THE UNIVERSITY OF CHICAGO

FOLDING AND MISFOLDING OF THE POTASSIUM CHANNEL PORE DOMAIN DURING ASSEMBLY AND TETRAMERIZATION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE PHYSICAL SCIENCES
AND
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN BIOPHYSICAL SCIENCES

BY
KEVIN C. SONG

CHICAGO, ILLINOIS
JUNE 2019

To my wife,

Louesa Song

To my parents,

Dr. Byungho Song and Jeungen Yu

To my sister,

Kelly Song

To my parents-in-law,

Bruce Akin and Dr. S. Renea Akin

"It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong" - Richard Feynman

"Ball don't lie" - Rasheed Wallace

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

First, I would like to thank my academic advisors, Dr. Benoît Roux and Dr. Tobin Sosnick, for their guidance throughout my thesis project. This thesis project was highly exploratory, forging a new collaboration between two labs. So, inherently, there were risks involved with the project potentially failing. The project was very trying and there were many times when I thought the project would not work. However, dicussing and brainstorming ideas with Benoît and Tobin kept me encouraged and provided me with many new ideas to pursue. One of my favorite things that Tobin says is "Argue with me". He really encouraged his students to discuss and argue against him. With his mentorship, I have really learned to become a critical thinker. I really appreciate them both for giving me the intellectual freedom to learn how to be creative and teach me how to be a good scientist. In addition, I would also like to thank my thesis committee members Drs. Eduardo Perozo and Andrei Tokmakoff for their insightful suggestions and discussions. Committee meetings can often be dreadful and scary for graduate students. However, I always looked forward to my committee meetings because I knew that I would come out of the meeting with new excellent ideas to try out.

I have also had the honor and privilege of working with some of the brightest people in Perozo, Roux and Sosnick labs. I would like to thank Michael Baxa, Lydia Blachowicz, Michael Clark, Isabelle Gagnon, Young Hoon Koh, Jing Li, Tian Li, Yilin Meng, Fabian Paul, Matthew Pond, Josh Riback, Joseph Sachleben, and Adam Zmyslowski for all the help with biochemistry troubleshooting and science discussions.

The University of Chicago Biophysics program has been invaluable to me, providing institutional support as well as many friendships that will last throughout my life. Michele Wittels, Julie Feder and Adam Hammond have provided constant assistance and support. Eugene Leypunskiy, Herman Gudjonson, Vaughn Spurrier, and Boleslaw Osinski provided me with support and friendships throughout my time at the University of Chicago.

Last but not least, I would like to thank my family members. My wife, Louesa Song,

has seen all the highs and lows in my journey through PhD, and she has supported me unconditionally. I would not have been able to finish my PhD as happily as I am without having her by my side, and I would also like to thank her for all of our engaging discussions about life, career, politics, podcasts and movies, which have been critical to my personal growth. I would like to thank my parents, Dr. Byung Ho Song and Jeungen Yu, who have taught me from young age that with hard work I can accomplish anything. I would like to thank them for their constant support and sacrifices they made in order for me to have the life I have today. I am forever grateful for their love and support. I would also like to thank my sister, Kelly Song, for sending me funny Youtube videos and introducing me to foodbabyNY to keep me laughing and smiling. Bruce and Dr. S. Renea Akin, my parents-in-law, have taken me into their family from the first day I met them, and I thank them for all the good BBQ they fed me.

ABSTRACT

This thesis connects the dynamics of Kv1.2 and KcsA potassium channel pore domain monomers to the kinetics of tetramerization. In simulations, monomers adopt multiple conformations with the three helices folded. NMR studies also find the monomers to be dynamic and structurally heterogeneous. However, a KcsA construct with a disulfide bridge engineered between the two transmembrane helices has an NMR spectrum with well-dispersed peaks, suggesting that the monomer can be locked into a native-like conformation. During tetramerization, FRET results indicate that monomers rapidly oligomerize upon insertion into liposomes, forming a dense protein-rich phase. Folding within this protein-rich phase occurs along separate fast and slow routes, with $\tau_f \sim 40$ and 1500 seconds, respectively. In contrast, constructs bearing the disulfide bond mainly fold via the faster pathway, suggesting that maintaining the TM helices in their native orientation reduces misfolding. Interestingly, folding is concentration independent in spite of the tetrameric nature of the channel, indicating that the rate-limiting step is unimolecular and occurs after monomer association in protein-rich phase. Finally, despite its name, the addition of KcsA's C-terminal "tetramerization" domain does not improve the kinetics of tetramerization.

CHAPTER 1

INTRODUCTION

While the field of soluble protein folding has greatly advanced in the past few decades, membrane protein folding is still in its infancy [1, 2]. The difficulty working with membrane proteins and the lack of available protein structures has limited the amount of studies on the folding of membrane proteins. However, we have seen numerous recent advances in technology and insights into membrane protein folding in the past decade. With better spatio-temporal resoultion of atomic force microscopy (AFM), we can monitor folding and unfolding of a single turn of a helix in bacteriorhodopsin [3]. With steric-trapping methods, ΔG_U° of GlpG can be measured in near native environment in bicelles without denaturants [4]. With optical tweezers, we can watch GlpG fold reversibly in a bicelle environment [5]. This thesis attempts to further advance the field of membrane protein folding by understanding the folding of potassium channels, an oligomeric membrane protein.

1.1 Membrane protein folding

Membrane proteins are the cell's gateway to its environment. Using a variety of membrane proteins, a cell transmits information, receives nutrients, shapes itself and responds to external signals [1, 2, 6, 7]. Membrane proteins reside in lipid membranes, and they constitute more than 30% of the entire proteome [8, 9]. Despite their functional importance and their abundance, our understanding of biochemical and biophysical properties of membrane proteins are still immature because of the difficulties working with membrane proteins experimentally. Membrane proteins easily aggregate because of their high hydrophobicity, and without the correct membrane mimetics (e.g., detergents, bicelles, and lipid vesicles), the protein structures can be easily disturbed [6]. While it is difficult to study membrane protein folding experimentally, significant advances have been made in the field since the

introduction of a two-stage model in 1990, which are highlighted in the following sections.

1.1.1 α-helical versus β-barrel membrane protein folding

Generally speaking, membrane proteins can be sub-divided into 2 classes (**Fig. 1.1**): α-helical and β-barrel membrane proteins [6, 10, 11]. α-helical membrane proteins are mainly found in the plasma membrane of eukaryotes, the inner membranes of bacterial cells, and sometimes in the outer membranes of bacteria [8, 9]. They constitute between 20 - 25% of all open reading frames [8, 9]. On the other hand, β-barrel membrane proteins are mostly found in the outer membranes of bacteria, mitochondria and chloroplasts, and they constitute only a few percent of all open reading frames [8, 9]. The folding processes for α-helical and β-barrel membrane proteins are very different utilizing different sets of chaperones and mechanisms.

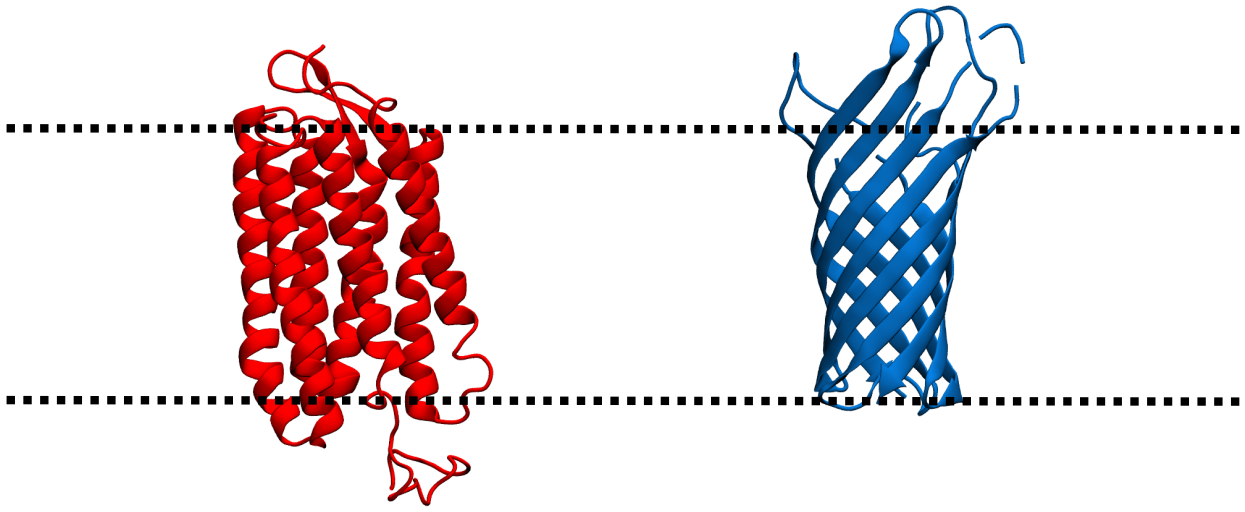

Figure 1.1: **α-helical versus β-barrel membrane protein.** Two representative structures of α-helical and β-barrel membrane proteins are shown. Structure of bacteriorhodopsin (PDB ID: 1X0S) monomer is on the left in red, and OmpA (PDB ID: 1BXW) is shown on the right in blue. Theoretical membrane boundaries are drawn in dashed black lines.

Individual α-helices can be stable in the membrane as long as sidechains are sufficiently hydrophobic; however, β-strands alone are not stable in the membrane because of exposed

unsatisfied hydrogen donors and acceptors from the peptide backbone. Due to this basic difference in backbone exposure, the mechanism for folding of α-helical and β-barrel membrane proteins are drastically different. β-barrel membrane proteins insert and fold concurrently whereas α-helical membrane proteins can insert one helix or in pairs with the helices rearranging to fold within the bilayer. In addition, β-barrel membrane proteins have been shown to fold reversibly from solution containing traditional denaturants such as urea and guanadine hydrochloride (GdnHCl) into lipids [12, 13, 14, 15, 16, 17, 18]. Because β-barrel membrane proteins can be monomeric and soluble in traditional denaturants, several proteins' folding thermodynamics have been measured (e.g., OmpA, OmpLA, OmpW, PagP) [12, 13, 14, 16]. They found that β-barrel membrane proteins fold cooperatively by forming β-sheets at the lipid-water interface and insert simultaneously.

On the other hand, α-helical membrane proteins are highly resistant to urea and GdnHCl, largely due to their hydrophobicity. For example, bacteriorhodopsin has been shown to be stable in 7 M urea with dimyristoyl phosphatidylcholine/3-[(3-cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate (DMPC/CHAPSO) membrane, and this condition is considered to be a reasonable mimic of folding condition [10, 19]. For folding studies of α-helical membrane proteins, the denaturant of choice is sodium dodecyl sulfate (SDS). However, with SDS, the unfolded state of α-helical membrane proteins still retains the majority of its helicity [20, 21, 22, 23]. Because the SDS-unfolded state still retains a lot of helicity, SDS folding studies mimic *in vivo* folding within the membrane, where the association of helices is the primary driving force for folding. However, no studies have ever studied the thermodynamics of folding of membrane proteins from completely unfolded chain in solution to fully folded in the membrane [10].

The difference in folding mechanism is also reflected in how chaperones help these two classes of proteins fold. For β-barrel membrane proteins, BAM complexes have been shown to help them insert and fold by thinning the lipid bilayer thickness hence lowering the energy

barrier to insert and fold [24, 25]. On the other hand, for α-helical membrane proteins, ribosomes dock onto the SEC translocons in the membrane and single helices are synthesized and inserted into the membrane one at a time as shown in Figure 1.2. The exact mechanism for how single helices partition from inside the translocon to the membrane is still hotly debated [26]. Many advances have been made in the field of β-barrel membrane proteins and there are excellent reviews available [10, 27, 28]. Here in this thesis, the focus will be on the folding of α-helical membrane proteins.

1.1.2 The two stage model and beyond

In 1990, the two stage model was proposed to provide a general framework for α-helical membrane protein folding as shown in Figure 1.2 [29]. The model was proposed based on refolding studies of bacteriorhodopsin, where Popot and Engelman showed that bacteriorhodopsin can still re-assemble even after being cut in different loop positions [30]. This seminal experiment was the first to show that helix-helix association in the membrane drives the folding of membrane proteins once all the helices are inserted in the membrane. Based on these results, Popot and Engelman proposed that α-helical membrane proteins fold in two stages: first helices must insert into the membrane, then helices associate laterally to fold into its native state, which is governed by thermodynamics.

Stage 1: Insertion

The first stage of membrane protein folding is called insertion, where helices insert normal to the membrane with close to its native-like secondary structures [29]. This process is very different for *in vivo* and *in vitro*. *In vivo*, the protein is secreted from the ribosome directly into SEC translocons, then the helices are transferred from the translocon into the lipid bilayer [2]. However, *in vitro*, a peptide transfers directly from solution into the lipid bilayer. For *in vivo* insertion, the exact mechanism on how a single helix transfers from inside the

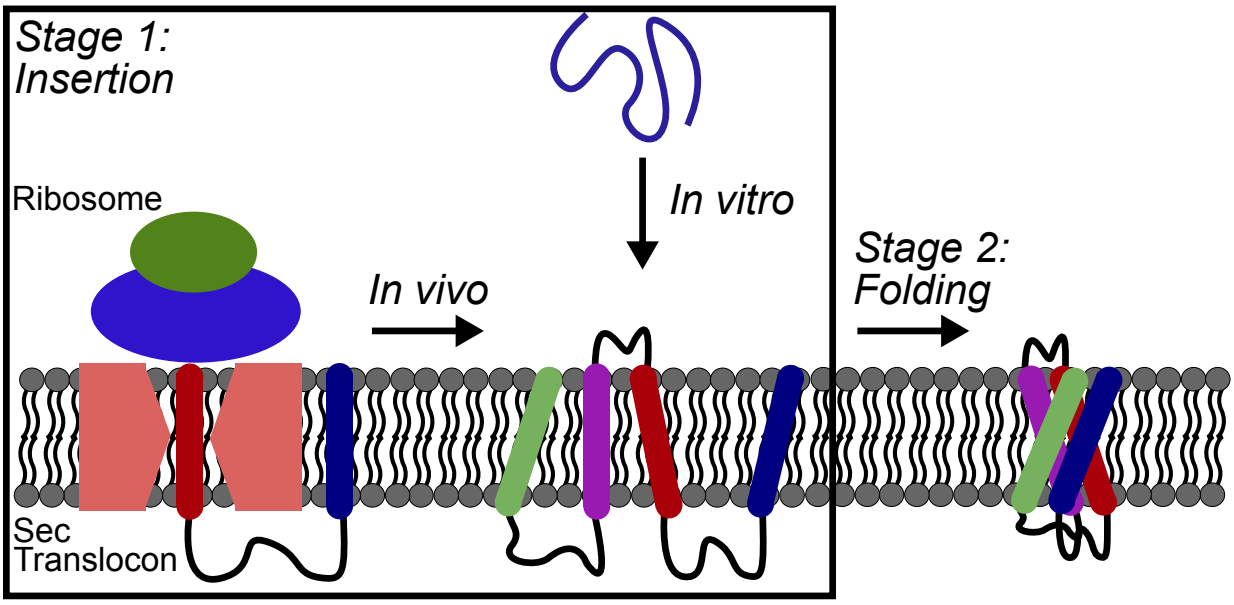

Figure 1.2: **Two stage folding process for membrane protein folding.** The first stage of the membrane protein folding process is insertion, where single helices are synthesized and inserted into the bilayer one at a time. As noted above, the insertion process is different for *in vivo* and *in vitro*. However, upon insertion of helices into the membrane, folding within the membrane is thought to process via the same pathway.

SEC translocon into the bilayer is not fully understood yet, but the transfer is thought to happen via a lateral gate that opens in the translocon [31, 32]. A more recent model suggests that helices move between inside of the translocon, lipid-water interface, and the hydrophobic core of the membrane [26]. Regardless, the insertion process is largely determined by the hydrophobicity of the peptide segment, and several different hydrophobicity scales were developed to understand the insertion process better [13, 33, 34, 35].

Most notably, the Wimley-White hydrophobicity scale is often used as the gold standard in the field of membrane protein fodling [33, 34]. They determined the partitioning free energies of all amino acid residues to the water-lipid interface using 1-Palmitoyl-2-Oleoyl-sn-glycero-phosphatidylcholine (POPC) as a model membrane system and to octanol, which is taken as a mimic of the hydrophobic core of the membrane. More recently, "biological" hydrophobic scales were developed using an *in vitro* translation system with the endoplasmic

reticulum (ER) Sec61 translocon for α-helical membrane proteins. For β-barrel membrane proteins, OmpLA was used as a guest-host system to determine a hydrophobicity scale for β-barrel membrane proteins [13]. Using these amino acid hydrophobicity scales and sliding window averages, transmembrane segments can be predicted well.

Stage 2: Folding within the membrane

After successful insertion of helices into the membrane, the helices begin to undergo folding within the membrane, with the helices starting to associate laterally. In addition, for some systems like LacY, the overall topologies can be flipped in *in vivo* and *in vitro* when the lipid composition changes [36, 37, 38]. Folding within the membrane is a complex process. Protein-protein and protein-lipid interactions act against each other. Hydrogen bonding, van der Waals interaction, polar side-chain burial, and salt-bridges could compete or complement each other to drive folding [2]. For soluble proteins, hydrophobic burial is one of the major driving forces for folding [39]. However, membrane proteins reside in a hydrophobic environment, so hydrophobic burial is not driving folding. For membrane proteins, without the hydrophobic effect, this implies that folding is likely in a delicate balance between all the weak forces mentioned above [2].

No studies have directly measured the thermodynamics of α-helical membrane protein folding from a completely unfolded state to folded state in the membrane because of its high hydrophobicity and our inability to unfold helical membrane proteins using conventional denaturants [10]. However, many studies have utilized SDS as membrane protein denaturant and measured thermodynamic stability of membrane proteins going from SDS-denatured state to folded state. This process likely mimics the folding within the membrane because in the SDS-unfolded state, most membrane proteins retain their native helicity but lose the helix-helix interactions [20, 21, 40]. Using SDS, many studies have been successful in measuring the unfolding free energies using SDS as denaturant. ΔG_U^0 have been measured

for bacteriorhodopsin [21, 41], diacylglycerol kinase (DGK) [42], DsbB [40], GlpG [4, 5, 43], peripheral myelin protein 22 (PMP22) [44], and others.

Recent advances in technology have also allowed measurements of ΔG_U^0 in native-like conditions. For example, the steric-trapping method, pioneered by Hong and Bowie, couples dissociation of helices with binding of streptavidin to the protein to measure the unfolding rate [4, 45]. Atomic force microscopy (AFM) and optical tweezers have been used to pull bacteriorhodopsin and GlpG reversibly. Perkins lab has significantly improved the spatio-temporal resolution of AFM, in which we can begin to observe reversible folding of a single turn of a helix. With these new techniques in place, we can begin to explore the folding of a variety of α-helical membrane proteins, which will help us answer the question of what is driving the folding of α-helical membrane proteins.

Stage 3: Additional folding and oligomerization

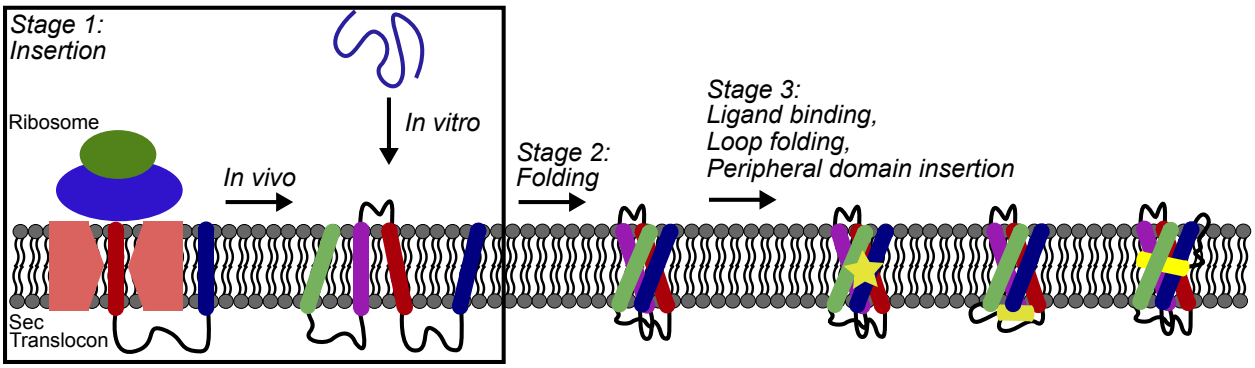

Figure 1.3: **Three stage model for membrane protein folding.** The first two stages are identical to the two stage model as shown in Figure 1.2. The third stage can be a variety of events, such as ligand binding, folding of the loops and/or peripheral domain insertion.

In the two stage model, insertion and lateral association of helices describe the entire folding process. However, many membrane proteins require more intricate folding even after the lateral association of helices. Therefore, a more detailed three stage model was proposed as shown in Figure 1.3 [46]. This model proposes that further folding happens

after lateral helix association, which can include binding of a ligand, folding of loop regions and insertion of peripheral domains into the bilayer. While the model itself did not mention oligomerization as part of Stage 3, for oligomeric membrane proteins, another step is required where monomeric subunits must associate to form its native structure. This is relevant for many different types of ion channels, and in particular, this thesis will focus on the folding of potassium channels.

1.2 Potassium channels

1.2.1 Potassium channel function and diseases

Potassium channels are found in all kingdoms of life [47]. From bacteria to humans, the basic function of potassium channels is to selectively regulate the flow of K^+ ions across cell membranes [47, 48, 49, 50]. While the basic functions are identical across organisms, potassium channels have many different uses in different organisms. In prokaryotes, potassium channels are responsible for cell growth and survival. For biofilms, the community of microorganisms use potassium channels to communicate with each other to coordinate their feeding cycles [51, 52].

In humans, potassium channels are used to drive muscle contraction and neuronal signaling [48, 49, 50]. Because of their significant role in human physiology, the proper functioning of potassium channels are critical to human health. [53] There are over 50 known genes encoded for potassium channels in humans, and mutations to any of these potassium channels presumably results in a disease [53]. Some examples include long-QT syndromes, *Weaver* syndrome, familial convulsions, hearing and vestibular diseases, Bartter's syndrome, and familial persistent hyperinsulinemic hypoglycemia of infancy.

Two of the most well-known diseases related to potassium channel mutations are long-QT (LQT) syndromes and *Weaver* syndromes. LQT syndromes are cardiac diseases that

affect the rhythm of the heart beat [54, 55]. This genetically inherited disease can cause heart to suddenly beat faster and more chaotically. With prolonged chaotic beating of the heart, this disease can cause sudden death in people with LQT syndromes. Mutations in KCNQ1 and hERG genes can cause a variety of LQT syndromes [53]. Studies of some of the more common mutations in these genes showed that assembly and proper trafficking of these potassium channels to plasma membrane of the cell is diminished in these mutants, which causes the heart to beat erratically.

Whereas LQT syndromes affect the function of the heart, *Weaver* syndromes affect the function of neurons with similar symptoms as Parkinson's diseases, resulting in severe tremors and loss of control of motor functions. In some cases, symptoms can also include male infertility, and sporadic seizures. The *Weaver* was mapped to a single mutation (G156S) in the pore region of Kir3.2 potassium channel. This mutation renders Kir3.2 potassium channel to be non-selective, allowing sodium ions through as well as potassium ions.

The importance of potassium ion channels in physiology is indisputable. Because of the importance, potassium channels are common drug targets, and many studies have focused on structure and function of potassium channels. However, the folding process of potassium channel is largely uncharacterized. Understanding the folding of potassium ion channels will be important as LQT and *Weaver* syndromes are all implicated to be caused by folding defects of K^+ ion channels.

1.2.2 Potassium channel structure

The first structure of potassium channel (**Fig. 1.4**) was solved in 1998 by Roderick MacKinnon's lab, for which he was awarded the Nobel Prize in Chemistry for in 2003 [56]. The structure showed that the pore is formed by the juxtaposition of 4 subunits along the symmetry axis. The pore is lined with oxygen atoms from the backbone of the selectivity filter residues (**Fig. 1.4C**). These oxygen atoms provide 4 binding sites for K^+ ions as they pass

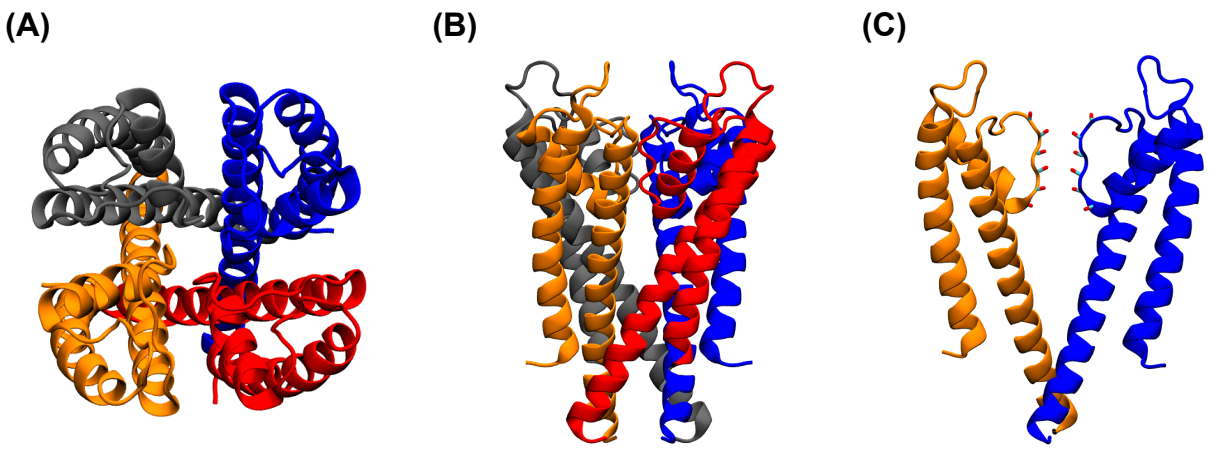

Figure 1.4: **Structure of KcsA (PDB ID: 1R3J) shown from different angles**. **(A)** KcsA shown from the extracellular side of the membrane. **(B)** KcsA shown from the side. **(C)** KcsA viewed from the side with 2 monomers removed for better visualization of the selectivity filter. The oxygen atoms lining the selectivity region is rendered in Licoriche mode and the protein structures are rendered in New Cartoon mode in VMD. Each monomer of KcsA is colored in grey, blue, orange and red.

through the pore. The pore selectively allows K^+ ions to flow through excluding other ions including Na^+, which has similar metallic structure as K^+ ions. Because of this important function of the potassium channel pore, the pore loop region is highly conserved amongst its family members. In particular, the **TVGYG** sequence in the selectivity filter region is highly conserved, because of its functional importance.

In terms of potassium channel folding, several biophysical properties of K^+ channels are relevant to their dynamics. The ion-conducting pore of K^+ channels, formed by the re-entrant segment located between the two hydrophobic TM helices, is lined with polar and charged residues. If the helices retained the same structure and orientation in the monomeric form as they do in the tetramer, the polar residues along with the non-hydrogen-bonded backbone atoms in the loops would be exposed to the hydrophobic core of the membrane. As these interactions are energetically unfavorable, the dominant conformation of individual monomers prior to tetramerization is unclear. Previous thiol-labeling experiments have shown that for Kv1.3, the monomers maintain the helical structure of the re-entrant p-helix in

a native-like orientation [57, 58]. Likewise, molecular dynamics (MD) simulations indicated that the monomers native conformation was stable on a timescale of ~ 1 μs [57]. These results demonstrated that the p-helix remains helical and lies at the water-lipid interface, however, the relative orientation of the 2 transmembrane helices and the p-helix is not known. The dynamics of monomer will be further explored in Chapter 2 of this thesis.

1.3 Potassium channel folding

1.3.1 Folding of KcsA

Reversible folding studies of KcsA

One of the first evidence that KcsA can be refolded from a completely unfolded peptide came from Valiyaveetil et al. [59]. The study showed that the N-terminal fragment corresponding to residues 1 – 73 of KcsA can be recombinantly expressed in *E. coli*, and a synthetic C-terminal fragment corresponding to residues 74 – 125 can be chemically ligated to the N-terminal fragment. Once, the two fragments are ligated, KcsA can be refolded into tetramers by dilution of the monomers into liposome solutions. This study was the first to show that membrane proteins can be chemically synthesized and refolded from SDS-unfolded state. While the study demonstrated the power of chemical synthesis for membrane protein production, it did not explore the folding kinetics of KcsA.

In the following years, several papers detailing unfolding and folding KcsA using 2,2,2-trifluoroethanol (TFE) were published [60, 61, 62]. The studies showed that at lower TFE concentrations, tryptophan fluorescence can be monitored to measure folding and unfolding of KcsA reversibly. However, at higher TFE concentrations, KcsA could not be refolded and formed irreversible aggregates. TFE is known to promote helicity in proteins, so for a α-helical membrane protein, it was not clear what this TFE-unfolded state would exactly look like. However, these studies suggested that at low TFE concentrations, TFE unfolds the

pore region but does not affect the transmembrane helices, and at high TFE concentration, everything dissociates including the pore region and the transmembrane helices, leading to irreversible aggregation of KcsA. Nevertheless, the interesting takeaway from these studies is that the pore region seems to be parts that unfold first, and then the chamber created by the 8 transmembrane helices is disrupted by high concentration of TFE. While these studies are the first to show that TFE can be a good reagent to monitor reversible refolding of KcsA, the exact mechanism for KcsA folding was still unclear and the studies did not examine the dynamics of KcsA monomers, which would surely affect the folding kinetics of KcsA tetramers. Interestingly, there are no known kinetics studies of the folding of potassium channels. The TFE studies measured thermodynamics of native KcsA tetramers with TFE, but did not measure the kinetics of folding.

Role of C-terminal stalk domain

Kv channels have N-terminal tetramerization domains and KcsA has a C-terminal helical bundle that has been thought to assist with the tetramerization process (**Fig. 1.5**). For KcsA, when KcsA is expressed in *E. coli* without the C-terminal stalk domain, the expression level is highly diminished [60]. In addition, the presence of C-terminal stalk has been shown to increase the thermostability of KcsA, and pH-sensitivity of the C-terminal stalk has been known to help KcsA function as a pH-sensitive potassium channel [20, 63, 64, 65].

Because of these attributes, many have thought that the KcsA C-terminal helical bundle must assist with the assembly of the tetramers in the membrane. The low expression level due to deletion of C-terminal stalk is the main evidence for such hypothesis. In addition, for Kv channels, the presence of their N-terminal tetramerization domain has been shown to directly help with the tetramerization process [66]. However, there is no direct evidence of the C-terminal stalk assisting with KcsA folding in part because there are no known studies of KcsA folding kinetics in the first place. In Chapter 5 of this thesis, we will evaluate the

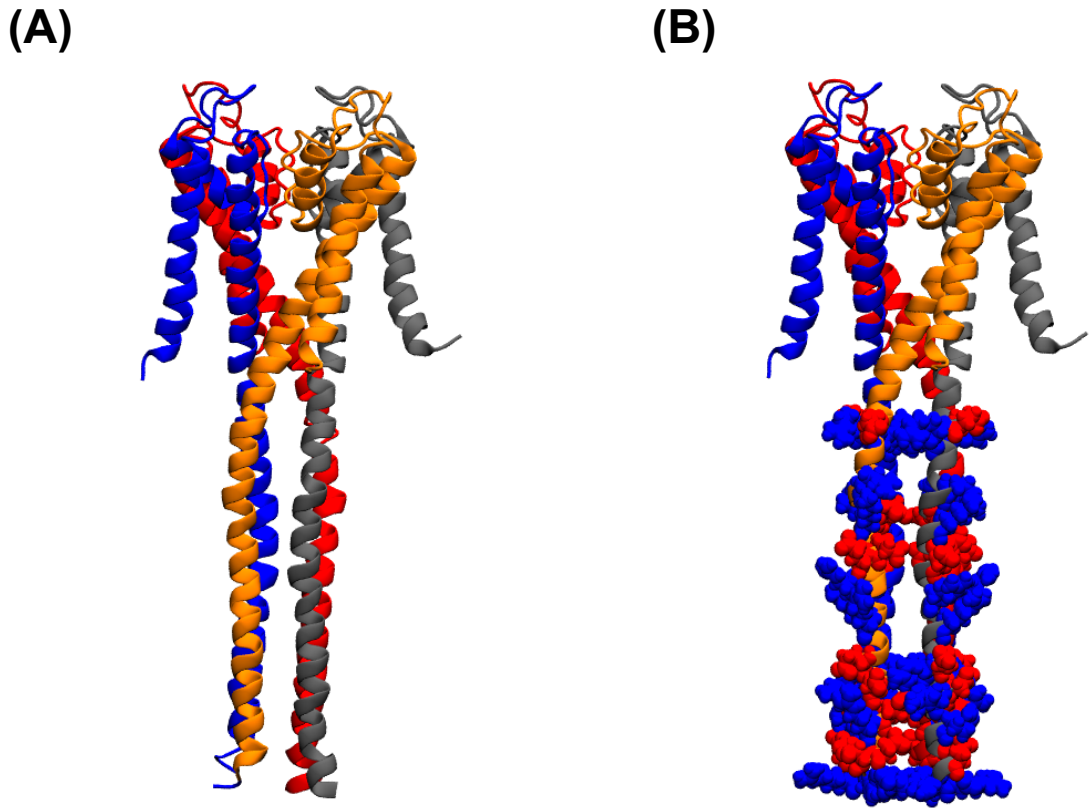

Figure 1.5: **Structure of KcsA full-length (PDB ID: 3EFF)**. **(A)** KcsA full-length is rendered in cartoon mode in VMD. Each monomeric subunit is colored in orange, grey, red and blue. The C-terminal stalk forms a helical bundle. **(B)** On the C-terminal stalk, charged residues are rendered in van der Waals mode in VMD, and the positive and negative residues are colored in blue and red, respectively.

folding kinetics of KcsA with and without the C-terminal stalk domain, and to my knowledge, this is the first to study the direct influence of KcsA C-terminal stalk on the folding kinetics of KcsA pore domain.

1.3.2 Folding of voltage-sensing potassium (Kv) channels

The first folding studies of *Shaker*-type voltage-gated potassium channel studied the effect of deleting the tetramerization (T1) domain on the kinetics of tetramerization process [66]. The

study showed that while the mutant without the T1 domain can still function as a tetramer, channel formation was dramatically reduced [66]. Upon introducing GCN1, a parallel helical tetramer, into the T1 region, channel formation was rescued. Based on this result, the T1 domain was suggested to help channels form by first forming tetramers and letting the pore domain assemble in sequence. This result provided a glimpse of how potassium channels fold. They hypothesized that the monomeric subunits first partition into oligomers with the help of T1 domains associating with each other, then the transmembrane pore domains fold subsequently.

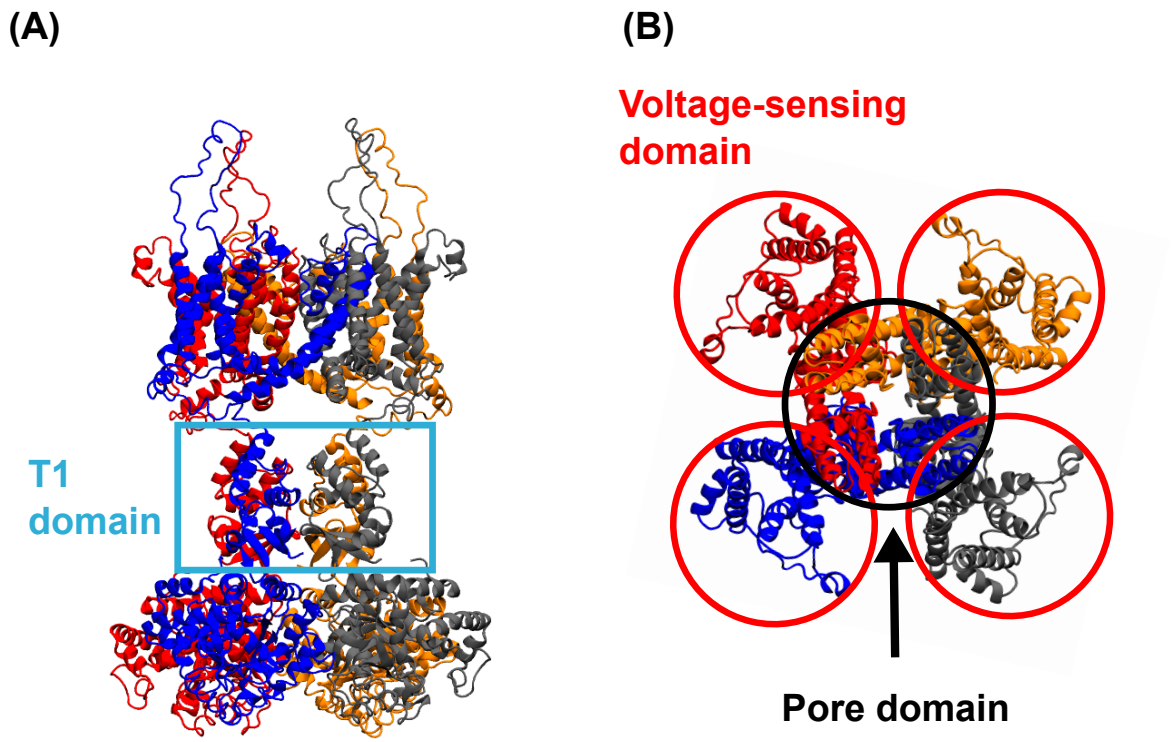

Figure 1.6: **Structure of Kv1.2 full-length (PDB ID: 3LUT)**. **(A)** Kv1.2 full-length is rendered in cartoon mode in VMD and visualized from the side. Each monomeric subunit is colored in orange, grey, red and blue. The T1 domain is highlighted in cyan. **(B)** Kv1.2 is viewed from the top. Inside the red circles are the 4 helix voltage-sensing domain and inside the black circle is the pore domain.

The first *in vitro* folding of Kv channels was done with KvAP, which is a voltage-sensitive

potassium channel found in archeabacterium, *Aeropyrum Pernix* [67]. The study showed that KvAP subunits can insert quickly into the bilayer at any given temperature; however, folding into tetramers is much more efficient at 80 °C than it is at room temperature. Using gel-based refolding assay in conjunction with SDS-quenching, they showed the first kinetic trace of potassium channel tetramerization [67]. The authors showed that KvAP folding kinetics depend on lipid composition and temperature; however, the study did not include any protein concentration dependence. The concentration dependence study could have provided some mechanistic insight into how potassium channels fold in general. Regardless, the study showed that these Kv channels can be refolded from SDS-solubilized unfolded state, and the refolding kinetics has a strong dependence on presence of phosphatidylethanolamine (PE) lipids and high temperature of 80 °C. We will explore the concentration dependence of tetramerization with KcsA in Chapter 5.

For Kv1.3, using a novel mass-tagging and molecular tape measure, Gajewski et al. showed that the Kv1.3 monomers can immediately begin to form secondary structures as they exit the ribosome-translocon complex [57]. In addition, with 650 ns of simulations and van der Waals energy calculations of Kv1.2 monomers, the native-like sturcture of Kv1.2 pore domain was thought to be stable in the membrane [57]. Using mass-tagging thiol-labeling methods, the structure of pore loop region was probed further [58]. By scanning the pore loop region with cysteines and comparing the rate of pegylation with the helical pattern, the pore helix was found to retain its native-like secondary structures and be buried [58]. The rate of pegylation varied with a helical periodicity suggesting that one side of the helix is more buried than the other [58]. Then, using a fixed cysteine mutation and introducing mutations at different sites along the p-helix, some of the key residues that change the orientation of the pore helix were found. In summary, what these thiol-labeling and short simulations suggested was that the pore domain of Kv1.3 channels retain their native-like structure prior to tetramerization. Although the experiments showed that the pore domain

monomers can maintain its native-like secondary structures, there are no evidence that the monomers would retain their native-like tertiary structures. The simulations suggested that the pore helix monomer is capable of holding a native-like structure for 650 ns, but the simulations may not be long enough to see more drastic changes in the tertiary structure. The dynamics of potassium channel monomers will be further explored in Chapter 2 with longer microsecond timescale simulations and millisecond timescale Markov state modeling analysis.

1.4 Aims and significance

1.4.1 Understanding the dynamics of potassium channel monomers

The structure and dynamics of potassium channel monomer in the membrane should affect its ability to tetramerize. To study the structure and dynamics of potassium channel monomers, we use molecular dynamics (MD) simulations, Markov state modeling (MSM), and nuclear magnetic resonance (NMR) spectroscopy. Previous studies established that the native-like secondary structures of Kv1.3 pore domain monomers are intact; however, the degree to how disordered the tertiary structures are remained elusive [57, 58]. The goal of my project is to characterize the dynamics of the potassium channel pore domain monomers in order to gain some insight into its effect in tetramerization.

1.4.2 Biochemical prepartion of wild-type KcsA and a more native-like mutant monomers

KcsA is a good model system for studying potassium channels [48, 56, 64]. KcsA expresses well in heterologous *E. coli* expression systems, and purification protocol has been well optimized. In addition, KcsA is an extremely stable membrane protein in its native tetrameric form. For biochemical purposes, being stable is a good feature; however, if the tetramer is

too stable, how do we isolate monomers for biophysical studies? In order to prepare stable KcsA monomer samples, we adopt several different biochemical prepartion protocols to prepare monomeric potassium channels. In addition, we designed a KcsA mutant that retains its native-like structure more than the WT KcsA monomer does. The protocols and results verifying the biochemical prepartion of WT and mutant KcsA monomer are presented.

1.4.3 Connecting the dynamics of potassium channel monomers to the assembly process

According to simulations and experiments, WT KcsA seems to be more structurally diverse than the more native-like variant having a disulfide bond holding the two transmembrane helices together. With the use of gel-based refolding assay and FRET, we investigate the folding kinetics of WT KcsA and the more native-like variant. Concentration dependent studies of the refolding kinetics show that the tetramerization process does not have a concentration dependence despite the native state of the protein being a tetramer. These results suggested that the rate-limiting step is unimolecular and led us to hypothesize that the proteins might form a dense protein-rich phase in lipid bilayers prior to formation of native tetramers. We verify this hypothesis using FRET and show that KcsA forms an early protein-rich phase in lipid bilayers and slowly form native tetramers.

1.4.4 Significance

Here in this thesis, we connect the dynamics of potassium channel monomers to the kinetics of tetramerization. Far from being a simple association of pre-formed monomers, they initially exist as a heterogeneous ensemble in a protein-rich phase in the membrane, which may be a general phenomenon for membrane proteins. Tetramerization kinetics indicate that the rate-limiting step is unimolecular despite the native state being a tetramer. Folding can occur along fast or slow (misfolded) pathways that can be modulated with mutants that

trap monomers in a native-like state. In spite of its name, the C-terminal tetramerization domain in KcsA does not enhance tetramerization, suggesting it plays another role in channel function.

CHAPTER 2

MOLECULAR DYNAMICS SIMULATIONS AND MARKOV STATE MODELING OF POTASSIUM CHANNEL MONOMERS

2.1 Introduction

For oligomeric membrane proteins like potassium channels, the dynamics of monomers are likely to affect the oligomerization process. *In vivo*, the monomers are inserted one at a time from the ribosome-translocon complex, and they must find one another and associate to form native tetramers. Prior to tetramerization, the monomers could be pre-folded in the native-like configuration, or the monomers could exist in a messy heterogeneous ensemble of structure. While the answer to this problem is fundamentally important to addressing the folding of potassium channels, no studies have looked at the dynamics of potassium channel monomers at the atomic-level.

Recently, thiol-labeling experiments demonstrated that the Kv1.3 pore domain monomers immediately begin to form its native-like secondary structures coming out of the ribosome-translocon complex [57]. Through cysteine-scanning experiments, the pore-helix was also shown to maintain its native-like secondary structures [58]. These studies are the first to show that the pore domain monomers can retain its native-like secondary structures; however, the experiments were not able to resolve what the tertiary structure of the monomer looks like.

In order to address the question of how dynamic is the potassium channel monomer, we use molecular dynamics (MD) simulations and Markov state modeling (MSM) analysis of Kv1.2 and KcsA monomers.

2.2 Methods

2.2.1 Preparation of Kv1.2 and KcsA monomer simulations

For all systems, the initial structures for the monomers were taken from the tetrameric Kv1.2 X-ray crystal structure (PDB ID: 2A79) using only the pore domain of 99 amino acids corresponding to residues 323 - 421. [68] The starting structure for KcsA monomer simulations was taken from the tetrameric KcsA X-ray crystal structure (PDB ID: 1R3J). [69] All systems were prepared by using CHARMM-GUIs Membrane Builder module (www.charmm-gui.org). [70, 71, 72, 73, 74] For Kv1.2, each system contained 70 1-palmitoyl-2-oleoyl-sn-phosphatidylcholine (POPC) molecules per leaflet totaling up to 140 POPC molecules in total, and for KcsA, each system contained 70 1,2-dimyristoyl-sn-glycero-3-phosphatidylcholine (DMPC) molecules per leaflet totaling up to 140 DMPC molecules to match the NMR sample conditions. All systems were hydrated by creating a water box 17.5 Å above and below the proteins maximum and minimum Z-positions. In addition, all systems were neutralized with 150 mM KCl. Each system comprised a total of approximately 45,000 atoms.

For Kv1.2 simulations, a number of different simulations are aggregated for analysis and they are all summarized in the table below.

Table 2.1: Simulation summary for Kv1.2

Simulation	Temperature (K)	Number of Trajectories	Trajectory length
Anton simulation	353.15	1	16.2 μs
GPU	303.15	10	$\sim$9.7 μs
Adaptive sampling	303.15	4685	0.06μs
Total:	**N/A**	**4696**	**394 μs**

For KcsA simulations, 5 independent simulations of KcsA WT monomer were carried out at T = 353 K to enhance sampling. All simulations were run with the parameters described in *Amber16 simulations* section below, and each simulation was run up to 6 μs. In total 30

μs of KcsA WT simulations were accumulated.

2.2.2 Anton simulations

A long-time simulation run on Anton [75] was prepared through energy minimization and equilibrating the initial simulation system using CHARMM36 force field with NAMD. [76] Each system was energy minimized for 1,000 steps and was slowly equilibrated by reducing restraints on protein backbone, sidechain, and lipids gradually over 5 ns as recommended by CHARMM-GUI. After initial equilibrations with the restraints, the system was freed from all restraints and run for 100 ns before preparing the system for running simulations on Anton. The temperature was set to 353 K to increase sampling, which has been shown to capture thermodynamics more computationally efficiently. [77] All bonds involving hydrogen atoms were constrained by M-SHAKE. Cut-off values for van der Waals and short-range electrostatic interactions were optimized by Anton guesser script. For long-range electrostatic interaction calculations, k-space Gaussian split Ewald method with a 64 x 64 x 64 Å mesh was used, and the long-range electrostatics were calculated every 6 fs. The integration time step was 2 fs for all simulations, and the r-RESPA integration method was employed. [78]

2.2.3 Amber16 simulations

For all systems that were simulated with Amber16 GPU, hydrogen mass repartitioning (HMR) scheme was applied, where all hydrogen masses were increased to 3, allowing simulation timestep to increase from 2 to 4 fs. The systems were prepared as described in System Preparation section using CHARMM-GUI, and *parmed.py* program was used to modify the system to accommodate the HMR scheme. For temperature control, Langevin thermostat was used with friction coefficient of 0.3 and T=303 K, and coordinates were printed out every 50 ps. van der Waals force switching function was used with switch on at 10 Å and cutoff at 12 Å. Berensden barostat was used with semi-isotropic at pressure 1 atm. The

Berensden coupling constant was set to 0.5 ps.

2.2.4 Markov state model (MSM) analysis

For all molecular dynamics (MD) simulation analysis, MDTraj 1.9.0 and MSMbuilder 3.8.0 were used [79, 80]. In the first steps of constructing a Markov State Model (MSM), proper set of reaction coordinates, and discretized states within the chosen reaction coordinates are needed. For these purposes, time-lagged independent component analysis (TICA) and K-Center clustering algorithm were used.

Time-lagged Independent Component Analysis (TICA)

TICA is a method that was first used in the signal processing community [81] and was later introduced to the molecular dynamics community by two independent research groups [82, 83]. TICA was introduced as a method to extract slow order parameters, and because most interesting protein motions are thought to be the slowest in timescaale, TICA is excellent for discovering important reaction coordinates.

TICA shares a lot of similarity to principal component analysis (PCA). In PCA, the projection vectors that maximize the explained variance is found by solving the generalized eigenproblem of the covariance matrix.

$$\Sigma\nu = \lambda\nu \tag{2.1}$$

where the covariance matrix is given by:

$$\Sigma_{ij} = \mathbb{E}\Big[X_i(t)X_j(t)\Big] \tag{2.2}$$

In biomolecular simulations, PCA can be useful when the system of interest needs a reaction coordinate that corresponds to the largest motions in the simulations. However,

often times, the largest motions in the biomolecular simulations do not necessarily correspond to the most interesting protein motions. TICA complements this problem by solving the generalized eigenproblem of the auto-correlation matrix of the dynamic system. For TICA, we solve the generalized eigenproblem defined as:

$$C^{(\Delta t)}\nu = \lambda\Sigma\nu \tag{2.3}$$

where

$$C_{ij}^{(\Delta t)} = \mathbb{E}\Big[X_i(t)X_j(t+\Delta t)\Big] \tag{2.4}$$

In practice, for TICA, a set of relevant features must be inputted to generate a set of low-dimensional linear combination of the input features. Common input features include pairwise Cα distances, inverse Cα distances, and dihedral angles. For our system, pairwise distances of all Cα atoms of all residues were used, which resulted in 4,831 total features. Then through TICA, dimensionality reduction was performed. Based on implied timescale analysis, 3 TICs with the slowest relaxation timescale was found to adequately describe our system (**Fig. 2.1**). Then, all simulations were transformed and projected onto the 3 TICs found.

K-Center clustering

To discretize a set of relevant conformations within the TIC space, K-Center clustering algorithm was used. K-Center clustering algorithm is an unsupervised machine learning method, where given k number of states, the algorithm learns to find k cluster centers that minimize the distances between the data point and its corresponding cluster center. Initially, the cluster centers are randomly chosen, and each data point is assigned to its closest cluster center based on distance. Then, new cluster centers are updated by finding data point closest

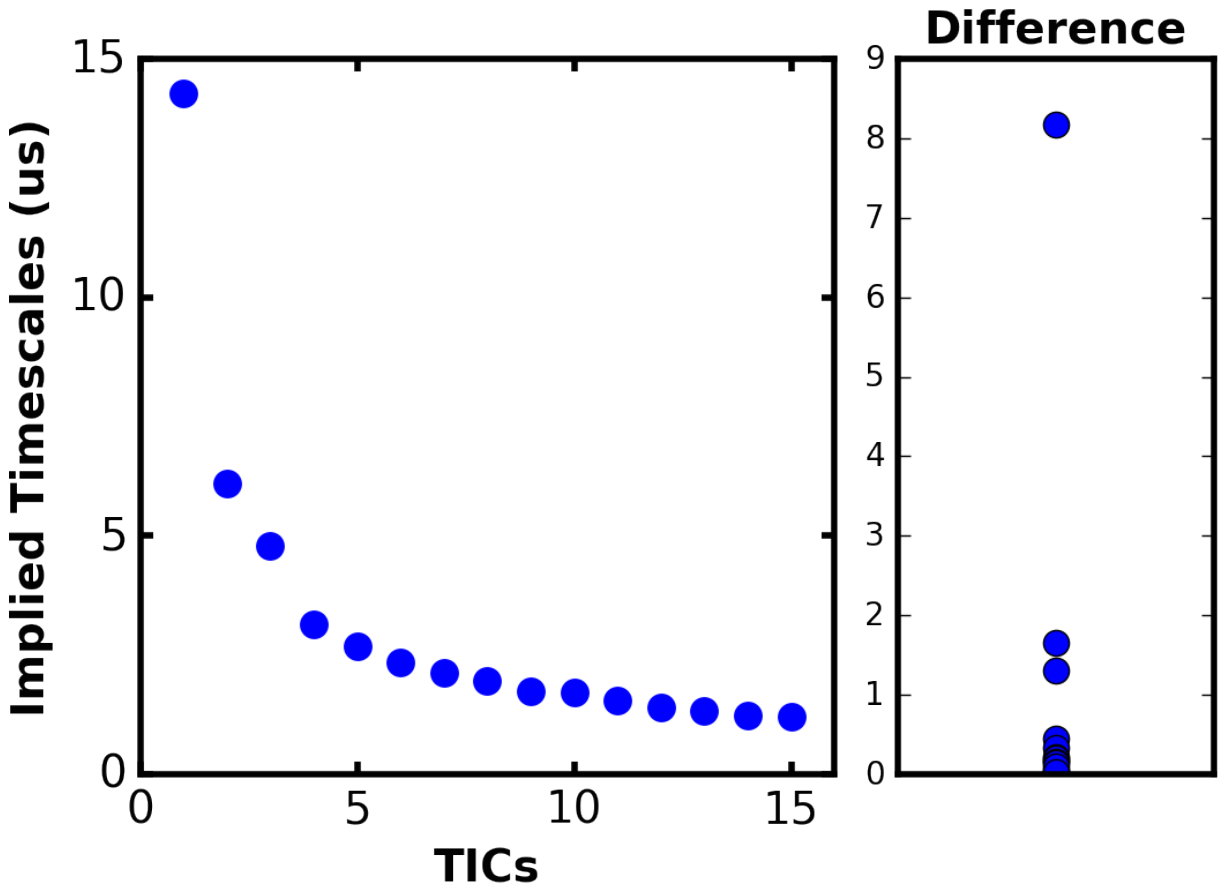

Figure 2.1: **TICA analysis of Kv1.2 simulations.** The implied relaxation timescale of each TIC is plotted versus the number of TICs. The first 3 TICs are chosen as reaction coordinates to represent the folding coordinates of the simulations.

to the average of each cluster. The process of assigning data to clusters and updating cluster centers are repeated until convergence.

In practice, finding the optimal number of states for K-Center clustering is empirically determined. The number of states is determined using the elbow method, where the momentum (sum of squared distances of samples to their closest centers) is plotted against the number of states. The number of states where the decrease in inertia became marginally less than a certain threshold is found. For Kv1.2, k=150 states was found to be appropriate for clustering (**Fig. 2.2**).

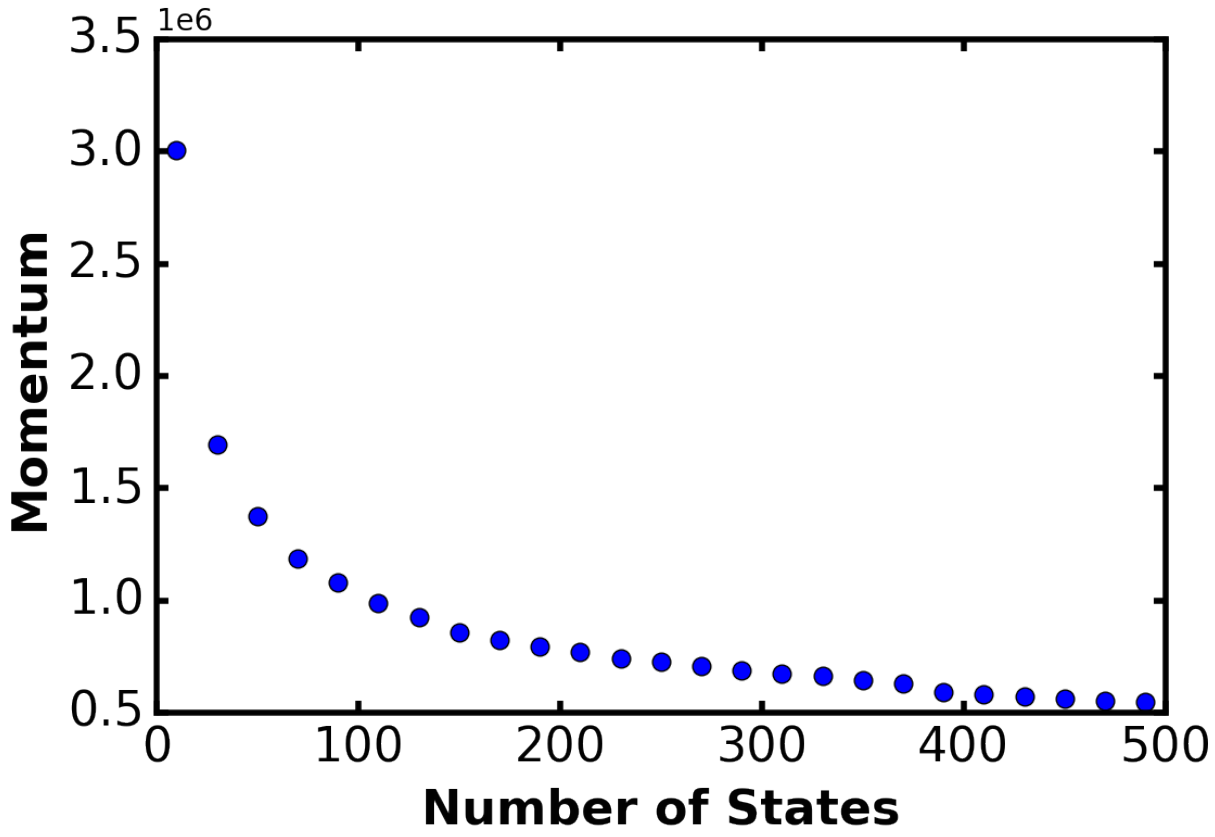

Figure 2.2: **Change in momentum over number of states.** Momentum is calculated as the sum of distances between each data point and its nearest centroid. To determine the optimal number of states for K-Center clustering method, the number of states where the momentum value begins to plateau is chosen. For our studies, k=150 states deemed to be optimal.

Markov State Model (MSM) analysis

Markov state models (MSMs) are a class of discrete, kinetic model that is largely based on Markov chains. MSMs assume that the system is memoryless and jump from one state to another purely based on its transition probabilities. In order to create a robust MSM, one must find a set of discrete conformational states and estimate the transition probabilities between each state. Previously through TICA and K-Center clustering algorithm, a set of microstates were found. Next, we must estimate the transition probability matrix. In theory, this process should just be a matter of counting the number of transitions. With infinite

data, one can covert the count matrix into transition matrix using maximum likelihood estimation:

$$T_{ij}(\tau) = \frac{C_{ij}}{\sum_k C_{ik}} \tag{2.5}$$

where τ is the lag time of the model, C_{ij} is the number of transitions between state i and state j given lag time of τ, and $\sum_k C_{ik}$ is the total number of outgoing transitions starting from state i.

Another key assumption in constructing a Markov state model is detailed balance. If the system is in equilibrium, then transitions from state i to j is as likely as transitions from state j to state i:

$$\pi_i T_{ij} = \pi_j T_{ji} \tag{2.6}$$

where π represents the population at a given state and T_{ij} represents the transition probability from state i to j. In practice, detailed balance is enforced by assuming that $C_{ij} = C_{ji}$ and the detailed balance is enforced by:

$$\hat{C}_{ij}(\tau) = \frac{C_{ij} + C_{ji}}{2} \tag{2.7}$$

After estimating the transition probability matrix, eigen decomposition of the transition probability matrix gives information about the transitions between microstates of the model and at what timescale these processes occur at. The eigenvalues are related to the implied timescale of a transition by the following equation:

$$k = -\frac{\tau}{ln(\lambda)} \tag{2.8}$$

One of the ways to test the validity of the model is using the Chapman-Kolmogorov equation:

$$T(n\tau) = T(\tau)^n \tag{2.9}$$

where n is the number of steps and τ is the lag time. Based on the Chapman-Kolmogorov equation, the implied timescale of a Markov model with lag time nτ should have the same implied time scale as Markov model with lag time of τ. Therefore, a common test for Markov model validity is by plotting the implied timescale as a function of lag time as shown in Figure 2.5A.

2.2.5 Adaptive sampling scheme

For adaptive sampling, initial MSM was constructed using only the Anton trajectory. Adaptive sampling scheme described previously by Bowman et. al was used [84]. In each round of simulations, 50 states with the lowest MSM populations are chosen and those chosen centroid conformations are used as seeds for additional 60 ns of AMBER16 simulations using the CHARMM36 force field with Hydrogen Mass Repartitioning (HMR) scheme. After each round of simulations, new sets of TICs are calculated and a new MSM is constructed. This method allows better sampling of states with low transition probabilities and help refine the total free energy landscape. In total, we accumulated 394 μs of simulations which are then used to create the final MSM for WT Kv1.2 monomer.

2.3 Results and Discussion

2.3.1 Potassium channel monomers are dynamical

In a previous 650 ns MD simulation, the wild-type (WT) Kv1.2 pore domain monomer in POPC lipid bilayer was stable in a native-like state with a Cα-RMSD below 3 Å. [57] However, this simulation is relatively short compared with the micro- to the millisecond timescale of membrane protein dynamics [21]. To further explore the monomers conforma-

tional flexibility, we carried out 16.2 μs of simulations at T=353 K using the specialized Anton computer designed for simulations [75]. The relatively high temperature was chosen to accelerate sampling while still reproducing the thermodynamics of membrane protein folding.

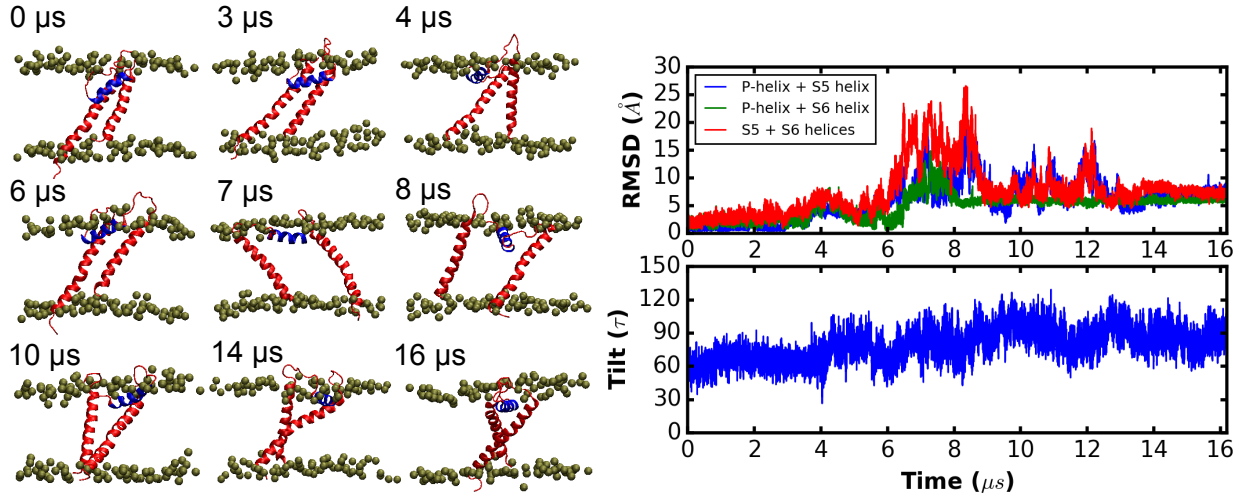

Figure 2.3: **Molecular dynamics simulation of Kv1.2 pore domain monomer in POPC lipid bilayer ran for 16.2 μs on Anton. Left**: Selected snapshots. **Right**: Pair-wise Cα-RMSD values for pairs of helices and the tilting angle of the pore helix.

During the first 3 μs of the simulation, the structure was stable with a Cα-RMSD to the initial (native-like) state below 4 Å (**Fig. 2.3**). Nevertheless, the p-helix became more parallel to the membrane during the first microsecond with the tilt angle increasing from 47° to 80° relative to the surface normal (**Fig. 2.3**). The average tilt angle over the entire 16.2 μs was 80° ± 10°. At 4 μs, the 2 transmembrane helices started to separate laterally with slightly different tilt angles due to their different lengths, leading to overall Cα-RMSD values greater than 4 Å.

Between 6 and 8 μs, all three helices separated (**Fig. 2.3**), but the p-helix remained helical and stayed nearly parallel to the membrane surface. This orientation allowed the polar and charged residues in the pore loop region to become more solvent exposed, while segregating the nonpolar residues. This result is qualitatively consistent with the previous

thiol-labeling results, which indicated that the p-helix remains helical and resides at the water-lipid interface [58]. Based on short simulations (650 ns) and thiol-labeling experimental data, the Kv1.3 monomer was inferred to retain native-like tertiary contacts [57]. In the long Anton simulation, however, the monomer is stable up to 3 μs with RMSD less than 4 Å but also explores many other non-native states.

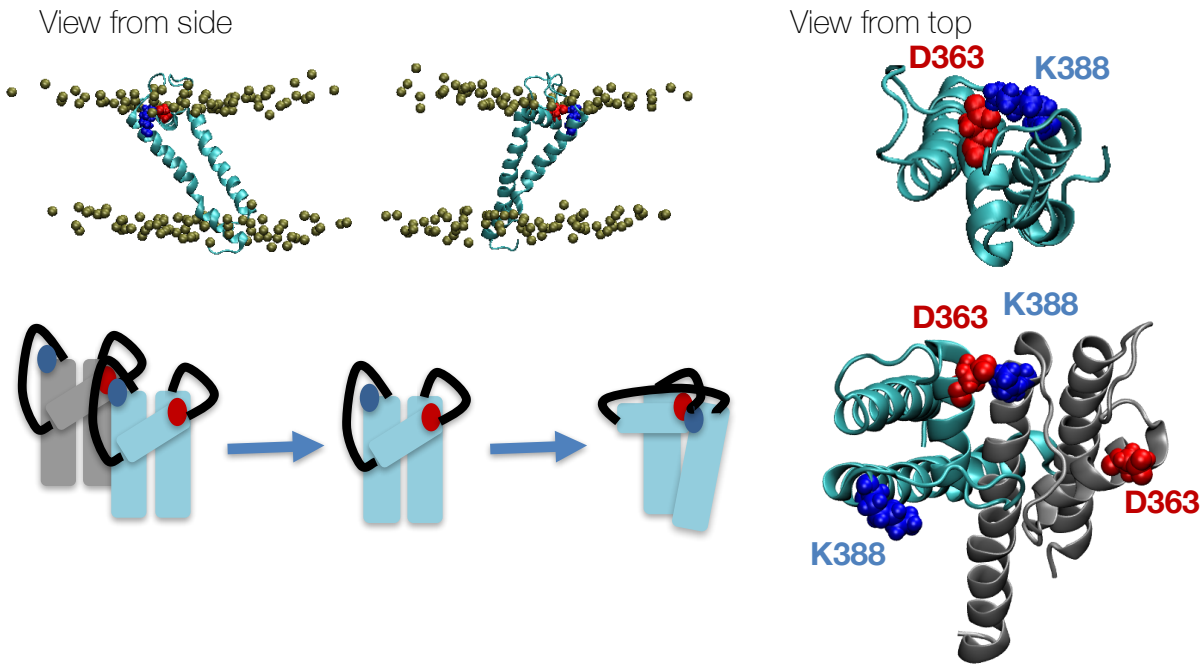

Figure 2.4: **Intra-monomer salt-bridge mimics a native inter-monomer salt-bridge. Top**: Snapshots from the simulation are shown from the side and from the top with D363 highlighted in red and K388 highlighted in blue. **Bottom**: left cartoon shows the rearrangement of the inter-monomer salt-bridge to an intra-monomer salt-bridge. Right figure shows the native inter-monomer salt-bridge interaction.

After 8 μs of simulations, residue D381 formed a salt bridge with a K398 that resides on top of the carboxy-terminal S6 helix (**Fig. 2.4**). This salt bridge persisted for the last 8 μs and stabilized the interactions between the S6 helix and the p-helix. The amino terminal S5 helix eventually drifted towards the complex formed by the S6 with the p-helix, to produce an alternate structure. Interestingly, the salt bridge formed between the p-helix and S6 helix mimics a salt-bridge that is formed between p-helix of one monomer with the S6 helix of an

adjacent monomer in the native tetramer (which does exist in the simulations). Salt bridges are known to be overly stabilized in simulations [85] so the longevity of the D381-K398 bridge may be unrealistic. Nevertheless, the monomer spent a majority of time with the three helices in a non-native arrangement.

2.3.2 Markov state modeling of Kv1.2 and KcsA monomer simulations

To obtain a better estimate of the conformation ensemble of Kv1.2 monomers, an analysis based on Markov state models (MSM) was performed using 3 sets of simulations (**Table. 2.1**): (1) 16.2 μs Anton simulation at T=353 K; (2) 10 independent 9 μs long simulations starting from the native state at T=303 K; (3) 100 rounds of adaptive sampling simulations (SI Methods) at T=303 K. A total of 394 μs of simulations was accumulated and used for time-structure Independent Component Analysis (TICA) and MSM analysis [81, 82, 83, 86, 87].

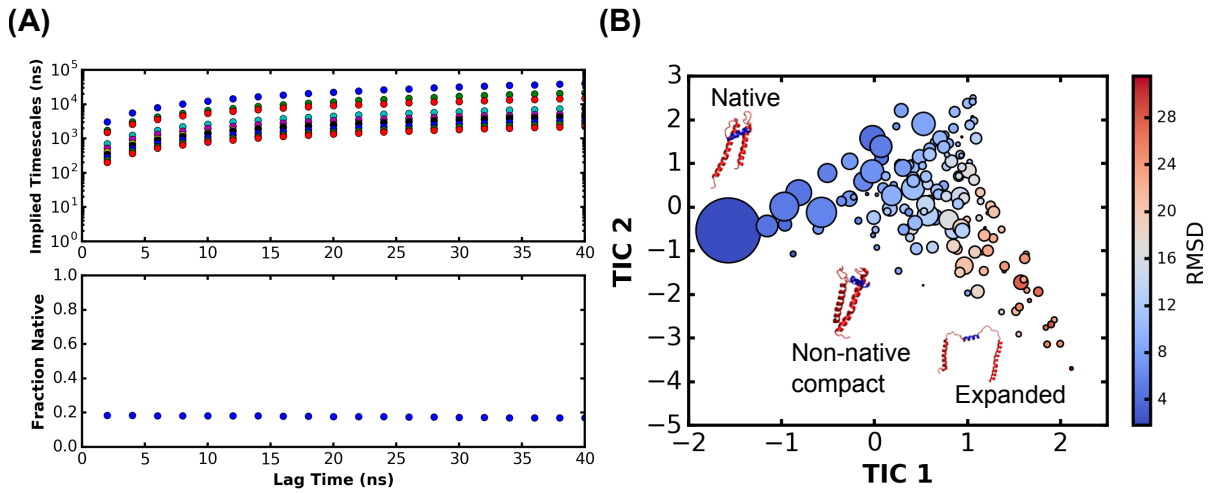

Figure 2.5: **MSM built on 394 μs of total simulation time at 303 K. (A) Top**: Implied timescale analysis of Markov State Model. **Bottom**: Lag-time analysis of fraction native indicates that 18% of the population remains native-like. **(B)** Markov state model built with lag time of 20 ns is projected onto TIC1 and TIC2. The size of the circle is proportional to the population of each microstate and the color of the circle represents the RMSD of each microstate to the native structure (in the tetramer).

Based on the MSM analysis, the population of native-like structures (Cα-RMSD < 4 Å) converged at 18% (**Fig. 2.5A Bottom**). This result further supports the finding that the native-like monomer conformation is present but is not predominant in lipid bilayer. While the two transmembrane helices and the p-helix retained their helicity in all structures, the monomer prefers to be partially disordered. The p-helix remained parallel to the water-lipid interface, which is consistent with the thiol-labeling results [57]. Overall, the monomer existed as a heterogeneous ensemble of contacting and non-contacting helices.

2.3.3 Kv1.2 and KcsA monomers show similar behavior

To compare with the Kv1.2 simulations, and to enable a direct comparison with our experiments, simulations were conducted on KcsA pore domain monomers without the C-terminal tetramerization domain (PDB ID: 1R3J, residues 22 – 124). The pore domain of KcsA and Kv1.2, comprising the pore loop and the 2 TMs, have 31% sequence identity. This suggests that the gross dynamical features of the two proteins should be qualitatively similar. Five 5 μs trajectories were initiated from the native state at T=353 K and ran for 6 μs each.

The number of contacts between 2 transmembrane helices and RMSD of all Cα atoms were calculated, where the number of contacts is calculated by counting the number of residues with heavy atoms that are within 5 Å of each other (**Fig. 2.6**). These plots show that KcsA displays qualitatively similar behavior to Kv1.2.

To compare with the Kv1.2 simulations, the KcsA trajectories were first projected onto the same set of TICs obtained from the Kv1.2 simulations and also aggregated with the Kv1.2 simulations to create a new set of common microstates. Although the sampling was less extensive as compared to Kv1.2, the general behavior of KcsA was similar. The KcsA monomer adopted a variety of native-like (44%) and non-native structures, albeit with the three helices separated less than Kv1.2 helices (**Fig. 2.6**). Transmembrane helices from 2 of the trajectories lose complete contact, which was seen from Kv1.2 simulations.

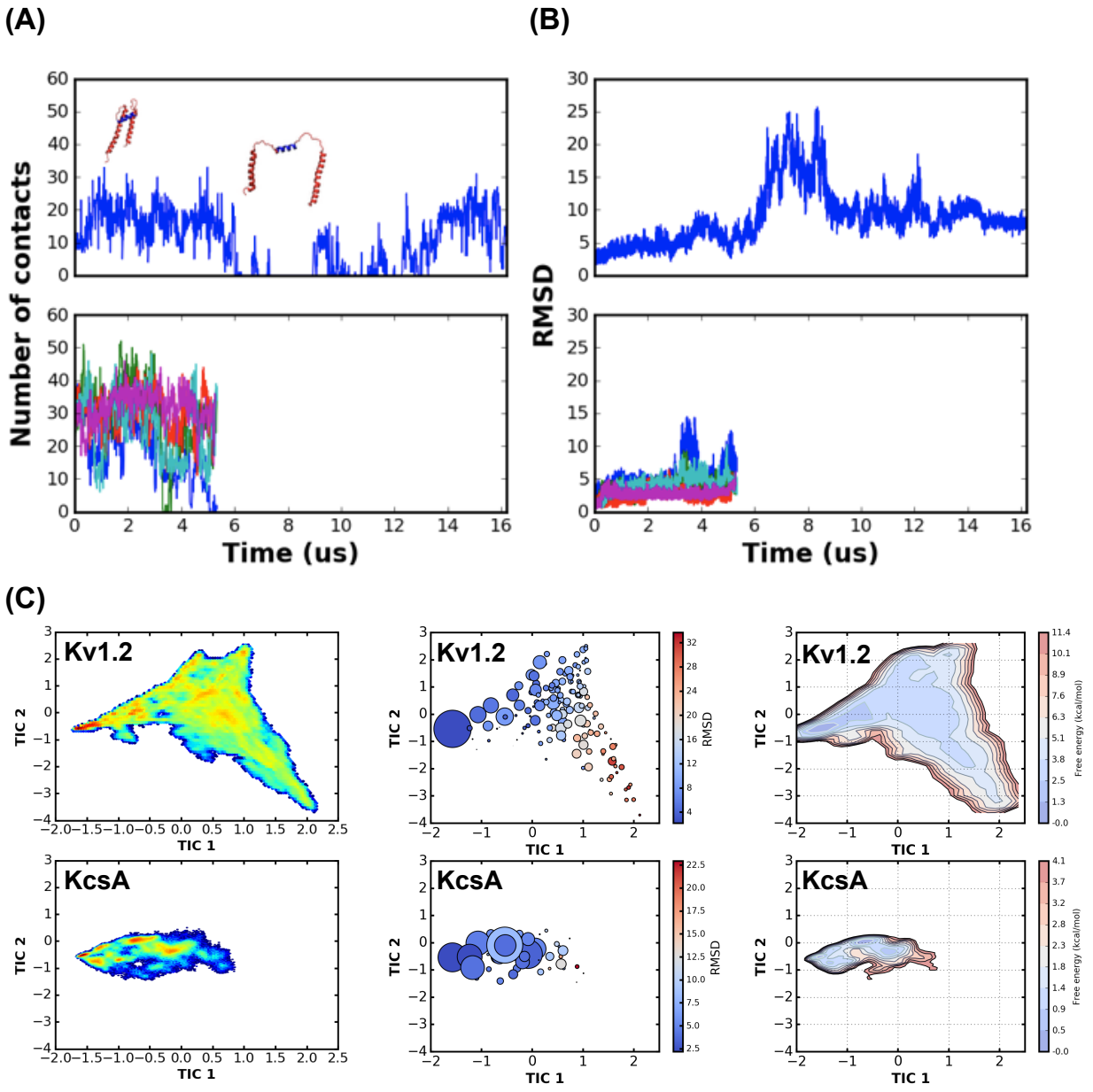

Figure 2.6: **Simulation comparison between Kv1.2 and KcsA.** **(A)** Number of contacts between the 2 transmembrane helices are plotted as a function of simulation time. **(B)** RMSD of the overall monomer is plotted as a function of time. **(C)** Kv1.2 and KcsA simulations are projected onto the same TIC space. Markov state model is made using the same set of microstates and the correponding free energy surface is plotted

In addition, while the KcsA simulations are highly undersampled compared to Kv1.2 simulations, the KcsA trajectories are projected onto the same TIC space as Kv1.2 simulations

to compare the MSMs between these two variants (**Fig. 2.6C**). For direct comparisons, microstates are created using K-Center clustering algorithm on the combined dataset of Kv1.2 and KcsA. This allows the two systems to share the same set of microstates and their resulting MSMs can be directly compared. MSMs were constructed separately for Kv1.2 and KcsA using the same set of microstates. Interestingly, even though KcsA simulations are highly undersampled, the simulations seem to suggest that KcsA follows a similar trajectory as Kv1.2 in that the protein moves away from its native basin and begins to explore other non-native states.

2.4 Conclusion

Through long-time unbiased molecular dynamics simulations and Markov state modeling, we found that the Kv1.2 and KcsA monomers both prefer to be partially disordered in lipid membranes. Like the thiol-labeling experiments [58], the simulations find that the potassium channel monomers retain its native-like secondary structures. However, the pore helix becomes more parallel to the membrane with a tilt angle near 80°. The pore helix acts as a tape connecting the two transmembrane helices together; however, with the change in pore-helix orientation, the transmembrane helices seem to lose contacts. This change allows the helices to dissociate and explore many other non-native states. One important thing to note is that even though the simulations show that the monomer explores many other non-native states, the monomer still has some propensity to form the native-like monomer structure. For Kv1.2 and KcsA, the population level calculated for native-like state was 18% and 44%, respectively. In the next chapter, we will discuss preparing KcsA monomers experimentally and how their dynamics compare to what we found through simulations and MSM analysis.

CHAPTER 3

BIOCHEMICAL PREPARATION OF KCSA MONOMERS AND DESIGN OF FAST FOLDING KCSA MUTANT

3.1 Introduction

Bacterial potassium channel, KcsA, is an extremely stable tetramer [59, 61, 62, 65, 88]. Conventional denaturing reagents such as urea and guanadine do not unfold KcsA [88]. Sodium dodecyl sulfate (SDS) detergent is a harsh denaturing reagent for most soluble and membrane proteins; however, even SDS alone does not unfold KcsA. Only with SDS and boiling of the protein sample at 90 °C for more than 10 minutes can disrupt the KcsA tetramer. So, the question is, how do we get stable monomers for biophysical studies?

In the first part of Chapter 3, a protocol for preparing KcsA monomer is presented, which are adopted from [89] and [67]. WT KcsA monomers were inserted into lipid nanodiscs as well as lipid bicelles for NMR studies. Through NMR and simulations, we find that the WT KcsA monomer is highly dynamic and exists in a structurally diverse ensemble of states. To stabilize the native-like state in monomers, we engineer a disulfide bond at the end of the two TM helices. This disulfide-bonded (CC) KcsA mutant monomer displays a much more native-like structure.

3.2 Methods

3.2.1 KcsA expression and purification

Expression

For KcsA expression with no isotope labeling, the protocol was adopted from [90]. XL10-GOLD *E. coli* competent cells were transformed with pQE70 KcsA WT with C-terminal

6xHis-tag using the heat shock method and grown overnight at 37 °C with 1% glucose and ampicillin (200 μg/mL). This overnight culture was used to inoculate LB media supplemented with 0.5% glycerol, 0.2% glucose and ampicillin (200 μg/mL) at a final concentration of 1% v/v. Once this culture reached $OD_{600} \sim 0.6$ and they were cooled down to 29 °C for 1 hour. Then, protein expression was induced with 0.1 mM Isopropyl-β-D-1-thiogalactopyranoside (IPTG), 10mM BaCl2 (K+ channel blocker) and 0.2 μg/mL ampicillin. Cells were incubated overnight at 29 °C and harvested after 14 – 18 hours of growth by centrifuging at 9,000 RCF for 15 minutes at 4 °C. The pellet was stored in -70 °C until purification.

For expressions with isotope labeling for NMR studies, the protocol was adopted from [91]. JM83 *E. coli* competent cells were transformed with pASK90 KcsA with N-terminal 6xHis-tag and grown overnight at 37 °C with ampicillin (200 μg/mL). The overnight culture was used to inoculate a 100 mL of LB media with ampicillin (200 μg/mL) and was grown until OD_{600} reached 0.5. This culture was then used to inoculate 4 L of LB media by adding 1% of the total volume. The 4 L of LB were grown until OD_{600} reached $\sim$ 0.9 and were spun down at 4,000 RCF for 10 minutes at 4 °C. The biomass was quadrupled by combining cell pellets from 4 L of LB media into 1 L of M9 media. The pellet was resuspended with 40 mL of M9 media and returned to 1 L of M9 media. The cells were allowed to acclimate for an hour at 37 °C and then were induced with 0.2mg/L anhydrotetracycline (aTC). Pellets were harvested 12 - 14 hours later by spinning down at 9,000 RCF for 15 minutes at 4 °C. After the pellet was spun down, the pellet was frozen in -70 °C until purification.

For the M9 media, the ingredients are shown in Tables 3.1, 3.2 and 3.3. For vitamin stock solution, one multi-purpose vitamin purchased from CVS is crushed using mortar and pestle. The crushed vitamin is then solubilized in deionized water in 20 mL volume. This solution is then sterile filtered and stored in -20 °C. When making the M9 media, Na_2HPO_4, KH_2PO_4, NaCl are first mixed together and autoclaved. Rest of the ingredients are added after autoclave.

Table 3.1: Ingredients for 1L of M9 media

Ingredients	Amount
Na_2HPO_4	12.8 g
KH_2PO4	3 g
NaCl	0.5 g
Solution C	20 mL
$MgSO_4$	0.1 g
$CaCl_2$	0.01 g
$FeCl_2$	0.01 g
Thiamine	0.01 g
Glucose	3 g
NH_4Cl	1 g
L-Proline	1 g
Vitamin Stock	2 mL
Ampicillin	0.2 g

Table 3.2: 1 L of Solution C
adjust pH to 6.7

Ingredients	Amount
KOH	7.3 g
Metal 44	50 mL
Nitrilotriacetic acid	10 g
$MgCl_2$ $6H_2O$	24 g
$CaCl_2$ $2H_2O$	3.335 g

Table 3.3: 100 mL of Metal 44 Solution, store in dark glass bottle at 4 °C

Ingredients	Amount
K_2EDTA $2H_2O$	0.327 g
$ZnCl_2$	0.522 g
$FeCl_2$ $4H_2O$	0.502 g
$MnCl_2$ $4H_2O$	0.18 g
$(NH_4)_6Mo_7O_{24}$ $6H_2O$	0.0185 g
$CuCl_2$ $2H_2O$	0.0156 g
$Co(NO_3)_2$ $6H_2O$	0.0248 g
Boric Acid	0.0114 g

Purification

For purification, the frozen pellets were resuspended in buffer A (50 mM HEPES, 200 mM KCl, pH 7.0) with 1 mM phenylmethylsulfonyl fluoride (PMSF), 3 mg of DNAse A (Goldbio) per liter of culture, and 0.4 mM $MgSO_4$. The cells were lysed through 3x passage through french press, and the lysed cells were spun down at 158,000 RCF for 30 minutes at 4 °C. The pellet was resuspended in buffer A with 10 mM n-dodecyl β-D-maltoside (DDM) and 1 mM PMSF. The mixture was rotated at room temperature for 1 hour to extract and solubilize KcsA in DDM. Then, the mixture was spun down at 185,000 RCF for 1 hour at 4

°C. Supernatant was incubated with Talon Metal Affinity Co2+ resins and rotated at room temperature for 1 hour. The resins were then collected by gravity and the flow through was discarded. The columns were washed with 15 bed volumes of Buffer A + 1 mM DDM + 10 mM imidazole and the proteins were eluted with Buffer A + 1 mM DDM + 500 mM imidazole.

For full-length KcsA constructs, protein in elution solution was further purified by size-exclusion chromatography on Superdex 200 Increase column that was pre-equilibrated with Buffer containing 50 mM NaPi, 100 mM NaCl, 1 mM DDM, pH 7. For KcsA Δ125 constructs, full length KcsA constructs were first trypsinized with α-chymotrypsin (Sigma Aldrich) at 4 °C overnight with 1:50 (chymotrypsin:KcsA) ratio. Then, the digested KcsA was further purified through size-exclusion chromatography (SEC) with Superdex 200 Increase column.

3.2.2 Design and production of fast folding KcsA mutant

Based on previous simulations of WT Kv1.2 and KcsA monomers as well as NMR results shown in *Results* section of this chapter, potassium channel pore domain monomers seemed to be highly dynamic and structurally heterogeneous. In order to limit the dynamical nature of WT KcsA monomer, we introduce a disulfide bridge at the end of the two TM helices to limit its dynamical nature. The residues chosen for mutations are A29 and A109, and both were mutated to cysteines using the QuickChange protocol. For expression and purification, the protocol is identical to expression and purification of WT KcsA discussed in the previous section.

3.2.3 KcsA monomer preparation

For preparing monomer samples, we follow the protocol highlighted in Figure 3.1. Purified tetrameric KcsA after SEC step are all pooled together, and they are precipitated by adding 15% v/v tricholoroacetic acid (TCA) to the samples. The precipitated samples were frozen

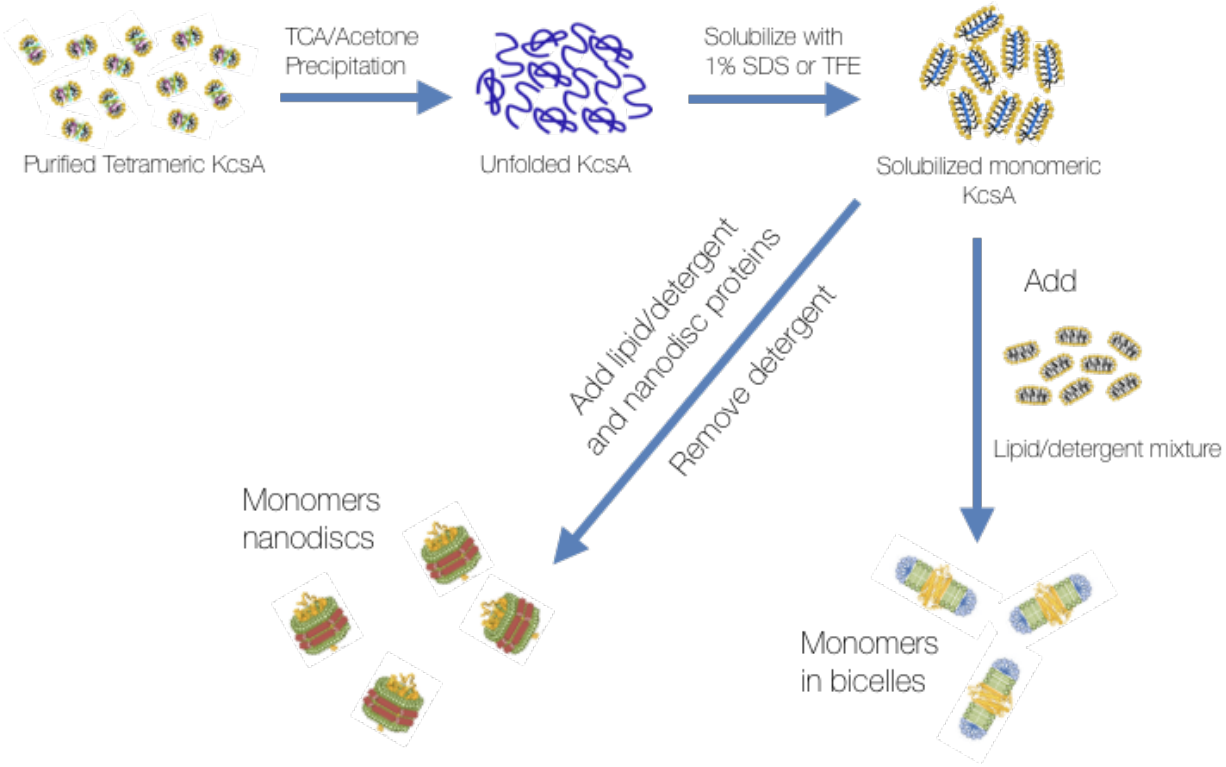

Figure 3.1: **Protocol for preparing monomeric KcsA in nanodisc or bicelle.** Starting from purified tetrameric KcsA, the protein is first precipitated using TCA/Acetone precipitation method. Then, the precipitated pellet is resolubilized using either 1% SDS buffer or TFE to obtain monomeric KcsA. These monomeric KcsA can be kinetically trapped in either nanodiscs or bicelles for biochemical studies of the monomers.

at -70 °C for 30 minutes. Then, the frozen samples were spun down at 16,200 RCF for 30 minutes at 4 °C. Supernatant was removed, and washed with chilled (-20 °C) acetone. Precipitates were resuspended, vortexed for 30 seconds, and incubated at -20 °C for 30 minutes until spin down at 16,200 RCF for 15 minutes at 4 °C. This acetone wash step was repeated 3 times. The final precipitate was dried under air until all acetone evaporated, and the precipitate was stored in -80 °C until usage. For solubilizing the precipitated monomers, either 1% SDS or 100% TFE is used.

KcsA monomer in nanodisc

The protocol for kinetically trapping KcsA monomer in nanodisc was adopted from [92] and [93]. First, 0.5% SDS-solubilized KcsA monomers is mixed with 1,2-dimyristoyl-sn-glycero-3-phosphoglycerol (DMPG) and sodium cholate. Typically, DMPG is mixed with cholate at ratio of 1:2. Optimal protein to lipid ratio was determined empirically at 1:5 by checking the monodispersity of size-exlusion chromatography (SEC) elution profile. After an hour of mixing at room temperature, 1 g of Biobead SM-2 is added per mL of nanodisc mixture. Then, the mixtures with Biobeads is further mixed for overnight at room temperature. Biobeads were removed by centrifugation and nanodiscs were further purified by SEC using Superdex 200 Increase column.

KcsA monomer in bicelle

KcsA monomers in bicelles were prepared in q=0.3 DMPC:DHPC bicelles using following procedure. Bicelle mixture was first prepared by adding 1:3 molar ratio of DMPC or POPC to DHPC (both solubilized in chloroform) at final lipid concentration of 10% w/v in the NMR sample. Chloroform was evaporated under stream of nitrogen and further removed under vacuum for 3 - 4 hours. This procedure resulted in a thin lipid film, which was resolubilized in 2,2,2-trifluoroethanol (TFE). WT KcsA Δ125 precipitate is solubilized in TFE and mixed into the lipid TFE mixture. TFE was evaporated under a stream of nitrogen until a thin film formed around the glass tube and was further evaporated by placing the tube under vacuum overnight. Then, the thin lipid film was rehydrated with buffer of choice and vortexed for 30 seconds. Typically for NMR, the buffer contained 20 mM NaPi, 50 mM NaCl, 0.03% NaN_3, and pH 6.5.

3.2.4 Nuclear magnetic resonance (NMR) measurements of KcsA

NMR spectra were acquired on Bruker AVANCE IIIHD 600 MHz NMR spectrometer equipped with a room temperature TXI probe. [^{15}N, ^{1}H]-TROSY-HSQC experiments were run at T=323 K. To obtain rotational correlation times, τ_c, [^{15}N, ^{1}H]-TRACT experiments were performed [94, 95, 96, 97, 98].

3.2.5 KcsA CC simulations and MSM analysis

The starting structure for KcsA monomer simulations was taken from the tetrameric KcsA X-ray crystal structure (PDB ID: 1R3J). Mutations were made at A29C and A109C for the disulfide mutant simulations, and disulfide bond was created using CHARMM-GUI's PDBReader module. All systems were prepared by using CHARMM-GUI's Membrane Builder module (www.charmm-gui.org) [70, 71, 72, 73, 74]. Each system contained 70 1,2-dimyristoyl-sn-glycero-3-phosphatidylcholine (DMPC) molecules per leaflet totaling up to 140 DMPC molecules to match the NMR sample conditions. All systems were hydrated by creating a water box 17.5 Å above and below the proteins maximum and minimum Z-positions. In addition, the system was neutralized with 150 mM KCl. The system comprised a total of approximately 45,000 atoms. 5 independent simulations of KcsA WT monomer and 5 simulations of KcsA CC mutant monomer were carried out at T = 353 K to enhance sampling. All simulations were run with the parameters described in Methods section with AMBER16 GPU, and each simulation was run up to 6 μs. In total 30 μs of KcsA WT simulations and 30 μs of KcsA CC simulations were accumulated.

For MSM analysis, all KcsA simulations were first projected onto the TIC space created by Kv1.2 simulations. KcsA monomer has 103 residues in total whereas Kv1.2 pore domain has 99 residues. In order to project KcsA monomer simulations, Cα atoms were aligned to Kv1.2 structure. The final corresponding residues in KcsA were from residues 24 to 122, which were used for TICA projection. After projecting KcsA simulations onto the same

TIC space as Kv1.2, all simulations were combined in order to create the same number of microstates using K-Center clustering algorithm. After the same microstates were created, Markov state model constructions were done as described in *Chapter 2 Methods*. However, MSM was constructed separately for Kv1.2, KcsA WT and KcsA CC using the same set of microstates.

3.3 Results and Discussion

3.3.1 NMR also suggest WT KcsA monomers are structurally diverse

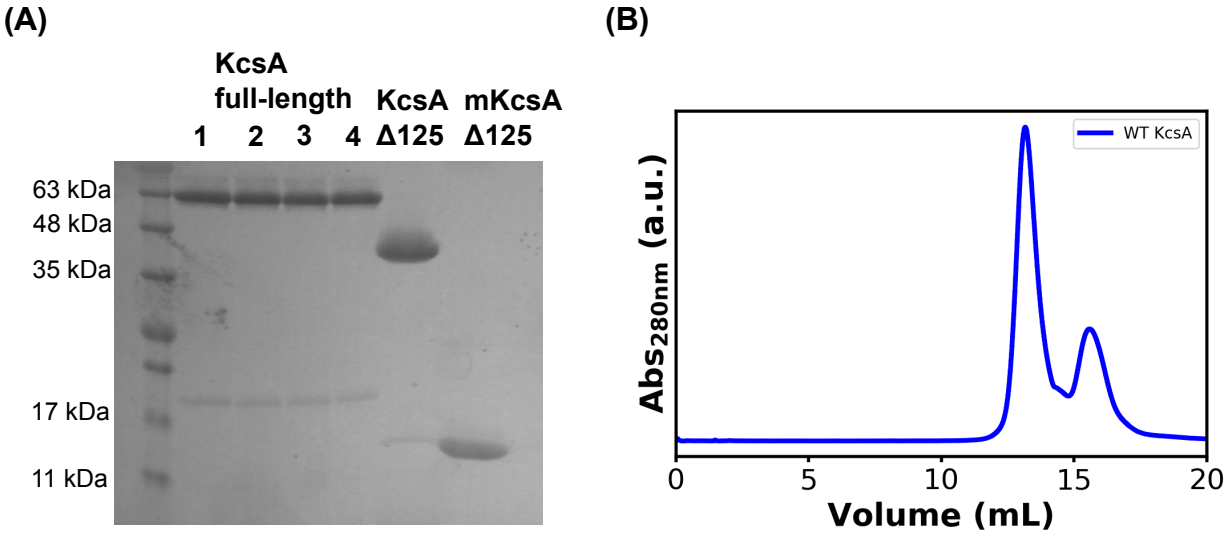

Figure 3.2: **Purification of KcsA (A)** SDS-PAGE gel with KcsA full-length (FL), KcsA Δ125, monomeric KcsA (mKcsA) Δ125. **(B)** Elution profile of WT KcsA Δ tetramer with chymotrypsin. KcsA Δ125 elutes around 13 mL and chymotrypsin elutes near 16 mL.

Expression and purification of KcsA tetramers is straightforward. Many labs can reliably produce large amounts of KcsA for biophysical studies [20, 59, 61, 90, 91, 99, 64]. KcsA Δ125 constructs are prepared by incubating KcsA full-length (FL) constructs with α-chymotrypsin at 1:50 ratio overnight at 4 °C. Then, the KcsA Δ125 are purified from chymotrypsin by size-exclusion chromatography (SEC) with Superdex 200 Increase column as shown in Figure

3.2B.

KcsA monomers are prepared by trichloroacetic acid (TCA)/Acetone precipitation protocol as discussed in *Methods* section. The precipitate monomers can be solubilized in 0.5% SDS and is shown to run as 13 kDa on SDS-PAGE (**Fig. 3.2**). The monomers were then transferred to nanodisc or bicelle for NMR studies.

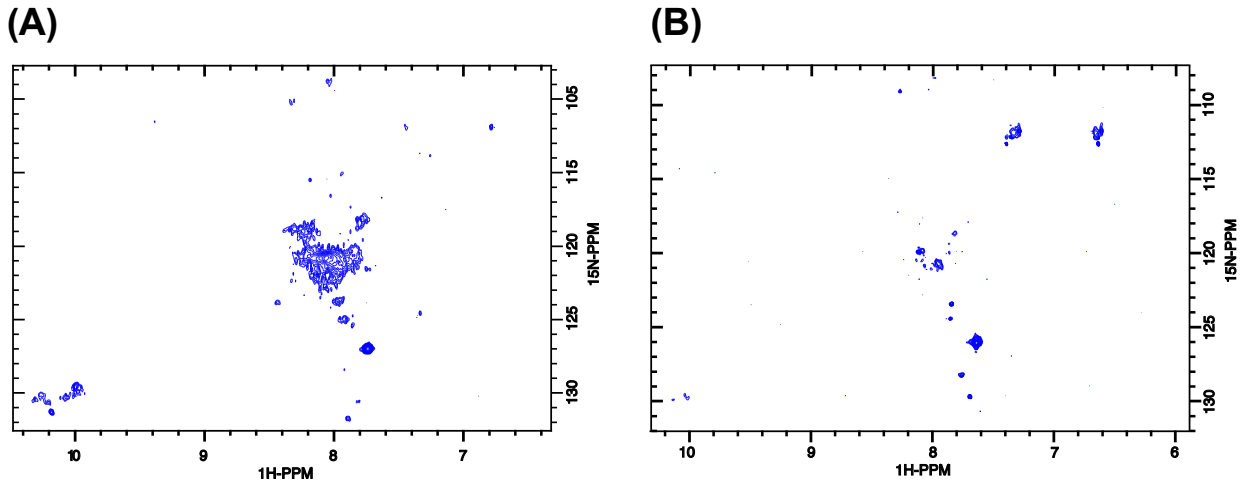

Figure 3.3: **[^{15}N-^{1}H]-TROSY-HSQC spectra of KcsA WT. (A)** HSQC of KcsA Δ125 solubilized in MSP1D1 ΔH5 nanodisc is shown. **(B)** HSQC of KcsA Δ125 solubilized in q=0.3 DMPC:DHPC bicelle is shown.

NMR spectra of WT KcsA in either MSP1D1 ΔH5 or q=0.3 DMPC:DHPC bicelle do not show well-resolved peaks in general. In fact, for WT KcsA solubilized in bicelle easily forms soluble aggregates in the NMR sample conditions. In Chapter 2, we found that Kv1.2 and KcsA monomers seem to be structurally diverse through MD simulations and Markov state model analysis. Given our NMR spectra and the computational modelling results, we suspect that the WT KcsA monomers are structurally heterogeneous in lipid bilayers.

3.3.2 Designing more native-like KcsA mutant

Our previous MD simulations and NMR results indicate that the KcsA monomers are partially disordered, existing in an ensemble of diverse structures. In order to stabilize the

native-like orientation of the 2 transmembrane helices, we introduce a disulfide bond at the end of the 2 transmembrane helices. The residues chosen for mutations are alanine 29 and alanine 109 as shown in Figure 3.4. Expression and purification protocols for CC KcsA is the same as the WT KcsA.

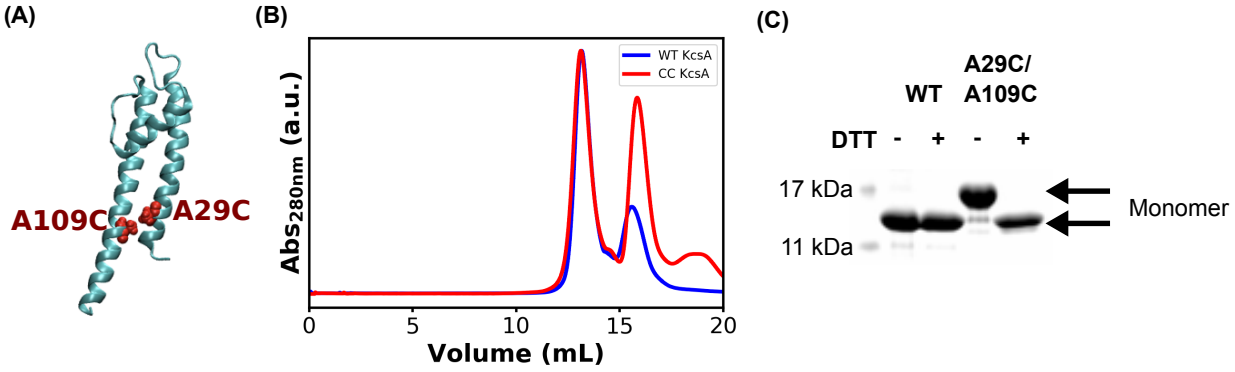

Figure 3.4: **Biochemical characterization of disulfide-bonded (CC) KcsA mutant.** **(A)** Mutations A29C and A109C are highlighted in red on KcsA monomer. **(B)** KcsA CC tetramer elutes identical to WT KcsA tetramers. The larger second peak corresponds to higher presence of chymotrypsin. **(C)** SDS-PAGE gel with and without reducing agent, DTT for WT and the disulfide-bonded (CC) KcsA monomers.

Generally, yield of CC KcsA is lower than WT KcsA. The CC variant assembles the same as WT KcsA and elutes the same as WT KcsA on Superdex 200 Increase column (**Fig. 3.4B**). The disulfide bond forms naturally, and the formation of disulfide bond is confirmed by running an SDS-PAGE gel with and without dithiothreitol (DTT) of the monomers (**Fig. 3.4C**). With the reducing agent, the CC mutant runs at the same molecular weight as the WT KcsA monomers; however, without the reducing agent, the CC monomer runs at a higher molecular weight.

Compared to the WT KcsA, the CC variant seems to be more folded (**Fig. 3.5**). The NMR spectrum of CC KcsA in DMPC:DHPC bicelles show well-dispersed and well-resolved peaks compared to the WT KcsA HSQC spectrum. This indicates that the CC variant is more folded and the monomer is more structurally homogeneous. This is another evidence that the WT KcsA monomer exists in a heterogeneous ensemble of structures. By locking

the 2 transmembrane helices with a disulfide bond, the monomer itself becomes more limited in conformational search space. An obvious next step with this construct will be to determine the structure using NMR. With the structure and peak assignments, the dynamics of monomer in bicelles can be studied as well. The dynamics of the pore loop region will be particularly interesting.

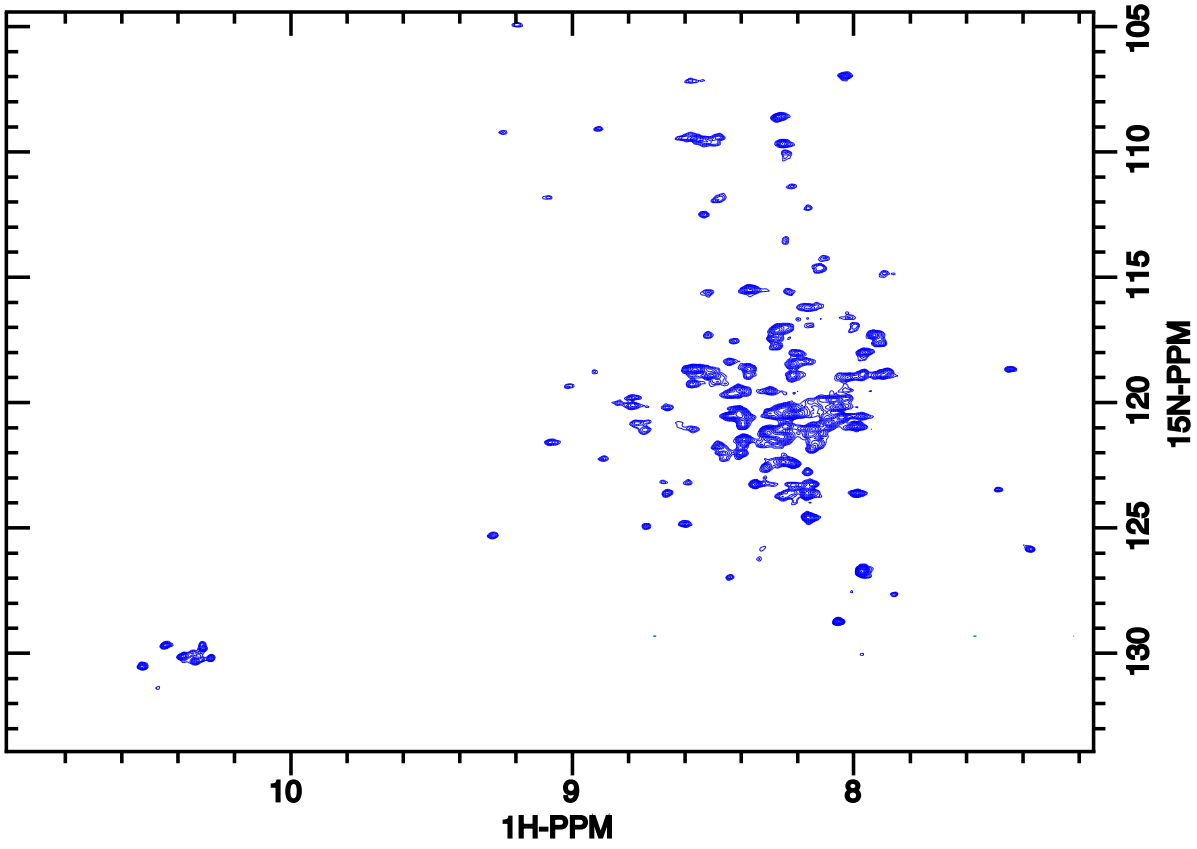

Figure 3.5: **[^{15}N-^{1}H]-TROSY-HSQC spectrum of CC KcsA.** Compared to Figure 3.3, this HSQC spectrum of CC KcsA displays well-dispersed and well-resolved peaks, indicating that the CC KcsA variant retains a more native-like structure in bicelles.

3.3.3 Simulations of CC KcsA

To compare with experiments and simulations of Kv1.2 and WT KcsA monomers, 5 independent simulations of CC KcsA were carried in DMPC bilayers (**Fig. 3.6**). Compared

to Kv1.2 and WT KcsA monomer simulations, the number of contacts in CC KcsA never reaches 0. This is expected as the disulfide bond at the bottom of the 2 transmembrane helices will keep them close to each other. In addition, the RMSD of KcsA CC monomer does not reach above 6 Åunlike Kv1.2 and WT KcsA monomers. Most of the RMSD resulting in CC KcsA simulations come from changes near the pore loop region, where the p-helix reorients itself to be more parallel to the water-lipid interface.

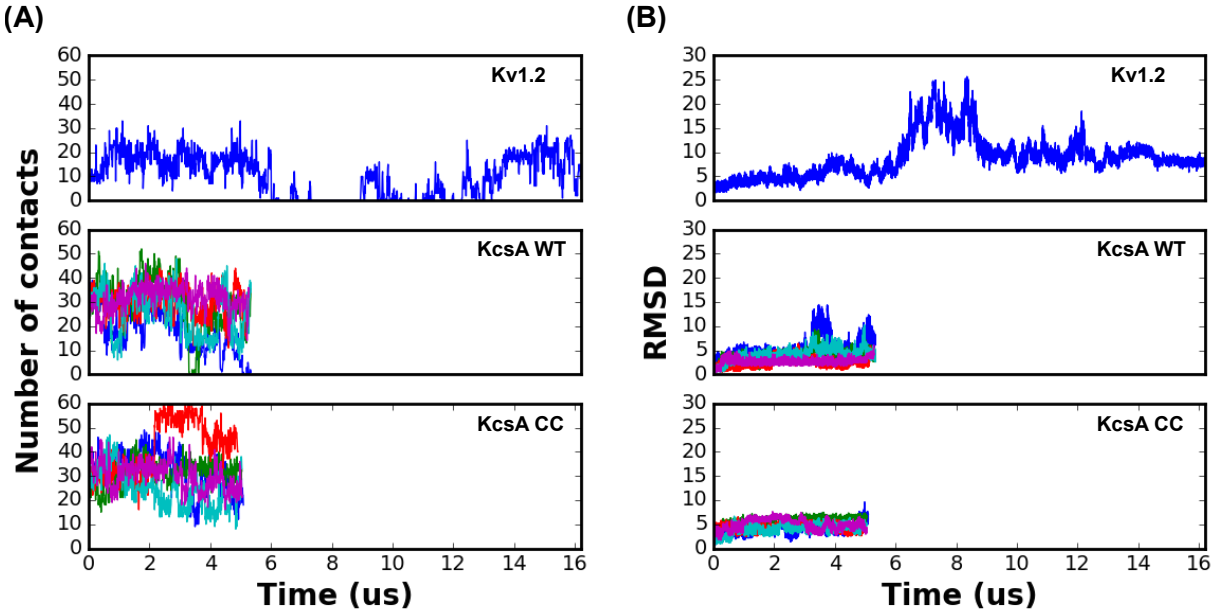

Figure 3.6: **Number of contacts and RMSD of Kv1.2, WT KcsA and CC KcsA plotted over time.** **(A)** Number of contacts is calculated by finding the number of residues with heavy atoms within 4 Å between the two transmembrane helices. The number of contacts is plotted over time. **(B)** The overall RMSD referenced to the initial frame is plotted over time.

When the pore helix snorkels up to the water-lipid interface, the 2 transmembrane helices can undergo a scissoring motion, which also contributes to changes in RMSD; however, the overall RMSD of CC KcsA monomer never reaches beyond 10 Å because the 2 transmembrane helices are locked together. A direct comparison of the simulation results with NMR data in the future will be necessary and helpful for understanding of the dynamics in the pore loop region.

3.4 Conclusion

The protocol outlined here shows how to prepare KcsA monomers in nanodisc, bicelle or detergent for biophysical studies. With our understanding of the dynamics of WT KcsA monomer, we introduced a more native-like disulfide-bonded mutant named CC KcsA. This mutant displays an NMR spectrum with well-dispersed peaks, suggesting that it is more folded than the WT KcsA monomers. This mutant can be purified in the same way as WT KcsA and the disulfide bond forms spontaneously.

In addition, the simulations of CC KcsA monomers show that the disulfide bond at the bottom of the 2 transmembrane helices keep the native-like state more stable. The pore loop region still changes its conformation, leading to scissoring motion of the 2 transmembrane helices. Yet, the overall RMSD never exceeds 6 Å and the number of contacts between the 2 transmembrane helices never reaches 0.

In the future, determining the structure of CC KcsA monomer in bicelles will be critical to understanding the dynamics of the pore loop region. With both experimental and computational modelling of CC KcsA monomer, we will be able to understand the dynamics of the pore-loop region better.

CHAPTER 4

FOLDING KINETICS OF POTASSIUM CHANNEL PORE DOMAIN

4.1 Introduction

In Chapter 3, we developed a KcsA monomer mutant that retains its native-like structure more than the wild-type (WT) does by engineering a disulfide bridge near the ends of the 2 transmembrane helices. Here in this chapter, we compare the folding kinetics of WT and the more native-like disulfide-bonded (CC) KcsA mutant with and without the C-terminal "tetramerization" domain.

The refolding studies reveal several interesting aspects about potassium channel folding. First, the WT KcsA folding displays a biphasic kientic behavior with fast and slow processes with $\tau_f \sim 50$ and 1400 seconds, respectively. However, in the more native-like CC mutant, the fast process dominates, suggesting that locking the 2 transmembrane helices with a disulfide bond minimizes its proclivity to misfold. Secondly, WT and CC without the C-terminal "tetramerization" domain does not have any concentration dependence despite the fact that the native state is a tetramer, implying that the rate-limiting step is unimolecular. Thirdly, both WT and CC with the C-terminal "tetramerization" domain fold with concentration dependence in the same concentration as without the C-terminal domain. Lastly, through ensemble FRET measurements, we propose that KcsA monomers form a dense protein-rich phase in the membrane first and fold into native structures within this phase.

4.2 Methods

4.2.1 Folding KcsA in liposomes

Liposome solution preparation

Desired amount of SoyPC lipids solubilized in chloroform at 25 mg/mL (Avanti) was pipetted into a glass tube and was dried under a stream of nitrogen gas to form a thin lipid film. Then, the thin lipid film in glass tube was further dried under vacuum overnight in a lyophilizer. Next day, the dried lipid film was brought up to 10 mg/mL concentration by rehydrating with a buffer of choice (50 mM NaPi, 100 mM NaCl, pH 6.5). This mixture was first rotated for 30 minutes at room temperature. Then, the mixture was sonicated in a bath sonicator for 30 minutes to form small unilamellar vesicles (SUVs).

Refolding KcsA in liposomes

Refolding experiments were adapted from [100]. The overall refolding protocol with liposome is summarized in Figure 4.1. Prior to refolding, precipitated KcsA Δ125 monomers were solubilized in 0.5% SDS in PBS buffer (50 mM NaPi, 100 mM NaCl, pH 6.5) with roughly protein concentrations of $\sim$1.8 mg/mL measured with NanoDrop. The buffer cannot contain any K^+ ions as they preciptate SDS, which can interfere with the refolding reaction.

To initiate refolding, the protein mixture was diluted 10-fold into the refolding buffer containing SoyPC liposomes prepared as discussed above. The mixture was vortexed for 10 seconds. Then, at each time point a small aliquot of the refolding mixture was taken out and the reaction was quenched by diluting into 10% SDS in PBS buffer (50 mM NaPi, 100 mM NaCl, pH 6.5) at 1:2 ratio. The final concentration of SDS was approximately 6% or 220 mM. SDS concentration at this point should be high enough to disrupt any liposomes and break any weakly associated KcsA monomers. These quenched reaction mixtures were then run on a Novex Tris-glycine 4 - 20% SDS-PAGE gel (ThermoFisher).

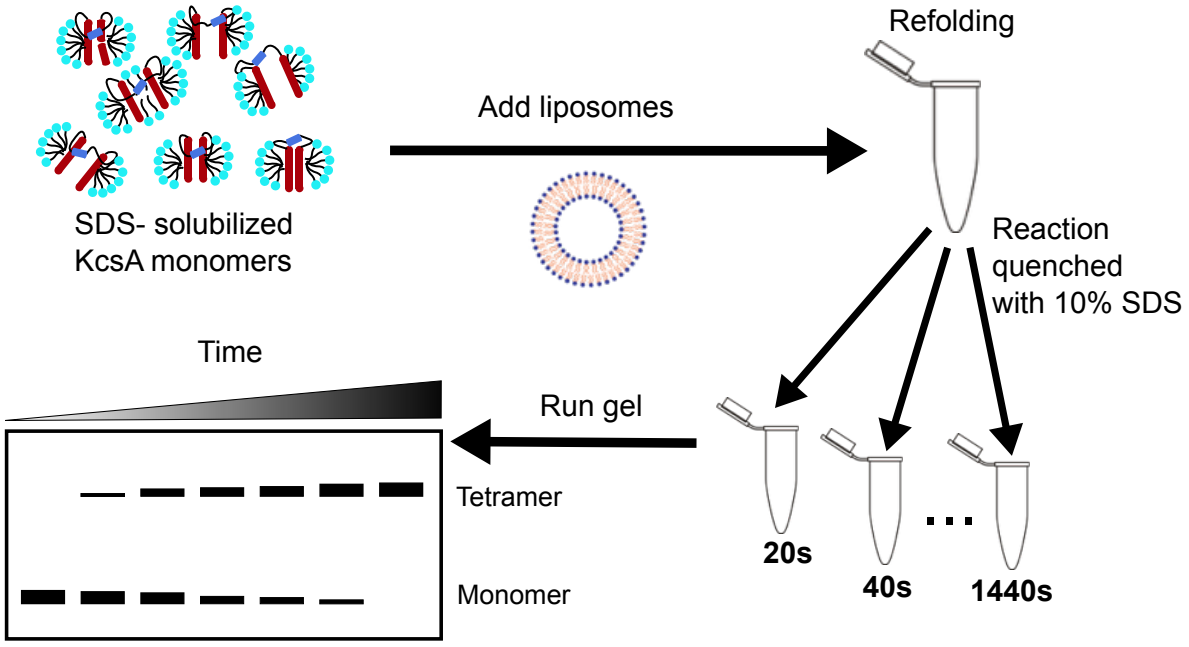

Figure 4.1: **Protocol for studying the folding kinetics of KcsA.** SDS-solubilizied KcsA monomers are diluted into solutions containing liposomes to initiate tetramerization. At different time points, an aliquot from reaction mixture is taken out and quenched with 10% SDS. Then the quenched aliquots are run on gel to monitor the formation of tetramers over time.

4.2.2 Kinetics analysis

For analysis, gel images were taken with a Bio-rad ChemiDoc instrument. Then, the image was processed using ImageLab software to enhance the contrast between the protein bands and the background. For quantifying tetramer to monomer ratio, ImageJ software's gel analysis tool was used. Each lane was normalized to itself by calculating the tetramer fraction in each lane. First, in each lane, band intensities that correspond to tetramers and monomers were integrated and fraction of tetramers were calculated by:

$$\frac{[Tetramer]}{[Tetramer] + [Monomer]} \tag{4.1}$$

For fitting, a double exponential function was used:

$$a - be^{-\frac{t}{c}} - de^{-\frac{t}{e}} \tag{4.2}$$

where the time constants c and e were fixed globally across all kinetic traces and the amplitudes a, b and d were allowed to vary for each kinetic trace. Our main assumptions with this analysis were that both the WT and CC mutants have two competing kinetic phases, and the rate constants are the same between WT and CC variant. All data were fitted using *Scipy.optimize* package with Python3.6.

4.2.3 Förster resonance energy transfer (FRET) of KcsA in liposomes

Biochemical preparation of FRET samples

For the site of dye conjugation, Leucine 86 in KcsA was chosen (**Fig. 4.2**). L86 was mutated to a cysteine using the QuickChange protocol on either WT KcsA or the CC KcsA variant. L86C KcsA or CC L86C KcsA were expressed and purified as tetramers as discussed Chapter 2. After the purification of either full-length or Δ125 constructs using size-exclusion chromatography, all tetramer fractions were pooled and dyes (Cy3 or Cy5) were conjugated to these mutants by overnight incubation at 4 °C. Next day, excess dyes were removed by size-exclusion chromatography using Superdex 200 Increase column (**Fig. 4.3**). For Cy3 conjugated KcsA systems, monitoring absorbance at 280 nm and 550 nm allows you to monitor the elution of KcsA (280 nm) and Cy3 (550 nm) during SEC (**Fig. 4.3A**). For Cy5, elution of protein and dye together can be monitored at 280 nm and 650 nm, respectively (**4.3B**).

With the sample collected from size-exclusion chromatography, conjugation efficiency was calculated by measuring the absorbance spectra with NanoDrop and using the following equation for calculation:

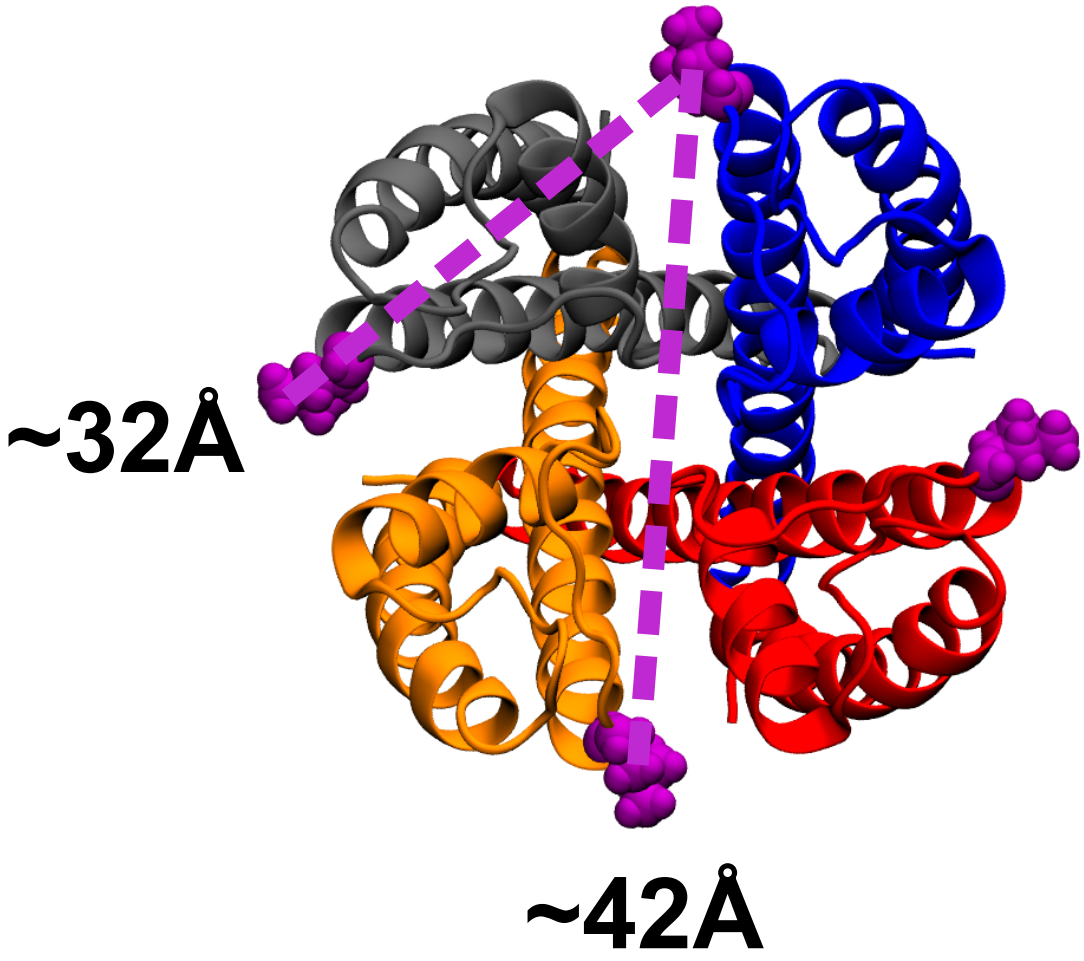

Figure 4.2: **Position of dyes in KcsA (PDB ID: 1R3J) shown from top**. L86 is colored in purple and each monomeric subunit is colored in grey, blue, red and orange. There are 2 different possible distances for FRET interaction, which are 32 Åand 42 Å. Both distances are well within R_0 value of Cy3-Cy5 FRET distance, which is 56 Å.

$$\text{Conjugation Efficiency} = \frac{\frac{\alpha_{dye}}{\epsilon_{dye}}}{\frac{\alpha_{280nm}-(CF*\alpha_{dye})}{\epsilon_{prot}} + \frac{\alpha_{dye}}{\epsilon_{dye}}} \tag{4.3}$$

where ϵ_{prot} for KcsA is 33,460 cm^{-1} M^{-1}, ϵ_{dye} = 150,000 cm^{-1} M^{-1} for Cy3 and 250,000 cm^{-1} M^{-1} Cy5, respectively, CF is the correction factor for dye absorbance at 280 nm, which is 0.08 and 0.05 for Cy3 and Cy5, respectively, and absorbances (α_{280nm}, α_{dye})

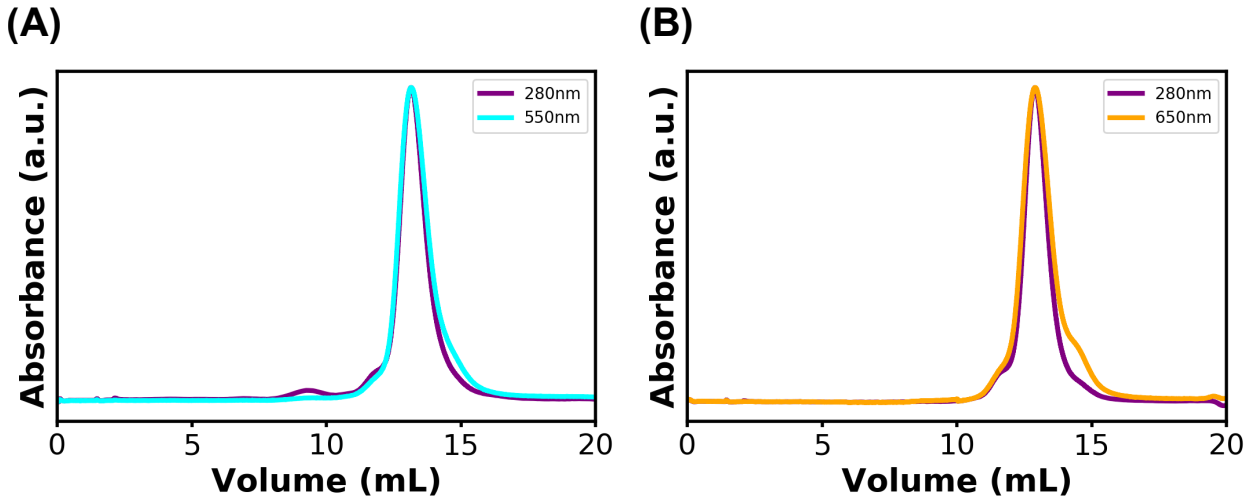

Figure 4.3: **Size-exclusion chromatography elution profile of KcsA L86C with Cy3 and Cy5. (A)** KcsA Δ125 L86C conjugated with Cy3. Absorbances are monitored at 280 nm (purple) for protein and 550 nm (cyan) for Cy3. The elution profile for both overlap well suggesting successful conjugation reaction. **(B)** KcsA Δ125 L86C is conjugated with Cy5. Absorbances are monitored at 280 nm (purple) for protein and 650 nm (orange) for Cy5.

were measured at 280 nm for KcsA, 550 nm for Cy3 and 650 nm for Cy5. The conjugation efficiency for both Cy3 and Cy5 samples were calculated to be ∼50%.

FRET refolding studies

All FRET measurements were conducted with a Horiba Fluorolog-3 machine equipped with Synapse OE-CCD Array Detector. All FRET samples were excited at 500 nm to avoid directly exciting the acceptor dye (Cy5) and to directly excite only the donor dye (Cy3). All samples contained 10 mg/mL SoyPC liposomes as prepared as described above with varying protein monomer concentrations from 1 to 10 μM.

For FRET refolding studies, two different types of experiments were carried out. First, a similar experiment to the gel-refolding studies (**Fig. 4.1**) was carried out by quenching refolding reaction of the dye-labeled samples at different time points. However, instead of running each time point on a gel, each time point's emission spectrum was measured on

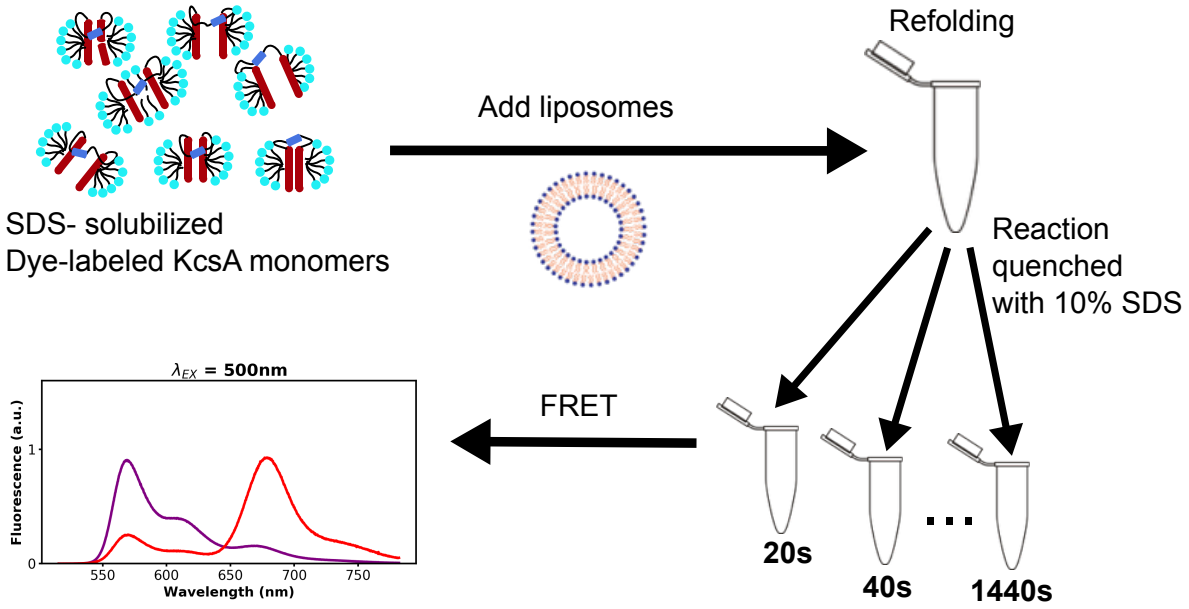

Figure 4.4: **Size-exclusion chromatography elution profile of KcsA L86C with Cy3 and Cy5. (A)** KcsA Δ125 L86C conjugated with Cy3. Absorbances are monitored at 280 nm (purple) for protein and 550 nm (cyan) for Cy3. The elution profile for both overlap well suggesting successful conjugation reaction. **(B)** KcsA Δ125 L86C is conjugated with Cy5. Absorbances are monitored at 280 nm (purple) for protein and 650 nm (orange) for Cy5.

a fluorometer (**Fig. 4.4**). This type of experiment allowed us to monitor the build-up of native tetramers as the SDS-quenching step will break up the liposomes to stop any refolding reactions and dissociate any non-native tetramers in the refolding mixture.

In the second type of experiment, instead of utilizing a quenching step, the FRET level of the sample was continuously monitored in the liposome without any quenching (**Fig. 4.5**). This measurement allowed us to monitor any non-native as well as native oligomerization happening in liposomes over time.

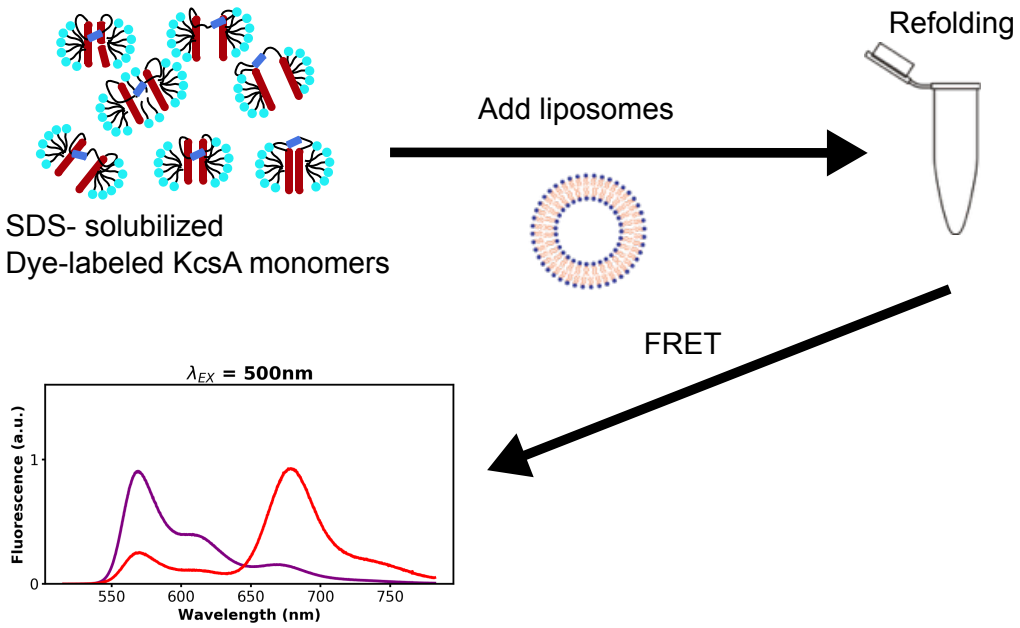

Figure 4.5: **Continuous FRET measurement in liposome.** Dye-labeled KcsA monomers are first solubilized in 0.5% SDS buffer. At concentrations ranging from 1 – 10 μM, protein is diluted into SoyPC liposome buffer to initiate refolding reactions. After manual mixing, this mixture's emission spectrum is continuously monitored using a fluorometer.

4.3 Results and Discussion

4.3.1 Kinetics of folding

Folding Kinetics of WT versus CC KcsA Δ125

The kinetics of tetramerization of KcsA Δ125 channels were examined by tracking the formation of native tetramers using an SDS resistance assay [65]. The refolding protocol started with tricholoroacetic acid (TCA)-precipitated monomers solubilized in 14 mM (∼0.5% w/v) SDS at pH 6.5 (**Fig. 4.1A**). To initiate refolding, these monomers were diluted 10-fold into refolding buffer containing 15 mM asolectin liposomes. Control experiments employing dynamic light scattering verified that the liposomes remain intact after mixing with the

SDS-solubilized KcsA Δ125 or the 14 mM SDS buffer (**Table. 4.1**).

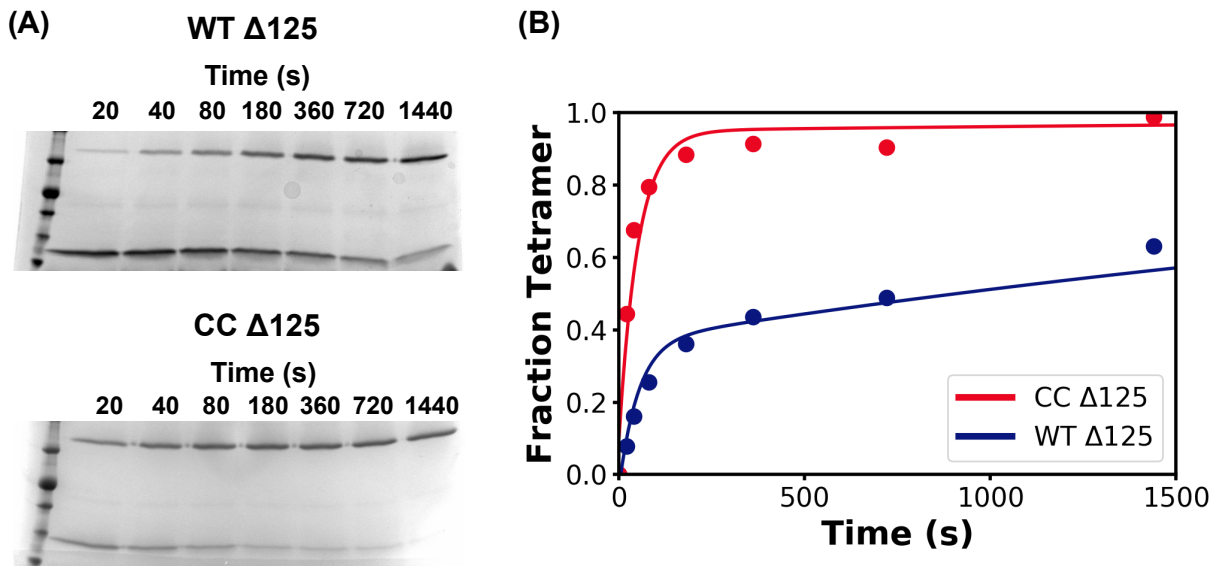

Figure 4.6: **The presence of constraining disulfide bond results in faster folding and higher yields**. **(A)** Examples of raw data for the refolding of WT Δ125 and CC Δ125 are shown. **(B)** Fraction tetramer is plotted against time for WT Δ125 and CC Δ125 constructs. The solid lines represent the best fit double exponential global fits to both WT Δ125 and CC Δ125 kinetic traces assuming common fast and slow rates.

Table 4.1: Liposome size measured using dynamic light scattering.

	13 mM SoyPC	13 mM SoyPC + 1 mM SDS	13 mM SoyPC + 1 mM SDS-solubilized KcsA
R_H (nm)	91.3 ± 2.1	86.0 ± 1.7	84.9 ± 1.3
Polydispersity	24.0 ± 0.4	23.1 ± 0.8	20.1 ± 0.2

To measure refolding, aliquots of the protein-liposome mixture were removed over 25 minutes and quenched in 220 mM (∼6%) SDS buffer to arrest tetramer formation. At this high SDS concentration, liposomes were disrupted and only native tetramers persisted whereas weakly associated species were broken up and ran as monomers on SDS-PAGE gels [59, 64, 65, 88]. The folding kinetics were quantified from gels by integrating the band intensities for tetramer and monomer fractions, and calculating the fraction tetramer.

For WT KcsA Δ125 monomers at 10 μM monomer concentration, two nearly equal refolding populations were observed, one that folded on sub-minute and another that folded on the 10 minute time scale (**Fig. 4.6**). Since the buildup of native tetramers is directly measured, the fast and slow appearance of native tetramers implies that there are multiple routes to the native state, rather than each phase representing a step on a sequential pathway, which would have resulted in a 10 minute lag in the buildup of tetramers.

On the other hand, for the disulfide bonded CC construct, most of the monomers tetramerized at the fast rate (**Fig. 4.6B**). This difference between the WT and CC variant implied that the presence of unconstrained transmembrane helices enabled the formation of a stably misfolded, slow folding species for WT KcsA. Data for the WT and CC were fit globally, assuming the rates of the two phases were the same for the two versions. The resulting time constants were τ_{fast} = 40 ± 2 s and τ_{slow} = 1500 ± 100 s, with a 28 ± 6% and 74 ± 6% fast folding population for the WT and CC constructs, respectively.

Concentration dependent folding kinetics

Next, the concentration dependence of the folding was studied. Interestingly, the tetramerization process, both in rate and branching ratio, appeared to be concentration independent from 1 to 10 μM within the accuracy of our measurements (**Fig. 4.7, Table. 4.2**). This striking result indicates that the rate-limiting step in the folding process is a unimolecular process, despite the native state being a tetramer. If tetramerization was limited by the association of four monomers, one would have observed a 1000-fold slowing for a 10-fold decrease in concentration. Even if dimerization was rate-limiting, one would have expected a 10-fold slowing in the apparent kinetic rate. As no measurable slowing was found, we hypothesized that the oligomerization process occurs early and is fast on both routes, with the rate-limiting step representing a productive folding (fast pathway) or error-correction step (slow pathway). This type of folding pathways have been observed before in folding of

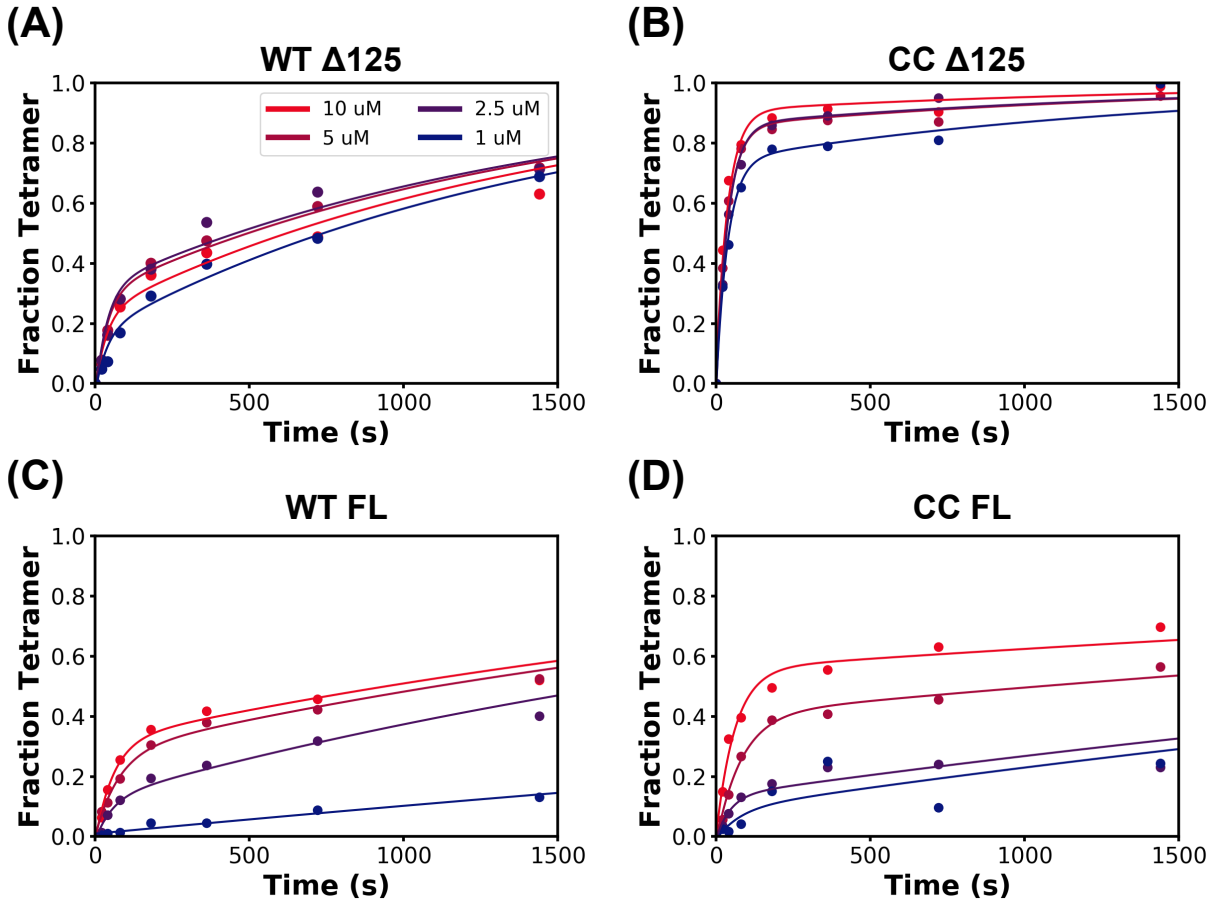

Figure 4.7: **Concentration dependent folding kinetics of WT and CC construct either with (FL, full length) and without (Δ125) carboxy-terminal tetramerization domain.** The concentration dependence between 1 – 10 μM of the folding kinetics is shown for (A) WT Δ125 (B) CC Δ125 (C) WT Full-Length (D) CC Full-Length.

soluble proteins [101, 102, 103]. This proposal and the nature of this fast oligomerization process are investigated further with FRET below.

4.3.2 FRET measurements suggest a formation of protein-rich phase

The SDS folding assay indicated that the folding of both the WT and CC constructs has a minimal concentration dependence implying that the rate-limiting step is a unimolecular process. This step seems to occur likely after oligomerization in the liposomes. In principle,

Table 4.2: Folding monitored by SDS-Page gel.

	WT Δ125	CC Δ125	WT FL	CC FL
Fast time constant (s)	40 ± 2		90 ± 10	
Fast population	0.25 ± 0.05 (0.33 ± 0.05 0.36 ± 0.05 0.20 ± 0.05)	0.86 ± 0.05 (0.86 ± 0.05 0.83 ± 0.05 0.87 ± 0.05)	0.42 ± 0.04 (0.38 ± 0.04 0.26 ± 0.04 0 ± 0.04)	0.55 ± 0.04 (0.44 ± 0.04 0.16 ± 0.04 0.13 ± 0.04)
Slow time constant (s)	1500 ± 100		8400 ± 1600	
Slow population	0.77 ± 0.03 (0.70 ± 0.03 0.68 ± 0.03 0.83 ± 0.03)	0.09 ± 0.03 (0.14 ± 0.03 0.14 ± 0.03 0.26 ± 0.03)	0.59 ± 0.02 (0.63 ± 0.02 0.76 ± 0.02 1 ± 0.02)	0.40 ± 0.02 (0.57 ± 0.02 0.83 ± 0.02 0.88 ± 0.02)

however, the insertion of monomers into the liposomes could be rate limiting. To test whether oligomerization is fast and occurs before the rate-limiting step, ensemble FRET measurements were carried out.

Dyes were first attached using thiol-labeling at a single position using a KcsA Δ125 L86C variant (**Fig. 4.2**). Equal mixtures of donor (Cy3) and acceptor (Cy5) labeled KcsA Δ125 were mixed and diluted 10-fold into the liposome mixture to initiate folding and the transfer of fluorescence was monitored (**Fig. 4.5**). Initially, the emission spectrum was dominated by that of the donor implying that most molecules began as isolated monomers. Upon dilution into a liposome mixture, the donor emission maximum at 570 nm was quenched by 71% within the 10 second manual mixing dead-time while the acceptors emission at 680 nm increased by 7.6-fold. During the next 25 minutes over which tetramer formation occurred, only a 4% increase in intensity was observed across the entire emission spectrum.

This biphasic FRET signal is interpreted as follows. The initial FRET increase indicates that the monomers labeled with donor and acceptor dyes rapidly associate in the liposomes prior to tetramerization. The minimal subsequent change implies that the FRET level in the rapidly associated monomers is similar to the level of folded tetramers in liposomes. The observation that most of the change in FRET signal occurred before significant tetramer

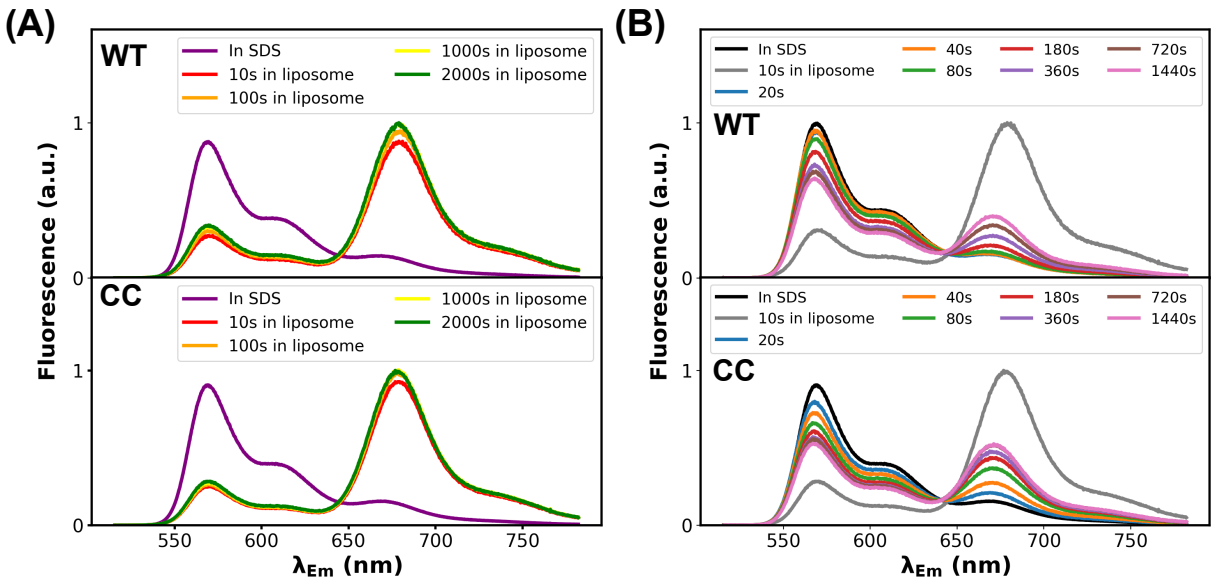

Figure 4.8: **FRET measurements of KcsA tetramerization.** **(A)** Ensemble FRET measurement of the KcsA monomer in SDS and in liposome as a function of time **(B)** Double-jump (unfold-fold-unfold) FRET measurements of KcsA refolding quenched with 10% SDS overlaid with in sds and 10s in liposome time points from **(A)**. Spectra are normalized to have the same value at 641 nm, an empirical iso-emissive point. Measurements are conducted at a monomer concentration of 10 μM.

formation, along with an estimate of $\sim$ 30 - 300 monomers per liposome at 1 - 10 μM monomer concentration, argues that most of the population forms a non-native oligomeric state upon insertion into liposomes with a FRET level comparable to that of native tetramers. The oligomers may become part of a protein-rich phase within the membrane.

To examine the possibility that FRET occurred within protein aggregates forming outside of liposomes, SDS-solubilized KcsA monomers were diluted 10-fold into water in the absence of liposomes. In this control, the overall donor fluorescence decreased 4%, but the acceptor fluorescence spectrum only had a minimal indication of FRET (**Fig. 4.9**), especially as compared to refolding in liposomes. The signal decrease indicates that some fraction of monomers was no longer in solution; however, more importantly, the lack of significant FRET indicates that the large observed FRET changes in the presence of liposome only

came from membrane-solubilized proteins.

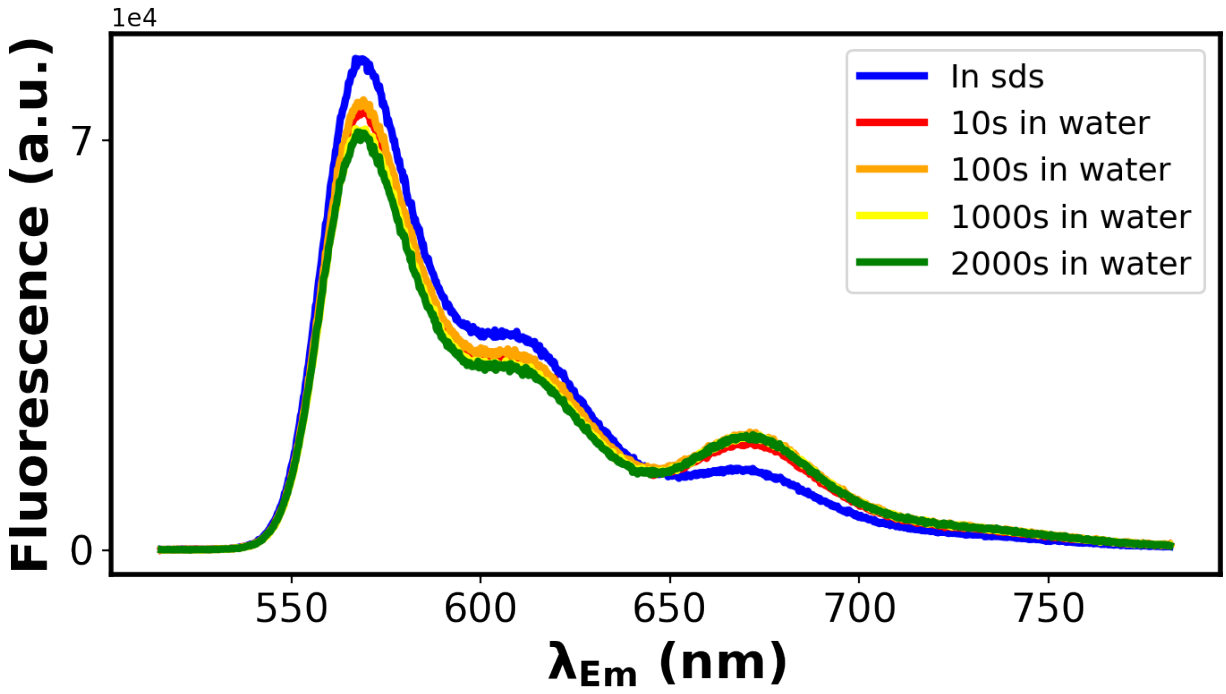

Figure 4.9: **SDS-solubilized KcsA monomer dilution into water does not FRET.** Minor changes in FRET level is observed over time for SDS-solubilized WT KcsA monomers diluted into water 10-fold.

In addition, the tetramerization kinetics data collected with FRET (**Fig. 4.8B**) with SDS-quenching was compared to the measurements collected with SDS-PAGE gel (**Fig. 4.10**). Tetramerization level from FRET was calculated using the following formula:

$$\text{Fraction Tetramer} = \frac{\epsilon_{Cy5} I_{650nm}}{\epsilon_{Cy3} I_{550nm} + \epsilon_{Cy5} I_{650nm}} \tag{4.4}$$

where the ϵs are the extinction coefficients for Cy3 (150,000 M^{-1} cm^{-1}) and Cy5 (250,000 M^{-1} cm^{-1}), and I is the average intensity from 545 nm to 555nm for Cy3 and 645 to 655 nm for Cy5.

Comparing the kinetics traces with gel results, the two agree qualitatively well. Because we only have 2 kinetic traces for the FRET data and we are fitting them to a double

exponential, there are many possible solutions. We assumed that FRET data would have the same rate constants as the gel data and were fitted. The result was plotted in Figure 4.10 and qualitatively speaking, the CC mutant still tetramerizes much more efficiently than the WT confirming our gel results.

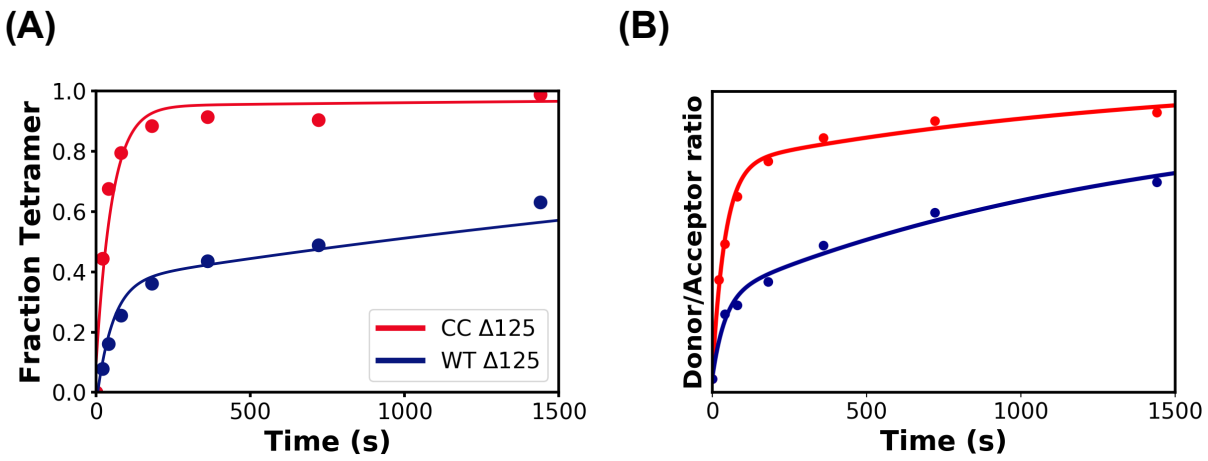

Figure 4.10: **Gel-based and FRET refolding assays agree well with each other.** Qualitatively, the two measurements show similar behavior. CC KcsA folds more efficiently than the WT KcsA monomers.

Previously, we found that the full length WT KcsA and CC KcsA both fold in a concentration dependent manner. The fast and slow time constants were both slowed for the full length constructs compared to the Δ125 constructs. We hypothesized that the monomers are folding in concentration dependent manner for the full-length constructs because they're not forming a protein-rich phase as we found in the Δ125 constructs. The C-terminal "tetramerization" domain contains 11 positive residues and 7 negative residues, resulting in net 4 positive charge. We hypothesized that this highly charged region could potentially prevent the monomers from forming a dense protein-rich phase. This hypothesis was tested with continuous FRET to see if this was indeed the case (**Fig. 4.11**). Interestingly, the FRET results seem to suggest that a large population of the monomers remain monomeric in the case of the full-length. Compared to the Δ125 constructs, where large fraction of FRET

happens very early and change little, the full-length constructs seem to show a large portion of monomers remaining monomeric and some population forming oligomers. Although a more careful investigations into the role of C-terminal stalk will be required, we hypothesize that the C-terminal domain does not help with the tetramerization process, but rather helps the monomers become more selective in finding other correct monomeric subunits.

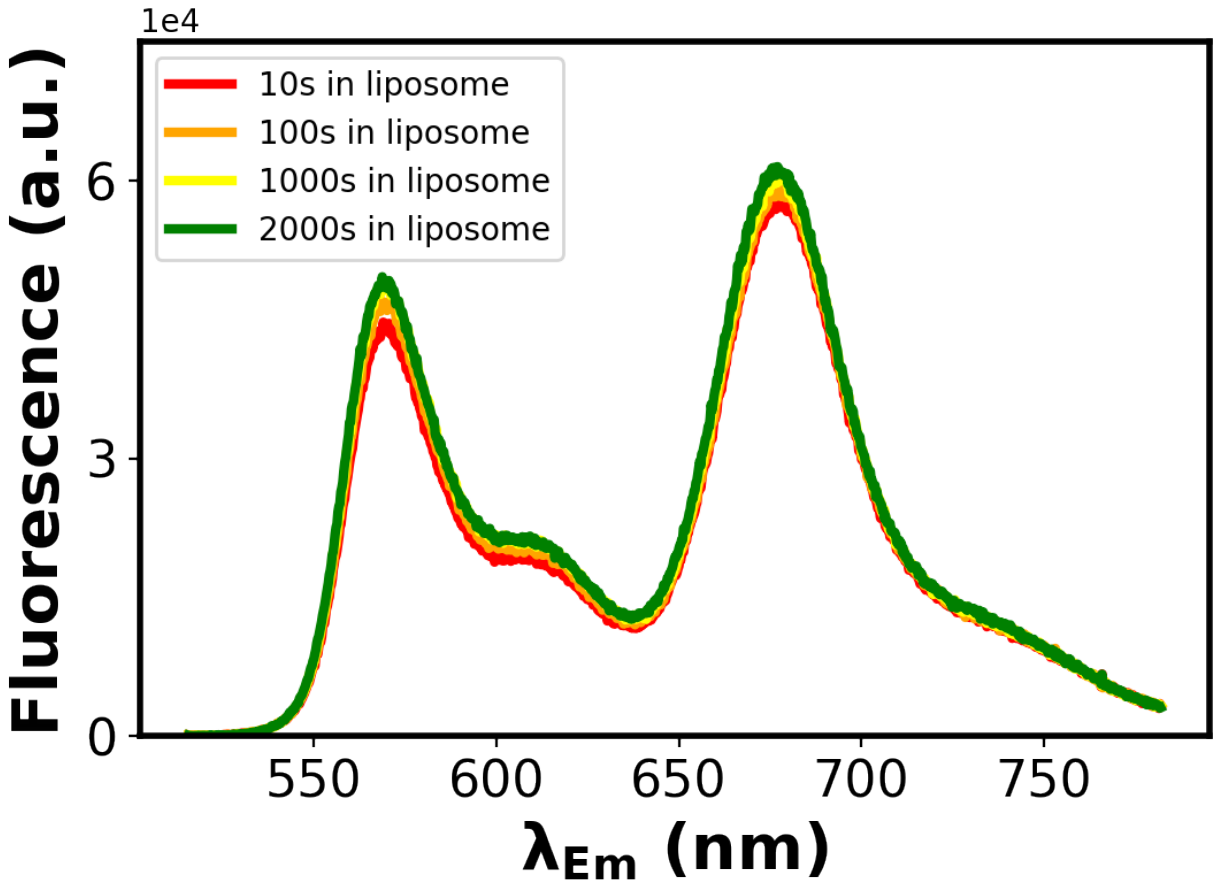

Figure 4.11: **Continuous FRET of WT KcsA full-length.** WT KcsA full-length monomers were diluted into liposomes and the emission spectra were monitored over time.

4.4 Conclusion

Our investigation of the folding of potassium channel monomers and their role in assembly of potassium channel pores revealed a number of salient features. The MD simulations and MSM analysis from Chapter 2 indicated that for both wild-type Kv1.2 and KcsA pore domains, monomers form a heterogeneous ensemble of native and non-native states with all three helices folded and lying within the membrane. While the population of native-like states of the monomer is non-negligible (18% for Kv1.2 and 44% for KcsA), it is nevertheless striking that a considerable number of non-native states exists despite the substantial conformational restriction imposed by the environment; namely, the secondary structure of all three helices is retained, and the two transmembrane helices remain inserted within the planes of the bilayer. Based on prior studies of hydrophobic matching [104, 105], the different lengths of transmembrane helices presumably contribute to their tendency to separate.

This overall picture is consistent with our NMR data combined with prior thiol-labeling studies [57, 58]. Our attempts to generate well-resolved NMR spectra for the WT KcsA monomer inserted in nanodisc and bicelles were unsuccessful. Suspecting that the separation of the transmembrane helices was the critical feature, a double cysteine variant was engineered to have a disulfide bond at the bottom of the two transmembrane helices locking them into a native-like arrangement. This CC variant yielded a well-dispersed ^{15}N-^{1}H-TROSY-HSQC spectrum supporting the view that the WT's transmembrane helices were often separated in the monomers and formed a heterogeneous ensemble.

FRET-monitored refolding measurements of monomers passing from SDS into liposomes indicated that both WT Δ125 and CC Δ125 variants oligomerized well before the appearance of native tetramers. According to both SDS-resistance assays and FRET measurements, WT KcsA monomers assembled into native tetramers via two distinct kinetic paths with $\tau \sim 40 \pm 2$ and 1500 ± 100 s. The CC variant largely if not fully lacked the slow phase. We posit that slow folding is the result of non-native packing arrangements of the transmembrane helices

that must ultimately be corrected for tetrameric assembly to proceed. The observation of slow and fast folding routes has been seen in many soluble proteins where the initial collapse step leads to species with native-like topology on a direct pathway, or to a species containing a partially misfolded structure that is slow to correct (**Fig. 4.7**) [101, 102, 103].

In spite of the channel being a tetramer, folding of the WT Δ125 and CC Δ125 constructs was concentration independent from 1 - 10 μM for both the fast and slow phases, both in rate and amplitude. This observation implies that the rate-limiting step on both pathways is unimolecular. Because the FRET measurements indicated that the monomer insertion and oligomerization was relatively quick, the rate-limiting step on the fast pathway likely occurs going from an oligomer to a native tetramer. This complex transition requires the formation of the slightly twisted arrangement of the 8 transmembrane helices, and the energetically costly opening of an aqueous channel. Final formation of native tetramers may occur in a concerted step with the association of correctly folded monomers already having the pore helices and selectivity filters positioned in a native or near-native orientation. Alternatively, this event may occur in two distinct steps, with the p-helices and selectivity filter segments folding into position only after the 8 transmembrane helices have adopted their correct native arrangement. Further studies are needed to resolve this question.

The carboxy-terminal stalk domain in KcsA was tested to see if it enhanced tetramerization. In eukaryotic channels such as the Shaker voltage-activated potassium channel, the presence of the tetramerization T1 domain improves the rate of successful folding and assembly [66]. Whereas the carboxy-terminal domain of KcsA forms a four-helix bundle that contributes to the stability of the folded tetrameric channel, [106] its role in the assembly process is unclear. In other studies, the KcsA tetramerization domain alone has been shown to oligomerize in pH and concentration dependent manners [63, 107] as well as increase the stability of KcsA tetramers [60].

However, the folding kinetics of KcsA pore domain with the tetramerization domain has

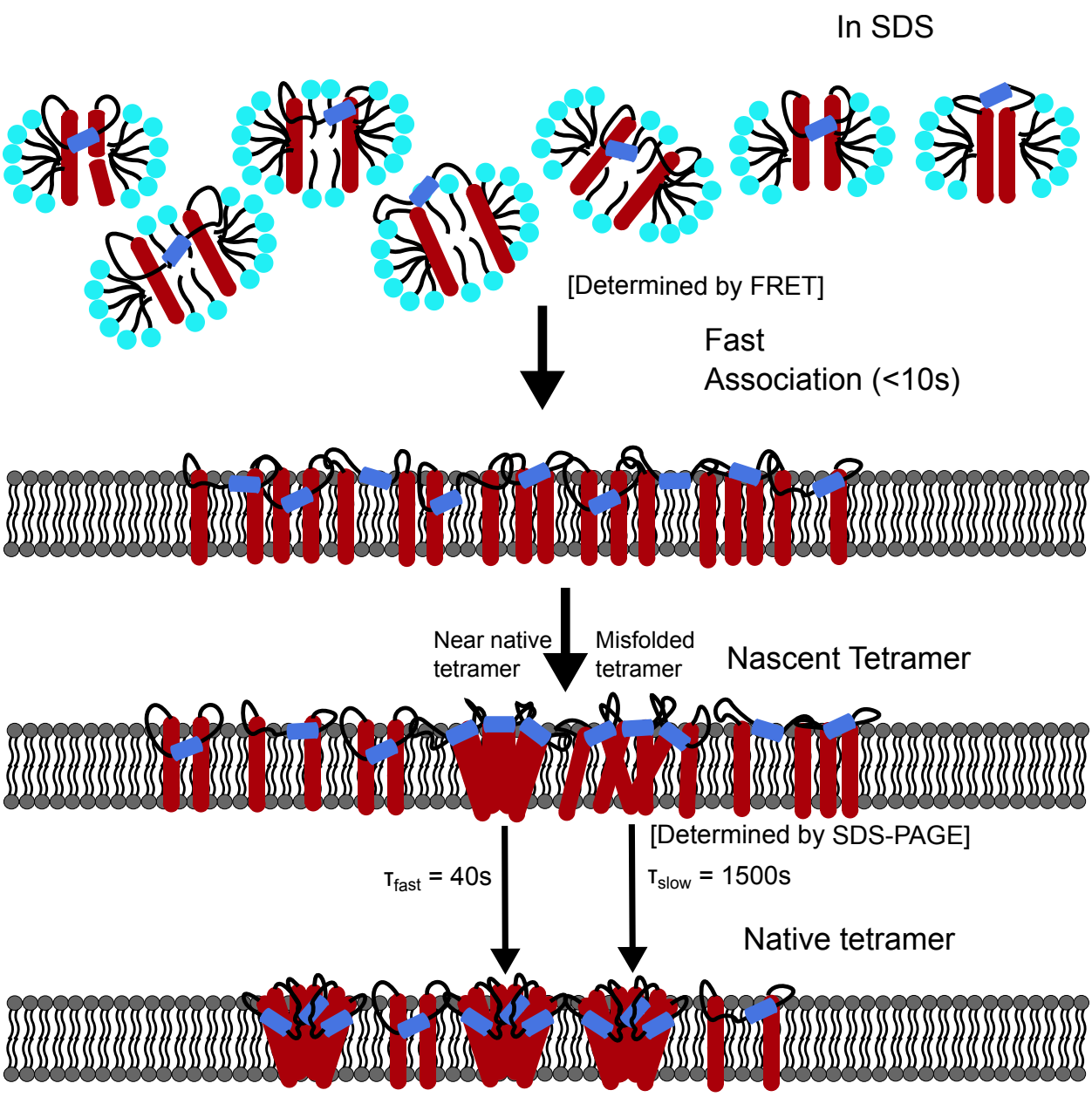

Figure 4.12: **Proposed KcsA folding and tetramerization.** SDS-solubilized monomers enter the liposomes and undergo rapid association into a protein-rich phase within the membrane prior to tetramerization. Oligomers can form with a native or non-native transmembrane helical arrangement, which fold on the minute or 20 minute time scale, respectively, The rate limiting step on the faster pathway is proposed to involve the insertion of the pore helix to stabilize the tetramer in its native conformation.

not been extensively studied in the past. To our surprise, the presence of the tetramerization domain did not improve the channels folding behavior. For both the WT and CC variants, tetramerization became concentration dependent in a complex manner with both a reduced fast phase and lower overall yield of tetramers. One may consider that our refolding protocol with insertion into liposomes starting from an SDS-solubilized state does not properly mimic the biological context and so preclude the tetramerization domain for assisting folding. Also, the C-terminal region carries a net positive charge, which may cause a repulsion at large distance between monomers prior to tetramerization. Potentially the domain lends specificity and is more important for finding other KcsA subunits or improving the stability of tetramers once they are formed.

The folding of the potassium channels displays similar behavior to other α-helical membrane proteins, having a transition state close to the native state [1, 29, 46]. For example, in the force unfolding studies of GlpG pulling parallel to the bicelle surface, the transition state was closer to the native state than that observed in SDS-driven folding/unfolding studies in solution [4, 5]. In other SDS-based refolding studies on bacteriorhodopsin, DsbB and GlpG, the transition states were expanded [22, 23, 108]. The observed difference between the force- and the SDS-driven unfolding studies may be due to the difference in the mode of denaturation as well as folding conditions (micelle or bicelle versus liposomes). Regardless, KcsA in liposomes appear to have a transition state near the native state.

The two-stage model, proposed nearly 30 years ago, has provided a useful framework for discussing membrane protein folding. The model, which posits insertion of all the transmembrane helices followed by lateral packing, was proposed based on studies which observed that bacteriorhodopsin cut in a few pieces could still be re-assembled [30, 109, 110]. With the observation of more diverse folding behaviors, a more sophisticated three-stage model10 was proposed where insertion and lateral packing of the transmembrane helices could be followed by ligand binding, loop folding or peripheral domain insertion. This extra step is relevant

to potassium channels, where the insertion of the four pore helices and the formation of the selectivity filter may be required to finalize the folding process.

For KcsA, the first two stages seem to correspond to insertion and formation of a protein-rich phase before folding into native tetrameric channels. This result brings up an interesting question in regards to membrane protein folding in general, namely, should the membrane be considered to be a good or poor solvent with respect to transmembrane helices, defined as one where helixlipid interactions are stronger or weaker, respectively, than helixhelix interactions. For potassium channels, a mixed behavior is observed. A protein-rich phase is detected by FRET, but the two transmembrane helices in a monomer can dissociate from one other according to NMR (in bicelles) and MD simulations (POPC bilayers).

Potentially in the folding of other membrane proteins, non-specific or quinary interactions between transmembrane helices also may give rise to a protein dense phase. These interactions can also occur in native structures, as seen with KcsA tetramers undergoing lateral association both in our and earlier studies [62, 111]. Several studies have shown that membrane proteins can alter lipid packing in the fluid liquid crystalline phase, which can in turn cause proteins to associate non-specifically in order to minimize the perturbation on the membrane [26, 112]. Generally, lipids may act as a marginally poor solvent for transmembrane helices especially if the helices contain polar amino acids that are less soluble in the hydrophobic bilayer [113, 114, 115, 116, 117, 118, 119, 120]. These mixed results, along with variability of the hydrophobicity of the transmembrane helices and the properties of the bilayer (e.g., composition, curvature, lateral pressure), suggest that solvent quality is likely system dependent. These issues have important implications to in vivo folding where helices partition between chamber of the translocon, lipid-water interface and the hydrophobic core of the membrane [120].

CHAPTER 5

FUTURE DIRECTIONS AND CONCLUSION

5.1 Future Directions

5.1.1 Structure determination of disulfide engineered fast folding KcsA mutant

The disulfide-bonded (CC) KcsA mutant was engineered to trap KcsA monomers in a native-like conformation. Through comparison of the folding kinetics of WT and CC KcsA variants we found that the misfolding of transmembrane helices contributes to slowing of folding kinetics in WT compared to the CC KcsA. While this information provides invaluable insight into the folding mechanics of potassium channels, we still do not fully understand how the arrangement of the pore helices affects overall folding.

Previously, NMR studies of KcsA monomer solubilized in SDS showed that the transmembrane helices as well as the pore helix retains its helicity, yet the pore helix was demonstrated to be highly dynamic residing outside the micelle environment [20]. However, in thiol-labeling experiments of Kv1.3, the pore helix was shown to be helical and buried in the bilayer [57, 58]. The discrepancy between the two proteins could be attributed to the difference in protein sequence itself and the use of different membrane mimetics (micelle versus ER membranes). In fact, the use of SDS as a membrane mimetic might attribute to loss of its native-like tertiary contacts in the KcsA studies. The two proteins have sequence identity of 31% and especially at the pore region, between residues 30 to 100 of KcsA and 328 to 398 of Kv1.2, RMSD is only 3.8 Å. Sequence and structure at the pore helix is especially conserved [48, 121], so we predict that the pore helices should behave qualitatively similar to each other when both are in similar membrane bilayer environment.

In Chapter 2, we showed that [^{15}N-^{1}H]-TROSY-HSQC spectrum of KcsA CC in bicelles

display well-dispersed peaks suggesting that this mutant is more folded and homogeneous on the NMR timescale. Given that this construct behaves stably in bicelle environment and possibly retain a more native-like secondary and tertiary structures, the major questions are what does the structure of this CC KcsA mutant look like in bicelles and what are its dynamics?

We hypothesize that the CC KcsA constructs will behave similarly to what was found in Kv1.3 thiol-labeling experiments, in which they found that the pore helix maintains its helicity and lay at the water-lipid interface. The pore helix may partition between multiple different environments (solvent-exposed, lipid-water interface, and buried) and states (helical and unfolded). With NMR structure, we can address the question of how native-like is the CC mutant compared to the monomer in its tetrameric state. With assignments and use of hydrogen-exchange methodologies, we can also begin to answer the questions of how stably folded is the pore helix and at what rate does it fold and unfold. Addressing these questions will ultimately help us to understanding the folding and assembly process of potassium channels in large.

5.1.2 Direct visualization of the protein-rich phase

The lipid raft hypothesis suggests that cholesterol and saturated lipids form preferential association in membranes and drive phase separation within the membrane [122, 123, 124, 125]. In return, this phase separation recruits certain types of lipids and proteins for biological functions such as signal transduction and act as a way to compartmentalize within the membrane [126, 127, 128, 129]. Phase separations with soluble proteins has also been shown to important for many biological functions. Phase separations initiated by proteins and RNA molecules can create membraneless organelles, signaling complexes, and help with cell's survival [130, 131, 132, 133].

Interestingly in our FRET studies, we found that KcsA can non-specifically aggregate

and form a protein-rich phase in the membrane. The existence of the protein-rich phase seems to facilitate the folding of KcsA in lipid bilayers. In the case of full-length KcsA, this protein-rich phase does not seem to exist, causing KcsA to assemble in complicated concentration dependent manner.

In order to address the question of the existence of protein-rich phase, we propose to use Total internal reflection fluorescence (TIRF) microscopy to study the formation of protein-rich phase. Measuring FRET under a microscope, we can directly visualize whether a protein-rich phase forms in lipid bilayers, which will can validate our hypothesis regarding the formation of protein-rich phase. The direct observation of such phase can open up many questions regarding phase-phase separations in membranes and physiological role of lipid rafts for protein folding.

5.1.3 Determining solvent quality for membrane proteins

Our observation of protein-rich phase for KcsA was found through SDS-PAGE and FRET studies using soybean lipids as membrane mimetic. Soybean lipids, which is a heterogeneous mixture of lipids, are generally considered great mimetics for plasma membranes. However, using pure single component lipids as well as simple mixtures of these pure lipid vesicles could allow further dissection of the driving force behind the formation of protein-rich phase or more interestingly, are some membrane components better solvent for membrane proteins than some others?

Good or poor solvents are defined as one where helix–lipid interactions are stronger or weaker, respectively, than helix-helix interactions. For KcsA in soybean lipids, the environment seems to be in a theta solvent environment, which is in between the good and poor solvent regime. In the monomeric state, the transmembrane helices explore both compact and highly expanded structures. The existence of highly expanded structures suggest that the solvent is good, hence helices dissociating from each other. However, in our FRET stud-

ies, when several monomer species are in a liposome, the monomers seem to associate with each other to form a protein-rich phase, suggesting that the membrane environment is a poor solvent. This phenomenon is likely lipid-dependent and protein-dependent, and should be further investigated. Perhaps, the addition of anionic lipids or unsaturated lipids can give rise to better solubilization of KcsA, which would eliminate the formation of protein-rich phase.

In addition, lipid rafts have been postulated and shown to recruit some proteins more than others [122, 123, 124, 125]. With the use of microscopy, we can study whether KcsA displays a preferential partitioning into a one lipid phase than the other. While not conclusive, observation of such phenomenon will let us hypothesize whether lipid rafts can help proteins fold *in vivo*.

5.1.4 Determination of the rate-limiting step in KcsA folding

Through our concentration dependent folding kinetics studies, the rate-limiting step in KcsA folding was found to be unimolecular. This was quite surprising given the fact that KcsA is a tetramer in its native state. While we were able to show that by locking the 2 transmembrane helices in near-native state reduces KcsA's proclivity to misfold, the rate-limiting step in the reaction was not resolved. We hypothesize that the rate-limiting step correponds to the insertion of pore helices into the chamber created by 8 transmembrane helices in the membrane.

One of the ways to test our hypothesis is by introducing perturbations at the pore region to see if these perturbations affect the kinetics of the rate-limiting step. For example, potassium (K^+), barium (Ba^+), and rubidium (Rb^+) ions have been known to stabilize the pore structure in KcsA. In our refolding assays, none of these ions are present because they are known to strongly interact with SDS and precipitate. However, a double jump experiment could be done where SDS-solubilized KcsA monomers are first diluted into liposome solu-

tion, and then, a high concentration of potassium, barium or rubidium is injected into the refolding solution soon afterwards. This protocol would give SDS-solubilized KcsA time to insert into the liposomes, so that SDS precipitation will not affect the protein. The slower tetramerization process would occur in the presence of K^+, Ba^+ and Rb^+ and could increase the rate of folding by stabilizing the pore structure. If the rate-limiting folding kinetics is affected by the existence of these cations, we would conclude that insertion of the pore helices is indeed rate-limiting. In addition, we could design a disulfide bond between the pore helix and one of the 2 transmembrane helices to stabilize the native-like pore structure. This could also affect the rate-limiting step which would show that the pore region is involved in the rate-limiting step.

5.2 Conclusion

This thesis is the first comprehensive study on potassium channel folding. The dynamics of potassium channel monomers are first studied through MD simulations and NMR spectroscopy, through which we found that the wild-type potassium channel monomers are dynamical and exists in a heterogeneous ensemble of native and non-native structures. Although the exact population level is system dependent and speculated based on simulations, we found that the native state is found in 18% and 44% for Kv1.2 and KcsA, respectively. While these numbers may not be accurate, the fact that the native-like state is populated at significant level seems plausible. Without the monomers having any proclivity to fold into its native-like state, forming a tetramer with four monomers in native-like state seems almost impossible. Regardless, the WT KcsA monomers seem to prefer to exist in a heterogeneous ensemble of structural states.

In order to design a mutant that remains its near-native state when monomerized, a disulfide-bridge was engineered at the ends of the 2 transmembrane helices in KcsA (A29C and A109C). By MD simulations and NMR spectroscopy, this disulfide-bonded (CC) KcsA

mutant was shown to be more native-like and homogeneous in NMR timescale. This mutant exhibited [^{15}N-^{1}H]-TROSY-HSQC spectrum with well-dispersed peaks, opening up future structural and dynamic studies of this particular mutant. The state of the pore helix is unclear whether it is buried or exposed to aqueous solution and helical or unfolded. I hypothesize that understanding the pore helix dynamics will uncover the folding mechanism of potassium channel in the future.

In our comparative folding kinetics studies of WT and CC KcsA, we found that WT KcsA folds via 2 distinct pathways, whereas the CC variant mostly undergoes through the faster pathway. This result suggests that locking the 2 transmembrane helices in native-like arrangement reduces KcsA monomers to misfold, which allows this CC variant to fold only through the faster process. This type of biphasic folding kinetics has been seen in soluble proteins as well [101, 102, 103], where the faster kinetic process represents productive folding and the slower kinetic process represents error-prone or misfolding process. This result is in agreement with our studies of the monomeric KcsA through MD simulations and NMR. WT KcsA was found to be heterogeneous in structure whereas CC KcsA was found to be more native-like. It seems plausible to think that folding kinetics of WT KcsA seems to display a biphasic behavior because they are more prone to misfolding even at the monomeric level.

In addition, concentration dependent folding kinetics of WT and CC KcsA was studied, for which we found that both WT and CC KcsA folding is concentration independent. This was quite surprising given that the native state of KcsA is a tetramer. For a tetramerization process assuming binding is rate-limiting, $4[M] \rightarrow [T]$, one would expect a 1000-fold decrease in k_{app} over 10-fold decrease in the monomer concentration. Even if dimerization was rate-limiting, one would expect 10-fold decrease when monomer concentration is changed 10-fold. Both slow and fast folding rates were concentration independent indicating that the rate-limiting step in both pathways is unimolecular.

Our concentration dependent folding kinetics study led us to hypothesize that the oligomer-

ization process occurs during or right after insertion into liposomes and folding within this near-native state is what is rate-limiting in the overall folding reaction. In order to investigate whether if oligomerization indeed is quick, we utilized FRET. FRET results indicated that the oligomerization process is much quicker than folding reactions indicating that KcsA might form a non-specific interaction with each other to form a dense protein-rich phase.

REFERENCES

[1] J. U. Bowie. Solving the membrane protein folding problem. *Nature*, 438(7068):581–589, 2005.

[2] Heedeok Hong. Toward understanding driving forces in membrane protein folding. *Archives of biochemistry and biophysics*, 564:297–313, 2014.

[3] H. Yu, M. G. W. Siewny, D. T. Edwards, A. W. Sanders, and T. T. Perkins. Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins. *Science*, 355(6328):945–949, 2017.

[4] R. Q. Guo, K. Gaffney, Z. Y. Yang, M. Kim, S. Sungsuwan, X. F. Huang, W. L. Hubbell, and H. Hong. Steric trapping reveals a cooperativity network in the intramembrane protease glpg. *Nature Chemical Biology*, 12(5):353–+, 2016.

[5] D. Min, R. E. Jefferson, J. U. Bowie, and T. Y. Yoon. Mapping the energy landscape for second-stage folding of a single membrane protein. *Nature Chemical Biology*, 11(12):981–987, 2015.

[6] S. H. White and W. C. Wimley. Membrane protein folding and stability: Physical principles. *Annual Review of Biophysics and Biomolecular Structure*, 28:319–365, 1999.

[7] Jonathan N Sachs and Donald M Engelman. Introduction to the membrane protein reviews: the interplay of structure, dynamics, and environment in membrane protein function. *Annu. Rev. Biochem.*, 75:707–712, 2006.

[8] Markus Sällman Almén, Karl JV Nordström, Robert Fredriksson, and Helgi B Schiöth. Mapping the human membrane proteome: a majority of the human membrane proteins can be classified according to function and evolutionary origin. *BMC biology*, 7(1):50, 2009.

[9] Erik Wallin and Gunnar Von Heijne. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Science*, 7(4):1029–1038, 1998.

[10] Karen G Fleming. Energetics of membrane protein folding. *Annual review of biophysics*, 43:233–255, 2014.

[11] Robert E Jefferson, Duyoung Min, Karolina Corin, Jing Yang Wang, and James U Bowie. Applications of single-molecule methods to membrane protein folding studies. *Journal of molecular biology*, 430(4):424–437, 2018.

[12] C Preston Moon, Sarah Kwon, and Karen G Fleming. Overcoming hysteresis to attain reversible equilibrium folding for outer membrane phospholipase a in phospholipid bilayers. *Journal of molecular biology*, 413(2):484–494, 2011.

[13] C Preston Moon and Karen G Fleming. Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proceedings of the National Academy of Sciences*, 108(25):10174–10177, 2011.

[14] C Preston Moon, Nathan R Zaccai, Patrick J Fleming, Dennis Gessmann, and Karen G Fleming. Membrane protein thermodynamic stability may serve as the energy sink for sorting in the periplasm. *Proceedings of the National Academy of Sciences*, page 201212527, 2013.

[15] Cosmin L Pocanschi, Geetika J Patel, Derek Marsh, and Jörg H Kleinschmidt. Curvature elasticity and refolding of ompa in large unilamellar vesicles. *Biophysical journal*, 91(8):L75–L77, 2006.

[16] Katheryn M Sanchez, Jonathan E Gable, Diana E Schlamadinger, and Judy E Kim. Effects of tryptophan microenvironment, soluble domain, and vesicle size on the thermodynamics of membrane protein folding: lessons from the transmembrane protein ompa. *Biochemistry*, 47(48):12844–12852, 2008.

[17] Gerard HM Huysmans, Stephen A Baldwin, David J Brockwell, and Sheena E Radford. The transition state for folding of an outer membrane protein. *Proceedings of the National Academy of Sciences*, 107(9):4099–4104, 2010.

[18] Heedeok Hong and Lukas K Tamm. Elastic coupling of integral membrane protein stability to lipid bilayer forces. *Proceedings of the National Academy of Sciences*, 101(12):4065–4070, 2004.

[19] Guo Qiang Chen and Eric Gouaux. Probing the folding and unfolding of wild-type and mutant forms of bacteriorhodopsin in micellar solutions: evaluation of reversible unfolding conditions. *Biochemistry*, 38(46):15380–15387, 1999.

[20] Jordan H Chill, John M Louis, Frank Delaglio, and Ad Bax. Local and global structure of the monomeric subunit of the potassium channel kcsa probed by nmr. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1768(12):3260–3270, 2007.

[21] P. J. Booth, S. L. Flitsch, L. J. Stern, D. A. Greenhalgh, P. S. Kim, and H. G. Khorana. Intermediates in the folding of the membrane-protein bacteriorhodopsin. *Nature Structural Biology*, 2(2):139–143, 1995.

[22] P. Curnow and P. J. Booth. Combined kinetic and thermodynamic analysis of alpha-helical membrane protein unfolding. *Proceedings of the National Academy of Sciences of the United States of America*, 104(48):18970–18975, 2007.

[23] W. Paslawski, O. K. Lillelund, J. V. Kristensen, N. P. Schafer, R. P. Baker, S. Urban, and D. E. Otzen. Cooperative folding of a polytopic alpha-helical membrane protein involves a compact n-terminal nucleus and nonnative loops. *Proceedings of the National Academy of Sciences of the United States of America*, 112(26):7978–7983, 2015.

[24] Dennis Gessmann, Yong Hee Chung, Emily J Danoff, Ashlee M Plummer, Clifford W Sandlin, Nathan R Zaccai, and Karen G Fleming. Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and bama. *Proceedings of the National Academy of Sciences*, page 201322473, 2014.

[25] Ashlee M Plummer and Karen G Fleming. From chaperones to the membrane with a bam! *Trends in biochemical sciences*, 41(10):872–882, 2016.

[26] F. Cymer, G. von Heijne, and S. H. White. Mechanisms of integral membrane protein insertion and folding. *J Mol Biol*, 427(5):999–1022, 2015.

[27] Lukas K Tamm, Heedeok Hong, and Binyong Liang. Folding and assembly of β-barrel membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1666(1-2):250–263, 2004.

[28] William C Wimley. The versatile β-barrel membrane protein. *Current opinion in structural biology*, 13(4):404–411, 2003.

[29] J. L. Popot and D. M. Engelman. Membrane protein folding and oligomerization: the two-stage model. *Biochemistry*, 29(17):4031–7, 1990.

[30] J. L. Popot, S. E. Gerchman, and D. M. Engelman. Refolding of bacteriorhodopsin in lipid bilayers - a thermodynamically controlled 2-stage process. *Journal of Molecular Biology*, 198(4):655–676, 1987.

[31] Tina Junne, Lucyna Kocik, and Martin Spiess. The hydrophobic core of the sec61 translocon defines the hydrophobicity threshold for membrane integration. *Molecular biology of the cell*, 21(10):1662–1670, 2010.

[32] David Pitonzo, Zhongying Yang, Yoshihiro Matsumura, Arthur E Johnson, and William R Skach. Sequence-specific retention and regulated integration of a nascent membrane protein by the endoplasmic reticulum sec61 translocon. *Molecular biology of the cell*, 20(2):685–698, 2009.

[33] William C Wimley and Stephen H White. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature Structural and Molecular Biology*, 3(10):842, 1996.

[34] William C Wimley, Trevor P Creamer, and Stephen H White. Solvation energies of amino acid side chains and backbone in a family of host- guest pentapeptides. *Biochemistry*, 35(16):5109–5124, 1996.

[35] Tara Hessa, Hyun Kim, Karl Bihlmaier, Carolina Lundin, Jorrit Boekel, Helena Andersson, IngMarie Nilsson, Stephen H White, and Gunnar von Heijne. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*, 433(7024):377, 2005.

[36] Mikhail Bogdanov, Philip Heacock, Ziqiang Guan, and William Dowhan. Plasticity of lipid-protein interactions in the function and topogenesis of the membrane protein lactose permease from escherichia coli. *Proceedings of the National Academy of Sciences*, 107(34):15057–15062, 2010.

[37] Heidi Vitrac, Mikhail Bogdanov, and William Dowhan. In vitro reconstitution of lipid-dependent dual topology and postassembly topological switching of a membrane protein. *Proceedings of the National Academy of Sciences*, 110(23):9338–9343, 2013.

[38] Heidi Vitrac, David M MacLean, Vasanthi Jayaraman, Mikhail Bogdanov, and William Dowhan. Dynamic membrane protein topological switching upon changes in phospholipid environment. *Proceedings of the National Academy of Sciences*, 112(45):13874–13879, 2015.

[39] Charlotte W Pratt and Kathleen Cornely. *Essential biochemistry*. Wiley Hoboken, NJ, 2004.

[40] Daniel E Otzen. Folding of dsbb in mixed micelles: a kinetic analysis of the stability of a bacterial membrane protein. *Journal of molecular biology*, 330(4):641–649, 2003.

[41] Kuo-S Huang, H Bayley, Mei-J Liao, E London, and HG Khorana. Refolding of an integral membrane protein. denaturation, renaturation, and reconstitution of intact bacteriorhodopsin and two proteolytic fragments. *Journal of Biological Chemistry*, 256(8):3802–3809, 1981.

[42] Francis W Lau and James U Bowie. A method for assessing the stability of a membrane protein. *Biochemistry*, 36(19):5884–5892, 1997.

[43] Rosanna P Baker and Sinisa Urban. Architectural and thermodynamic principles underlying intramembrane protease function. *Nature chemical biology*, 8(9):759, 2012.

[44] Jonathan P Schlebach, Dungeng Peng, Brett M Kroncke, Kathleen F Mittendorf, Malathi Narayan, Bruce D Carter, and Charles R Sanders. Reversible folding of human peripheral myelin protein 22, a tetraspan membrane protein. *Biochemistry*, 52(19):3229–3241, 2013.

[45] Heedeok Hong, Tracy M Blois, Zheng Cao, and James U Bowie. Method to measure strong protein–protein interactions in lipid bilayers using a steric trap. *Proceedings of the National Academy of Sciences*, 107(46):19802–19807, 2010.

[46] D. M. Engelman, Y. Chen, C. N. Chin, A. R. Curran, A. M. Dixon, A. D. Dupuy, A. S. Lee, U. Lehnert, E. E. Matthews, Y. K. Reshetnyak, A. Senes, and J. L. Popot. Membrane protein folding: beyond the two stage model. *Febs Letters*, 555(1):122–125, 2003.

[47] Mario M-C Kuo, W John Haynes, Stephen H Loukin, Ching Kung, and Yoshiro Saimi. Prokaryotic k+ channels: from crystal structures to diversity. *FEMS microbiology reviews*, 29(5):961–985, 2005.

[48] S. Choe. Potassium channel structures. *Nature Reviews Neuroscience*, 3(2):115–121, 2002.

[49] Roderick MacKinnon. Potassium channels. *FEBS letters*, 555(1):62–65, 2003.

[50] Christopher Miller. An overview of the potassium channel family. *Genome biology*, 1(4):reviews0004–1, 2000.

[51] Arthur Prindle, Jintao Liu, Munehiro Asally, San Ly, Jordi Garcia-Ojalvo, and Gürol M Süel. Ion channels enable electrical communication in bacterial communities. *Nature*, 527(7576):59, 2015.

[52] Jintao Liu, Arthur Prindle, Jacqueline Humphries, Marçal Gabalda-Sagarra, Munehiro Asally, D Lee Dong-yeon, San Ly, Jordi Garcia-Ojalvo, and Gürol M Süel. Metabolic co-dependence gives rise to collective oscillations within biofilms. *Nature*, 523(7562):550, 2015.

[53] Char-Chang Shieh, Michael Coghlan, James P Sullivan, and Murali Gopalakrishnan. Potassium channels: molecular defects, diseases, and therapeutic opportunities. *Pharmacological reviews*, 52(4):557–594, 2000.

[54] Warren M Jackman, Karen J Friday, Jerome L Anderson, Etienne M Aliot, Mel Clark, and Ralph Lazzara. The long qt syndromes: a critical review, new clinical observations and a unifying hypothesis. *Progress in cardiovascular diseases*, 31(2):115–172, 1988.

[55] Stephen M Modell and Michael H Lehmann. The long qt syndrome family of cardiac ion channelopathies: a huge review. *Genetics in Medicine*, 8(3):143, 2006.

[56] Declan A Doyle, Joao Morais Cabral, Richard A Pfuetzner, Anling Kuo, Jacqueline M Gulbis, Steven L Cohen, Brian T Chait, and Roderick MacKinnon. The structure of the potassium channel: molecular basis of k+ conduction and selectivity. *science*, 280(5360):69–77, 1998.

[57] C. Gajewski, A. Dagcan, B. Roux, and C. Deutsch. Biogenesis of the pore architecture of a voltage-gated potassium channel. *Proceedings of the National Academy of Sciences of the United States of America*, 108(8):3240–3245, 2011.

[58] E. Delaney, P. Khanna, L. W. Tu, J. M. Robinson, and C. Deutsch. Determinants of pore folding in potassium channel biogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, 111(12):4620–4625, 2014.

[59] F. I. Valiyaveetil, R. MacKinnon, and T. W. Muir. Semisynthesis and folding of the potassium channel kcsa. *J Am Chem Soc*, 124(31):9113–20, 2002.

[60] M. L. Molina, J. A. Encinar, F. N. Barrera, G. Fernandez-Ballester, G. Riquelme, and J. M. Gonzalez-Ros. Influence of c-terminal protein domains and protein-lipid interactions on tetramerization and stability of the potassium channel kcsa. *Biochemistry*, 43(47):14924–14931, 2004.

[61] F. N. Barrera, M. L. Renart, M. L. Molina, J. A. Poveda, J. A. Encinar, A. M. Fernandez, J. L. Neira, and J. M. Gonzalez-Ros. Unfolding and refolding in vitro of a tetrameric, alpha-helical membrane protein: The prokaryotic potassium channel kcsa. *Biochemistry*, 44(43):14344–14352, 2005.

[62] F. N. Barrera, M. L. Renart, J. A. Poveda, B. De Kruijff, J. A. Killian, and J. M. Gonzalez-Ros. Protein self-assembly and lipid binding in the folding of the potassium channel kcsa. *Biochemistry*, 47(7):2123–2133, 2008.

[63] G. Kamnesky, H. Shaked, and J. H. Chill. The distal c-terminal region of the kcsa potassium channel is a ph-dependent tetramerization domain. *Journal of Molecular Biology*, 418(3-4):237–247, 2012.

[64] D. M. Cortes and E. Perozo. Structural dynamics of the streptomyces lividans k+ channel (skc1): oligomeric stoichiometry and stability. *Biochemistry*, 36(33):10343–52, 1997.

[65] L. Heginbotham, E. Odessey, and C. Miller. Tetrameric stoichiometry of a prokaryotic k+ channel. *Biochemistry*, 36(33):10335–10342, 1997.

[66] N. Zerangue, Y. N. Jan, and L. Y. Jan. An artificial tetramerization domain restores efficient assembly of functional shaker channels lacking t1. *Proceedings of the National Academy of Sciences of the United States of America*, 97(7):3591–3595, 2000.

[67] Prasanna K Devaraneni, Jordan J Devereaux, and Francis I Valiyaveetil. In vitro folding of kvap, a voltage-gated k+ channel. *Biochemistry*, 50(48):10442–10450, 2011.

[68] S. B. Long, E. B. Campbell, and R. MacKinnon. Voltage sensor of kv1.2: Structural basis of electromechanical coupling. *Science*, 309(5736):903–908, 2005.

[69] Yufeng Zhou and Roderick MacKinnon. The occupancy of ions in the k+ selectivity filter: charge balance and coupling of ion binding to a protein conformational change underlie high conduction rates. *Journal of molecular biology*, 333(5):965–975, 2003.

[70] S. Jo, T. Kim, and W. Im. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *Plos One*, 2(9), 2007.

[71] S. Jo, T. Kim, V. G. Iyer, and W. Im. Software news and updates - charnim-gui: A web-based grraphical user interface for charmm. *Journal of Computational Chemistry*, 29(11):1859–1865, 2008.

[72] S. Jo, J. B. Lim, J. B. Klauda, and W. Im. Charmm-gui membrane builder for mixed bilayers and its application to yeast membranes. *Biophysical Journal*, 97(1):50–58, 2009.

[73] J. Lee, X. Cheng, J. M. Swails, M. S. Yeom, P. K. Eastman, J. A. Lemkul, S. Wei, J. Buckner, J. C. Jeong, Y. F. Qi, S. Jo, V. S. Pande, D. A. Case, C. L. Brooks, A. D.

MacKerell, J. B. Klauda, and W. Im. Charmm-gui input generator for namd, gromacs, amber, openmm, and charmm/openmm simulations using the charmm36 additive force field. *Journal of Chemical Theory and Computation*, 12(1):405–413, 2016.

[74] E. L. Wu, X. Cheng, S. Jo, H. Rui, K. C. Song, E. M. Davila-Contreras, Y. F. Qi, J. M. Lee, V. Monje-Galvan, R. M. Venable, J. B. Klauda, and W. Im. Charmm-gui membrane builder toward realistic biological membrane simulations. *Journal of Computational Chemistry*, 35(27):1997–2004, 2014.

[75] D. E. Shaw. Anton: A specialized machine for millisecond-scale molecular dynamics simulations of proteins. *Abstracts of Papers of the American Chemical Society*, 238, 2009.

[76] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, and K. Schulten. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[77] M. B. Ulmschneider, J. P. Ulmschneider, N. Schiller, B. A. Wallace, G. von Heijne, and S. H. White. Spontaneous transmembrane helix insertion thermodynamically mimics translocon-guided insertion. *Nature Communications*, 5, 2014.

[78] M. Tuckerman, B. J. Berne, and G. J. Martyna. Reversible multiple time scale molecular-dynamics. *Journal of Chemical Physics*, 97(3):1990–2001, 1992.

[79] M. P. Harrigan, M. M. Sultan, C. X. Hernandez, B. E. Husic, P. Eastman, C. R. Schwantes, K. A. Beauchamp, R. T. McGibbon, and V. S. Pande. Msmbuilder: Statistical models for biomolecular dynamics. *Biophysical Journal*, 112(1):10–15, 2017.

[80] R. T. McGibbon, K. A. Beauchamp, M. P. Harrigan, C. Klein, J. M. Swails, C. X. Hernandez, C. R. Schwantes, L. P. Wang, T. J. Lane, and V. S. Pande. Mdtraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophysical Journal*, 109(8):1528–1532, 2015.

[81] L. Molgedey and H. G. Schuster. Separation of a mixture of independent signals using time-delayed correlations. *Physical Review Letters*, 72(23):3634–3637, 1994.

[82] G. Perez-Hernandez, F. Paul, T. Giorgino, G. De Fabritiis, and F. Noe. Identification of slow molecular order parameters for markov model construction. *Journal of Chemical Physics*, 139(1), 2013.

[83] C. R. Schwantes and V. S. Pande. Improvements in markov state model construction reveal many non-native interactions in the folding of ntl9. *Journal of Chemical Theory and Computation*, 9(4):2000–2009, 2013.

[84] G. R. Bowman, D. L. Ensign, and V. S. Pande. Enhanced modeling via network theory: Adaptive sampling of markov state models. *Journal of Chemical Theory and Computation*, 6(3):787–794, 2010.

[85] M. C. Ahmed, E. Papaleo, and K. Lindorff-Larsen. How well do force fields capture the strength of salt bridges in proteins? *Peerj*, 6, 2018.

[86] F. Noe and C. Clementi. Kinetic distance and kinetic maps from molecular dynamics simulation. *Journal of Chemical Theory and Computation*, 11(10):5002–5011, 2015.

[87] F. Noe, R. Banisch, and C. Clementi. Commute maps: Separating slowly mixing molecular configurations for kinetic modeling. *Journal of Chemical Theory and Computation*, 12(11):5620–5630, 2016.

[88] F. I. Valiyaveetil, Y. Zhou, and R. MacKinnon. Lipids in the structure, folding, and function of the kcsa k+ channel. *Biochemistry*, 41(35):10771–7, 2002.

[89] Zakhar O Shenkarev, Ekaterina N Lyukmanova, Ivan O Butenko, Lada E Petrovskaya, Alexander S Paramonov, Mikhail A Shulepko, Oksana V Nekrasova, Mikhail P Kirpichnikov, and Alexander S Arseniev. Lipid–protein nanodiscs promote in vitro folding of transmembrane domains of multi-helical and multimeric membrane proteins. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1828(2):776–784, 2013.

[90] Cholpon Tilegenova, Spandana Vemulapally, Doris M Cortes, and Luis G Cuello. An improved method for the cost-effective expression and purification of large quantities of kcsa. *Protein expression and purification*, 127:53–60, 2016.

[91] Manasi P Bhate, Benjamin J Wylie, Ameer Thompson, Lin Tian, Crina Nimigean, and Ann E McDermott. Preparation of uniformly isotope labeled kcsa for solid state nmr: Expression, purification, reconstitution into liposomes and functional assay. *Protein expression and purification*, 91(2):119–124, 2013.

[92] Franz Hagn, Manuel Etzkorn, Thomas Raschle, and Gerhard Wagner. Optimized phospholipid bilayer nanodiscs facilitate high-resolution structure determination of membrane proteins. *Journal of the American Chemical Society*, 135(5):1919–1925, 2013.

[93] TK Ritchie, YV Grinkova, TH Bayburt, IG Denisov, JK Zolnerciks, WM Atkins, and SG Sligar. Reconstitution of membrane proteins in phospholipid bilayer nanodiscs. *Methods in enzymology*, 464:211–231, 2009.

[94] D. Lee, C. Hilty, G. Wider, and K. Wuthrich. Effective rotational correlation times of proteins from nmr relaxation interference. *Journal of Magnetic Resonance*, 178(1):72–76, 2006.

[95] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, M. Rance, and N. J. Skelton. Protein nmr spectroscopy: Principles and practice, 2nd edition. *Protein Nmr Spectroscopy: Principles and Practice, 2nd Edition*, pages 1–888, 2007.

[96] C. H. Cho, J. Urquidi, S. Singh, and G. W. Robinson. Thermal offset viscosities of liquid h2o, d2o, and t2o. *Journal of Physical Chemistry B*, 103(11):1991–1994, 1999.

[97] October, 13th 2011.

[98] M. P. Pond, A. Majumdar, and J. T. J. Lecomte. Influence of heme post-translational modification and distal ligation on the backbone dynamics of a monomeric hemoglobin. *Biochemistry*, 51(29):5733–5747, 2012.

[99] K. Takeuchi, H. Takahashi, S. Kawano, and I. Shimada. Identification and characterization of the slowly exchanging ph-dependent conformational rearrangement in kcsa. *Journal of Biological Chemistry*, 282(20):15179–15186, 2007.

[100] A. G. Komarov, C. A. Costantino, and F. I. Valiyaveetil. Engineering k+ channels using semisynthesis. *Methods Mol Biol*, 995:3–17, 2013.

[101] T. R. Sosnick, L. Mayne, R. Hiller, and S. W. Englander. The barriers in protein-folding. *Nature Structural Biology*, 1(3):149–156, 1994.

[102] T. R. Sosnick, L. Mayne, and S. W. Englander. Molecular collapse: The rate-limiting step in two-state cytochrome c folding. *Proteins-Structure Function and Bioinformatics*, 24(4):413–426, 1996.

[103] B. A. Krantz, L. Mayne, J. Rumbley, S. W. Englander, and T. R. Sosnick. Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *Journal of Molecular Biology*, 324(2):359–371, 2002.

[104] T. Kim and W. Im. Revisiting hydrophobic mismatch with free energy simulation studies of transmembrane helix tilt and rotation. *Biophys J*, 99(1):175–83, 2010.

[105] T. Kim, K. I. Lee, P. Morris, R. W. Pastor, O. S. Andersen, and W. Im. Influence of hydrophobic mismatch on structures and dynamics of gramicidin a and lipid bilayers. *Biophys J*, 102(7):1551–60, 2012.

[106] S. Uysal, V. Vasquez, V. Tereshko, K. Esaki, F. A. Fellouse, S. S. Sidhu, S. Koide, E. Perozo, and A. Kossiakoff. Crystal structure of full-length kcsa in its closed conformation. *Proc Natl Acad Sci U S A*, 106(16):6644–9, 2009.

[107] G. Kamnesky, O. Hirschhorn, H. Shaked, J. F. Chen, L. S. Yao, and J. H. Chill. Molecular determinants of tetramerization in the kcsa cytoplasmic domain. *Protein Science*, 23(10):1403–1416, 2014.

[108] D. E. Otzen. Mapping the folding pathway of the transmembrane protein dsbb by protein engineering. *Protein Engineering Design and Selection*, 24(1-2):139–149, 2011.

[109] T. W. Kahn and D. M. Engelman. Bacteriorhodopsin can be refolded from 2 independently stable transmembrane helices and the complementary 5-helix fragment. *Biochemistry*, 31(26):6144–6151, 1992.

[110] T. Marti. Refolding of bacteriorhodopsin from expressed polypeptide fragments. *Journal of Biological Chemistry*, 273(15):9312–9322, 1998.

[111] K. M. Visscher, J. Medeiros-Silva, D. Mance, J. P. G. L. M. Rodrigues, M. Daniels, A. M. J. J. Bonvin, M. Baldus, and M. Weingarth. Supramolecular organization and functional implications of k+ channel clusters in membranes. *Angewandte Chemie-International Edition*, 56(43):13222–+, 2017.

[112] P. Lague, M. J. Zuckermann, and B. Roux. Lipid-mediated interactions between intrinsic membrane proteins: Dependence on protein size and lipid composition. *Biophysical Journal*, 81(1):276–284, 2001.

[113] S. Katira, K. K. Mandadapu, S. Vaikuntanathan, B. Smit, and D. Chandler. Pretransition effects mediate forces of assembly between transmembrane proteins. *Elife*, 5, 2016.

[114] F. X. Zhou, H. J. Merianos, A. T. Brunger, and D. M. Engelman. Polar residues drive association of polyleucine transmembrane helices. *Proceedings of the National Academy of Sciences of the United States of America*, 98(5):2250–2255, 2001.

[115] H. Gratkowski, J. D. Lear, and W. F. DeGrado. Polar side chains drive the association of model transmembrane peptides. *Proceedings of the National Academy of Sciences of the United States of America*, 98(3):880–885, 2001.

[116] M. Hermansson and G. von Heijne. Inter-helical hydrogen bond formation during membrane protein integration into the er membrane. *Journal of Molecular Biology*, 334(4):803–809, 2003.

[117] J. P. Dawson, J. S. Weinger, and D. M. Engelman. Motifs of serine and threonine can drive association of transmembrane helices. *Journal of Molecular Biology*, 316(3):799–805, 2002.

[118] M. Eilers, S. C. Shekar, T. Shieh, S. O. Smith, and P. J. Fleming. Internal packing of helical membrane proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 97(11):5796–5801, 2000.

[119] K. R. MacKenzie and K. G. Fleming. Association energetics of membrane spanning alpha-helices. *Curr Opin Struct Biol*, 18(4):412–9, 2008.

[120] J. D. Lear, H. Gratkowski, L. Adamian, J. Liang, and W. F. DeGrado. Position-dependence of stabilizing polar interactions of asparagine in transmembrane helical bundles. *Biochemistry*, 42(21):6400–6407, 2003.

[121] R. T. Shealy, A. D. Murphy, R. Ramarathnam, E. Jakobsson, and S. Subramaniam. Sequence-function analysis of the k(+)-selective family of ion channels using a comprehensive alignment and the kcsa channel structure. *Biophysical Journal*, 84(5):2929–2942, 2003.

[122] Daniel Lingwood and Kai Simons. Lipid rafts as a membrane-organizing principle. *science*, 327(5961):46–50, 2010.

[123] Ken Jacobson, Ole G Mouritsen, and Richard GW Anderson. Lipid rafts: at a crossroad between cell biology and physics. *Nature cell biology*, 9(1):7, 2007.

[124] John F Hancock. Lipid rafts: contentious only from simplistic standpoints. *Nature Reviews Molecular Cell Biology*, 7(6):456, 2006.

[125] Linda J Pike. The challenge of lipid rafts. *Journal of lipid research*, 50(Supplement):S323–S328, 2009.

[126] John A Allen, Robyn A Halverson-Tamboli, and Mark M Rasenick. Lipid raft microdomains and neurotransmitter signalling. *Nature reviews neuroscience*, 8(2):128, 2007.

[127] Yang-Yi Fan, Lan H Ly, Rola Barhoumi, David N McMurray, and Robert S Chapkin. Dietary docosahexaenoic acid suppresses t cell protein kinase cθ lipid raft recruitment and il-2 production. *The Journal of Immunology*, 173(10):6151–6160, 2004.

[128] Kai Simons and Derek Toomre. Lipid rafts and signal transduction. *Nature reviews Molecular cell biology*, 1(1):31, 2000.

[129] Susan K Pierce. Lipid rafts and b-cell activation. *Nature Reviews Immunology*, 2(2):96, 2002.

[130] Steven Boeynaems, Simon Alberti, Nicolas L Fawzi, Tanja Mittag, Magdalini Polymenidou, Frederic Rousseau, Joost Schymkowitz, James Shorter, Benjamin Wolozin, Ludo Van Den Bosch, et al. Protein phase separation: a new phase in cell biology. *Trends in cell biology*, 2018.

[131] Clifford P Brangwynne, Christian R Eckmann, David S Courson, Agata Rybarska, Carsten Hoege, Jöbin Gharakhani, Frank Jülicher, and Anthony A Hyman. Germline p granules are liquid droplets that localize by controlled dissolution/condensation. *Science*, 324(5935):1729–1732, 2009.

[132] Edward WJ Wallace, Jamie L Kear-Scott, Evgeny V Pilipenko, Michael H Schwartz, Pawel R Laskowski, Alexandra E Rojek, Christopher D Katanski, Joshua A Riback, Michael F Dion, Alexander M Franks, et al. Reversible, specific, active aggregates of endogenous proteins assemble upon heat stress. *Cell*, 162(6):1286–1298, 2015.

[133] Joshua A Riback, Christopher D Katanski, Jamie L Kear-Scott, Evgeny V Pilipenko, Alexandra E Rojek, Tobin R Sosnick, and D Allan Drummond. Stress-triggered phase separation is an adaptive, evolutionarily tuned response. *Cell*, 168(6):1028–1040, 2017.