C M Deane & Blundell, 2001).

3) Side-Chain Modeling

The prediction of side-chains' orientation for conserved residues is often a straightforward procedure. This is because in most cases they have well-defined torsion angles about their Cα-Cβ bond and, thus, the conserved residues can be entirely copied from the template to the model. However, this does not hold true in the case of low sequence similarity regions, where neighbouring residues are usually different and rotamers of the conserved residues show up differently in up to 45% of cases (Sanchez & Sali, 1997).

The main approaches to side-chain modeling are knowledge-based methods (Johnson, Srinivasan, Sowdhamini, & Blundell, 1994). They use libraries of common rotamers extracted from high-resolution X-ray structures. These are tested sequentially and scored using energy functions. The choice of one rotamer intuitively affects the choice of neighbouring ones, which in turn affects their neighbours and so on. This makes the process computationally demanding due to the combinatorial explosion involved. Finding techniques and mechanisms to reduce this huge search space is an active area of research (Francis-Lyon & Koehl, 2014). However, it is known that certain backbone conformations strongly favour certain rotamers and this can be used as a solution to simplify the problem. Such position-specific rotamer libraries are widely used (Dunbrack & Karplus, 1994; Stites, Meeker, & Shortle, 1994).

SCWRL3 (Canutescu, Shelenkov, & Dunbrack, 2003) is one of the first successful tools for side-chain packing, that is based on graph-theory algorithms. A newer version of this tool; SCWRL4 (Krivov et al., 2009) has been introduced by re-implementing the

algorithm described by Xu et al. (J. Xu & Berger, 2006). These methods have been compared to other side-chain modeling tools in the comparative study done by Wallner and Elofsson in 2005 (Björn Wallner & Elofsson, 2005). Other available tools are SCAP (Z Xiang & Honig, 2001), SMOL (Liang & Grishin, 2002) and RAMP (Samudrala & Moult, 1998).

Model Refinement

Although the initially generated model through the pervious steps resembles the overall structure and fold of a particular protein, the model is still slightly far from the global minimum for the native configuration. This means that subsequent refinement is required to bring the structure closer to this minimum. The refinement step adjusts the atomic coordinates of the predicted model through appropriate refinement techniques and potentials. Each homology modeling package employs a different method of refinement, a number of which are discussed here, in brief.

One of the main goals of the refinement step is to optimize the model through accurate modification of backbone and side-chain rotamers and their packing. A common approach is an iterative method of predicting a configuration of the side-chains' rotamers, which minimizes any existing steric clashes. This is followed by a shift in the backbone, which in turn requires further adjustment of the side chains' rotamers for the newly shifted backbone, and so on, until the procedure converges (Krieger et al., 2003). The refinement can also involve the use of energy functions (force fields) and several steps of energy minimization for the loops and side-chains (Xun, Jiang, & Wu, 2015). These steps are applied repeatedly in order to relax the backbone, bringing it closer to the native structure. In many cases, during this process, many small errors can accumulate, shifting the model away from the template. Restraining the position of the atoms or reducing the number of minimization steps to only a few hundred, could solve this problem. However, high-resolution refinement of homology models remains a challenging problem because of the incapability of sampling huge numbers of alternatively packed conformations and the lack of high precision in the current force fields (Ginalski, 2006).

In addition, Monte Carlo simulation and all-atom molecular dynamics (MD) simulations that follow the motions of the protein and mimic the true folding process, have been frequently employed for model refinement (Fan, 2004; Han et al., 2008;

Ishitani, Terada, & Shimizu, 2008; Kannan & Zacharias, 2010; Lu & Skolnick, 2003). Other types of enhanced sampling and MD simulations such as Replica Exchange Molecular Dynamics have also been reported to be useful in refining protein models (J. Zhu, Fan, Periole, Honig, & Mark, 2008). Combinations of free energy optimizations along with evolutionary favoured principal components of the backbone structure within a homologous family have also produced enhanced results, shown in (Qian, Ortiz, & Baker, 2004).

The refinement category of the biennial Critical Assessment of Structure Prediction (CASP) experiment, known as CASPR, represents a de facto benchmark for testing refinement tools and techniques. A number of target proteins are selected, along with a homology model for each. These homology models are selected from the many models that are submitted to other categories of the CASP experiment and signify the best efforts of protein structure prediction. These structures have already been refined by the modellers throughout the homology modeling procedure and hence are considered to be sufficiently challenging to test the refinement tools (Raval, Piana, Eastwood, Dror, & Shaw, 2012). The results from such benchmarking studies can provide good hints regarding which tool or technique is to be used for refining homology models. ModRefiner (D. Xu & Zhang, 2011) is an example of an online tool that has been used for model refinement. However, it should be noted that refinement tools heavily depend on the type of protein being tested and, occasionally, they perform better for certain typical types of proteins (against which they have been trained), compared to others.

Model Validation

Every homology model contains errors introduced at various stages of modeling. The number of errors could be related to different factors. An important determinant for the number of errors is the percentage sequence identity between the template and the target; as the similarity decreases the errors in the model normally increase. In addition, the experimentally determined 3D structures are prone to several types of errors and pitfalls, extensively reviewed in (Laskowski & Swaminathan, 2013). The inaccurate 3D structures when used as templates for comparative modeling can be troublesome and can cause huge errors in the modeled structure and, eventually, in the quality of the specific study performed and their results. Also, it is the quality and accuracy of the predicted models

that define which kind and level of information can be extracted from them. The latter topic of what sort of application should be expected from different models is more elaborated in the next sections.

Other types of errors can be associated with the alignment; mainly arising from unaligned regions (loops and low sequence similarity regions), misalignment and distortions or shifts in correctly aligned regions (Dalton & Jackson, 2007). Many of the alignment-related errors could be minimized by the simultaneous use of several related templates belonging to the same family/class of proteins. Yet other errors could be introduced in the models during the assembly or the building stage. A critical example of this type of errors is the incorrect side-chain packing, especially if they occur in the active or binding sites of the protein (Sali et al., 1995). Therefore the verification of the model and estimation of the likelihood and magnitude of errors is an important step in homology modeling (Rodriguez, Chinea, Lopez, Pons, & Vriend, 1998).

The predicted model can be assessed as a whole or for individual regions/domains. Sometimes, performing the evaluation of both built model and template and comparing the results can be more useful in deriving better conclusions regarding the model quality. Tools such as PROCHECK (Laskowski et al., 1993) and WHAT IF (Vriend, 1990) can check a number of structural variables in the predicted homology model against expected/reference values. VERIFY3D (Bowie et al., 1991) is also a widely used tool that evaluates the model structure by taking into account the location and environmental profile of each residue in relation to a set of reference structures. ProSA-Web (Wiederstein & Sippl, 2007) uses knowledge-based potentials of mean force (PMF) to evaluate the accuracy of models. It provides a Z-score that is in most cases, indicative of the overall model quality.

Other factors that are checked for the assessment of the predicted model are bond lengths, bond angles, peptide bond and side-chain ring planarities, chirality, steric clashes, main-chain and side-chain torsion angles (phi and psi angles). A more detailed list of available programs and tools used for model evaluation is given in Table B.1.

B.5. Homology Modeling Tools

Several programs and servers are available for homology modeling, both in the commercial and public domains (Table B.1). The MODELLER program developed by Andrej Sali and colleagues (31) remains one of the most powerful tools of homology modeling. The most popular servers, as reported by the Continuous Automated Model Evaluation (CAMEO) community project are SwissModel (40,101), Phyre2 (39), RaptorX (38) and ROBETTA (D. E. Kim et al., 2004).

Automation of these tools makes homology modeling accessible and easy to handle for both experts and non-specialists alike. However, many stages including the process of model calculation, model refinement and optimization as well as visualization can involve several scripts, local programs and servers. In addition, manual intervention is generally still needed to maximize the accuracy of the predicted models in difficult cases (4). Some of the most commonly used tools for homology modeling are described in the following sections.

MODELLER

MODELLER is a widely used computer program for homology modeling (Šali & Blundell, 1993), which predicts protein structures by satisfying specific spatial restraints. These restraints include: (1) homology-derived restraints on the distances and dihedral angles in the target sequence, extracted from the homologous template structures; (2) stereochemical restraints, e.g. bond lengths and bond angle preferences, obtained from the CHARMM-22 molecular mechanics force field (MacKerell et al., 1998); (3) statistical preferences for dihedral angles and non-bonded interatomic distances, obtained from a representative set of known protein structures (Sali & Overington, 1994); and (4) user-defined restraints, e.g. from NMR spectroscopy, rules of secondary structure packing, cross-linking experiments, fluorescence spectroscopy, image reconstruction from electron microscopy, site-directed mutagenesis and intuition. The spatial restraints, expressed as probability density functions (pdfs), are combined into an objective function that is optimized by a combination of conjugate gradients and molecular dynamics with simulated annealing. The pdfs restrain CA-CA and backbone N-O distances, and backbone and side-chain dihedral angles for different residues. In the simplest case, the

input is an alignment of a sequence to be modeled with the template structure(s), the atomic coordinates of the templates and a short script file. MODELLER then automatically calculates a model containing all non-hydrogen atoms, without any user intervention and within minutes on a Pentium processor. This model building procedure is regarded as being similar to NMR spectroscopy structure determination (Sali et al., 1995).

MODELLER is also capable of performing tasks other than homology modeling. These include protein sequence and profile alignment, multiple sequence/structure alignment, phylogenetic tree calculations, de novo loop modeling and model refinement (A. Fiser et al., 2000).

SwissModel

SwissModel is a web accessible server that performs comparative modeling in a very user-friendly manner (Schwede et al., 2003). In contrast to MODELLER, SwissModel follows the standard protocol of homology modeling shown in Figure B.2. For the last stage of model building, it performs a search in a loop library or the conformational space using constrains space programming. The best loop is then chosen through a scoring scheme that takes into account the force-field energy, steric hindrance, hydrogen bond formation, etc. For the side-chain modeling, the conformations are selected from a rotamer library through a scoring function. This scoring function provides the highest score for favourable interactions of the side chains such as H-bonds, S-S bridges and penalizes unfavourable interactions of making close contacts (e.g. steric clashes, overlaps, bumps).

SwissModel offers different modes for user intervention and extended functionality. For example, the Automated mode is a fully automated process wherein the only input required is a protein sequence (target) or a SwissProt / UniProt accession code (Kiefer, Arnold, Kunzli, Bordoli, & Schwede, 2009). Another mode, which requires more user intervention, is the “alignment mode”, where the user provides the input in the form of a multiple sequence alignment between the template and the target(s). The alignment can be prepared using any alignment tool (e.g. T-Coffee) in an acceptable format (FASTA, MSF, CLUSTALW, PFAM, SELEX). A third and more recent mode for

SwissModel is the "project mode", which enables the submission of a SwissModel project file (Kiefer et al., 2009; Schwede et al., 2003). This file is created by the program DeepView (Swiss-PdbViewer) (Guex & Peitsch, 1997) and contains superimposed template structures, and the alignment between the target and template. The project mode allows the user to have more control over a number of modeling parameters, such as the selection of the template structures, the correct alignment of the residues, and the ability to modify insertions and deletions in the context of the 3D structure.

3D-JIGSAW

3D-Jigsaw is also an online tool for homology modeling (Bates et al., 2001). After the alignment is done, the perfectly aligned regions in the sequence allow the generation of an initial model. 3D-JIGSAW performs a database fragment search to model poorly aligned gaps, loops and incompatible backbone angles. The backbone is selected from an ensemble of secondary structure elements and connecting loops using a self-consistent mean field approach. For the side-chains, the rotamers of the template structure as well as a rotamer library along with a second mean field calculation are taken into account. Loops are trimmed by adjusting torsion angles within each loop to give a good geometry. Finally, to remove steric clashes and reduce the number of structural errors, 100 steps of steepest descent energy minimization is run using CHARMM package (Brooks et al., 1983).

ROBETTA (Rosetta Comparative Modeling and ab initio Modeling)

Robetta is another publicly available server for protein structure prediction based on the ROSETTA standalone package (D. E. Kim et al., 2004; Song et al., 2013). It combines the Robetta de novo (Rohl, Strauss, Chivian, & Baker, 2004) and a number of homology modeling methods to facilitate the modeling of proteins even in the absence of homologous templates. Robetta follows the typical homology modeling protocol similar to the one shown in Figure B.2 in building a model. The initial step called 'Ginzu', involves BLAST, PSI-BLAST (S F Altschul et al., 1997) and 3D-Jury (Ginalski et al., 2003). This is followed by multiple sequence alignment using locally installed versions of HHsearch/HHpred, RaptorX, and Sparks-X. Alignments are clustered and comparative

models are generated using the RosettaCM protocol. The procedure is fully automated and is relatively fast in providing the user with the results.

In case of de novo prediction (when there is no homologous protein of known structure), Robetta uses a PDB-derived library of fragments that represent the range of accessible local structures for all short segments of the protein chain. Structures are then assembled randomly by fragment insertion using a scoring function that favours native protein-like features. In this way, it generates a large number of models, which are clustered together and the final models are selected based on RMSD over all un-gapped positions. Finally, using MAMMOTH (Ortiz, Strauss, & Olmea, 2002), the top models are compared to the PDB structures to find potential similarities.

Experimental nuclear magnetic resonance (NMR) constraints data can also be submitted with a query sequence for RosettaNMR de novo structure determination (Rohl et al., 2004). ROSIE, the RosettaAntibody protocol (Sircar, Kim, & Gray, 2009) is also a homology modeling program within the Rosetta suite, also available through a server that is optimized for predicting high-resolution antibody FV-structures.

B.6. Choosing a Homology Modeling Tool

Each modeling tool discussed so far has its own strengths as well as its limitations. A fair way to compare these different tools is by testing them against the same set of data using exactly the same set of parameters with an identical amount of elapsed time. This has been the case employed by the bi-annual critical assessment of techniques for protein structure prediction (CASP) that started in 1994 (J Moult, Pedersen, Judson, & Fidelis, 1995) and provides periodical comparison data for the different tools. CASP has attracted significant attention of the scientific community and the results have led to great improvements to the existing tools. It is perhaps one of the best sources for up-to-date information on current methods and algorithms. In addition, Continuous Automated Model Evaluation (CAMEO) community project (Haas et al., 2013), LiveBench (Bujnicki, Elofsson, Fischer, & Rychlewski, 2001) and EVA (Eyrich et al., 2001; Koh et al., 2003) are other sources that continually provide assessment of automated structure prediction servers.

According to the overall results of CASP meetings over years, I-TASSER (J. Yang et al., 2014), ROBETTA (D. E. Kim et al., 2004), HHpred (A. Hildebrand, Remmert, Biegert, & Soding, 2009), 3D-Jury (Ginalski et al., 2003), RaptorX (Källberg et al., 2012), SwissModel (Kiefer et al., 2009) and MODELLER (Šali & Blundell, 1993) were amongst the best protein structure prediction tools (http://www.predictioncenter.org/index.cgi?page=links). Based on the CAMEO results (http://cameo3d.org/sp), the most popular online servers are SwissModel (Schwede et al., 2003), RaptorX (Källberg et al., 2012), ROBETTA (D. E. Kim et al., 2004), IntFold servers (McGuffin, Atkins, Salehe, Shuid, & Roche, 2015) and Phyre2 (Kelley et al., 2015). More extensive comparative studies and benchmarking of available modeling programs and servers for high-accuracy homology modeling have been captured in (Eswar, 2003; Nayeem et al., 2006; Saxena et al., 2013; Björn Wallner & Elofsson, 2005).

B.7. Applications

As previously mentioned, the usefulness of the homology models and the level of information they provide is directly related to the quality and accuracy of the models. This in turn is dependent on the sequence similarity of the target to the template. Applications of homology modeling span over a wide range of fields, some of which are briefly described in the following sections along with appropriate case studies.

Structural Genomics

The aim of structural genomics is to determine or accurately predict the 3D structure of all the proteins encoded in the genome. Topologically similar proteins or gene families are great assets for further progress in the development of new drugs. They can also answer many challenging questions related to disease mechanisms and signalling pathways (Lundstrom, 2007). This can be achieved by a focused, large-scale determination of protein structures using X-ray crystallography and/or NMR spectroscopy, combined efficiently with accurate protein structure modeling techniques. As homology modeling is relatively easier compared to these experimental methods, it has been playing a pivotal role in this field (Marti-Renom et al., 2000) by speeding up the

process of structure determination using a representative of a class as a template and by predicting the structures of other related members.

One of the major efforts taken in this regard is the creation of the Structural Genomics Consortium (Williamson, 2000). This consortium attempts to obtain X-ray structures for broad representative structures across different families of human proteins and to deposit their structural coordinates in the public domain. Such representatives can then act as templates to use homology modeling and deriving 3D models for the structurally and functionally related proteins. These worldwide projects have accelerated the pace of protein-function analysis as well as drug discovery applications to a great extent.

Drug Design and Discovery

Perhaps the broadest applications of homology modeling are concentrated around the field of drug design and discovery. Numerous successful applications and case studies of homology modeling in drug discovery are extensively described in (Cavasotto & Phatak, 2009; Charlotte M Deane & Blundell, 2003; Hillisch et al., 2004; Vyas et al., 2012). In the absence of experimental structures of drug target proteins, homology models have supported the design of several potent pharmacological agents (Hillisch et al., 2004). Typical modeled targets span a wide range of protein structures including antibodies, viruses and many enzymatic, transport and structural proteins that are involved in human biology and pathophysiology (Rajapaksha & Petrovsky, 2014; K. Zhu et al., 2014). Searching for ligands of a given binding site (database mining), designing novel ligands, modeling substrate specificity and predicting antigenic epitopes are just a few examples.

Furthermore, homology models could be used to rationalize known experimental observations; for example, to support the hypotheses of medicinal chemists on how to generate biologically active and selective compounds in the early stages of drug design. *In silico* structure-based prediction of metabolism and toxicity of small molecules are yet other applications of such models (Vyas et al., 2012). However, typical applications of a homology model in drug discovery require a very high accuracy of the local side chain positions in the binding site and loop structures. Thus, the inaccuracy of modeling tools remains a challenge in the face of this important field.

G protein-coupled receptors (GPCRs) are a vastly studied class of proteins, for which many models were built using homology modeling. Given the complexity in experimentally determining their structure, the predicted structures have formed a crucial component of our current structural data on these proteins. This is mainly because experimental purification, structure determination and characterization for these proteins are extremely difficult. A complete overview of the major challenges in GPCR crystallography is discussed in (E. Ghosh et al., 2015). During recent years, better templates have become available for homology modeling of related GPCRs that have allowed structure-based drug design (SBDD) and virtual screening of agonists and antagonists for their activity (Heifetz, James, Morao, Bodkin, & Biggin, 2016; Ngo et al., 2016). Studies of the Alpha1A adrenergic receptor (Evers & Klabunde, 2005), protease activated receptor 2 (PAR2) (Perry et al., 2015), P2Y14 receptor (Trujillo, Paoletta, Kiselev, & Jacobson, 2015) and GPR132 receptor (Shehata et al., 2015), are just a few representatives of showing how homology modeling has facilitated progress in designing and discovering agonists and antagonists for GPCRs. Recently, Trujillo et al. (Trujillo et al., 2015) used a homology model for P2Y14R receptor along with the application of other powerful computational tools for the subsequent refinement of these models. The final model they generated was further used to qualitatively explain structure–activity relationships of existing analogues and to evaluate the quality of P2Y14R homology modeling as a template for structure-based design. This study revealed the structural components responsible for binding of the agonists and valuable data regarding the selectivity of agonists for this specific GPCR. Their study proves that homology modeling can allow the detailed prediction of interactions to facilitate the design of selective agonists with high affinity, as pharmacological tools to study the P2Y14R.

In another study by Arafat et al. (Arafat et al., 2014), the 3D structure of snake venom 5′ nucleotidase (SV-5′ NUC) was predicted using a comparative homology modeling approach. This model enabled them to study the key interactions of SV-5′ NUC using experimental studies/molecular docking analysis of a number of inhibitors. Further, atomic level docking interaction studies using inhibitors of the SV-5′ NUC active site has revealed details about its interactions as well as pharmacophoric features useful in

management of snakebites. Studies of this sort play a guiding role in the experimental design of new inhibitors against various important physiological targets.

Characterization of protein function and interaction

Functional annotation of protein sequences is merely based on their sequence similarities with other proteins. This is considered as a vague concept due to its lack of direct and clear relationship between sequence similarities and function (Tian & Skolnick, 2003). Addressing the function and mechanism of action of a particular protein is of utmost importance for drug discovery purposes. With the increase in availability of 3D protein structures, several methods have been developed to supplement sequence information with structural information in order to explore the biological function of proteins (Cavasotto & Phatak, 2009).

Another well-studied feature in this context stems from the fact that proteins interact both with indigenous ligands and/or other proteins to mediate their effect. The structure of protein-protein complexes can be constructed using the complexes of known structure as a template (Launay & Simonson, 2008). The complex structure templates can be detected either through homology-based sequence alignments (template based prediction) or given the structure of monomer components by structure-based comparisons. An excellent review of these methods and available tools are given in (Szilagyi & Zhang, 2014). Models built through homology modeling can help in understanding the details of protein-protein interactions and mechanisms at the atomistic level (Villoutreix et al., 2014; Waksman & Sansom, 2005). Furthermore, they also aid in identifying active and binding sites on proteins and to eventually facilitate the design of drugs (small molecules) either to stabilize these complexes or to block their interactions

One case study along this line is the successful work by Xu et al., where a homology model of the KCNQ1 potassium ion channel was built to characterize its interaction with an interacting protein partner, KCNE1 (Y. Xu et al., 2013b). This model has been validated and proved to be accurate enough to extract useful information regarding their interaction and how their interaction affects ion passage through the ion channel pore. Furthermore, the same model has been used to understand the mechanism of action of a channel activator (Y. Xu et al., 2015).

Dhanavade et al. were able to study the details and mechanism of function of cysteine protease, a protein that degrades amyloid beta peptide which is a causative agent of Alzheimer's disease (Dhanavade, Jalkute, Barage, & Sonawane, 2013). Docking simulations of this model revealed that specific residues of cysteine protease form an active site pocket similar to human cathepsin B. The predicted model might be useful in further studies of amyloid beta peptide degradation as well as to design new novel lead structures in the treatment of Alzheimer's disease.

An example of using homology modeling to understand disease mechanism is the work done by Merlino et al. (Merlino, Vieites, Gambino, & Laura Coitiño, 2014). Fumarate reductase is a central enzyme in the conversion of fumarate to succinate, an energy-releasing path essential for the survival of Leishmania major and Trypanosoma cruzi, causing leishmaniasis and Chagas disease, respectively. This enzyme has been considered as a good candidate for targeting by new drugs designed against these pathogens. They built a homology model for the NADH-dependent fumarate reductase of the two pathogens. Further studies on this model revealed structural features relevant to understanding the mechanism of action of the enzyme with special attention to the hydrogen bond network involving the cofactor and water. In addition, the model enabled them to test a set of inhibitory molecules at the binding sites of the enzymes revealing their binding modes and mechanism of action.

In addition to the above applications, homology modeling approaches can provide excellent starting models and facilitate molecular replacement in X-ray crystallography and NMR spectroscopy and refining models based on NMR constraints (Johnson et al., 1994).

Design of mutagenesis experiments using 3D structures

Mutagenesis studies play an important role in the identification of amino acids with relevant biological function in a protein (Antikainen & Martin, 2005; Brannigan & Wilkinson, 2002). This structural information helps in rationalizing the selection of amino acids for mutagenesis experimental approaches. There are several reports where homology models have aided mutagenesis experiments to study ligand-receptor interactions (Gagnidze et al., 2008; Silvestrov, Müller, Clark, Hausinger, & Cisneros,

2014; Szklarz & Halpert, 1997), analyze the role of non-conserved residues in active sites and suggest ligand binding modes (Anwar-Mohamed et al., 2014; Lerche et al., 2007).

Cytochrome P450 (CYP450), for example, is a very important protein for which homology modeling has become a crucial tool, especially in conjunction with site-directed mutagenesis approaches (Szklarz, Ornstein, & Halpert, 1994). Molecular models of various CYP450s can be constructed based on the available P450 crystal structures and modified to mimic the mutated enzymes in physiological conditions. Such modifications could address the altered binding of substrates and/or drugs (inhibitors or activators) to the P450 enzymes as well as alterations in inhibition and activation due to residue replacement. Overall, they help to identify or confirm key residues, evaluate enzyme-substrate interactions and explain changes in protein stability and/or regio- and stereospecificity of substrate oxidation upon residue substitution by site-directed mutagenesis.

In another study by Ismail et al. (Ismail, Sharma, Kumar, Kannangai, & Abraham, 2013), a homology model of hepatitis B virus polymerase protein was built which helped in confirming the experimental findings related to a mutation in this protein and its importance for the action of an antiviral drug, adefovir. These findings provided evidence for the selection and counterselection of the mutant virus resistance in patients treated by adefovir during antiviral therapy. This example clearly illustrates how homology modeling can be applied right from a research question up to solving clinical problems.

Li et al. (M. Li & Wang, 2007) investigated the effect of an important mutation, D92E in the NS1 protein of the H5N1 strains of influenza virus. For this purpose, they made an extensive use of homology modeling to develop 3D models for the H5N1 NS1 protein. This NS1 mutation has been correlated with an increased virulence and/or cytokine resistance. The NS1 protein is also a potential target for the development of novel antiviral agents against H5N1 strains. Furthermore, they explored the structural changes brought about by the D92E mutation involving weakened interactions with host targets leading to changes of the protein structure and function. The detailed structural understanding achieved may help structure-based design of novel antiviral agents against the H5N1 avian influenza virus.

B.8. Limitations and Challenges

As previously discussed, homology modeling has noteworthy potential as a tool in different fields. However, there are certain limitations that impede a more extensive use of the generated models in important applications, specifically when it comes to drug discovery process and characterization of pharmacological targets. Some of the clear shortcomings are linked to template identification and accurate sequence alignment that can lead to false results and low quality models.

Furthermore, high-resolution refinement of homology models remains a challenging problem because of the incapability of sampling huge numbers of alternatively packed conformations and lack of high precision in the current force fields (Ginalski, 2006). Refinement tools heavily depend on the type of protein being tested and occasionally they perform better for certain typical types of proteins compared to others, depending on the training set and algorithms used to develop the specific tools. Another important feature of the homology modeling process is accurate scoring and evaluation of the models. Currently, there are very few tools for evaluation of modelled protein structures and there is definitely a need for more tools with higher levels of precision and sensitivity to detect errors in the structures.

Models become more error-prone and produce poor predictions, when target and template proteins share lower sequence similarities (Rodrigues et al., 2013). Thus, there is a clear need to develop better protocols of template identification and sequence alignment to reduce modeling errors to rally the ultimate model. However, there will always be structural variances between the target and its templates and these variances have to be recognized and recompensed for by *ab initio* modeling (Hardin, Pogorelov, & Luthey-Schulten, 2002) or by post-modeling optimization approaches (Krieger et al., 2009; Y. Li & Zhang, 2009). Therefore, in the current state, a platform consisting of a combination of protein structure prediction techniques would help to counteract the existing limitations of each technique.

Despite all the advances and efforts, the modeling of certain classes of proteins still remains a challenge. Particularly, these include membrane proteins (Carpenter et al., 2008) and intrinsically disordered proteins (IDPs) (181). Membrane proteins are associated with difficulties in purification and crystallization, for example G-protein

coupled receptors (GPCRs). This in turn causes a scarcity of templates and makes them nevertheless difficult to model using homology modeling and other template based prediction methods. IDPs on the other hand, are proteins that lack a fixed or ordered 3D structure and possess distinct properties in terms of function, structure, sequence, interactions, evolution and regulation (Fisher et al., 2010). These extraordinary properties make them difficult to be widely studied and furthermore the existing homology modeling techniques are not optimized to model the highly dynamic and turbulent structural behaviour of these proteins.

B.9. Conclusion

In the absence of experimental structures, homology modeling is a reliable alternative to predict the 3D structure of proteins. Increased availability of reliable and fast computational resources as well as improvements in the prediction algorithms has enabled more exhaustive and accurate prediction in less time.

The process of comparative protein structure modeling usually requires the simultaneous use of many programs and tools, to identify template structures, to generate sequence– structure alignments, to build the models and to evaluate them. In addition, various sequence and structure databases that are accessed by these programs are needed. Once an initial model is calculated, it is generally refined and finally analyzed in the context of many other related proteins and their functional annotations. In this regard, the CASP meeting is an excellent resource for evaluating the current state of the art tools and programs in the field of protein structure prediction. The results of CASP evaluations are available and published in the official website (http://predictioncentre.org/), for the scientific community.

The various applications of predicted protein models have different requirements with regards to precision and resolution. High-resolution models are useful for atomistic molecular modeling (for instance pharmacological targets), whereas lower resolution models might still be practical for designing site-directed mutagenesis experiments, epitope mapping, or supporting experimental structure determination. In general, successful applications include the use of such models in rational drug design, guiding

experimental design, protein annotation, structural genomics and mutagenesis studies. Currently, most modeling cases fall in within the 20-30% sequence similarity, where the majority of new information is generated and further progress is required to overcome the existing limitations and challenges of homology modeling to move the field forward.

The current review not only covers the fundamentals of homology modeling like the existing reviews, but it also covers many more aspects of the topic in hand. This review helps the reader in understanding the operational basics behind homology modeling technique and deciding whether it is the right tool for solving their problem of interest. Furthermore, choosing a suitable tool for modeling and identifying reliable resources to gain up to date information on benchmarking and comparison of different tools and software are also discussed. The reader is familiarized with the existing limitations and possible aspects of modeling which might need a closer monitoring during the modeling process. And finally, possible applications and successful examples are introduced to provide a wider perspective of homology modeling.

B.10. References

Abagyan, R., Totrov, M., & Kuznetsov, D. (1994). ICM - A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, *15*(5), 488–506. http://doi.org/10.1002/jcc.540150503

Allison, J. R., Hertig, S., Missimer, J. H., Smith, L. J., Steinmetz, M. O., & Dolenc, J. (2012). Probing the structure and dynamics of proteins by combining molecular dynamics simulations and experimental NMR data. *Journal of Chemical Theory and Computation*, *8*(10), 3430–3444. http://doi.org/10.1021/ct300393b

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. http://doi.org/http://dx.doi.org/10.1016/S0022-2836(05)80360-2

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, *25*(17), 3389–3402.

Antikainen, N. M., & Martin, S. F. (2005). Altering protein specificity: techniques and applications. *Bioorganic & Medicinal Chemistry*, *13*(8), 2701–2716. Journal Article, Research Support, U.S. Gov't, P.H.S., Review. http://doi.org/10.1016/j.bmc.2005.01.059

Anwar-Mohamed, A., Barakat, K. H., Bhat, R., Noskov, S. Y., Tyrrell, D. L., Tuszynski, J. a., & Houghton, M. (2014). A human ether-á-go-go-related (hERG) ion channel atomistic model generated by long supercomputer molecular dynamics simulations and its use in predicting drug cardiotoxicity. *Toxicology Letters*, *230*(3), 382–392. http://doi.org/10.1016/j.toxlet.2014.08.007

Arafat, A. S. Y., Arun, A., Ilamathi, M., Asha, J., Sivashankari, P. R., D'Souza, C. J. M., … Dhananjaya, B. L. (2014). Homology modeling, molecular dynamics and atomic level interaction study of snake venom 5' nucleotidase. *Journal of Molecular Modeling*, *20*(3), 2156. http://doi.org/10.1007/s00894-014-2156-1

Aszodi, A., & Taylor, W. R. (1994). Secondary structure formation in model polypeptide chains. *Protein Eng*, *7*(5), 633–644. http://doi.org/10.1093/protein/7.5.633

Aszodi, A., & Taylor, W. R. (1996). Homology modelling by distance geometry. *Folding & Design*, *1*(5), 325–334. http://doi.org/10.1016/S1359-0278(96)00048-X

Bairoch, A., & Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Research*, *28*(1), 45–48.

Bates, P. A., Kelley, L. A., MacCallum, R. M., & Sternberg, M. J. (2001). Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins*, *Suppl 5*, 39–46. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11835480

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2005). GenBank. *Nucleic Acids Research*, *33*(Database issue), D34–D38. http://doi.org/10.1093/nar/gki063

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., … Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. http://doi.org/10.1093/nar/28.1.235

Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E., & Thornton, J. M. (1987). Knowledge-based prediction of protein structures and the design of novel molecules. *Nature*, *326*(6111), 347–352. Retrieved from http://dx.doi.org/10.1038/326347a0

Bowie, J. U., Luthy, R., & Eisenberg, D. (1991). A method to identify protein sequences that fold into a known three-dimensional structure. *Science (New York, N.Y.)*, *253*(5016), 164–170.

Brannigan, J. a, & Wilkinson, A. J. (2002). Protein engineering 20 years on. *Nature Reviews. Molecular Cell Biology*, *3*(12), 964–70. http://doi.org/10.1038/nrm975

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *Journal of Computational Chemistry*, *4*(2), 187–217. http://doi.org/10.1002/jcc.540040211

Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., & Jones, D. T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Research*, *41*(Web Server issue), W349-57. http://doi.org/10.1093/nar/gkt381

Bujnicki, J. M., Elofsson, A., Fischer, D., & Rychlewski, L. (2001). LiveBench-2: large-scale automated evaluation of protein structure prediction servers. *Proteins*, *Suppl 5*, 184–91. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/11835496
Buxbaum, E. (2007). Cell skeleton. In *Fundamentals of Protein Structure and Function* (pp. 175–184). Boston, MA: Springer US. http://doi.org/10.1007/978-0-387-68480-2_11
Canutescu, A. A., Shelenkov, A. A., & Dunbrack, R. L. J. (2003). A graph-theory algorithm for rapid protein side-chain prediction. *Protein Science : A Publication of the Protein Society*, *12*(9), 2001–2014. http://doi.org/10.1110/ps.03154503
Carpenter, E. P., Beis, K., Cameron, A. D., & Iwata, S. (2008). Overcoming the challenges of membrane protein crystallography. *Current Opinion in Structural Biology*, *18*(5), 581–586. http://doi.org/10.1016/j.sbi.2008.07.001
Cavasotto, C. N., & Phatak, S. S. (2009). Homology modeling in drug discovery: current trends and applications. *Drug Discovery Today*, *14*(13–14), 676–83. http://doi.org/10.1016/j.drudis.2009.04.006
Chen, C. Y.-C., & Tou, W. leong. (2013). How to design a drug for the disordered proteins? *Drug Discovery Today*, *18*(19–20), 910–915. http://doi.org/10.1016/j.drudis.2013.04.008
Coleman, J. E. (1992). Zinc proteins: enzymes, storage proteins, transcription factors, and replication proteins. *Annual Review of Biochemistry*, *61*, 897–946. http://doi.org/10.1146/annurev.biochem.61.1.897
Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Science : A Publication of the Protein Society*, *2*(9), 1511–1519. http://doi.org/10.1002/pro.5560020916
Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research*, *16*(22), 10881–10890.
Dalton, J. A. R., & Jackson, R. M. (2007). An evaluation of automated homology modelling methods at low target-template sequence similarity. *Bioinformatics*, *23*(15), 1901–1908. http://doi.org/10.1093/bioinformatics/btm262
Deane, C. M., & Blundell, T. L. (2001). CODA: a combined algorithm for predicting the structurally variable regions of protein models. *Protein Science : A Publication of the Protein Society*, *10*(3), 599–612. http://doi.org/10.1110/ps.37601
Deane, C. M., & Blundell, T. L. (2003). 27 - Protein Comparative Modelling and Drug Discovery. In C. G. B. T.-T. P. of M. C. (Second E. Wermuth (Ed.), (pp. 445–458). London: Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-012744481-9/50031-3
Dhanavade, M. J., Jalkute, C. B., Barage, S. H., & Sonawane, K. D. (2013). Homology modeling, molecular docking and MD simulation studies to investigate role of cysteine protease from Xanthomonas campestris in degradation of Aβ peptide. *Computers in Biology and Medicine*, *43*(12), 2063–2070. http://doi.org/http://dx.doi.org/10.1016/j.compbiomed.2013.09.021
Dunbrack, R. L. J., & Karplus, M. (1994). Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nature Structural Biology*, *1*(5), 334–340.
Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics (Oxford, England)*, *14*(9), 755–763.
Eswar, N. (2003). Tools for comparative protein structure modeling and analysis. *Nucleic Acids Research*, *31*(13), 3375–3380. http://doi.org/10.1093/nar/gkg543
Evers, A., & Klabunde, T. (2005). Structure-based Drug Discovery Using GPCR Homology Modeling: Successful Virtual Screening for Antagonists of the Alpha1A Adrenergic Receptor. *Journal of Medicinal Chemistry*, *48*(4), 1088–1097. http://doi.org/10.1021/jm0491804
Eyrich, V. A., Marti-Renom, M. A., Przybylski, D., Madhusudhan, M. S., Fiser, A., Pazos, F., … Rost, B. (2001). EVA: continuous automatic evaluation of protein structure prediction servers. *Bioinformatics (Oxford, England)*, *17*(12), 1242–1243.
Fallis, A. . (2013). *Fundamentals of Protein Structure and Function*. *Journal of Chemical Information and Modeling* (Vol. 53). http://doi.org/10.1017/CBO9781107415324.004
Fan, H. (2004). Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Science*, *13*(1), 211–220. http://doi.org/10.1110/ps.03381404
Ferreira, P. M. (1995). Protein structure: By N J Darby and T E Creighton. pp 97. IRL Press, Oxford University Press. 1993. SBN 0-19-963310-X. *Biochemical Education*, *23*(1), 46. http://doi.org/10.1016/0307-4412(95)90200-7
Fiser, A., Do, R., & Sali, A. (2000). Modeling of loops in protein structures. *PRS*, *9*(9), 1753–1773.
Fiser, a, Fiser, a, Do, R. K., Do, R. K., Sali, a, & Sali, a. (2000). Modeling of loops in protein structures. *Protein Science : A Publication of the Protein Society*, *9*(9), 1753–73. http://doi.org/10.1110/ps.9.9.1753

Fisher, C. K., Huang, A., & Stultz, C. M. (2010). Modeling intrinsically disordered proteins with Bayesian statistics. *Journal of the American Chemical Society*, *132*(42), 14919–14927. http://doi.org/10.1021/ja105832g

Francis-Lyon, P., & Koehl, P. (2014). Protein side-chain modeling with a protein-dependent optimized rotamer library. *Proteins: Structure, Function and Bioinformatics*, *82*(9), 2000–2017. http://doi.org/10.1002/prot.24555

Gagneux, P. (2004). Protein Structure and Function. *Journal of Heredity*, *95*(3), 274–274. http://doi.org/10.1093/jhered/esh040

Gagnidze, K., Sachchidanand, Rozenfeld, R., Mezei, M., Zhou, M.-M., & Devi, L. A. (2008). Homology Modeling and Site-directed Mutagenesis to Identify Selective Inhibitors of Endothelin-Converting Enzyme-2. *Journal of Medicinal Chemistry*, *51*(12), 3378–3387. http://doi.org/10.1021/jm7015478

Gerstein, M. (1997). A structural census of genomes: comparing bacterial, eukaryotic, and archaeal genomes in terms of protein structure. *Journal of Molecular Biology*, *274*(4), 562–576. http://doi.org/10.1006/jmbi.1997.1412

Ghosh, E., Kumari, P., Jaiman, D., & Shukla, A. K. (2015). Methodological advances: the unsung heroes of the GPCR structural revolution. *Nat Rev Mol Cell Biol*, *16*(2), 69–81. Retrieved from http://dx.doi.org/10.1038/nrm3933

Ginalski, K. (2006). Comparative modeling for protein structure prediction. *Current Opinion in Structural Biology*, *16*(2), 172–177. http://doi.org/10.1016/j.sbi.2006.02.003

Ginalski, K., Elofsson, A., Fischer, D., & Rychlewski, L. (2003). 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics (Oxford, England)*, *19*(8), 1015–1018.

Godzik, A., Kolinski, A., & Skolnick, J. (1992). Topology fingerprint approach to the inverse protein folding problem. *Journal of Molecular Biology*, *227*(1), 227–238. http://doi.org/http://dx.doi.org/10.1016/0022-2836(92)90693-E

Gribskov, M., McLachlan, A. D., & Eisenberg, D. (1987). Profile analysis: detection of distantly related proteins. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(13), 4355–4358.

Guex, N., & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, *18*(15), 2714–2723. http://doi.org/10.1002/elps.1150181505

Haas, J., Roth, S., Arnold, K., Kiefer, F., Schmidt, T., Bordoli, L., & Schwede, T. (2013). The Protein Model Portal--a comprehensive resource for protein structure and model information. *Database : The Journal of Biological Databases and Curation*, *2013*, bat031. http://doi.org/10.1093/database/bat031

Han, R., Leo-Macias, A., Zerbino, D., Bastolla, U., Contreras-Moreira, B., & Ortiz, A. R. (2008). An efficient conformational sampling method for homology modeling. *Proteins*, *71*(1), 175–88. http://doi.org/10.1002/prot.21672

Hardin, C., Pogorelov, T. V., & Luthey-Schulten, Z. (2002). Ab initio protein structure prediction. *Current Opinion in Structural Biology*, *12*(2), 176–181. http://doi.org/10.1016/S0959-440X(02)00306-8

Hasani, H. J., & Barakat, K. (n.d.). Protein-Protein Docking: *Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery*, 173–195. http://doi.org/10.4018/978-1-5225-0115-2.ch007

Heifetz, A., James, T., Morao, I., Bodkin, M. J., & Biggin, P. C. (2016). Guiding lead optimization with GPCR structure modeling and molecular dynamics. *Current Opinion in Pharmacology*, *30*, 14–21. http://doi.org/http://dx.doi.org/10.1016/j.coph.2016.06.004

Hildebrand, A., Remmert, M., Biegert, A., & Soding, J. (2009). Fast and accurate automatic structure prediction with HHpred. *Proteins*, *77 Suppl 9*, 128–132. http://doi.org/10.1002/prot.22499

Hille, B. (2001). Ion Channel Excitable Membranes. *Sunderland Massachusetts USA*. http://doi.org/10.1007/3-540-29623-9_5640

Hillisch, A., Pineda, L. F., & Hilgenfeld, R. (2004). Utility of homology models in the drug discovery process. *Drug Discovery Today*, *9*(15), 659–69. http://doi.org/10.1016/S1359-6446(04)03196-4

Holm, L., & Rosenström, P. (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Research* , *38*(suppl 2), W545–W549. http://doi.org/10.1093/nar/gkq366

Holm, L., & Sander, C. (1991). Database algorithm for generating protein backbone and side-chain co-ordinates from a C alpha trace application to model building and detection of co-ordinate errors. *Journal of Molecular Biology*, *218*(1), 183–194.

Hongmao, S. (2016). Chapter 4 - Homology Modeling and Ligand-Based Molecule Design. In S. B. T.-A.

P. G. to R. D. D. Hongmao (Ed.), *A Practical Guide to Rational Drug Design* (pp. 109–160). Woodhead Publishing. http://doi.org/http://dx.doi.org/10.1016/B978-0-08-100098-4.00004-1

Hooft, R. W., Vriend, G., Sander, C., & Abola, E. E. (1996, May). Errors in protein structures. *Nature*. ENGLAND. http://doi.org/10.1038/381272a0

Ishitani, R., Terada, T., & Shimizu, K. (2008). Refinement of comparative models of protein structure by using multicanonical molecular dynamics simulations. *Molecular Simulation*, *34*(3), 327–336. http://doi.org/10.1080/08927020801930539

Ismail, A. M., Sharma, O. P., Kumar, M. S., Kannangai, R., & Abraham, P. (2013). Impact of rtI233V mutation in hepatitis B virus polymerase protein and adefovir efficacy: Homology modeling and molecular docking studies. *Bioinformation*, *9*(3), 121–125. http://doi.org/10.6026/97320630009121

Iwata, Y., Kasuya, A., & Miyamoto, S. (2002). An efficient method for reconstructing protein backbones from alpha-carbon coordinates. *Journal of Molecular Graphics & Modelling*, *21*(2), 119–128.

Johnson, M. S., & Overington, J. P. (1993). A structural basis for sequence comparisons. An evaluation of scoring methodologies. *Journal of Molecular Biology*, *233*(4), 716–738. http://doi.org/10.1006/jmbi.1993.1548

Johnson, M. S., Srinivasan, N., Sowdhamini, R., & Blundell, T. L. (1994). Knowledge-based protein modeling. *Critical Reviews in Biochemistry and Molecular Biology*, *29*(1), 1–68. http://doi.org/10.3109/10409239409086797

Jones, D. T. (1999). Protein secondary structure prediction based on position-specific scoring matrices. *Journal of Molecular Biology*, *292*(2), 195–202. http://doi.org/10.1006/jmbi.1999.3091

Jones, D. T., Taylor, W. R., & Thornton, J. M. (1992). A new approach to protein fold recognition. *Nature*, *358*(6381), 86–89. http://doi.org/10.1038/358086a0

Kahraman, A., Herzog, F., Leitner, A., Rosenberger, G., Aebersold, R., & Malmström, L. (2013). Cross-Link Guided Molecular Modeling with ROSETTA. *PLoS ONE*, *8*(9), e73411. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0073411

Källberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H., & Xu, J. (2012). Template-based protein structure modeling using the RaptorX web server. *Nat. Protocols*, *7*(8), 1511–1522. Retrieved from http://dx.doi.org/10.1038/nprot.2012.085

Kann, M. G. (2007). Protein interactions and disease: Computational approaches to uncover the etiology of diseases. *Briefings in Bioinformatics*, *8*(5), 333–346. http://doi.org/10.1093/bib/bbm031

Kannan, S., & Zacharias, M. (2010). Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent. *Proteins*, *78*(13), 2809–19. http://doi.org/10.1002/prot.22796

Karplus, K., Barrett, C., & Hughey, R. (1998). Hidden Markov models for detecting remote protein homologies. *Bioinformatics (Oxford, England)*, *14*(10), 846–856.

Karplus, K., Karchin, R., Draper, J., Casper, J., Mandel-Gutfreund, Y., Diekhans, M., & Hughey, R. (2003). Combining local-structure, fold-recognition, and new fold methods for protein structure prediction. *Proteins*, *53 Suppl 6*, 491–496. http://doi.org/10.1002/prot.10540

Kelley, L. A., MacCallum, R. M., & Sternberg, M. J. E. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM1. *Journal of Molecular Biology*, *299*(2), 501–522. http://doi.org/http://dx.doi.org/10.1006/jmbi.2000.3741

Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protocols*, *10*(6), 845–858. Retrieved from http://dx.doi.org/10.1038/nprot.2015.053

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., & Schwede, T. (2009). The SWISS-MODEL Repository and associated resources. *Nucleic Acids Research*, *37*(Database issue), D387-92. http://doi.org/10.1093/nar/gkn750

Kim, D. E., Chivian, D., & Baker, D. (2004). Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Research*, *32*(Web Server issue), W526-31. http://doi.org/10.1093/nar/gkh468

Koh, I. Y. Y., Eyrich, V. A., Marti-Renom, M. A., Przybylski, D., Madhusudhan, M. S., Eswar, N., … Rost, B. (2003). EVA: Evaluation of protein structure prediction servers. *Nucleic Acids Research*, *31*(13), 3311–3315.

Korasick, D. A., & Jez, J. M. (2016). Protein Domains: Structure, Function, and Methods. In R. A. Bradshaw & P. D. B. T.-E. of C. B. Stahl (Eds.), (pp. 91–97). Waltham: Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-0-12-394447-4.10011-2

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., … Karplus, K. (2009). Improving physical

realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins: Structure, Function and Bioinformatics*, *77*(SUPPL. 9), 114–122. http://doi.org/10.1002/prot.22570

Krieger, E., Nabuurs, S. B., & Vriend, G. (2003). Homology Modeling. In *Structural Bioinformatics* (pp. 509–523). John Wiley & Sons, Inc. http://doi.org/10.1002/0471721204.ch25

Krissinel, E., & Henrick, K. (2004). Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallographica Section D*, *60*(12 Part 1), 2256–2268. http://doi.org/10.1107/S0907444904026460

Krivov, G. G., Shapovalov, M. V., & Dunbrack, R. L. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins: Structure, Function and Bioinformatics*, *77*(4), 778–795. http://doi.org/10.1002/prot.22488

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., … Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics*, *23*(21), 2947–2948. http://doi.org/10.1093/bioinformatics/btm404

Larsson, P., Wallner, B., Lindahl, E., & Elofsson, A. (2008). Using multiple templates to improve quality of homology models in automated homology modeling. *Protein Science : A Publication of the Protein Society*, *17*(6), 990–1002. http://doi.org/10.1110/ps.073344908

Laskowski, R. A., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, *26*(C), 283–291. http://doi.org/10.1107/S0021889892009944

Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, *8*(4), 477–486.

Laskowski, R. A., & Swaminathan, G. J. (2013). *Problems of Protein Three-Dimensional Structures*. *Reference Module in Chemistry, Molecular Sciences and Chemical Engineering*. Elsevier Inc. http://doi.org/10.1016/B978-0-12-409547-2.02535-X

Launay, G., & Simonson, T. (2008). Homology modelling of protein-protein complexes: a simple method and its possibilities and limitations. *BMC Bioinformatics*, *9*, 427. http://doi.org/10.1186/1471-2105-9-427

Lerche, C., Bruhova, I., Lerche, H., Steinmeyer, K., Wei, A. D., Strutz-Seebohm, N., … Seebohm, G. (2007). Chromanol 293B binding in KCNQ1 (Kv7.1) channels involves electrostatic interactions with a potassium ion in the selectivity filter. *Molecular Pharmacology*, *71*, 1503–1511. http://doi.org/10.1124/mol.106.031682

Lesk, A. M. (1997). CASP2: report on ab initio predictions. *Proteins*, *Suppl 1*, 151–166. http://doi.org/10.1002/(sici)1097-0134(1997)1+<151::aid-prot20>3.3.co;2-j

Levitt, M. (1992). Accurate modeling of protein conformation by automatic segment matching. *Journal of Molecular Biology*, *226*(2), 507–533. http://doi.org/http://dx.doi.org/10.1016/0022-2836(92)90964-L

Li, M., & Wang, B. (2007). Homology modeling and examination of the effect of the D92E mutation on the H5N1 nonstructural protein NS1 effector domain. *Journal of Molecular Modeling*, *13*(12), 1237–1244. http://doi.org/10.1007/s00894-007-0245-0

Li, Y., & Zhang, Y. (2009). REMO: A new protocol to refine full atomic protein models from C-alpha traces by optimizing hydrogen-bonding networks. *Proteins: Structure, Function and Bioinformatics*, *76*(3), 665–676. http://doi.org/10.1002/prot.22380

Liang, S., & Grishin, N. V. (2002). Side-chain modeling with an optimized scoring function. *Protein Science : A Publication of the Protein Society*, *11*(2), 322–331. http://doi.org/10.1110/ps.24902

Lu, H., & Skolnick, J. (2003). Application of statistical potentials to protein structure refinement from low resolution ab initio models. *Biopolymers*, *70*(4), 575–584. http://doi.org/10.1002/bip.10537

Lundstrom, K. (2007). Structural genomics and drug discovery. *Journal of Cellular and Molecular Medicine*, *11*(2), 224–238. http://doi.org/10.1111/j.1582-4934.2007.00028.x

Ma, J., & Wang, S. (2014). Chapter Five - Algorithms, Applications, and Challenges of Protein Structure Alignment. In R. D. B. T.-A. in P. C. and S. Biology (Ed.), (Vol. Volume 94, pp. 121–175). Academic Press. http://doi.org/http://dx.doi.org/10.1016/B978-0-12-800168-4.00005-6

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., … Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry. B*, *102*(18), 3586–3616. http://doi.org/10.1021/jp973084f

Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F., & Sali, A. (2000). Comparative protein

structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure*, *29*, 291–325. http://doi.org/10.1146/annurev.biophys.29.1.291

McGuffin, L. J., Atkins, J. D., Salehe, B. R., Shuid, A. N., & Roche, D. B. (2015). IntFOLD: an integrated server for modelling protein structures and functions from amino acid sequences. *Nucleic Acids Research*, *43*(W1), W169-73. http://doi.org/10.1093/nar/gkv236

Melo, F., Devos, D., Depiereux, E., & Feytmans, E. (1997). ANOLEA: a www server to assess protein structures. *Proceedings / ... International Conference on Intelligent Systems for Molecular Biology ; ISMB. International Conference on Intelligent Systems for Molecular Biology*, *5*, 187–190.

Merlino, A., Vieites, M., Gambino, D., & Laura Coitiño, E. (2014). Homology modeling of T. cruzi and L. major NADH-dependent fumarate reductases: Ligand docking, molecular dynamics validation, and insights on their binding modes. *Journal of Molecular Graphics and Modelling*, *48*, 47–59. http://doi.org/http://dx.doi.org/10.1016/j.jmgm.2013.12.001

Mills, C. L., Beuning, P. J., & Ondrechen, M. J. (2015). Biochemical functional predictions for protein structures of unknown or uncertain function. *Computational and Structural Biotechnology Journal*, *13*, 182–191. http://doi.org/10.1016/j.csbj.2015.02.003

Moult, J., Pedersen, J. T., Judson, R., & Fidelis, K. (1995). A large-scale experiment to assess protein structure prediction methods. *Proteins*, *23*(3), ii–v. http://doi.org/10.1002/prot.340230303

Muller, A., MacCallum, R. M., & Sternberg, M. J. (1999). Benchmarking PSI-BLAST in genome annotation. *Journal of Molecular Biology*, *293*(5), 1257–1271. http://doi.org/10.1006/jmbi.1999.3233

Murzin, A. G., Brenner, S. E., Hubbard, T., & Chothia, C. (1995). SCOP: A structural classification of proteins database for the investigation of sequences and structures. *Journal of Molecular Biology*, *247*(4), 536–540. http://doi.org/http://dx.doi.org/10.1016/S0022-2836(05)80134-2

Nayeem, A., Sitkoff, D., & Krystek, S. (2006). A comparative study of available software for high-accuracy homology modeling: from sequence alignments to structural models. *Protein Science : A Publication of the Protein Society*, *15*(4), 808–824. http://doi.org/10.1110/ps.051892906

Ngo, T., Kufareva, I., Coleman, J. L., Graham, R. M., Abagyan, R., & Smith, N. J. (2016). Identifying ligands at orphan GPCRs: Current status using structure-based approaches. *British Journal of Pharmacology*. http://doi.org/10.1111/bph.13452

Notredame, C., Higgins, D. G., & Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, *302*(1), 205—217. http://doi.org/10.1006/jmbi.2000.4042

Orengo, C. A., Bray, J. E., Buchan, D. W. A., Harrison, A., Lee, D., Pearl, F. M. G., … Thornton, J. M. (2002). The CATH protein family database: a resource for structural and functional annotation of genomes. *Proteomics*, *2*(1), 11–21.

Ortiz, A. R., Strauss, C. E. M., & Olmea, O. (2002). MAMMOTH (matching molecular models obtained from theory): an automated method for model comparison. *Protein Science : A Publication of the Protein Society*, *11*(11), 2606–2621. http://doi.org/10.1110/ps.0215902

Pabo, C. O., & Sauer, R. T. (2003). Transcription Factors: Structural Families and Principles of DNA Recognition. *Annual Review of Biochemistry*. http://doi.org/10.1146/annurev.biochem.61.1.1053

Pearson, W. R. (1990). Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods in Enzymology*, *183*, 63–98.

Pearson, W. R. (1998). Empirical statistical estimates for sequence similarity searches. *Journal of Molecular Biology*, *276*(1), 71–84. http://doi.org/10.1006/jmbi.1997.1525

Peng, J., & Xu, J. (2011). A multiple-template approach to protein threading. *Proteins: Structure, Function, and Bioinformatics*, *79*(6), 1930–1939. http://doi.org/10.1002/prot.23016

Perry, S. R., Xu, W., Wirija, A., Lim, J., Yau, M.-K., Stoermer, M. J., … Fairlie, D. P. (2015). Three Homology Models of PAR2 Derived from Different Templates: Application to Antagonist Discovery. *Journal of Chemical Information and Modeling*, *55*(6), 1181–1191. http://doi.org/10.1021/acs.jcim.5b00087

Perutz, M. (2012). Protein Structure, 1–18. http://doi.org/10.5772/2335

Petrey, D., Xiang, Z., Tang, C. L., Xie, L., Gimpelev, M., Mitros, T., … Honig, B. (2003). Using multiple structure alignments, fast model building, and energetic analysis in fold recognition and homology modeling. *Proteins*, *53 Suppl 6*, 430–435. http://doi.org/10.1002/prot.10550

Pieper, U., Eswar, N., Davis, F. P., Braberg, H., Madhusudhan, M. S., Rossi, A., … Sali, A. (2006). MODBASE: a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Research*, *34*(Database issue), D291-5. http://doi.org/10.1093/nar/gkj059

Qian, B., Ortiz, A. R., & Baker, D. (2004). Improvement of comparative model accuracy by free-energy optimization along principal components of natural structural variation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(43), 15346–15351. http://doi.org/10.1073/pnas.0404703101

Rajapaksha, H., & Petrovsky, N. (2014). <italic>In Silico</italic> Structural Homology Modelling and Docking for Assessment of Pandemic Potential of a Novel H7N9 Influenza Virus and Its Ability to Be Neutralized by Existing Anti-Hemagglutinin Antibodies. *PLoS ONE*, *9*(7), e102618. Retrieved from http://dx.doi.org/10.1371%2Fjournal.pone.0102618

Rappsilber, J. (2011). The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of Structural Biology*, *173*(3), 530–540. http://doi.org/http://dx.doi.org/10.1016/j.jsb.2010.10.014

Raval, A., Piana, S., Eastwood, M. P., Dror, R. O., & Shaw, D. E. (2012). Refinement of protein structure homology models via long, all-atom molecular dynamics simulations. *Proteins*, *80*(8), 2071–9. http://doi.org/10.1002/prot.24098

Rodrigues, J. P. G. L. M., Melquiond, A. S. J., Karaca, E., Trellet, M., Van Dijk, M., Van Zundert, G. C. P., … Bonvin, A. M. J. J. (2013). Defining the limits of homology modeling in information-driven protein docking. *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2119–2128. http://doi.org/10.1002/prot.24382

Rodriguez, R., Chinea, G., Lopez, N., Pons, T., & Vriend, G. (1998). Homology modeling, model and software evaluation: three related resources. *Bioinformatics (Oxford, England)*, *14*(6), 523–528. http://doi.org/10.1093/bioinformatics/14.6.523

Rohl, C. A., Strauss, C. E. M., Chivian, D., & Baker, D. (2004). Modeling structurally variable regions in homologous proteins with rosetta. *Proteins*, *55*(3), 656–677. http://doi.org/10.1002/prot.10629

Roy, A., Kucukural, A., & Zhang, Y. (2010). I-TASSER: a unified platform for automated protein structure and function prediction. *Nature Protocols*, *5*(4), 725–738. http://doi.org/10.1038/nprot.2010.5

Rychlewski, L., Jaroszewski, L., Li, W., & Godzik, A. (2000). Comparison of sequence profiles. Strategies for structural predictions using sequence information. *Protein Science : A Publication of the Protein Society*, *9*(2), 232–241. http://doi.org/10.1110/ps.9.2.232

Šali, A., & Blundell, T. L. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology*, *234*(3), 779–815. http://doi.org/http://dx.doi.org/10.1006/jmbi.1993.1626

Sali, A., & Overington, J. P. (1994). Derivation of rules for comparative protein modeling from a database of protein structure alignments. *Protein Science : A Publication of the Protein Society*, *3*(9), 1582–1596. http://doi.org/10.1002/pro.5560030923

Sali, A., Potterton, L., Yuan, F., van Vlijmen, H., & Karplus, M. (1995). Evaluation of comparative protein modeling by MODELLER. *Proteins*, *23*(3), 318–326. http://doi.org/10.1002/prot.340230306

Samudrala, R., & Moult, J. (1998). Determinants of side chain conformational preferences in protein structures. *Protein Engineering*, *11*(11), 991–997.

Sanchez, R., & Sali, A. (1997). Evaluation of comparative protein structure modeling by MODELLER-3. *Proteins*, *Suppl 1*, 50–58.

Saxena, A., Sangwan, R. S., & Mishra, S. (2013). Fundamentals of Homology Modeling Steps and Comparison among Important Bioinformatics Tools: An Overview. *Science International*. http://doi.org/10.5567/sciint.2013.237.252

Schwede, T. (2013). Protein Modeling: What Happened to the "Protein Structure Gap"? *Structure*, *21*(9), 1531–1540. http://doi.org/http://dx.doi.org/10.1016/j.str.2013.08.007

Schwede, T., Kopp, J., Guex, N., & Peitsch, M. C. (2003). SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* , *31*(13), 3381–3385. http://doi.org/10.1093/nar/gkg520

Shehata, M. A., Belcik Christensen, H., Isberg, V., Sejer Pedersen, D., Bender, A., Brauner-Osborne, H., & Gloriam, D. E. (2015). Identification of the first surrogate agonists for the G protein-coupled receptor GPR132. *RSC Advances*, *5*(60), 48551–48557. http://doi.org/10.1039/C5RA04804D

Silvestrov, P., Müller, T. A., Clark, K. N., Hausinger, R. P., & Cisneros, G. A. (2014). Homology modeling, molecular dynamics, and site-directed mutagenesis study of AlkB human homolog 1 (ALKBH1). *Journal of Molecular Graphics and Modelling*, *54*, 123–130. http://doi.org/http://dx.doi.org/10.1016/j.jmgm.2014.10.013

Simon, M. I., Strathmann, M. P., & Gautam, N. (1991). Diversity of G proteins in signal transduction.

Science (New York, N.Y.), *252*(5007), 802–8. http://doi.org/10.1126/science.1902986
Sircar, A., Kim, E. T., & Gray, J. J. (2009). RosettaAntibody: antibody variable region homology modeling server. *Nucleic Acids Research*, *37*(Web Server issue), W474-9. http://doi.org/10.1093/nar/gkp387
Soding, J., Biegert, A., & Lupas, A. N. (2005). The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research*, *33*(Web Server issue), W244-8. http://doi.org/10.1093/nar/gki408
Song, Y., DiMaio, F., Wang, R. Y.-R., Kim, D., Miles, C., Brunette, T., … Baker, D. (2013). High-resolution comparative modeling with RosettaCM. *Structure (London, England : 1993)*, *21*(10), 1735–1742. http://doi.org/10.1016/j.str.2013.08.005
Stites, W. E., Meeker, A. K., & Shortle, D. (1994). Evidence for strained interactions between side-chains and the polypeptide backbone. *Journal of Molecular Biology*, *235*(1), 27–32.
Sutcliffe, M. J., Haneef, I., Carney, D., & Blundell, T. L. (1987). Knowledge based modelling of homologous proteins, Part I: Three-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Engineering*, *1*(5), 377–384.
Szilagyi, A., & Zhang, Y. (2014). Template-based structure modeling of protein–protein interactions. *Current Opinion in Structural Biology*, *24*, 10–23. http://doi.org/http://dx.doi.org/10.1016/j.sbi.2013.11.005
Szklarz, G. D., & Halpert, J. R. (1997). Use of homology modeling in conjunction with site-directed mutagenesis for analysis of structure-function relationships of mammalian cytochromes P450. *Life Sciences*, *61*(26), 2507–2520.
Szklarz, G. D., Ornstein, R. L., & Halpert, J. R. (1994). Application of 3-dimensional homology modeling of cytochrome P450 2B1 for interpretation of site-directed mutagenesis results. *Journal of Biomolecular Structure & Dynamics*, *12*(1), 61–78.
Tappura, K. (2001). Influence of rotational energy barriers to the conformational search of protein loops in molecular dynamics and ranking the conformations. *Proteins*, *44*(3), 167–179.
Taylor, W. (2004). Protein Structure Folding and Prediction. In *Compact Handbook of Computational Biology* (pp. 223–240). CRC Press. http://doi.org/doi:10.1201/9780203021415.ch6
Teichmann, S. A., Chothia, C., Church, G. M., & Park, J. (2000). Fast assignment of protein structures to sequences using the intermediate sequence library PDB-ISL. *Bioinformatics (Oxford, England)*, *16*(2), 117–124.
Tian, W., & Skolnick, J. (2003). How well is enzyme function conserved as a function of pairwise sequence identity? *Journal of Molecular Biology*, *333*(4), 863–882.
Trujillo, K., Paoletta, S., Kiselev, E., & Jacobson, K. A. (2015). Molecular modeling of the human P2Y14 receptor: A template for structure-based design of selective agonist ligands. *Bioorganic & Medicinal Chemistry*, *23*(14), 4056–4064. http://doi.org/http://dx.doi.org/10.1016/j.bmc.2015.03.042
van Gelder, C. W., Leusen, F. J., Leunissen, J. A., & Noordik, J. H. (1994). A molecular dynamics approach for the generation of complete protein structures from limited coordinate data. *Proteins*, *18*(2), 174–185. http://doi.org/10.1002/prot.340180209
van Vlijmen, H. W., & Karplus, M. (1997). PDB-based protein loop prediction: parameters for selection and methods for optimization. *Journal of Molecular Biology*, *267*(4), 975–1001. http://doi.org/10.1006/jmbi.1996.0857
Venselaar, H., Joosten, R. P., Vroling, B., Baakman, C. A. B., Hekkelman, M. L., Krieger, E., & Vriend, G. (2010). Homology modelling and spectroscopy, a never-ending love story. *European Biophysics Journal* , *39*(4), 551–563. http://doi.org/10.1007/s00249-009-0531-0
Villoutreix, B. O., Kuenemann, M. a., Poyet, J. L., Bruzzoni-Giovanelli, H., Labbé, C., Lagorce, D., … Miteva, M. a. (2014). Drug-like protein-protein interaction modulators: Challenges and opportunities for drug discovery and chemical biology. *Molecular Informatics*, *33*(6–7), 414–437. http://doi.org/10.1002/minf.201400040
Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *Journal of Molecular Graphics*, *8*(1), 29,52-56.
Vyas, V., Ukawala, R., Chintha, C., & Ghate, M. (2012). Homology modeling a fast tool for drug discovery: Current perspectives. *Indian Journal of Pharmaceutical Sciences*, *74*(1), 1. http://doi.org/10.4103/0250-474X.102537
Waksman, G., & Sansom, C. (2005). Introduction : Proteomics and Protein – Protein Interactions : Biology , Chemistry , Bioinformatics , and Drug Design. In G. Waksman (Ed.), *Proteomics and Protein – Protein Interactions : Biology , Chemistry , Bioinformatics , and Drug Design* (pp. 1–18). New York:

Springer.
Wallner, B., & Elofsson, A. (2005). All are not equal: a benchmark of different homology modeling programs. *Protein Science : A Publication of the Protein Society*, *14*(5), 1315–27. http://doi.org/10.1110/ps.041253405
Wiederstein, M., & Sippl, M. J. (2007). ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Research*, *35*(Web Server issue), W407-10. http://doi.org/10.1093/nar/gkm290
Williamson, A. R. (2000). Creating a structural genomics consortium. *Nat Struct Mol Biol*, *7*(953). http://doi.org/10.1038/80726
Xiang, Z. (2006). Advances in homology protein structure modeling. *Current Protein & Peptide Science*, *7*(3), 217–227.
Xiang, Z., & Honig, B. (2001). Extending the accuracy limits of prediction for side-chain conformations. *Journal of Molecular Biology*, *311*(2), 421–430. http://doi.org/10.1006/jmbi.2001.4865
Xiang, Z., Soto, C. S., & Honig, B. (2002). Evaluating conformational free energies: The colony energy and its application to the problem of loop prediction. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(11), 7432–7437. http://doi.org/10.1073/pnas.102179699
Xu, D., & Zhang, Y. (2011). Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization. *Biophysical Journal*, *101*(10), 2525–2534. http://doi.org/10.1016/j.bpj.2011.10.024
Xu, J., & Berger, B. (2006). Fast and accurate algorithms for protein side-chain packing. *Journal of the ACM*, *53*(4), 533–557. http://doi.org/10.1145/1162349.1162350
Xu, Y., Wang, Y., Meng, X., Zhang, M., Jiang, M., Cui, M., & Tseng, G. (2013). Building KCNQ1 / KCNE1 Channel Models and Probing their Interactions by Molecular-Dynamics Simulations, *105*(December), 2461–2473. http://doi.org/10.1016/j.bpj.2013.09.058
Xu, Y., Wang, Y., Zhang, M., Jiang, M., Rosenhouse-Dantsker, A., Wassenaar, T., & Tseng, G.-N. (2015). Probing Binding Sites and Mechanisms of Action of an IKs Activator by Computations and Experiments. *Biophysical Journal*, *108*(1), 62–75. http://doi.org/10.1016/j.bpj.2014.10.059
Xun, S., Jiang, F., & Wu, Y. D. (2015). Significant refinement of protein structure models using a residue-specific force field. *Journal of Chemical Theory and Computation*, *11*(4), 1949–1956. http://doi.org/10.1021/acs.jctc.5b00029
Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., & Zhang, Y. (2014). The I-TASSER Suite: protein structure and function prediction. *Nature Methods*, *12*(1), 7–8. http://doi.org/10.1038/nmeth.3213
Yang, Y., Faraggi, E., Zhao, H., & Zhou, Y. (2011). Improving protein fold recognition and template-based modeling by employing probabilistic-based matching between predicted one-dimensional structural properties of query and corresponding native properties of templates. *Bioinformatics (Oxford, England)*, *27*(15), 2076–2082. http://doi.org/10.1093/bioinformatics/btr350
Yona, G., & Levitt, M. (2002). Within the twilight zone: a sensitive profile-profile comparison tool based on information theory. *Journal of Molecular Biology*, *315*(5), 1257–1275. http://doi.org/10.1006/jmbi.2001.5293
Yonath, A. (2011). X-ray crystallography at the heart of life science. *Current Opinion in Structural Biology*, *21*(5), 622–626. http://doi.org/http://dx.doi.org/10.1016/j.sbi.2011.07.005
Zhang, X., & Cheng, X. (n.d.). Structure of Protein, *34*, 978–981.
Zhang, Y. (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, *9*(1), 40. http://doi.org/10.1186/1471-2105-9-40
Zhu, J., Fan, H., Periole, X., Honig, B., & Mark, A. E. (2008). Refining Homology Models by Combining Replica-Exchange Molecular Dynamics and Statistical Potentials. *Proteins*, *72*(4), 1171–1188. JOUR. http://doi.org/10.1002/prot.22005
Zhu, K., Day, T., Warshaviak, D., Murrett, C., Friesner, R., & Pearlman, D. (2014). Antibody structure determination using a combination of homology modeling, energy-based refinement, and loop prediction. *Proteins*, *82*(8), 1646–55. http://doi.org/10.1002/prot.24551
Zorko, M. (2009). Protein Folding. In *Introduction to Peptides and Proteins* (pp. 101–122). CRC Press. http://doi.org/doi:10.1201/b15106-11

APPENDIX C: PROTEIN-PROTEIN DOCKING[10]

C.1. Introduction

Protein-protein docking has emerged as a powerful tool in obtaining data-driven and meaningful complexes between protein structures. Similarly, in this Thesis we have employed protein-protein docking simulations to build protein complexes for the KCNQ1/KCNE1 channel. As explained in Chapter 1, a functional KCNQ1 ion channel is essentially associated with its auxiliary beta subunit; KCNE1 protein. For this reason, an overview of the concept of protein-protein docking in general, with an emphasis on the different approaches and tools is provided in this chapter.

Protein-protein interactions play key roles in several biological processes. These processes involve many essential mechanisms ranging from signal transduction and cellular transport to gene expression and immune responses. All these processes are mediated by selective and potent protein-protein interactions (Waksman & Sansom, 2005). Furthermore, many diseases have been associated with either an over-activated or an under-regulated protein-protein interaction and the cure for these diseases has been focused on either inhibiting or stimulating these interactions, respectively. For example, the p53-MDM2 interaction is associated with a severe down regulation of the p53 pathway. An inhibitor for this interaction (e.g. nutlin3) can reactivate the p53 pathway, forcing cancer cells to undergo apoptosis (K Barakat, Gajewski, & Tuszynski, 2012; Khaled Barakat, Issack, Stepanova, & Tuszynski, 2011; Khaled Barakat, Mane, Friesen, & Tuszynski, 2010; Chène, 2003; Kojima et al., 2006). The more we know about such crucial interactions, the more we can build vital protein networks and apply this knowledge to identify treatments for many diseases. Moreover, characterizing these interactions at the atomic level can help in rationally designing new therapeutic agents that can either enhance or inhibit these interactions. Constructing a three dimensional

[10] A version of this Appendix has been published as Jalily Hasani H, Barakat K. Protein-Protein Docking: Are We There Yet?, Methods and Algorithms for Molecular Docking-Based Drug Design and Discovery. Dastmalchi S, Hamzeh-Mivehroud M, Sokouti B, editors. IGI Global; 173-195 p. 2016.

structure of such protein complexes is an essential step toward identifying their binding interface and recognizing any hot spots that can be targeted for their regulation (Elcock, Sept, & McCammon, 2001; Kann, 2007; Kortemme & Baker, 2004).

For the last few decades X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy and electron microscopy (EM) have been the main source to predict such structures. Despite their accuracy, efficiency and the huge amount of details they can provide, they are expensive and very labour and skill demanding. A simple comparison between protein structure and gene sequence databases would simply reveal the great discrepancy between the two. That is, although hundreds of thousands of gene expressions have been characterized, only less than thirty thousand protein structures have been determined so far and most of these structures are either redundant or describe only apo (unbound) proteins (Villoutreix et al., 2014). Moreover, protein complexes are more difficult to crystallize than the individual proteins, consequently, they are less represented in Protein Data Bank (PDB) (Berman et al., 2000) and constitute only a small fraction of the experimentally determined structures. This huge discrepancy and lack of structural details motivated many computational groups to fill this gap and suggest a new, rapid and cheap way to predict these interactions (Anwar-Mohamed et al., 2014; K. H. Barakat, Mane, & Tuszynski, 2011; Gógl et al., 2015; Nillegoda et al., 2015; Pedotti, Simonelli, Livoti, & Varani, 2011; D. W. Taylor et al., 2015). One solution they provided which is also the focus of this chapter is protein-protein docking.

Protein-protein docking simulations date back to the early 90s (Katchalski-Katzir et al., 1992), soon after the development of many protein-ligand docking techniques. Although the concept of docking in both cases is similar, there are still key differences that prevented the direct technology transfer between the two problems. For example, since proteins are continually dynamic and usually undergo huge structural changes, their modeling is not as easy as that of a small molecule and requires more precise predictions. Furthermore, proteins are large and have complex shapes making them harder to simulate than ligand-protein interactions (Ravikant & Elber, 2011; Sheinerman, Norel, & Honig, 2000). On the other hand, the concept of a binding site is usually inapplicable to protein-protein interactions and in many cases proteins interact through ‘hot spot’ areas on their surfaces. Identifying these binding locations is particularly difficult and requires highly

innovative methods. In this context, the search methods in protein-protein docking simulations are somehow different from those being used in small molecule docking (Keskin, Ma, & Nussinov, 2005). However, the reader should keep in mind that although protein-protein docking is in many ways different from protein-ligand docking, some of its reflections might be mentioned in this chapter, if applicable.

In this chapter, we will first outline the concept of docking in general, with an emphasis on the different approaches. We then briefly review the theoretical concepts, principles and specific features of the search strategies and scoring functions implemented in different programs, highlighting their strengths as well as their limitations. We will also discuss the different methods of incorporating protein flexibility within the docking procedure. After these descriptive review sections, we will guide the reader through a number of case studies and recent successes that illustrate some of the significant applications of protein-protein docking algorithms. We will also discuss the existing challenges and limitations in current algorithms and highlight some potential future directions and possible improvements in this field. We hope that this brief, yet comprehensive review chapter provides researchers interested in this area with current state-of-the-art technologies in this field.

C.2. The "docking problem"

The so-called protein-protein "docking problem" refers to the task of docking two individual protein structures and predicting their biologically relevant complex. It can be classified into two different approaches; "bound docking" where two proteins from a co-crystallized complex are separated and then reassembled using a computational docking algorithm or "unbound docking" (Wiehe, Peterson, Pierce, Mintseris, & Weng, 2008) where individually solved structures of two proteins are docked together to predict their most probable mode of binding. The bound docking approach has no investigational significance, but it is frequently employed to test and validate the different docking algorithms while developing a new method. Unbound docking, on the other hand, is the approach employed by the current docking programs and is a valuable tool that can predict complexes from individual proteins for producing new knowledge at the atomic level. Unbound docking is far more complicated than bound docking. This is because the

two proteins, which come together to form a complex, may undergo some conformational changes upon binding, which needs to be predicted correctly (Ravikant & Elber, 2011; I. A. Vakser, 2014). In this section, we will focus mainly on the unbound docking approach involving two different methods, namely rigid-body docking and flexible docking. Initially, we will have a closer look at these two approaches and then will focus more on rigid-body docking and its different aspects.

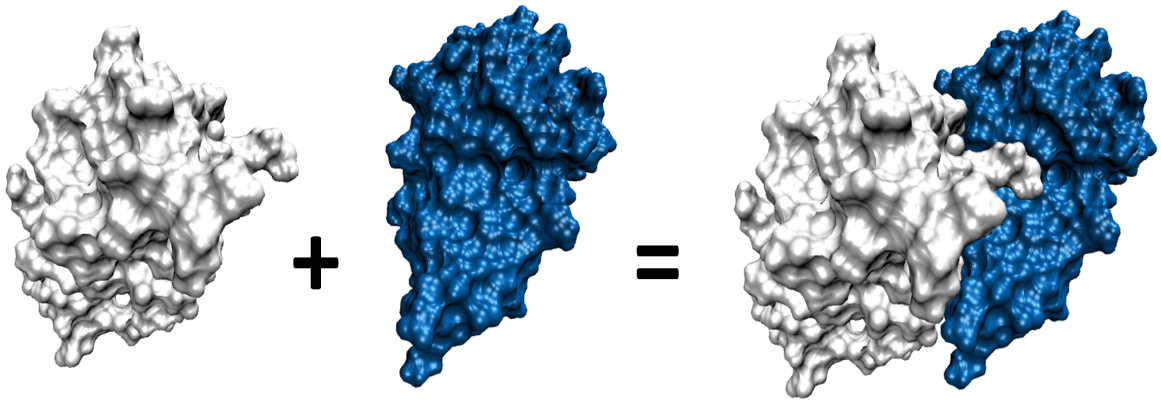


Figure C.1. A depiction of two proteins shown individually and when docked together in complex.

Rigid-body docking treats the two proteins as rigid entities. In this case, the side chains and backbone of the two interacting proteins are kept fixed with no bond angles or distances allowed to change upon docking. Although this method offers a rapid way to predict the formed complex, it completely ignores any possible conformational rearrangements that may occur upon complex formation. This sampling approach adopts a simplistic, yet structurally meaningful, technique. It projects the protein coordinates onto a low-resolution three-dimensional grid and then calculates the overlap between the two grids of the interacting proteins (Ehrlich & Wade, 2001). Flexible docking, on the other hand, takes into account such conformational changes by allowing minimal movements of the side chains or the backbone of the proteins at high resolution. Rigid-body docking includes six degrees of freedom; three translations and three rotational in the Cartesian coordinates, whereas flexible docking is usually preceded by transforming the Cartesian coordinates into internal coordinates, which involves more variables compared to rigid docking (Guo, Li, Ma, & Wang, 2013; I. A. Vakser, 2014).

It is noteworthy to mention that in many cases, no significant conformational

changes take place upon the binding of the two proteins and a "lock-and-key" approach would be sufficient to address their interaction (Betts & Sternberg, 1999). This usually occurs when the two protein structures are not fluctuating too much, particularly at their binding interface. In this case, employing a rigid docking protocol to study these interactions would be adequate and can predict, with an acceptable degree of accuracy, their mode of binding. However, to predict higher resolution structural details at the interface some degree of flexibility is required, and the more protein dynamics incorporated during the docking simulation, the more structural details can be revealed. In this case, flexible docking comes into play and more details on that are described in the next sections.

It is important to emphasize the significance of incorporating experimental data within the docking procedure. This type of data not only helps in validating the docking outcomes, but also can help in identifying the potential binding interface of the two proteins, reducing the search space to a manageable region (Grinter & Zou, 2014; Wodak & Méndez, 2004). In this context, a number of docking algorithms can restrict the search space to a specific region and therefore feeding these algorithms with a minimal knowledge of the binding site residues called, 'hot spots' can significantly improve the search process (Grosdidier & Fernández-Recio, 2012; Szilagyi, Grimm, Arakaki, & Skolnick, 2005).

C.3. Docking Steps: Sampling & Scoring

Within the scope of the two classes (rigid-body docking & flexible docking) discussed above, the process of docking is implemented through two main interconnected steps: *sampling* which will search for all possible binding modes between the two proteins and *scoring* for the differentiation between correct, near-correct or incorrect poses (Figure C.2). In the following two sections, we will describe these two processes in details, giving examples of different software packages employing the different methods. We will then describe how protein flexibility is incorporated during docking and post-docking approaches. Finally, a brief evaluation of the different docking techniques will follow.

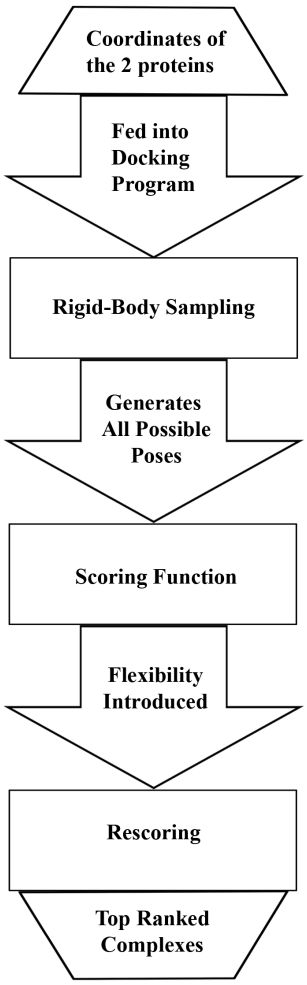


Figure C.2. General scheme of docking and the different stages involved. The docking process usually starts with rigid-body docking, covering only the six-dimensional space of the two proteins in 3D translational and 3D rotational. This is followed by ranking the generated complexes using a simple scoring function (e.g. shape complementarity). A more complex energy-based scoring method is next employed. Flexibility is then introduced to allow side chains or backbone dynamics. Incorporating protein flexibility usually leads to a more realistic and detailed picture of interaction. Rescoring stage is sometimes implemented for a more sophisticated ranking of poses.

C.3.1. Sampling Techniques

We will begin by briefly introducing the sampling stage. The reader is referred to the references included here and the references within for a more details on this area of research. The sampling stage of docking involves generating a large number of alternative poses for the proteins by sampling all the possible binding modes. The correct

prediction of the complex is more of a near-native approximation than an accurate native (co-crystalized) complex of the two proteins, which is computationally impossible to generate (G. R. Smith & Sternberg, n.d.). Standard search methods are computationally expensive and lead to a huge number of solutions; most of them are physiologically irrelevant. It is for this reason more sophisticated sampling approaches and strategies are usually incorporated to correctly predict realistic solutions at this stage. This includes exhaustive global search methods as adopted by fast Fourier transform (FFT) algorithms, local shape feature matching used in geometrical hashing and genetic algorithms, randomized search employed in Monte Carlo (MC) search or Boolean operations and effective energy gradients to determine new atomic positions (Ehrlich & Wade, 2001)(S. Huang, 2014). Table C.1 lists the most commonly used software packages used for protein-protein docking with their respective sampling techniques. Although the focus of this chapter is particularly on protein-protein docking, there are a number of small-molecule docking programs optimized to become applicable to protein-protein docking and therefore, they will be discussed in various sections to illustrate the different aspects of docking.

Some recently developed and/or optimized protein-protein docking programs are DockTrina (docking triangular protein trimmers) (Popov, Ritchie, & Grudinin, 2014), ATTRACT (Vries & Zacharias, 2013), MEGA-DOCK (Matsuzaki et al., 2013), F(2)Dock 2.0 and GB-Rerank (R. Chowdhury et al., 2013), SwarmDock (Moal, Torchala, Bates, & Fernández-Recio, 2013). In addition to the programs listed in Table C.1, there are numerous docking programs, which facilitate the use of docking algorithms by availing online servers. Some of them include: ClusPro (Comeau, Gatchell, Vajda, & Camacho, 2004), GRAMM-X (Tovchigrechko & Vakser, 2006), ZDOCK SERVER (B. G. Pierce et al., 2014), 3D-GARDEN (V. I. Lesk & Sternberg, 2008), PatchDock & SymmDock (Schneidman-Duhovny, Inbar, Nussinov, & Wolfson, 2005), HADDOCK (Dominguez, Boelens, & Bonvin, 2003), RosettaDock (Gray et al., 2003), FIBERDOCK (Mashiach, Nussinov, & Wolfson, 2010) , FIREDOCK (Mashiach, Schneidman-Duhovny, Andrusier, Nussinov, & Wolfson, 2008), pyDockWEB (Jiménez-García, Pons, & Fernández-Recio, 2013), pyDockSAXS (Jimenez-Garcia, Pons, Svergun, Bernado, & Fernandez-Recio, 2015) and ATTRACT web interface (S. J. De Vries, Schindler,

Chauvot de Beauchêne, & Zacharias, 2015).

Table C.1. The most commonly used protein-protein docking programs with their method of sampling.

Name of Software	Method of Sampling	References
ZDOCK, FTDOCK, GRAMM, F2DOCK, DOT, Cell-Dock	FFT-correlation	(Chen et al., 2003; Gabb et al.,1997; Vakser, 1995; Bajaj et al., 2011; Mandell et al., 2001; Pons et al., 2012)
HEX, FRODOCK	FFT- Spherical Harmonics	(Ritchie & Kemp, 2000; Garzon et al., 2009)
PatchDock, LZerD	Geometric Hashing	(Duhovny et al., 2002; Venkatraman et al., 2009)
BIGGER, SKE-DOCK, SoftDock	Direct Search in Cartesian Space	Palma et al., 2000; Terashi et al., 2007; Jiang & Kim 1991)
DOCK	Distance Geometry Algorithm	(Kuntz et al., 1982)
DARWIN	Genetic Algorithm	(Taylor & Burnett, 2000)
RosettaDock, ATTRACT, ICM	Monte Carlo Search	(Gray et al., 2003; Zacharias, 2003; Fernandez-Recio et al., 2002)
HADDOCK	Based on both Biochemical/Biophysical Information and MC Search	(Dominguez et al., 2003)

Fast Fourier Transform (FFT) correlation approach is regarded as one of the most popular and widely used techniques of rigid-body search, developed in 1992 by Katchalski-Katzir and coworkers (Katchalski-Katzir et al., 1992). This technique allows an accelerated exhaustive global search of the binding orientations over six degrees of freedom in 3D translational plus 3D rotational space. The popularity and wide usage of this technique is attributed to its ability to rapidly perform an exhausted sampling of all the possible binding modes of the interacting proteins. In this method, the two proteins (receptor protein and ligand protein) are discretized onto two different 3-dimensional grids each with the size N * N * N (where N is the number of grid points), and each point on the grid with coordinates *l,m,n* is assigned a value based on its geometric

characteristics (a $_{l,m,n}$). This includes a value of one for grid points on the surface of the protein, a value of zero for points outside of the protein surface and a value of ρ for points located within the protein core structure (where ρ usually has a negative value). The complementarity of shape function C (p,q,r) between the two proteins is then evaluated from the Equation 1, shown below:

$$C(p,q,r) = \sum_{l=1}^{N} \sum_{m=1}^{N} \sum_{n=1}^{N} R(l,m,n) \times L(l+p, m+q, n+r)$$

The translational space is then sampled by calculating a correlation between the two discrete grids. FFT correlation has been employed in a number of docking programs: FTDock (Gabb, Jackson, & Sternberg, 1997), GRAMM (Vakser, 1995), DOT (Mandell et al., 2001), ZDOCK (R. Chen & Weng, 2002), F2DOCK (Bajaj, Chowdhury, & Siddahanavalli, 2011) and Cell-Dock (Pons et al., 2012) . The basic approach is the same for all of them and the difference lies in the way the proteins are projected onto the grids.

A successful modification of the FFT algorithm is the spherical Fourier transform-based search, initially introduced in the program HEX (Ritchie & Kemp, 2000), which employs spherical harmonics to rapidly search the complete conformational space of the proteins. Another exhaustive sampling technique employs a direct search in Cartesian space to identify the correct binding mode between two proteins using 3D Cartesian grid space. This technique has been adopted in many docking packages including SOFTDOCK (F. Jiang & Kim, 1991), BIGGER (Palma et al., 2000), SKE-DOCK (Terashi et al.,2007). The program BIGGER (Palma et al., 2000) for example, employs a surface implicit method based on Boolean operations, in which the molecular shape of the two proteins is mapped onto a 3D grid in Cartesian space. Each point on the grid is assigned a value of '1' if it is on the protein core and a '0' if it is outside. Geometric fitting between the two proteins is then performed by shape matching directly in the Cartesian space. The advantage of this technique over FFT-based methods is the ability to have more control over the search process because of the direct operation in Cartesian space, which allows inclusion of protein flexibility and biological information (Huang, 2014).

Geometric hashing techniques are another type of global protein-protein docking search using a shape-explicit algorithm. The technique is based on finding a set of shape

descriptors on the molecular surface of the proteins (one of them called the receptor and the other the ligand). Following that, a set of local coordinate frames is defined, for which the positions of nearby critical points are used as indices hashed into a table. The hash table stores the current coordinate frame of each protein. After superimposing the local coordinate frames of the ligand and the receptor, the positions of the critical points of the ligand in proximity of the receptor are then used to scan the hash table for the corresponding ligand reference frames (G. R. Smith & Sternberg, n.d.). Several programs employ different shape descriptors. For example patch descriptors are used in PatchDock (Duhovny, Nussinov, & Wolfson, 2002) and 3D Zernike descriptors are adopted by LZerD (Venkatraman, Yang, Sael, & Kihara, 2009). High shape similarity is revealed by the frequency of correspondences to a specific frame. The method has been improved by introducing surface softness as a tolerance value of about 1.5 A in the coordinates.

Genetic algorithm is the basis for a number of other successful docking algorithms. This generally involves a real-space search method, which is based on a local shape feature matching. The DARWIN (J. S. Taylor & Burnett, 2000), for example, includes a solvent accessible area (SAA) for the proteins with normals, surface curvature and associated hydrogen bonding character. The complementarity between the curvatures and normals are scanned for the two proteins to find the most probable binding mode. The 'chromosome' of the genetic algorithm (GA) consists of the relative positions and orientations of the molecules (6D space).

Randomized search is another broad category of sampling techniques. Monte Carlo (MC) search is one of the approaches in which sampling of all possible conformations with a technique like FFT correlation is replaced with a random approach. As such, random translational and rotational orientations for the two proteins are performed and the final positions are selected based on a Metropolis criterion. This is then followed by a multistage optimization of the side-chains and rigid-body orientations based on MC simulations (Gray et al., 2003). The advantages of MC methods over FFT methods includes the allowance of more physical decoy distribution, arbitrary energy functions and structural flexibility to some additional extent (Lorenzen & Zhang, 2007). RosettaDock (Gray et al., 2003) is an efficient protein-protein docking software that is based on the MC method. Some other software based on randomized search are

HADDOCK (Dominguez et al., 2003), ICM-DISCO (Fernandez-Recio, Totrov, & Abagyan, 2002) and ATTRACT (Zacharias, 2003).

More recent practical successes in sampling techniques, include the recent work of Zhang et al. (Z. Zhang, Schindler, Lange, & Zacharias, 2015) who identified a variation of MC search for the RosettaDock package to improve the sampling stage towards a more efficient identification of the near native docking geometries. This variation involves combining the well-tempered ensemble method with a 2-dimensional temperature and Hamiltonian replica exchange scheme (WTE-REMC). A comparative study of the results from different docking protocols of Rosetta, using the same set of targets has shown the enhanced performance of WTE-REMC in exploring the phase space.

HADDOCK (High Ambiguity Driven Protein-Protein Docking) software (Dominguez et al., 2003) is implemented with a different scheme for the randomized search process. It employs a data-driven approach and supports various types of biochemical and/or biophysical data sets such as chemical shift perturbation data from NMR titration experiments or mutagenesis data. HADDOCK is able to handle almost full flexibility (side-chain and backbone) and incorporate explicit water during the modeling process. In addition, the software provides a user-friendly web server (S. J. De Vries, Van Dijk, & Bonvin, 2010), which is accessible through four interfaces: Easy, Expert, Guru and parameter file upload interfaces, for different levels of expertise and type of application allowing more control and modification on the docking process. Several CAPRI experiments (Janin, 2013), have confirmed the profound ability of HADDOCK in precise predictions of protein-protein complexes based on experimental information. In addition it is stated that HADDOCK has generated about 70 protein-protein complexes in PDB (Grosdidier & Fernández-Recio, 2012).

C.3.2. Scoring Functions

The large numbers of possible conformations (poses) generated from the sampling stage are ranked using a scoring function that involves the assessment of the overall favorability of a protein-protein complex and/or a score between the two proteins in the bound form. The final docking results comprise of the top ranked complexes, which are in order of their binding score (S. Huang, 2014). These different scoring methods can be

categorized into different groups based on the basis they are derived from; for example, the *force field-based* scoring functions employ non-bonded interaction terms as the scoring method. *Empirical* scoring functions, on the other hand make use of regression methods to differentiate between the different poses. *Knowledge-based* scoring functions use statistical atom pair potentials from the structural databases as the scoring method. Yet, another category is the *heuristic* scheme of scoring, which makes use of chemical scores or geometrical scores such as contact/surface scores, or shape complementarity scores (Ehrlich & Wade, 2001; Muegge & Rarey, 2001).

Docking programs use different scoring functions that largely vary from one algorithm to the other. Usually, a docking program makes use of a composition of different scoring functions to increase the efficacy and accuracy of the ranking process. For instance, although shape complementarity between the interacting proteins is regarded as a simple, inexpensive and powerful scoring function, it does not necessarily distinguish near-native from non-native complexes and is thus combined with other functions such as electrostatic field, hydrophobic complementarity, desolvation free energy, etc. (Szilagyi et al., 2005). An example of implementing such composite scoring functions is the ZDOCK program (R. Chen & Weng, 2002). FTDOCK (Gabb et al., 1997) and DOT (Mandell et al., 2001), which combine electrostatic and geometric scores but in different ways. FTDOCK uses geometric matching in the initial stage and then filters the docked poses based on their electrostatics scores in later stages. In the DOT algorithm, two types of scoring methods are used in conjunction. The GRAMM package performs the scoring through a geometric surface matching approach by involving the hydrophobic contacts between the two proteins. PyDock (Cheng, Blundell, & Fernandez-Recio, 2007) uses an energy function based on ICM potentials, composed of van der Waals, Coloumbic electrostatics and Atomic Solvation Parameter (ASP)-based desolvation energy. The scoring function in ICM-DISCO (Fernandez-Recio et al., 2002) takes into consideration the intermolecular grid-based energy terms and the Accessible Surface Area (ASA)-based desolvation energy. RosettaDock program also employs a similar scoring scheme.

C.4. Incorporating Protein Flexibility

Currently, all modern docking techniques use different approximations to include protein flexibility within the docking process. This may include rotating the protein side chains and possibly involve simple dynamics of the backbone. Flexibility in general is treated in two different ways: “implicit flexibility” and “explicit flexibility”. Rigid-body approaches incorporate protein flexibility in an implicit way, by either using soft docking (surface softening) or through more advanced techniques such as ensemble-based docking. Surface softening techniques include low resolution docking methods, simplified models for the side chains and surface layer thickening (Szilagyi et al., 2005). Rigid-body docking methods show successful results in many cases (Viricel, Ahmed, & Barakat, 2015; Y. Xu et al., 2013a). Nevertheless, an accurate and biologically relevant prediction of these protein-protein interactions which involve considerable conformational changes of backbone and/or side chains upon complex formation is merely not possible through rigid-body docking (Grosdidier & Fernández-Recio, 2012; C. H. U. Wang & Schueler-furman, 2005).

Explicit treatment of the protein flexibility is only possible in docking algorithms that involve the atomic coordinates of the different atoms, rather than those that implicitly represent the protein structure on a grid. Each approach treats flexibility to a specific extent in a different way, such as accounting for the dynamic changes of the protein domains, amino acid side-chains and less frequently the backbone of the protein. The different approaches to include protein flexibility are based on energy minimization, atomic-level simulations or optimization of the shape-based characteristics and/or complementarity. Others include considerations related to electrostatic interactions and hydrogen bonding which have been known for a long time. All the mentioned approaches are time-consuming and take several hours to days for an effective treatment in a docking process and this represents the main reason for which flexibility remains a major challenge in the field of protein-protein docking (Ehrlich & Wade, 2001). For example, FiberDock (Mashiach et al., 2010) includes protein flexibility through several energetic calculations and ranking through normal mode analysis (NMA) along with side-chain optimizations (using a rotamer library) and restrained energy minimization with limited degrees of freedom. Another randomized search algorithm, RosettaDock (Gray et al.,

2003) that is based on Monte Carlo principles, employs side-chain repacking, followed by optimizing the docked poses using a MC search with rigid-body displacements. Hinge-bending treatment is used in FlexDock (Schneidman-Duhovny, Ruth, & Wolfson, 2007) and FlexDoBi (Guo et al., 2013) programs in which the hinge regions (regions of the protein causing large-scale motions) are specified on the proteins prior to the docking simulation and the docking procedure is then performed, taking these specifications into account. HADDOCK (Dominguez et al., 2003) is regarded as an experimental data-driven program which performs a semi-flexible simulated annealing. Biochemical & biophysical interaction data from NMR experiments, mutagenesis investigations or bioinformatics predictions are used to introduce ambiguous interaction restraints (AIRs). These are then used to control the docking process with allowance for both backbone and side-chain flexibility on the interface. Average interaction energies are then used as a basis for the clustering and ranking of top predictions.

C.5. Post Docking Refinement

Docking procedures can be followed by a second ranking/scoring stage to refine the docking results and remove any steric clashes that may emerge from the docking algorithm. This is generally required due to the lack of precision in docking scoring functions. There are many different methods that have been used in the post-docking refinement stage. This includes a biased probability side-chain optimization approach as implemented in the ICM program (Abagyan et al., 1994) or side-chain minimization as employed by Multidock (Multiple copy side-chain refinement Dock) (R. M. Jackson, Gabb, & Sternberg, 1998) algorithm. The RosettaDock program (Gray et al., 2003) uses another effective method, which involves the correction of main-chain displacements. In addition, some post-docking algorithms are specialized in refining the results to a higher extent. These include ZRANK (B. Pierce & Weng, 2007), EMPIRE (Liang, Liu, Zhang, & Zhou, 2007), DARS (Chuang, Kozakov, Brenke, Comeau, & Vajda, 2008), DECK (Shiyong Liu & Vakser, 2011), RDOCK (L. Li, Chen, & Weng, 2003), pyDock (Cheng et al., 2007), Eigen-Hex (Venkatraman & Ritchie, 2012), RPScore (Moont, Gabb, & Sternberg, 1999), Multidock (R. M. Jackson et al., 1998) among others. DOCKSCORE (T.-T. Huang et al., 2015) and FiltRest3D (Gajda, Tuszynska, Kaczor, Bakulina, &

Bujnicki, 2010) are two of the online webservers which allow ranking of poses from protein-protein docking.

C.6. Docking Evaluation

When a new docking technique is developed, and is claimed to be good at predicting the structure of a protein-protein complex, it needs to be evaluated to provide reliable data and information to scientists and research groups to obtain the highest quality results from their investigations. The Critical Assessment of Predicted Interactions (CAPRI) experiment (Janin et al., 2003) is created to compare and evaluate the performance of various docking algorithms developed by different groups. CAPRI is modeled after Critical Assessment of Structural Prediction (CASP), which started in 1994 for comparing the performance of protein-folding algorithms (Moult et al., 2014). The way CAPRI aims at evaluating the docking algorithms is through using a set of standard benchmark proteins (unbound pairs), with the correct match known (co-crystallized complex) which is kept confidential. These are assigned to the computational research groups and the results obtained are further assessed and compared. CAPRI has attracted significant attention of the scientific community and the results have led to great improvements to the docking techniques (I. A. Vakser, 2014). Furthermore, systematic studies are being constantly conducted to provide up to date insights into different aspects of docking programs (Y. C. Chen, 2015; Moal et al., 2013).

The variability in the conformational sampling methods and scoring employed by the different docking algorithms, makes selecting a particular docking software a hard task. Each docking tool discussed so far has its own strengths as well as shortcomings. A fair way to compare these different tools is to test them against the same set of data using exactly the same set of parameters with identical amount of elapsed time (Cole, Murray, Nissink, Taylor, & Taylor, 2005). This has been the case employed by CAPRI evaluations, which provide reliable and periodical comparison among the different groups developing PPI docking software packages and servers.

Table C.2 summarizes a comparative overview of protein-protein docking software and/ servers with an emphasis on the most recent CAPRI experimental outcomes. For the complete references for each docking tool, please refer to Table C.1.

The latest CAPRI experiment (5th Evaluation) which was carried out in 2013 (Bonvin, 2013), reported the ClusPro server as the best in the automatic protein docking server category. In particular, the server's performance was equivalent to the best human predictor group. HADDOCK (De Vries et al., 2010), SwarmDock (Moal et al., 2013) and PIE-DOCK (Ravikant & Elber, 2010) servers were ranked next. In the human predictor category, HADDOCK (Dominguez et al., 2003) was ranked first followed by SwarmDock (Venkatraman & Ritchie, 2012), Vakser group, Vajda/Kozakov group (Camacho & Vajda, 2002; Chuang et al., 2008) and ICM in the 2nd to 5th position (Kozakov et al., 2013; Lensink & Wodak, 2013).

However, it is worth noting that a single docking algorithm cannot provide the ultimate answer for a particular problem and it would be better to combine at least two algorithms for the same research question. This will ensure a high quality of the final results, which should also make use of additional computational tools along with the docking methods to fulfill their gaps and shortcomings. Likewise, other numerical measures such as root-mean-square deviation (RMSD), careful analysis of interaction-based measures and visual examinations may provide additional credibility to the results (Hernández-santoyo, Tenorio-barajas, Altuzar, Vivanco-cid, & Mendoza-barrera, 2013).

Table C.2. Comparative overview of different docking software and tools with respect to 2013 CAPRI experiment.

Docking Software and /Server	Strength	Weakness
HADDOCK	• Semi-flexible simulated annealing permitting almost full protein flexibility • Allows for ambiguous interaction restraints • Efficient scoring performance • CAPRI 2013: software ranked 1st, server ranked 2nd	• Reduced accuracy without structural information
ClusPro	• Efficient prediction • CAPRI 2013: Server ranked 1st	• Rigid-body docking; no/less inclusion of flexibility
RosettaDock	• Flexible protein docking • Post-docking correction of main-chain displacements.	• Scoring functions require improvement
ZDOCK	• Efficient scoring through a combination of functions. • Allows for specifying residual restraints.	• Rigid-body docking; no/less inclusion of flexibility • Inaccurate for docking of homology models
SwarmDock	• Efficient prediction using normal modes and flexibility during docking • CAPRI 2013: Server ranked 3rd	• Inconsistency in scoring functions.
PIE-DOCK (software and server	• Efficient prediction • CAPRI 2013: server ranked 4th.	• Poor scoring as indicated by CAPRI 2013
PatchDock	• Scalable FFT-based algorithm • Efficient computer run time	• Poor scoring as indicated by CAPRI 2013 • Takes minimal flexibility into account • Poor refinement

Furthermore, we carried out a quick citations' survey for the top ranked docking algorithms (using Google Scholar; http://scholar.google.com), which resulted in the pie chart shown in Figure C.3. This chart includes the citations of only the software packages. All docking servers were excluded to avoid discrepancies in the presented data. Accordingly, HADDOCK software was considered to be one of the most popular protein-protein docking tools, which confirms its proven accuracy and credibility amongst scientific groups.

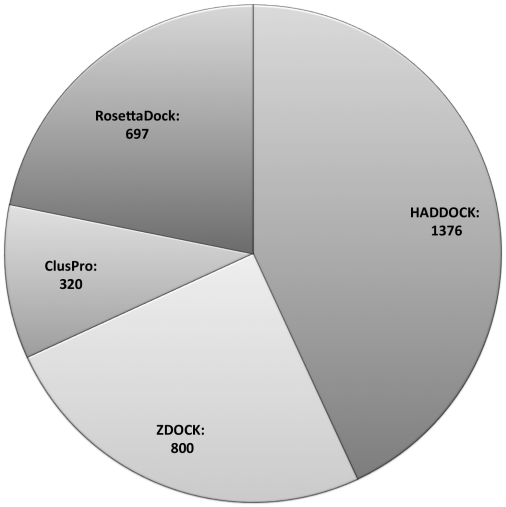


Figure C.3. Pie chart presenting the relative number of citations to each docking tool. For the sake of consistency, only the docking software are mentioned and servers are excluded.

C.7. Case Studies

In our recent work (Ahmed & Barakat, 2015; Viricel et al., 2015), we have applied ensemble-based docking technique to understand the interaction between Programmed Death- 1 (PD-1) and PD-Ligand-1 proteins using the ZDOCK software (R. Chen, Li, & Weng, 2003) and the FFT search. These proteins belong to the immune system and play important functions in the T-cell inhibition pathway, leading to immune tolerance and inactivity against cancer and infectious cells. The novel results produced in this study predicted how these proteins interact in human. This interesting modeling approach, which combined molecular dynamics simulations, docking and free energy calculations, provided a good example for how protein-protein docking techniques can help understand complex protein interactions at the atomic level. In another interesting study (Zhang et

al., 2005), the ZDOCK server (Pierce et al., 2014) was used for finding the possible protein-protein complex formed between the severe acute respiratory syndrome coronavirus (SARS-CoV) S1 protein and its receptor, human angiotensin-converting enzyme 2 (ACE2). This study illustrates successful identification of the interfacial key residues involved in the process of complex formation and has provided valuable information for further study of the mentioned system.

Protein-protein docking tools are frequently employed for investigating the structural details of various types of ion channels. These multimeric biological complexes are difficult to crystallize and the available literature indicates wide use of computational techniques for their study. A good example of such successful attempts is the work by Smith et al. (2007), who employed Rosetta-based docking for modeling the transmembrane domain of KCNE1 cardiac ion channel. The docking approach in this study has produced one of the most accurate starting models and has provided significant information on the structural basis for this ion channel.

The HADDOCK software (Dominguez et al., 2003) has been quite successful in resolving a large number of accurate models for protein-protein complexes. One good example is the study of the complex formed between plectasin, a member of the innate immune system and the bacterial cell-wall precursor Lipid II (Schneider et al., 2010). The bacterial cell wall biosynthesis is targeted by plectasin, through complex formation with Lipid II. The study has clearly identified the residues involved at the binding site between the two proteins by using the HADDOCK protein-protein docking tool, providing valuable information for the design of novel antibiotics.

C.8. Challenges & Future Directions

Despite the outstanding success of many protein-protein docking algorithms and their application throughout the last decade, there still a number of challenges that needs to be addressed within the next few years to advance the field for more realistic predictions. The initial sampling stage of most of docking algorithms is generally rapid and efficient, however there are certain shortcomings of protein-protein docking approaches, which fall into different categories mainly, related to scoring and flexibility issues.

The scoring functions are regarded as one of the major challenges of docking

algorithms and frequently fail to accurately rank the resultant complexes. Even with the considerable improvements so far, detection of correct conformations is still considered to be beyond the scope of current docking algorithms. Using post-docking filtering techniques usually conceals this drawback, although it complicates the overall docking procedure. Accordingly, it is crucial to reach a point where the docking algorithms themselves gain the ability of completely refining the complexes correctly. The scoring functions also need to focus on solvation, entropy and the effects of protonation states of the charged residues. Development of new implicit and explicit models for further investigation of the role of water in the proteins interfaces during protein-protein docking, requires more effort. (Moal et al., 2013)

The treatment of proteins' structural flexibility remains an active area of research and is still lagging behind. The problem is even more puzzling when dealing with the proteins' backbone flexibility and large structural changes as seen in antigen-antibody complexes (Chen & Tou, 2013).

Furthermore, advancements in the structural genomics studies are leading to more number of therapeutically significant 3D structures being determined and thus, opening doors for docking studies to explore such interactions at the atomic level. However, it is predicted that there would be more number of 3D structures obtained by homology modeling rather than experimental techniques. This would produce more controversies and errors in the resulting complexes, due to uncertainties and limitations of homology modeling in correct prediction of 3D structures (Wass, Fuentes, Pons, Pazos, & Valencia, 2011).

The success of the current docking algorithms is largely dependent on the availability of experimental data (such as binding sites, interaction partners and conformational changes), without which there are high chances of obtaining incorrect conformations. This reliance on experimental data would always remain a challenge for many docking studies (Smith & Sternberg, 2002).

Improvements to the docking techniques in general, may include the optimization of existing algorithms or the development of new ones to overcome these limitations. With the aid of community efforts such as CAPRI, docking algorithms are being successfully evaluated to identify the potential limitations and the benchmark testing has

led to overcoming a large number of them. The remaining challenges such as flexibility incorporation are still under investigation and with the accelerated pace of research in this area, the problem of flexibility will be soon tackled. Other tricky challenges like the prediction of multi-protein complex structures could be addressed and investigated more in the future (Grimm et al., 2005; Wodak & Méndez, 2004).

C.9. Conclusion

Understanding protein-protein interactions is an essential step to comprehend many important biological processes and to identify new treatments for many diseases. The huge discrepancy between the amount of genomic data deposited in the literature and the experimentally determined protein structures place a huge pressure on the computational groups to predict the structure of many proteins and understand how they interact. Amongst the current computational tools in use is protein-protein docking. This technique involves several algorithms that have been developed to analyze the protein-protein interactions at the atomic-level. They predict the mode of binding between two interacting proteins and involve two interconnected stages, namely sampling of protein conformations and scoring the docked poses. More advanced algorithms include the proper incorporation of protein flexibility and sophisticated scoring methods, referred to as post-docking refinement.

This chapter briefly introduced the concepts behind protein-protein docking, with an emphasis on the different approaches. It also reviewed the theoretical concepts, principles and specific features of the search strategies and scoring functions implemented in different programs, highlighting their strengths as well as their limitations. The main sampling techniques reviewed here include FFT correlation, direct search in Cartesian space, randomized search and MC algorithm. Scoring of the protein-protein docking solutions span different methods including force field-based methods, empirical scoring, knowledge-based and heuristic scoring functions.

Despite the great efforts of many research groups in this exciting area of research, there are still many challenges that need to be addressed to improve the method. This includes mainly developing more accurate scoring methods and the proper incorporation of protein flexibility. Although many algorithms are on the edge of overcoming these

problems, more work need to be done to fully solve these issues. With the accelerated pace of improvements in the field of computational approaches and their successful applications in understanding many biological systems, we believe that the field of protein-protein docking will evolve rapidly towards a more accurate method to predict protein-protein interactions. We hope this review provides a brief, yet comprehensive, overview to researchers working in this area.

C.10. References

Abagyan, R., Totrov, M., & Kuznetsov, D. (1994). ICM - A new method for protein modeling and design: Applications to docking and structure prediction from the distorted native conformation. *Journal of Computational Chemistry*, *15*(5), 488–506. http://doi.org/10.1002/jcc.540150503

Ahmed, M., & Barakat, K. (2015). Baby steps toward modelling the full human programmed Death-1 (PD-1) pathway. *Receptors & Clinical Investigation*, *1*, 1–5. http://doi.org/10.14800/rci.825

Anwar-Mohamed, A., Barakat, K. H., Bhat, R., Noskov, S. Y., Tyrrell, D. L., Tuszynski, J. a., & Houghton, M. (2014). A human ether-á-go-go-related (hERG) ion channel atomistic model generated by long supercomputer molecular dynamics simulations and its use in predicting drug cardiotoxicity. *Toxicology Letters*, *230*(3), 382–392. http://doi.org/10.1016/j.toxlet.2014.08.007

Bajaj, C. L., Chowdhury, R., & Siddahanavalli, V. (2011). F2dock: Fast fourier protein-protein docking. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *8*(1), 45–58. http://doi.org/10.1109/TCBB.2009.57

Barakat, K., Gajewski, M., & Tuszynski, J. (2012). DNA Repair Inhibitors: Our Last Disposal to Improve Cancer Therapy. *Current Topics in Medicinal Chemistry*.

Barakat, K. H., Mane, J. Y., & Tuszynski, J. A. (2011). Virtual Screening : An Overview on Methods and Applications. In *L. Liu, D. Wei, Y. Li, & H. Lei (Eds.) Handbook of Research on Computational and Systems Biology: Interdisciplinary Applications* (pp. 28–60). Hershey, PA: Medical Information Science Reference. http://doi.org/10.4018/978-1-60960-491-2.ch002

Barakat, K., Issack, B. B., Stepanova, M., & Tuszynski, J. (2011). Effects of temperature on the p53-DNA binding interactions and their dynamical behavior: comparing the wild type to the R248Q mutant. *PloS One*, *6*(11), e27651.

Barakat, K., Mane, J., Friesen, D., & Tuszynski, J. (2010). Ensemble-based virtual screening reveals dual-inhibitors for the p53-MDM2/MDMX interactions. *Journal of Molecular Graphics and Modelling*, *28*(6), 555–568. http://doi.org/10.1016/j.jmgm.2009.12.003

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., … Bourne, P. E. (2000). The Protein Data Bank. *Nucleic Acids Research*, *28*(1), 235–242. http://doi.org/10.1093/nar/28.1.235

Betts, M. J., & Sternberg, M. J. (1999). An analysis of conformational changes on protein-protein association: implications for predictive docking. *Protein Engineering*, *12*(4), 271–283. http://doi.org/10.1093/protein/12.4.271

Bonvin, A. (2013). Coming to peace with protein complexes? 5th CAPRI evaluation meeting, April 17–19th 2013 – Utrecht. *Proteins: Structure, Function, and Bioinformatics*, *81*(12), 2073–2074. http://doi.org/10.1002/prot.24431

Boyd, D. B., & Lipkowitz, K. B. (2001). *Reviews in Computational Chemistry* (Vol. 17). New York: Wiley-VCH.

Camacho, C. J., & Vajda, S. (2002). Protein-protein association kinetics and protein docking. *Current Opinion in Structural Biology*, *12*(1), 36–40. http://doi.org/10.1016/S0959-440X(02)00286-5

Chen, C. Y.-C., & Tou, W. leong. (2013). How to design a drug for the disordered proteins? *Drug Discovery Today*, *18*(19–20), 910–915. http://doi.org/10.1016/j.drudis.2013.04.008

Chen, R., Li, L., & Weng, Z. (2003). ZDOCK : An Initial-Stage Protein-Docking Algorithm. *Proteins*, *52*(November 2002), 80–87.

Chen, R., & Weng, Z. (2002). Docking unbound proteins using shape complementarity, desolvation, and electrostatics. *Proteins: Structure, Function and Genetics*, *47*(August 1999), 281–294. http://doi.org/10.1002/prot.10092

Chen, Y. C. (2015). Beware of docking! *Trends in Pharmacological Sciences*, *36*(2), 78–95. http://doi.org/10.1016/j.tips.2014.12.001

Chène, P. (2003). Inhibiting the p53–MDM2 interaction: an important target for cancer therapy. *Nat Rev Cancer*, *3*(2), 102–109. Retrieved from http://dx.doi.org/10.1038/nrc991

Cheng, T. M.-K., Blundell, T. L., & Fernandez-Recio, J. (2007). pyDock: Electrostatics and Desolvation for Effective Scoring of Rigid-Body Protein–Protein Docking. *Proteins*, *68*(2), 503–515.

Chowdhury, R., Rasheed, M., Keidel, D., Moussalem, M., Olson, A., Sanner, M., & Bajaj, C. (2013). Protein-Protein Docking with F2Dock 2.0 and GB-Rerank. *PLoS ONE*, *8*(3). http://doi.org/10.1371/journal.pone.0051307

Chuang, G.-Y., Kozakov, D., Brenke, R., Comeau, S. R., & Vajda, S. (2008). DARS (Decoys As the Reference State) potentials for protein-protein docking. *Biophysical Journal*, *95*(9), 4217–4227. http://doi.org/10.1529/biophysj.108.135814

Cole, J. C., Murray, C. W., Nissink, J. W. M., Taylor, R. D., & Taylor, R. (2005). Comparing protein-ligand docking programs is difficult. *Proteins: Structure, Function, and Bioinformatics*, *60*(3), 325–332. http://doi.org/10.1002/prot.20497

Comeau, S. R., Gatchell, D. W., Vajda, S., & Camacho, C. J. (2004). ClusPro: An automated docking and discrimination method for the prediction of protein complexes. *Bioinformatics*, *20*(1), 45–50. http://doi.org/10.1093/bioinformatics/btg371

De Vries, S. J., Schindler, C. E. M., Chauvot de Beauchêne, I., & Zacharias, M. (2015). A Web Interface for Easy Flexible Protein-Protein Docking with ATTRACT. *Biophysical Journal*, *108*(3), 462–465. http://doi.org/10.1016/j.bpj.2014.12.015

De Vries, S. J., Van Dijk, M., & Bonvin, A. M. J. J. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*, *5*(5), 883–897. http://doi.org/10.1038/nprot.2010.32

De Vries, S., & Zacharias, M. (2013). Flexible docking and refinement with a coarse-grained protein model using ATTRACT. *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2167–2174. http://doi.org/10.1002/prot.24400

Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A protein-protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–1737. http://doi.org/10.1021/ja026939x

Duhovny, D., Nussinov, R., & Wolfson, H. (2002). Efficient Unbound Docking of Rigid Molecules. *Algorithms in Bioinformatics*, 185–200. http://doi.org/10.1007/3-540-45784-4_14

Ehrlich, L. P., & Wade, R. C. (2001). Protein–Protein Docking. In *Reviews in Computational Chemistry* (17th ed., pp. 61–92). New York: Wiley-VCH.

Elcock, A. A. H., Sept, D., & McCammon, J. A. (2001). Computer simulation of protein-protein interactions. *The Journal of Physical Chemistry B*, *105*(8), 1504–1518. http://doi.org/10.1021/jp003602d

Fernandez-Recio, J., Totrov, M., & Abagyan, R. (2002). Soft protein – protein docking in internal coordinates. *Protein Science*, *11*(2), 280–291. http://doi.org/10.1110/ps.19202.ical

Gabb, H. a, Jackson, R. M., & Sternberg, M. J. (1997). Modelling protein docking using shape complementarity, electrostatics and biochemical information. *Journal of Molecular Biology*, *272*(1), 106–120. http://doi.org/10.1006/jmbi.1997.1203

Gajda, M. J., Tuszynska, I., Kaczor, M., Bakulina, A. Y., & Bujnicki, J. M. (2010). FILTREST3D: Discrimination of structural models using restraints from experimental data. *Bioinformatics*, *26*(23), 2986–2987. http://doi.org/10.1093/bioinformatics/btq582

Gógl, G., Schneider, K. D., Yeh, B. J., Alam, N., Nguyen Ba, A. N., Moses, A. M., … Weiss, E. L. (2015). The Structure of an NDR/LATS Kinase–Mob Complex Reveals a Novel Kinase–Coactivator System and Substrate Docking Mechanism. *PLoS Biology*, *13*(5), e1002146. http://doi.org/10.1371/journal.pbio.1002146

Gray, J. J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C. a., & Baker, D. (2003). Protein–Protein Docking with Simultaneous Optimization of Rigid-body Displacement and Side-chain Conformations. *Journal of Molecular Biology*, *331*(1), 281–299. http://doi.org/10.1016/S0022-2836(03)00670-3

Grinter, S. Z., & Zou, X. (2014). Challenges, applications, and recent advances of protein-ligand docking in structure-based drug design. *Molecules*, *19*(7), 10150–10176. http://doi.org/10.3390/molecules190710150

Grosdidier, S., & Fernández-Recio, J. (2012). Protein-protein docking and hot-spot prediction for drug discovery. *Current Pharmaceutical Design*, *18*(30), 4607–18. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22650255

Guo, F., Li, S. C., Ma, W., & Wang, L. (2013). Detecting protein conformational changes in interactions via scaling known structures. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7821 LNBI*(10), 58–74. http://doi.org/10.1007/978-3-642-37195-0_6

Hernández-santoyo, A., Tenorio-barajas, A. Y., Altuzar, V., Vivanco-cid, H., & Mendoza-barrera, C. (2013). *Protein-Protein and Protein-Ligand Docking, Protein Engineering - Technology and Application*. (D. T. Ogawa, Ed.). http://doi.org/10.5772/56376

Huang, S. (2014). Search strategies and evaluation in protein – protein docking : principles , advances and challenges. *Drug Discovery Today*, *19*(8), 1081–1096. http://doi.org/10.1016/j.drudis.2014.02.005

Huang, T.-T., Hwang, J.-K., Chen, C.-H., Chu, C.-S., Lee, C.-W., & Chen, C.-C. (2015). (PS)2: protein structure prediction server version 3.0. *Nucleic Acids Research*, *454*(Ext 56921), 1–5. http://doi.org/10.1093/nar/gkv454

Jackson, R. M., Gabb, H. a, & Sternberg, M. J. (1998). Rapid refinement of protein interfaces incorporating solvation: application to the docking problem. *Journal of Molecular Biology*, *276*(1), 265–285. http://doi.org/10.1006/jmbi.1997.1519

Janin, J. (2013). The targets of CAPRI rounds 20-27. *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2075–2081. http://doi.org/10.1002/prot.24375

Janin, J., Henrick, K., Moult, J., Eyck, L. Ten, Sternberg, M. J. E., Vajda, S., … Wodak, S. J. (2003). CAPRI: A critical assessment of PRedicted interactions. *Proteins: Structure, Function and Genetics*, *52*(1), 2–9. http://doi.org/10.1002/prot.10381

Jiang, F., & Kim, S. H. (1991). "Soft docking": Matching of molecular surface cubes. *Journal of Molecular Biology*, *219*(1), 79–102. http://doi.org/10.1016/0022-2836(91)90859-5

Jiménez-García, B., Pons, C., & Fernández-Recio, J. (2013). pyDockWEB: A web server for rigid-body protein-protein docking using electrostatics and desolvation scoring. *Bioinformatics*, *29*(13), 1698–1699. http://doi.org/10.1093/bioinformatics/btt262

Jimenez-Garcia, B., Pons, C., Svergun, D. I., Bernado, P., & Fernandez-Recio, J. (2015). pyDockSAXS: protein-protein complex structure by SAXS and computational docking. *Nucleic Acids Research*, 1–6. http://doi.org/10.1093/nar/gkv368

Kann, M. G. (2007). Protein interactions and disease: Computational approaches to uncover the etiology of diseases. *Briefings in Bioinformatics*, *8*(5), 333–346. http://doi.org/10.1093/bib/bbm031

Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, a a, Aflalo, C., & Vakser, I. a. (1992). Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(6), 2195–2199. http://doi.org/10.1073/pnas.89.6.2195

Keskin, O., Ma, B., & Nussinov, R. (2005). Hot regions in protein-protein interactions: The organization and contribution of structurally conserved hot spot residues. *Journal of Molecular Biology*, *345*(5), 1281–1294. http://doi.org/10.1016/j.jmb.2004.10.077

Kojima, K., Konopleva, M., McQueen, T., O'Brien, S., Plunkett, W., & Andreeff, M. (2006). Mdm2 inhibitor Nutlin-3a induces p53-mediated apoptosis by transcription-dependent and transcription-independent mechanisms and may overcome Atm-mediated resistance to fludarabine in chronic lymphocytic leukemia. *Blood*, *108*(3), 993–1000. http://doi.org/10.1182/blood-2005-12-5148

Kortemme, T., & Baker, D. (2004). Computational design of protein-protein interactions. *Current Opinion in Chemical Biology*, *8*(1), 91–97. http://doi.org/10.1016/j.cbpa.2003.12.008

Kozakov, D., Beglov, D., Bohnuud, T., Mottarella, S. E., Xia, B., Hall, D. R., & Vajda, S. (2013). How good is automated protein docking? *Proteins: Structure, Function and Bioinformatics*, *81*(12), 2159–2166. http://doi.org/10.1002/prot.24403

Lensink, M. F., & Wodak, S. J. (2013). Docking, scoring, and affinity prediction in CAPRI. *Proteins: Structure, Function, and Bioinformatics*, *81*(12), 2082–2095. http://doi.org/10.1002/prot.24428

Lesk, V. I., & Sternberg, M. J. E. (2008). 3D-Garden: A system for modelling protein-protein complexes based on conformational refinement of ensembles generated with the marching cubes algorithm. *Bioinformatics*, *24*(9), 1137–1144. http://doi.org/10.1093/bioinformatics/btn093

Li, L., Chen, R., & Weng, Z. (2003). RDOCK:Refinement of Rigid-body Protein Docking Predictions. *Proteins: Struc. Funct. Gen*, *53*(061/14), 693–707. http://doi.org/061/14

Liang, S., Liu, S., Zhang, C., & Zhou, Y. (2007). A simple reference state makes a significant improvement in near-native selections from structurally refined docking decoys. *Proteins*, *69*, 244–253.

Liu, S., & Vakser, I. a. (2011). DECK: Distance and environment-dependent, coarse-grained, knowledge-based potentials for protein-protein docking. *BMC Bioinformatics*, *12*(1), 280. http://doi.org/10.1186/1471-2105-12-280

Lorenzen, S., & Zhang, Y. (2007). Monte Carlo refinement of rigid-body protein docking structures with backbone displacement and side-chain optimization. *Protein Science : A Publication of the Protein Society*, *16*(12), 2716–2725. http://doi.org/10.1110/ps.072847207

Mandell, J. G., Roberts, V. a, Pique, M. E., Kotlovyi, V., Mitchell, J. C., Nelson, E., … Ten Eyck, L. F. (2001). Protein docking using continuum electrostatics and geometric fit. *Protein Engineering*, *14*(2),

105–113. http://doi.org/10.1093/protein/14.2.105
Mashiach, E., Nussinov, R., & Wolfson, H. J. (2010). FiberDock: Flexible induced-fit backbone refinement in molecular docking. *Proteins: Structure, Function and Bioinformatics*, *78*(6), 1503–1519. http://doi.org/10.1002/prot.22668
Mashiach, E., Schneidman-Duhovny, D., Andrusier, N., Nussinov, R., & Wolfson, H. J. (2008). FireDock: a web server for fast interaction refinement in molecular docking. *Nucleic Acids Research*, *36*(Web Server issue), W229-32. http://doi.org/10.1093/nar/gkn186
Matsuzaki, Y., Uchikoga, N., Ohue, M., Shimoda, T., Sato, T., Ishida, T., & Akiyama, Y. (2013). MEGADOCK 3.0: a high-performance protein-protein interaction prediction software using hybrid parallel computing for petascale supercomputing environments. *Source Code for Biology and Medicine*, *8*(1), 18. http://doi.org/10.1186/1751-0473-8-18
Moal, I. H., Torchala, M., Bates, P. A., & Fernández-Recio, J. (2013). The scoring of poses in protein-protein docking: current capabilities and future directions. *BMC Bioinformatics*, *14*, 286. http://doi.org/10.1186/1471-2105-14-286
Moont, G., Gabb, H. a., & Sternberg, M. J. E. (1999). Use of pair potentials across protein interfaces in screening predicted docked complexes. *Proteins: Structure, Function and Genetics*, *35*(3), 364–373. http://doi.org/10.1002/(SICI)1097-0134(19990515)35:3<364::AID-PROT11>3.0.CO;2-4
Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., & Tramontano, A. (2014). Critical assessment of methods of protein structure prediction (CASP) - round x. *Proteins: Structure, Function and Bioinformatics*, *82*(SUPPL.2), 1–6. http://doi.org/10.1002/prot.24452
Muegge, I., & Rarey, M. (2001). Small Molecule Docking and Scoring. In *Reviews in Computational Chemistry* (pp. 1–60). John Wiley & Sons, Inc. http://doi.org/10.1002/0471224413.ch1
Nillegoda, N. B., Kirstein, J., Szlachcic, A., Berynskyy, M., Stank, A., Stengel, F., … Bukau, B. (2015). Crucial HSP70 co-chaperone complex unlocks metazoan protein disaggregation. *Nature*, *524*(7564), 247–251. Retrieved from http://dx.doi.org/10.1038/nature14884
Palma, P. N., Krippahl, L., Wampler, J. E., & Moura, J. J. (2000). BiGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins*, *39*(4), 372–384. http://doi.org/10.1002/(SICI)1097-0134(20000601)39:4<372::AID-PROT100>3.0.CO;2-Q [pii]
Pedotti, M., Simonelli, L., Livoti, E., & Varani, L. (2011). Computational Docking of Antibody-Antigen Complexes, Opportunities and Pitfalls Illustrated by Influenza Hemagglutinin. *International Journal of Molecular Sciences*, *12*(1), 226–251. http://doi.org/10.3390/ijms12010226
Pierce, B. G., Wiehe, K., Hwang, H., Kim, B.-H., Vreven, T., & Weng, Z. (2014). ZDOCK server: interactive docking prediction of protein–protein complexes and symmetric multimers. *Bioinformatics*, *30*(12), btu097. http://doi.org/10.1093/bioinformatics/btu097
Pierce, B., & Weng, Z. (2007). ZRANK: Reranking Protein Docking Predictions With an Optimized Energy Function. *Proteins*, *67*, 1078–1086.
Pons, C., Jiménez-González, D., González-Álvarez, C., Servat, H., Cabrera-Benítez, D., Aguilar, X., & Fernández-Recio, J. (2012). Cell-Dock: High-performance protein-protein docking. *Bioinformatics*, *28*(18), 2394–2396. http://doi.org/10.1093/bioinformatics/bts454
Popov, P., Ritchie, D. W., & Grudinin, S. (2014). DockTrina: Docking triangular protein trimers. *Proteins: Structure, Function and Bioinformatics*, *82*(1), 34–44. http://doi.org/10.1002/prot.24344
Ravikant, D. V. S., & Elber, R. (2010). PIE-efficient filters and coarse grained potentials for unbound protein-protein docking. *Proteins*, *78*(2), 400–419. http://doi.org/10.1002/prot.22550
Ravikant, D. V. S., & Elber, R. (2011). Energy design for protein-protein interactions. *The Journal of Chemical Physics*, *135*(6), 65102. http://doi.org/10.1063/1.3615722
Ritchie, D. W., & Kemp, G. J. L. (2000). Protein docking using spherical polar Fourier correlations. *Proteins: Structure, Function and Genetics*, *39*(2), 178–194. http://doi.org/10.1002/(SICI)1097-0134(20000501)39:2<178::AID-PROT8>3.0.CO;2-6
Schneider, T., Kruse, T., Wimmer, R., Wiedemann, I., Sass, V., Pag, U., … Kristensen, H.-H. (2010). Plectasin, a Fungal Defensin, Targets the Bacterial Cell Wall Precursor Lipid II. *Science* , *328*(5982), 1168–1172. http://doi.org/10.1126/science.1185723
Schneidman-Duhovny, D., Inbar, Y., Nussinov, R., & Wolfson, H. J. (2005). PatchDock and SymmDock: Servers for rigid and symmetric docking. *Nucleic Acids Research*, *33*(SUPPL. 2), 363–367. http://doi.org/10.1093/nar/gki481
Schneidman-Duhovny, D., Ruth, N., & Wolfson, H. J. (2007). Automatic prediction of protein interactions with large scale motion. *Proteins*, *69*, 764–773.

Sheinerman, F. B., Norel, R., & Honig, B. (2000). Electrostatic aspects of protein-protein interactions. *Current Opinion in Structural Biology*, *10*(2), 153–159. http://doi.org/10.1016/S0959-440X(00)00065-8

Smith, G. R., & Sternberg, M. J. E. (n.d.). Prediction of protein – protein interactions by docking methods, 28–35.

Smith, J. a, Vanoye, C. G., Jr, a L., Meiler, J., & Sanders, C. R. (2007). Structural Models for the KCNQ1 Voltage-Gated Potassium Channel. *Biochemistry*, 14141–14152. http://doi.org/10.1021/bi701597s

Szilagyi, A., Grimm, V., Arakaki, K., & Skolnick, J. (2005). Prediction of physical protein – protein. *Most*, *2*, 1–16. http://doi.org/10.1088/1478-3967/2/0/000

Taylor, D. W., Zhu, Y., Staals, R. H. J., Kornfeld, J. E., Shinkai, A., van der Oost, J., … Doudna, J. A. (2015). Structures of the CRISPR-Cmr complex reveal mode of RNA target positioning. *Science* , *348*(6234), 581–585. http://doi.org/10.1126/science.aaa4535

Taylor, J. S., & Burnett, R. M. (2000). DARWIN: A program for docking flexible molecules. *Proteins: Structure, Function and Genetics*, *41*(2), 173–191. http://doi.org/10.1002/1097-0134(20001101)41:2<173::AID-PROT30>3.0.CO;2-3

Terashi, G., Takeda-Shitaka, Mayuko Kazuhiko, K., Iwadate, M., Takaya, D., & Umeyama, H. (2007). The SKE-DOCK server and human teams based on a combined method of shape complementarity and free energy estimation. *Proteins*, *69*, 866–872.

Tovchigrechko, A., & Vakser, I. a. (2006). GRAMM-X public web server for protein-protein docking. *Nucleic Acids Research*, *34*(WEB. SERV. ISS.), 310–314. http://doi.org/10.1093/nar/gkl206

Vakser, I. a. (1995). Protein docking for low-resolution structures. *Protein Engineering*, *8*(4), 371–377. http://doi.org/061/14

Vakser, I. A. (2014). Protein-Protein Docking: From Interaction to Interactome. *Biophysical Journal*, *107*(8), 1785–1793. http://doi.org/10.1016/j.bpj.2014.08.033

Venkatraman, V., & Ritchie, D. W. (2012). Flexible protein docking refinement using pose-dependent normal mode analysis. *Proteins: Structure, Function and Bioinformatics*, *80*(9), 2262–2274. http://doi.org/10.1002/prot.24115

Venkatraman, V., Yang, Y. D., Sael, L., & Kihara, D. (2009). Protein-protein docking using region-based 3D Zernike descriptors. *BMC Bioinformatics*, *10*, 407. http://doi.org/10.1186/1471-2105-10-407

Villoutreix, B. O., Kuenemann, M. a., Poyet, J. L., Bruzzoni-Giovanelli, H., Labbé, C., Lagorce, D., … Miteva, M. a. (2014). Drug-like protein-protein interaction modulators: Challenges and opportunities for drug discovery and chemical biology. *Molecular Informatics*, *33*(6–7), 414–437. http://doi.org/10.1002/minf.201400040

Viricel, C., Ahmed, M., & Barakat, K. (2015). Human PD-1 Binds Differently To Its Human Ligands: A Comprehensive Modelling Study. *Journal of Molecular Graphics and Modelling*, *57*, 131–142. http://doi.org/10.1016/j.jmgm.2015.01.015

Waksman, G., & Sansom, C. (2005). Introduction : Proteomics and Protein – Protein Interactions : Biology , Chemistry , Bioinformatics , and Drug Design. In G. Waksman (Ed.), *Proteomics and Protein – Protein Interactions : Biology , Chemistry , Bioinformatics , and Drug Design* (pp. 1–18). New York: Springer.

Wang, C. H. U., & Schueler-furman, O. R. a. (2005). Improved side-chain modeling for protein – protein docking. *Protein Science*, *14*(5), 1328–1339. http://doi.org/10.1110/ps.041222905.nent

Wass, M. N., Fuentes, G., Pons, C., Pazos, F., & Valencia, A. (2011). Towards the prediction of protein interaction partners using physical docking. *Molecular Systems Biology*, *7*(469), 469. http://doi.org/10.1038/msb.2011.3

Wiehe, K., Peterson, M. W., Pierce, B., Mintseris, J., & Weng, Z. (2008). Protein – Protein Docking : Overview and Performance Analysis. In M. Zaki & C. Bystroff (Eds.), *Protein Structure Prediction* (Second Edi, Vol. 413, pp. 283–314). Humana Press.

Wodak, S. J., & Méndez, R. (2004). Prediction of protein-protein interactions: The CAPRI experiment, its evaluation and implications. *Current Opinion in Structural Biology*, *14*(2), 242–249. http://doi.org/10.1016/j.sbi.2004.02.003

Xu, Y., Wang, Y., Meng, X. Y., Zhang, M., Jiang, M., Cui, M., & Tseng, G. N. (2013). Building KCNQ1/KCNE1 channel models and probing their interactions by molecular-dynamics simulations. *Biophysical Journal*, *105*(11), 2461–2473. http://doi.org/10.1016/j.bpj.2013.09.058

Zacharias, M. (2003). Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Science : A Publication of the Protein Society*, *12*(6), 1271–1282.

http://doi.org/10.1110/ps.0239303
Zhang, Y., Zheng, N., Hao, P., Cao, Y., & Zhong, Y. (2005). A molecular docking model of SARS-CoV S1 protein in complex with its receptor, human ACE2. *Computational Biology and Chemistry*, *29*, 254–257. http://doi.org/10.1016/j.compbiolchem.2005.04.008
Zhang, Z., Schindler, C. E. M., Lange, O. F., & Zacharias, M. (2015). Application of Enhanced Sampling Monte Carlo Methods for High-Resolution Protein-Protein Docking in Rosetta. *Plos One*, *10*(6), e0125941. http://doi.org/10.1371/journal.pone.0125941

APPENDIX D: KCNQ1/KCNE1 INTERACTION DATA

Table D.1. The H-bonding interaction between KCNE1-1 and KCNQ1 in Complex #154.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-1	KCNQ1	Segment	Occupancy %
PHE12 NT	SER217 C	S3	98
	LYS218 C	S3	94
	GLN220 C	S3-S4 L	42
	ALA149 C	S1-S2 L	49
	TYR148 C	S3-S4 L	31
	VAL221 C	S4	28
LEU13 NT	SER217 C	S3	78
	GLN220 C	S3-S4 L	19
	THR153 C	S1-S2 L	11
LEU16 NT	ALA149 C	S1-S2 L	96
	GLN220 C	S3-S4 L	37
TRP17 NT	THR155 C	S1-S2 L	38
	GLY154 C	S1-S2 L	15
GLU19 NT	GLN147 C	S1-S2 L	19
GLN23 NT	GLN147 C	S1-S2 L	57
MET27 NT	THR144 C	S1-S2 L	21
LEU30 NT	LEU142 C	S1	96
	LEU156 C	S2	43
	SER143 C	S1	12
ARG33 NT	THR144 C	S1-S2 L	89
	LEU142 C	S1	27
	LEU156 C	S2	55
	GLN147 C	S1-S2 L	43
	THR155 C	S1-S2 L	190
ASP39 NT	LYS326 A TRP323 A LYS318 D ARG259 D	S6 S6 S6 S5	31 11 65
LYS41 NT	GLU290 A	S5-S6 L	36
	ASN289 A	S5-S6 L	15
	VAL288 A	S5-S6 L	12
LEU45 TMD	TRP323 A SER298 D ALA300 D	S6 S5-S6 L S5-S6 L	102 46 13
LEU48 TMD	VAL141 C LEU142 C TYR299 D ALA300 D	S1 S1 S5-S6 L S5-S6 L	222 42 55 90
MET49 TMD	LEU142 C	S1	50
LEU51 TMD	LEU137 C	S1	14

	VAL141 C	S1	27
	LEU303 D	S5-S6 L	97
	TYR299 D	S5-S6 L	17
PHE54 TMD	PHE270 D	S5-S6 L	46
PHE56 TMD	LEU134 C	S1	140
LEU59 TMD	PHE127 C	S1	86
	PHE130 C	S1	151
	LEU131 C	S1	88
	LEU134 C	S1	85
MET62 TMD	PHE130 C	S1	120
	ILE263 D	S5	12
	TYR267 D	S5	33
LEU63 TMD	PHE127 C	S1	192
	PHE130 C	S1	37
	PHE123 C	S1	36
TYR65 TMD	ILE510 C	S5	191
	GLN507 C	S5	194
	ARG506 C	S5	131
ILE66 TMD	HSD126 C	S1	166
	PHE123 C	S1	96
	PHE130 C	S1	57
ARG67 TMD	PHE123 C	S1	141
LYS69 TMD	GLN260 D	S5	100
LYS70 TMD	PHE123 C	S1	94
	GLY119 C	S1	11
	TRP120 C	S1	23
GLU72 TMD	ARG506 C	S5	292

Table D.2. The H-bonding interaction between KCNE1-2 and KCNQ1 in Complex #154.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-2	KCNQ1	Segment	Occupancy %
PHE12 NT	SER217 A	S3-S4 L	11
GLU19 NT	GLN147 A	S1-S2 L	14
LEU30 NT	GLN147 A	S1-S2 L	142
	THR144 A	S1	36
	THR155 A	S2	82
	ALA149 A	S1-S2 L	45
	ALA150 A	S1-S2 L	67
	THR155 A	S2	33
	GLY154 A	S1-S2 L	10
ARG33 NT	GLN147 A	S1-S2 L	145
ARG36 NT	GLU146 A	S1-S2 L	282
	ASN289 B	S5-S6 L	38
	SER291 B	S5-S6 L	54
ASP39 NT	ARG293 B	S5-S6 L	232
	SER291 B	S5-S6 L	182
LYS41 NT	GLU146 A	S1-S2 L	113
	SER291 B	S5-S6 L	20
LEU42 TMD	GLY297 B	S5-S6 L	15

LEU45 TMD	SER298 B	S5-S6 L	14
	TRP323 C	S6	19
LEU48 TMD	LEU142 A	S1	25
	VAL141 A	S1	22
LEU51 TMD	LEU303 B	S5-S6 L	11
	THR327 C	S6	26
	TRP323 C	S6	18
PHE54 TMD	THR327 C	S6	84
	CYS331 C	S6	39
LEU59 TMD	LEU134 A	S1	69
	PHE130 A	S1	34
MET62 TMD	ILE263 B	S5	129
LEU63 TMD	PHE130 A	S1	22
TYR65 TMD	ARG259 B	S5	286
	GLN260 B	S5	76
	ILE263 B	S5	110
ILE66 TMD	GLN260 B	S5	106
	ILE263 B	S5	94
	THR264 B	S5	17
ARG67 TMD	PHE123 A	S1	33
	GLN260 B	S5	45
LYS69 TMD	THR247 A	S5	11
	GLN260 B	S5	251
	ILE257 B	S5	25
	PHE256 B	S5	27
LYS70 TMD	PHE256 B	S5	64
	ILE257 B	S5	32
LEU71 TMD	LYS358 A	S6	64
	GLN359 A	S6	31
	LYS362 A	S1	15
	ILE257 B	S5	28
GLU72 TMD	LYS358 A	S6	119
	GLN359 A	S6	39
	ARG360 A	S6	54
	LYS362 A	S6	162
	ILE257 B	S5	21

Table D.3. The H-bonding interaction between KCNE1-1 and KCNQ1 in Complex #117.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-1	KCNQ1	Segment	Occupancy %
GLU43 NT	TRP323 A	S6	17
LEU45 TMD	TRP323 A	S6	11
	SER330 A	S6	10
	PHE130 C	S1	48
	ILE257 D	S5	14
TYR46 TMD	LYS362 C	D	12
LEU48 TMD	PHE130 C	S1	8
MET49 TMD	SER298 D	S5-S6 L	8
PHE53 TMD	PHE256 D	S5	12

PHE56 TMD	PHE270 D	S5	11
PHE57 TMD	VAL141 C	S1	10
	ALA300 D	S5-S6 L	39
	ARG195 D	S2-S3 L	20
ILE66 TMD	SER298 D	S5-S6 L	12
ARG67 TMD	PHE127 C	S1	79
	PHE256 D	S5	224
	ARG259 D	S5	46
	ARG192 D	S2-S3 L	21
LYS69 TMD	TRP323 A	S6	14
LEU71 TMD	LYS115 D	S1	11
GLU72 TMD	ARG195 D	S2-S3 L	337
	ARG192 D	S2-S3 L	284

Table D.4. The H-bonding interaction between KCNE1-2 and KCNQ1 in Complex #117.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-2	KCNQ1	Segment	Occupancy %
MET1 NT	GLN220 A	S3-S4 L	10
PHE12 NT	GLN220 A	S3-S4 L	10
GLU19 NT	LYS218 A	S3-S4 L	82
	GLN147 A	S1-S2 L	36
GLN23 NT	SER217 A	S3-S4 L	17
ASP39 NT	THR144 A	S1	32
	GLN147 A	S1-S2 L	27
	LYS285 B	S5	20
LYS41 NT	GLU290 C	S5-S6 L	8
LEU42 NT	LEU142 A	S1	10
TYR65 TMD	SER199 B	S3	36
	ILE200 B		25
	PRO197 B		16
ARG67 TMD	GLN260 B	S5	11
	PHE256 B	S5	8
LYS69 TMD	ALA194 B	S2-S3 L	42
	ARG195 B		25
LYS70 TMD	ASP242 B	S4-S5 L	119
	PHE256 B	S5	15
LEU71 TMD	PHE256 B	S5	12
GLU72 TMD	ARG116 A	S1	209
	LYS362 A	S6	56
	ARG259 B	S2	215
	ARG195 B	S2-S3 L	87
	ARG192 B	S2-S3 L	19
	LYS196 B	S2-S3 L	18
	ARG249 B	S5	18
	ARG181 B	S2-S3 L	17

Table D.5. The hydrophobic interactions between KCNE1-1 and KCNQ1 in Complex #154.
NT: N-Terminal

CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-1	KCNQ1	Segment	Hydrophobic Interaction Energy (kcal/mol)
PHE12 NT	SER217 C	S3	-0.76
LEU13 NT	SER217 C	S3	-0.62
LEU16 NT	ALA149 C	S1-S2 L	-0.36
LEU30 NT	LEU142 C	S1	-0.51
LEU30 NT	LEU156 C	S1-S2 L	-0.45
SER34 NT	LEU142 C	S1	-0.57
LEU45 TMD	TRP323 A	S6	-0.72
LEU45 TMD	SER298 D	S5-S6 L	-0.38
VAL47 TMD	TRP323 A	S6	-0.92
LEU48 TMD	VAL141 C	S1	-0.73
LEU48 TMD	LEU142 C	S1	-0.47
LEU48 TMD	TYR299 D	S5-S6 L	-0.75
LEU48 TMD	ALA300 D	S5-S6 L	-0.49
MET49 TMD	LEU142 C	S1	-0.52
LEU51 TMD	LEU137 C	S1	-0.51
LEU51 TMD	LEU303 D	S5-S6 L	-0.44
PHE54 TMD	PHE270 D	S5	-0.42
PHE56 TMD	LEU134 C	S1	-0.86
LEU59 TMD	PHE127 C	S1	-0.68
LEU59 TMD	PHE130 C	S1	-0.93
LEU59 TMD	LEU131 C	S1	-0.76
LEU59 TMD	LEU134 C	S1	-0.51
MET62 TMD	PHE130 C	S1	-1.08
MET62 TMD	ILE263 D	S5	-0.49
MET62 TMD	TYR267 D	S5	-0.37
LEU63 TMD	PHE123 C	S1	-0.62
LEU63 TMD	PHE127 C	S1	-0.87
LEU63 TMD	PHE130 C	S1	-0.54
TYR65 TMD	ILE263 D	S5	-1.1
ILE66 TMD	PHE123 C	S1	-0.75
ILE66 TMD	PHE130 C	S1	-0.38

Table D.6. The hydrophobic interactions between KCNE1-2 and KCNQ1 in Complex #154.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-2	KCNQ1	Segment	Hydrophobic Interaction Energy (kcal/mol)
LEU30 NT	ALA149 A	S1-S2 L	-0.40
LEU30 NT	ALA150 A	S1-S2 L	-0.48
LEU30 NT	THR155 A	S1-S2 L	-0.47
VAL47 TMD	TRP323 C	S6	-1.33
LEU48 TMD	LEU142 A	S1	-0.34
PHE54 TMD	THR327 C	S6	-0.47
LEU59 TMD	LEU134 A	S1	-0.46
MET62 TMD	ILE263 B	S5	-1.00
TYR65 TMD	ILE263 B	S5	-0.66
ILE66 TMD	ILE263 B	S5	-0.73
SER68 TMD	GLN260 B	S5	-0.34
SER68 TMD	PHE256 B	S4-S5 L	-0.39
LEU71 TMD	GLN359 A	S6	-0.38

Table D.7. The hydrophobic interactions between KCNE1-1 and KCNQ1 in Complex #117.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-1	KCNQ1	Segment	Hydrophobic Interaction Energy (kcal/mol)
LEU42 NT	ILE138 C	S1	-0.53
LEU42 NT	LEU142 C	S1	-0.68
LEU45 TMD	TRP323 A	S6	-1.29
LEU45 TMD	SER298 D	S5-S6 L	-0.59
TYR46 TMD	VAL141 C	S1	-0.94
TYR46 TMD	TYR299 D	S5-S6 L	-0.91
TYR46 TMD	ALA300 D	S5-S6 L	-0.67
MET49 TMD	LEU303 D	S5-S6 L	-0.87
VAL50 TMD	ILE138 C	S1	-0.56
PHE53 TMD	LEU134 C	S1	-0.83
PHE53 TMD	PHE270 D	S5	-0.98
PHE54 TMD	LEU134 C	S1	-0.89
PHE56 TMD	PHE130 C	S1	-0.74

PHE56 TMD	ILE263 D	S5	-0.72
PHE56 TMD	LEU266 D	S5	-0.94
PHE56 TMD	TYR267 D	S5	-0.6
PHE56 TMD	PHE270 D	S5	-0.67
PHE57 TMD	PHE127 C	S1	-1.17
PHE57 TMD	PHE130 C	S1	-1.56
PHE57 TMD	LEU131 C	S1	-0.91
ILE61 TMD	PHE123 C	S1	-0.48
ILE61 TMD	PHE127 C	S1	-1.08
LEU63 TMD	VAL255 D	S4-S5 L	-0.54
LEU63 TMD	PHE256 D	S4-S5 L	-0.95
SER64 TMD	PHE123 C	S1	-0.46
SER64 TMD	ARG259 D	S5	-0.87
ILE66 TMD	PHE256 D	S4-S5 L	-1.08

Table D.8. The hydrophobic interactions between KCNE1-2 and KCNQ1 in Complex #117.
NT: N-Terminal
CT: C-Terminal
TMD: Transmembrane Domain
Letters A, B, C and D in the second column, indicate the KCNQ1 subunit.

KCNE1-2	KCNQ1	Segment	Hydrophobic Interaction Energy (kcal/mol)
LEU42 TMD	LEU142 A	S1	-0.88
LEU45 TMD	TRP323 C	S6	-0.58
LEU45 TMD	LEU142 A	S1	-0.52
TYR46 TMD	LEU142 A	S1	-0.76
MET49 TMD	ILE138 A	S1	-1.14
MET49 TMD	LEU142 A	S1	-0.75
MET49 TMD	VAL141 A	S1	-0.42
PHE53 TMD	LEU131 A	S1	-0.92
PHE53 TMD	LEU134 A	S1	-0.81
PHE53 TMD	VAL135 A	S1	-0.55
PHE53 TMD	ILE138 A	S1	-0.45
PHE56 TMD	LEU134 A	S1	-0.82
PHE56 TMD	PHE130 A	S1	-0.62
PHE56 TMD	PHE127 A	S1	-0.36
PHE56 TMD	LEU131 A	S1	-0.33
PHE57 TMD	PHE127 A	S1	-0.72
PHE57 TMD	LEU131 A	S1	-0.63
MET62 TMD	ILE201 B	S3	-0.42
TYR65 TMD	ILE200 B	S3	-0.82
LEU71 TMD	PHE256 B	S4-S5 L	-0.49