Molecular simulations of ion channel gating

Tanadet Pipatpolkai

Christ Church College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2021

To my dearest mother

who love me unconditionally from the very beginning.

Abstract

Ion channels are membrane proteins that facilitate ion permeation across the membrane down the concentration gradient. Recently, many structures of the ion channels were solved using cryo-electron microscopy. An advancement in structural determination and computational power has enabled ion channels to be thoroughly investigated at the molecular level using molecular dynamics (MD) simulation. Here, I used a range of methods, including coarse-grained MD, free energy perturbation (FEP) and atomistic MD simulation, to study the gating mechanisms of two ion channels, the ATP-sensitive K (K_{ATP}) channel and the Transmembrane 16A (TMEM16A) channel. I used CG-FEP to determine the binding affinity of phosphatidylinositol-4,5-bisphosphate (PIP_2) to the K_{ATP} channel and investigate how insulin secretory disorder effects channels PIP_2 affinity. In the light of collaborative experimental studies, I have also proposed the gating mechanism of how PIP_2, an activator of the channel, may compete, with ATP, an inhibitor of the K_{ATP} channel. Last, I have determined an additional gate at the outer pore of the TMEM16A channel, which is the binding site of the biphasic drug anthracene-9-carboxylate (A9C). Together, the work in this thesis has demonstrated the power of MD simulation to investigate the gating mechanisms of multiple ion channels and their interactions with the allosteric modulators.

Word count: 203 words

Declaration of authorship

I hereby certify that the work completed in this thesis is the original work of the author. The work in this thesis has also been previous published in:

1. **Pipatpolkai T.,** Usher S., Stansfeld P.J., Ashcroft F.M. **New insights into K_{ATP} channel gene mutations and neonatal diabetes mellitus.** *Nature Review Endocrinology* (2020). https://doi.org/10.1038/s41574-020-0351-y

2. **Pipatpolkai T.,** Corey R.A., Proks P., Ashcroft F.M. and Stansfeld P.J. **Evaluating Inositol phospholipid interactions with Inward Rectifier Potassium Channels and their Role in Disease**. *Communications Chemistry* **3,** 147 (2020). https://doi.org/10.1038/s42004-020-00391-0

3. **Pipatpolkai T.,** Usher S*., Vedovato N*. Corey R.A., Ashcroft F.M., and Stansfeld P.J., **The dynamic interplay of PIP_2 and ATP in the regulation of the K_{ATP} channel.** *bioRxiv*.(2021) https://doi.org/10.1101/2021.05.06.442933

4. **Pipatpolkai T**., Quetschlich D. and Stansfeld P.J. **From Bench to Biomolecular simulation: Potassium Channel Modulation by Lipids.** *Journal of Molecular Biology*. (2021) https://doi.org/10.1016/j.jmb.2021.167105

5. Dinsdale R.L., **Pipatpolkai T.***, Agostinelli E.*, Russell A.J., Stansfeld P.J. and Tammaro P. **The outer pore gate controls the pharmacology of the TMEM16A channel**. *Proceedings of the National Academy of Science* (2021). https://doi.org/10.1073/pnas.2023572118

The following projects were completed during the completion of DPhil but are not included in this thesis.

6. Moseng M.A.*, Lyu M.*, **Pipatpolkai T.**, Glaza P., Emerson C. C., Stewart P.L., Stansfeld P.J., Yu E.W. **Cryo-EM structures of CusA reveal a mechanism of metal-ion export.** *mBio* 12 (2) e00452-21 (2021) https://doi.org/10.1128/mBio.00452-21

7. Ansell T.B., Curran L., Horrell M. R., **Pipatpolkai T**., Letham S., Song W., Siebold C., Stansfeld P.J., Sansom M. S. P., Corey R. A. **Relative Affinities of Protein-Cholesterol Interactions from Equilibrium Molecular Dynamics Simulation**. *Journal of Chemical Theory and Computation*. (2021) https://doi.org/10.1101/2021.06.02.446704

8. Moseng M.A., Koumangoye R., **Pipatpolkai T.,** Lyu M., Klenotic P.A., Su C-C, Stansfeld P.J., Delpire E., Yu E.W. **Structural basis of transport and inhibition of the human NKCC1 cotransporter.** (manuscript in prep)

9. Panasawatwong A., **Pipatpolkai T.**[+], Tucker S. J.[+] **Molecular dynamics simulation reveals interaction between C-terminus of TREK-1 channel and phosphatidylinositol-4,5-bisphosphate.** (manuscript in prep)

* Authors contributed equally

[+] Corresponding authors

- Experimental studies completed in chapter 3 were carried out by Dr Peter Proks.
- Experimental studies completed in chapter 4 were carried out by Dr Sam Usher and Dr Natascia Vedovato
- Experimental studies completed in chapter 5 were carried out by Dr Ria Dinsdale and Dr Emilio Agostinelli.
- All data are reinterpreted by the author of the thesis.

No part of this thesis has been submitted for any other degree at this or any other university.

Acknowledgements

It was 2015 when I first walked into DPAG and chatted with Fran on the K_{ATP} channel. I was stunned by her intellect and wealth of knowledge. Since my undergraduate degree, working with Fran has been my goal, and I would like to thank her for being my supervisor. I remembered so well that I could not patch a single cell with K_{ATP} current for four months. I have found experimental work incredibly difficult. Yet, she has been one of my most incredible supporters throughout this journey. Her knowledge, energy and passion for science has influenced and brought me back up on many doomsdays. I would like to thank her experimental support throughout my DPhil for always listening to my bombardment of ideas in ion channels simulation. I would also like to thank her for her incredible patients in my written English and how she always tries her best to teach me how to write well. My English now is better than what I have started.

It was 2017 when I did my “Membrane Transport” option during my final year undergraduate degree, and I think that “This is just very fun”. Phill was given a lecture on ion channels, and with my passion for K_{ATP}, I think this could be an incredible journey. I approached Phill at the end of my undergraduate and asked him whether he would like to take me on board with the K_{ATP} channel project, and he said yes. Despite all that, I started with zero computational background. In fact, I was actually incredibly frustrated that by the end of my undergraduate, I couldn’t even render an image on PyMOL! I set sail on a journey to learn computational biochemistry, and Phill supports me from the

very beginning until the very end. I would always knock on his door, text him on Skype, drop him multiple e-mails and probably be quite annoying, but he would always be there to help me. His knowledge, ideas, and generosity in support are definitely among the key factors that have gotten me through this DPhil happily. I would like to thank Phill for all the knowledge, time and support he has provided. I would never found my enjoyment and career in computational biochemistry without Phill's generosity and kindness.

Collaborations have been one of the critical features of my DPhil, and I would like to thank Peter, Tasha and Sam for their support. I thank you all for the great conversation, ideas and all the fun scientific chat we have. I thank them for being such great friends, and we wouldn't have set sail on this fun journey without you. I would also like to thank Paolo, Ria and Emillio for our work on the TMEM16A channel. That was an enjoyable project, and I enjoyed every conversation and idea we have proposed and tested together.

SBCB and Ashcroft group have been my second home over the last 4 years. I am fortunate to find a place that I want to get up to work every day. I thank Mark Sansom in particular for his fruitful discussion in every JC/SysX, which has driven me out of my comfort zone and to think outside the box regularly. I would also like to thank Mabel Wong for her initiative and passion, which has got me to get into the world of computational biochemistry. I would like to thank Josh and Michael for being on this journey together with a lot of positive energy as a 4th year Stansfeld team. I would like to thank Robin for his mentorship in many projects – you have a lot of cool ideas! I would also like to

thank several current and ex-SBCB members: Vishal, Viwan, Gianni, Sophie, Nick, Fiona, Will Wu, Irfan, Owen, Keith, Shanlin, Anna, Wanling, Bertie and Patrick for their generous supports and all the fun conversation we had. I thank members of the Ashcroft group: Mike, Idoia, Gregor, Matthew, Lizzy and Tom, for all the great time and support I have when I was doing experimental work in the group. That has been very fun indeed!

Let ne admit one thing. I have struggled with my finance at the end of my undergraduates. I am very thankful for OXION Wellcome funding for their support, or otherwise, this whole thing would never have happened. I thank Stephen for his support of the programme, which has changed my life. I would like to thank all my OXION friend, Ali, Olof and Giacomo for always being there to support me and discuss fun science with some great wine and dinner. I also thank the Clarendon scholarship for the top-up funding for additional international fees.

Teaching has been the critical part both to keep my family's living comfortable and to maintain the state of my mentality. I thank Simon for the great opportunity and advice he has given to me, both in teaching and research. I would also like to thank Elena, Lynne, Maike, Jason, Nick, Naomi, Louis and John for their trust and support for me to mentor their students. I also thank Luke and Adisorn for being the great Part II student, yielding a publication for each of you. I thank all my students from 2017-2021 for being such fun and great mental support during my research. I learned how much I love mentoring and teaching alongside my research because of all of you.

I would like to thank all my Thai friends: Charlie, KB, Boom, Most, Nash, Ji, An, Jean, Pomme, Poon, Jane, Punpun, Ou and many others. I would like to thank Emma Louise Jennings, Simon Talbot and Kirstie Phipps for all their support since the very first day I arrived in England and remaining as my British family. Last, I would like to thank my mother, who has been my greatest supporter throughout this degree during such a difficult COVID-19 pandemic. This whole thing would have been impossible without her. Thank you.

Contents

List of figures

List of tables

Abbreviations

Å	Ångström
ABCC	Class C ATP-binding cassette
ADP	Adenosine diphosphate
ANAP	L-3-(6-acetylnaphthalen-2-ylamino)-2-aminopropionic acid
Ani9	2-(4-chloro-2-methylphenoxy)-N-[(2-methoxyphenyl)methylideneamino]-acetamide
ANO1	Anoctamin-1
atm	Atmosphere
AT-MD	Atomistic molecular dynamics
ATP	Adenosine triphosphate
ATPγS	Adenosine-5'-(γ-thio)-triphosphate
A9C	Anthracine-9-carboxylate
BAR	Bennett's Acceptance Ratio
BB	Backbone
ca.	circa (latin for *around*)
CaCC	Calcium-activated chloride channels
CF	Cystic fibrosis
CFTR	Cystic fibrosis transmembrane conductance regulator
CLC-1	Chloride channel 1
CG-MD	Coarse-grained molecular dynamics simulation
CG-FEP	Coarse-grained free-energy perturbation
CHAP	Channel Annotation Packages

CHI	Congenital hyperinsulinism
CPK	Corey, Pauling and Koltun (colouring scheme)
Cryo-EM	Cryo-electron microscopy
CV	Collective variable
ΔG	Gibb's free energy change
ΔH	Enthalpy change
ΔS	Entropy change
DEND	Developmental delay, epilepsy and neonatal diabetes
EC_{50}	Half maximal effective concentration
ER	Endoplasmic reticulum
FEP	Free-energy perturbation
FRET	Förster resonance energy transfer
fs	Femtosecond(s)
FVPP	Fluorovanadate and pyrophosphate
GDP	Guanosine diphosphate
GLUT2	Glucose transporter 2
GFP	Green fluorescence protein
GROMACS	Groningen machine for chemical simulations
HEK293T	Human embryonic kidney 293 cells containing the SV40 T-antigen
hKir channels	Human Kir channels
HT	Hypertension
I_a	Large current increase after A9C is removed
I_b	Fast current block after A9C application
IC_{50}	Half maximal inhibitory concentration

K	Potassium
K	Kelvin
K_{ATP} channel	ATP-sensitive potassium channel
k_b	Boltzmann's constant
K_d	Dissociation constant
K_d^{app}	Apparent dissociation constant
Kir	Inwardly rectifying potassium channel
Kir6.2*	Kir6.2 channel with W311ANAP co-expressed with SUR1
K_v	Voltage-gated potassium channel
K2P	Two-pore domain potassium channel
L0	Loop zero
LJ	Lennard-Jones
MBAR	Multistate Bennett's Acceptance Ratio
MD	Molecular dynamics
Mg	Magnesium
MONNA	N-((4-methoxy)-2-naphthyl)-5-nitroanthranilic acid
µs	Microsecond(s)
NBD	Nucleotide binding domain
nAChR	Nicotinamide acetylcholine receptor
NDM	Neonatal diabetes mellitus
nm	Nanometre
NPT	Constant number of molecules, pressure and temperature
ns	Nanosecond(s)
NVT	Constant number of molecules, volume and temperature
PA	Phosphatidic acid

PBC	Periodic boundary conditions
PC	Phosphatidylcholine
PDB	Protein data bank
PI	Phosphatidylinositol
PI4P	Phosphatidylinositol-4-phosphate
PIP_2	Phosphatidylinositol-4,5-bisphosphate
PLC	Phospholipase C
PME	Particle mesh Ewald
PNDM	Permanent neonatal diabetes mellitus
PO	Phosphate
P_{open}	Open probability
PS	Phosphatidylserine
POPC	Palmitoyl oleyl phosphatidylcholine
QM-MM	Quantum mechanics molecular mechanics
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
RRD	Remitting and relapsing diabetes
SC	Side chain
SEM	Standard error of the mean
SUR	Sulphonylurea receptor
T	Temperature
t	Time
TM	Transmembrane helix
TMD	Transmembrane domain
TMEM16	Transmembrane 16 protein

TNDM	Transient neonatal diabetes mellitus
TNP-ATP	2'-(or-3')-O-(Trinitrophenyl) Adenosine 5'-Triphosphate
VdW	Van-der-Waals
VSP	Voltage-sensitive phosphatase
WHAM	Weighted histogram analysis method
WT	Wild-type
WT-MetaD	Well-tempered metadynamics

Chapter 1

Introduction

"Nothing in life is to be feared, it is only to be understood. Now is the time to understand more, so that we may fear less."

Marie Salomea Skłodowska Curie

Part of the work in this chapters are published in:

Pipatpolkai, T., Usher, S., Stansfeld, P.J., Ashcroft F.M. **New insights into K_{ATP} channel gene mutations and neonatal diabetes mellitus.** *Nature Review Endocrinology* (2020). https://doi.org/10.1038/s41574-020-0351-y

Pipatpolkai T., Quetschlich, D. and Stansfeld P.J. **From Bench to Biomolecular simulation: Potassium Channel Modulation by Lipids.** *Journal of Molecular Biology*. (2021) https://doi.org/10.1016/j.jmb.2021.167105

Dinsdale R.L., **Pipatpolkai T.***, Agostinelli E.*, Russell A.J., Stansfeld P.J. and Tammaro P. **The outer pore gate controls the pharmacology of the TMEM16A channel**. *Proceedings of the National Academy of Science* (2021). https://doi.org/10.1073/pnas.2023572118

1.1 Plasma membrane

The cell is the basic unit of almost all life forms on Earth. The cell partitions its components from outside using a plasma membrane. This prevents leakage of the cellular components and regulates the transport of substances such as ions, small polar molecules or proteins (Alberts, 2015).

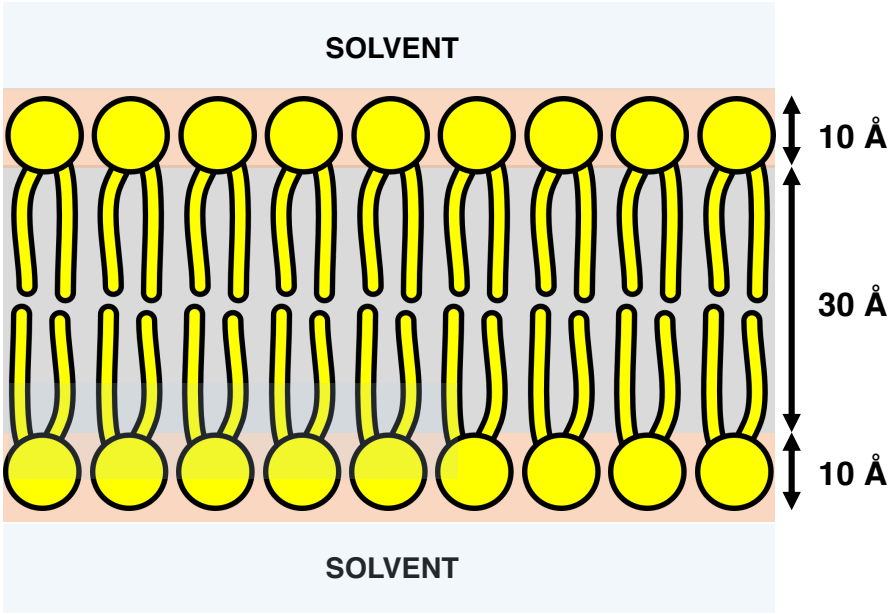


Figure 1.1 Phospholipid bilayer

Phospholipid molecules are shown in yellow. The solvent is represented using a light blue rectangle. The environment surrounding the phosphate headgroups is represented in orange. The centre of the bilayer is shown in dark grey. The length of the phosphate headgroup and the alkyl tail is approximated based on Nagle and Tristram-Nagle, 2000.

The plasma membrane is made of lipids. These lipids are classed into three groups; phospholipids, sphingolipids and cholesterol. The majority of the plasma membrane components are phospholipids, where the phosphate

headgroup faces the solvent, and the hydrophobic alkyl group forms a ~30 Å thick bilayer (Figure 1.1) (van Meer *et al.*, 2008). This specific arrangement of phospholipids is driven by entropy. Interactions between the hydrophobic tails of the lipid increase the disorder of the surrounding water molecules. An increase in the entropy drives the hydrophobic tails to pack against each other and form a phospholipid bilayer. An arrangement of the phospholipid bilayer provides three distinctive environments; water as a solvent, a phosphate headgroup (10 Å thick) (the water-oil emulsion layer) and a hydrophobic slab, which is ~30 Å (Nagle & Tristram-Nagle, 2000; Chiu *et al.*, 2001; Wang *et al.*, 2006).

The hydrophobic nature at the centre of the bilayer prevents the exchange of polar molecules across the membrane. The membrane hydrophobicity is crucial for the cell to regulate its osmolarity and its electrochemical gradient. However, small hydrophobic substances (such as dioxygen) can still permeate across the membrane (Hannesschlaeger *et al.*, 2019). For polar or charged molecules/ions to be transported, integral membrane proteins are required to mediate this process. In this chapter, we aim to address the composition of the phospholipid bilayer and how integral membrane proteins, especially potassium (K) channels and calcium-activated chloride channels (CaCC), transport their substrates across the bilayer.

1.1.1 Lipid composition

Membrane lipids provide a residence for membrane proteins and regulate their functions, such as ion channel gating and protein oligomerisation (Duncan *et al.*, 2019; Song *et al.*, 2021). The lipid composition of the eukaryotic plasma membrane includes glycerophospholipids (*ca.* 50%), sphingolipids (*ca.* 15%) and cholesterol (*ca.* 35%) (Casares *et al.*, 2019). However, lipid compositions are dynamically regulated through cellular activity and vary greatly among different cell types (Harayama & Riezman, 2018). Of the glycerophospholipids, phosphatidylcholine (PC) is usually the most abundant at the eukaryotic plasma membrane. However, other anionic phospholipids, such as phosphatidylinositol (PIP), phosphatidylserine (PS) and phosphatidic acid (PA), are also present in lower amounts (Boesze-Battaglia & Schimmel, 1997; van Meer *et al.*, 2008).

1.1.2 Integral membrane proteins

Integral membrane proteins are embedded in the lipid bilayer, spanning from the extracellular to the intracellular side of the membrane, and provide a communication platform from the outside to the inside of the cell. Commonly, these proteins are receptors, ion channels, transporters or porins and, therefore, involved in cellular communication and transport homeostasis. Because of their importance, membranes protein takes up to 25% of the mammalian genome. Their importance also makes them constitute up to 50% of the current pharmacological targets. (Overington *et al.*, 2006)

As the membrane environment is biphasic, specific chemical properties are needed for the protein to be favourably embedded. In this thesis, I will mainly be discussing the biophysical properties of α-helical membrane proteins.

The majority of the amino acids at the interface of the transmembrane region are hydrophobic as they are required to interact with the alkyl group of the phospholipid (Newport *et al.*, 2018). α-helical structures in the bilayer are stabilised by hydrogen bonding between amide groups within the α-helix backbone. The strength of an intramolecular hydrogen bond is stronger in the bilayer than in water due to a low dielectric environment in the hydrophobic core (Bowie, 2011). A lower dielectric constant in the membrane makes the formation of the α-helix in the bilayer more stable than a random coil or disordered structure.

Eukaryotic α-helical plasma membrane proteins are inserted in the bilayer using Sec translocase at the endoplasmic reticulum (ER) and trafficked to the plasma membrane via the Golgi apparatus and vesicles, respectively (Alberts, 2015). The transmembrane helical insertion is determined by the hydrophobicity of the peptide. In some cases, the hydrophobicity of the amino acid residues in an α-helix is large enough to allow insertion of positively charged sidechains, such as an arginine-rich S4 helix on voltage sensor domain in the voltage-gated potassium (K_v) channels (Bond & Sansom, 2007).

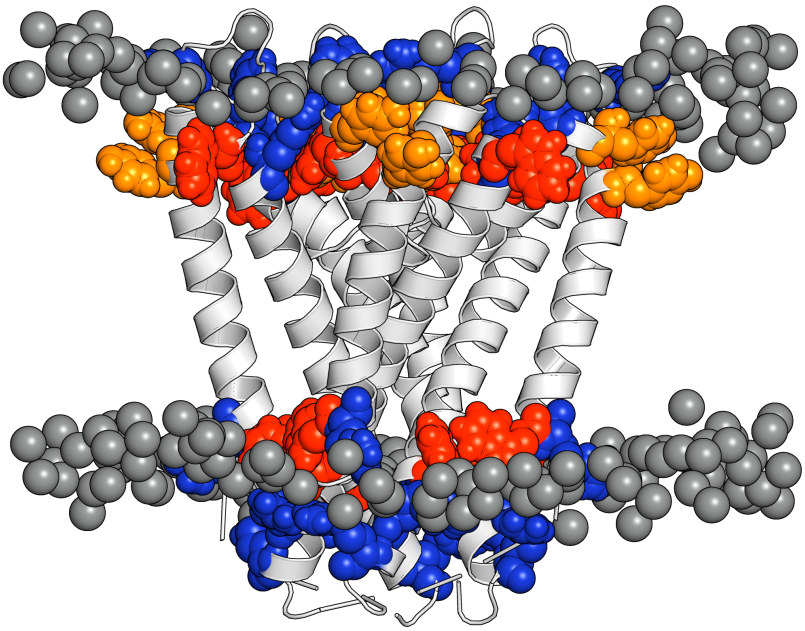


Figure 1.2 Interactions between KcsA and phosphate headgroups

Crystal structure of KcsA is shown in cartoon representation. The phosphorus atoms of the phosphate headgroup are shown as grey spheres. Arginine and lysine side chains are shown as blue spheres. Tryptophan side chains are shown as red spheres. Tyrosine side chains are shown as orange spheres.

The phosphate headgroups on phospholipids interact with amphipathic amino acids (tyrosine, tryptophan, arginine and lysine). Amphipathic amino acids contain both polar and non-polar functional groups in a single amino acid side chain. The hydrocarbons of the aromatic rings of tyrosine, tryptophan, arginine and lysine interact with the phospholipid alkyl tail and the positively/polar charged heads interact with the phosphate headgroups (Figure 1.2). These features of membrane proteins are helpful for initially predicting the topology of a protein in the phospholipid bilayer (Popot & Engelman, 2000; Newport *et al.*, 2018).

1.2 Ion channels

Ion channels are classes of protein that facilitate the transport of ions across the plasma membrane without an additional energy input to drive the transport process (Hille, 1978). Each ion channel exhibits three key characteristics; selectivity, permeation and gating. The ion permeation path has variations in charge and size in order to select specific ions. Ions can permeate through pore without a hydration shell, such as in pH gated bacterial K^+ channel (KcsA) (Doyle *et al.*, 1998), or with their hydration shell intact, as is the case of Transmembrane 16A (TMEM16A) channel (Le *et al.*, 2019*a*; Jia & Chen, 2021). I will discuss these channels in a later section.

Ion channels can exist in multiple conformations, but functionally they exist in two discrete states; opened and closed. Ion channels are opened and closed (gated) by ligands, pressure, temperature, membrane potential, drugs etc. (Alberts, 2015). The channel is closed either by a conformational change which then a) narrows the radius and/or b) increases the hydrophobicity of the pore (Lynch *et al.*, 2020).

The work in this thesis focuses on two membrane proteins; (i) the pancreatic ATP sensitive potassium (K_{ATP}) channel, which is a complex between an inwardly rectifying potassium (Kir) channel and an ABCC8/ABCC9 transporter (the sulfonylurea receptor, SUR1) and (2) the Anoctamin-1 or Transmembrane 16A (TMEM16A) calcium-activated chloride channel.

1.3 K_{ATP} channels

The ATP sensitive potassium (K_{ATP}) channel is critical for the regulation of glucose homeostasis as it is crucial in insulin secretion. Insulin is the only anabolic hormone in mammals and has a range of effects, such as an increase in muscle glycogen synthesis and inhibition of liver gluconeogenesis (Petersen & Shulman, 2018). The term “diabetes” is defined as elevated blood glucose due to insufficient insulin for the body’s needs. Thus, diabetes can be caused by insulin resistance at the target issue, or by failure at the pancreatic β cell to secrete insulin (due to β cell death, or defects in β cell electrical excitability) (Ashcroft & Rorsman, 2012).

At a stimulatory blood glucose concentration (>6 mM), glucose is transported into pancreatic β cells through the GLUT2 glucose transporter (Thorens, 2015). Glucose is metabolised to ATP through glycolysis and oxidative phosphorylation. Elevation of the cellular ATP concentration through glucose metabolism is sufficient to close the K_{ATP} channel. Closure of the K_{ATP} decreases cellular K^+ efflux, causing the membrane to depolarise (Ashcroft *et al.*, 1984). The depolarised membrane potential activates the voltage-gated calcium channel and triggers calcium influx. This then triggers the release of insulin from the insulin secretory vesicle into the bloodstream (Ashcroft, 2005).

Mutations in the K_{ATP} channel have been previously associated with neonatal diabetes (NDM) and congenital hyperinsulinism (CHI) (Ashcroft, 2007). Clinical mutations that prevent the channel from closing in response to

ATP are associated with NDM. The most severe mutations cause Developmental Delay, Epilepsy, Neonatal Diabetes, a condition known as DEND syndrome. The less severe mutations cause Permanent Neonatal Diabetes mellitus (PNDM) and Transient Neonatal Diabetes Mellitus (TNDM, or as Remitting and Relapsing Diabetes RRD) (Hattersley & Ashcroft, 2005). The symptoms of TNDM may not be apparent until adulthood (Gloyn *et al.*, 2005; Flanagan *et al.*, 2007).

Congenital hyperinsulinism (CHI) occurs when too much insulin is released from the pancreatic β-cell. Opposite to the NDM, CHI mutations cause the channel to be permanently closed, either because it is not present on the plasma membrane (Nessa *et al.*, 2015) or because it is not activated by MgADP (Ortiz & Bryan, 2015).

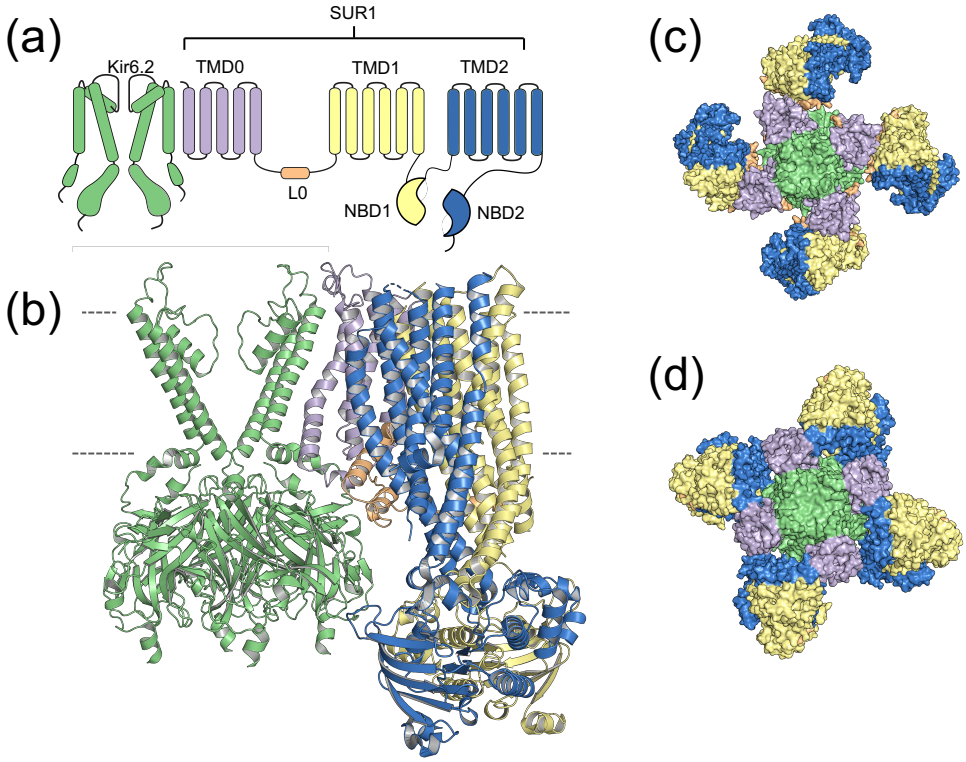


Figure 1.3 Structure of the K_{ATP} channel complex

(a) Transmembrane topology of Kir6.2 and SUR1. For clarity, only a single SUR1 subunit and two opposing Kir6.2 subunits are shown.
(b) Side view of the cryo-electron microscopy structures of the Kir6.2 tetramer (green) and a single SUR1 subunit (Protein Data Bank (PDB) accession number 6BAA). SUR1 is colour coded, with transmembrane domain 0 (TMD0) in purple, L0 in orange, TMD1 in yellow and TMD2 in blue. The dashed lines show the approximate location of the membrane.
(c),(d) Top view of the propeller (c) and quatrefoil (d) structures of the ATP-sensitive potassium channel (K_{ATP} channel) complex (PDB accession numbers 6BAA and 6C3O, respectively). Same colour code as (b).

Figures (a) and (b) are adapted from Puljung, 2018.

Figures are taken from Pipatpolkai et. al., 2020. The figures are produced by the author of the thesis.

The K_{ATP} channel is an octameric protein complex comprised of four copies of two different protein subunits: an inwardly rectifying potassium channel (Kir6.2), which forms a tetrameric pore complex, and a regulatory subunit, ABCC8 or sulfonylurea receptor 1 (SUR1) (Figures 1.3a,b) (Puljung, 2018). The Kir6.2 channel can only be trafficked to the membrane successfully when a) the Kir6.2 forms an octameric complex with SUR1 (Zerangue *et al.*, 1999); or experimentally when b) GFP is fused to the C-terminus of Kir6.2 (Schwappach *et al.*, 2000) or c) the retention signal peptide at the C-terminus of Kir6.2 is cleaved (ΔC construct) (Tucker *et al.*, 1997). The overall assembly of the K_{ATP} channel complex was initially identified at low-resolution by heavy metal electron microscopy (Mikhailov *et al.*, 2005). This was followed by atomic-resolution structures, determined by cryo-electron microscopy, of the complex in two major conformational states; quatrefoil and propeller (Figures 1.3c,d) (Lee *et al.*, 2017; Li *et al.*, 2017; Martin *et al.*, 2017*b*, 2017*a*, 2019; Wu *et al.*, 2018). The channel activity is regulated by nucleotides: binding of ATP to the Kir6.2 subunits inhibits channel activity, and the binding of MgATP and MgADP to the SUR1 subunits activates the channel (Puljung, 2018).

1.3.1 Kir6.2 or KCNJ11 - Inwardly rectifier K^+ channel

The pancreatic K_{ATP} channel was first identified as an inwardly rectifying potassium channel whose closure is induced by glucose (Ashcroft *et al.*, 1984), and subsequently, by ATP (Sakura *et al.*, 1995). The Kir6.2 subunit was subsequently shown to be a member of the inward rectifier K channel (Kir) family. All inwardly rectifying K channels conduct current at a negative potentials and cease to do so as the membrane becomes more depolarised, due to voltage-dependent block by intracellular polyamines and divalent cations (Matsuda *et al.*, 1987; Lopatin *et al.*, 1994). Kir6.2 has high amino acid sequence identity to its homologue, Kir6.1, a ubiquitously expressed Kir channel (Inagaki *et al.*, 1995; Sakura *et al.*, 1995). Kir6.2 has a long 72 amino acid cytoplasmic N-terminus domain, followed by a canonical potassium channel motif (M1 and M2 transmembrane helices and a pore helix) and a 218 amino acid cytoplasmic C-terminus (Sakura *et al.*, 1995). The canonical potassium channel motif is highly conserved amongst all K^+ channel. The pore-lining helix (H5) has a highly conserved selectivity filter TXGXG sequence where the carbonyl backbone forms a tight water-filled pore, similar to KcsA for potassium permeation (Doyle *et al.*, 1998). The sequence at the selectivity filter of Kir6.2 is slightly different from KcsA, as the tyrosine (TYGYG) is swapped with the phenylalanine in the Kir6.2 (TVGFG) (Sakura *et al.*, 1996; Doyle *et al.*, 1998).

The channel also exhibits three gates for ion permeation, one in the transmembrane region (a.k.a. the helix bundle crossing gate or HBC gate); one

in the cytoplasmic domain; and one at the selectivity filter. (1) The HBC gate blocks water and ion permeation using F168. Mutations at residues near this gate (such as C166F (Tammaro *et al.*, 2008) or I167L (Shimomura *et al.*, 2007) increases the channel open probability, decreases channel ATP sensitivity and are associated with the DEND syndrome. (2) The cytoplasmic gate blocks ion permeation using I296 (Martin *et al.*, 2017*a*). Mutations of I296 to L (Proks *et al.*, 2005) or at residues in the gating loops such as T293N (Shimomura *et al.*, 2009), increase channel open probability, reduce channel inhibition by ATP and are also associated with DEND syndrome. (3) The selectivity filter is a fast gate, which governs the opening and closing of the channel within the burst (Proks *et al.*, 2001). Mutations at the selectivity filter such as V127T do not affect the channel burst duration, or the long closed-time duration. To date, mutations at the fast gate are also not found to be associated with NDM (De Franco *et al.*, 2020). Currently, all atomic resolution structures solved had ATP bound to Kir6.2, a closed HBC gate in the transmembrane region, and an open cytoplasmic gate (Lee *et al.*, 2017; Martin *et al.*, 2017*a*; Wu *et al.*, 2018).

1.3.2 SUR1 or ABCC8 - ABC transporter as a regulator

Sulfonamides were first used as bacteriostatic agents in the Second World War. It was then discovered that they cause hypoglycaemia as the side effect, and thus, the drugs were used clinically to treat patients with type 2 diabetes (Ashcroft & Ashcroft, 1992). Later work then showed that tolbutamide stimulates depolarisation of the pancreatic β-cell as it blocks the K_{ATP} channel through the sulfonylurea receptor (Sturgess *et al.*, 1985) Later on, SUR1 was isolated by photolabeling of iodinated glibenclamide (Kramer *et al.*, 1988) and cloned successfully (Aguilar-Bryan *et al.*, 1995). The cloning of the gene suggested that SUR1 is the member of the ATP binding cassette (ABC) transporters with high sequence similarity to the cystic fibrosis transconductance regulator (CFTR) and the multidrug resistance protein (MRP1) (Aguilar-Bryan *et al.*, 1995). SUR1 contains 17 TM helices divided into three domains, denoted as TMD0, TMD1 and TMD2, where TMD1 and TMD2 form a canonical ABC transporter motif (Aguilar-Bryan *et al.*, 1995) (Figure 1.3a). Structurally, cryo-EM structures of SUR1 are very similar to those of other ABC transporters and they have two different nucleotide-binding sites (denoted as NBS1 and NBS2).

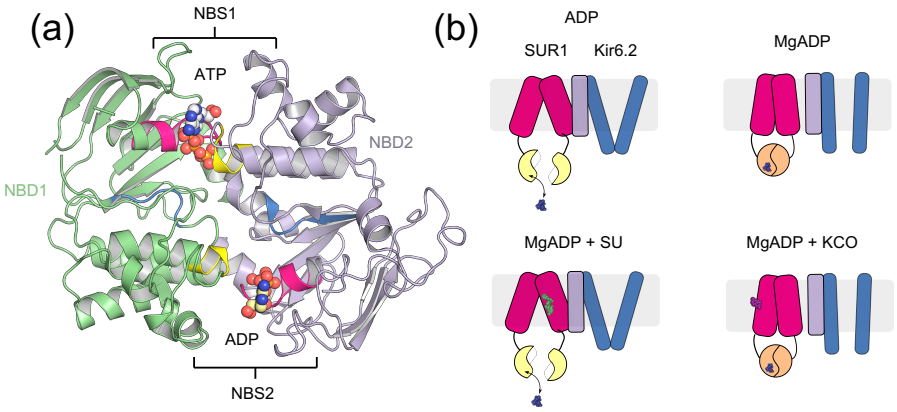


Figure 1.4 The nucleotide-binding sites on SUR1.

(a) The stimulatory nucleotide-binding sites (NBSs) on SUR1 are formed at the interface of nucleotide-binding domain 1 (NBD1) (green) and NBD2 (lilac). The Walker A motif is shown in magenta, the Walker B motif is shown in blue and the ATP-binding cassette signature sequence is shown in yellow. MgATP is shown bound to NBS1 and MgADP is shown bound to NBS2. Protein Data Bank accession number 6C3P.

(b) Activation of the ATP-sensitive potassium channel (K_{ATP}) channel complex by magnesium nucleotides. Top left: In the absence of Mg^{2+}, ADP rapidly binds and unbinds to SUR1, but has no effect on channel activity. Top right: The presence of Mg^{2+} causes NBD dimerization, which induces a conformational change that leads to opening of the Kir6.2 pore. Bottom left: Binding of sulfonylureas (SU) to SUR1 causes dimer dissociation and prevents MgADP stimulation. Bottom right: binding of K^+ channel openers (KCO) to SUR1 stabilizes the NBD dimer and supports MgADP stimulation.

Figures are taken from Pipatpolkai et. al., 2020. The figures are produced by the author of the thesis.

Based on the cryo-EM structure, when MgATP and MgADP bind to their nucleotide binding sites on the SUR1, SUR1 transits to an occluded state (Lee *et al.*, 2017). SUR1 is required for channel activation by MgADP (Tucker *et al.*, 1997). Both nucleotide binding domains contain a canonical WALKER-A motif, a WALKER-B motif and a P-loop, responsible for ATP binding similar to other ABC transporters (Figure 1.4a). The binding of K channel openers (KCOs) such as diazoxide, slow down the dissociation of the nucleotides from SUR1 (Puljung *et al.*, 2019). This means that diazoxide stabilises the occluded state of the SUR1 (Figure 1.4b) (Puljung *et al.*, 2019). In contrast, binding of sulfonylureas such as tolbutamide speeds up the dissociation of MgATP from SUR1. This infers that tolbutamide stabilises an inward open conformation of the SUR1 subunit and closes the K_{ATP} channel (Figure 1.4b) (Puljung *et al.*, 2019).

1.3.3 Inhibition of K_{ATP} channel by ATP

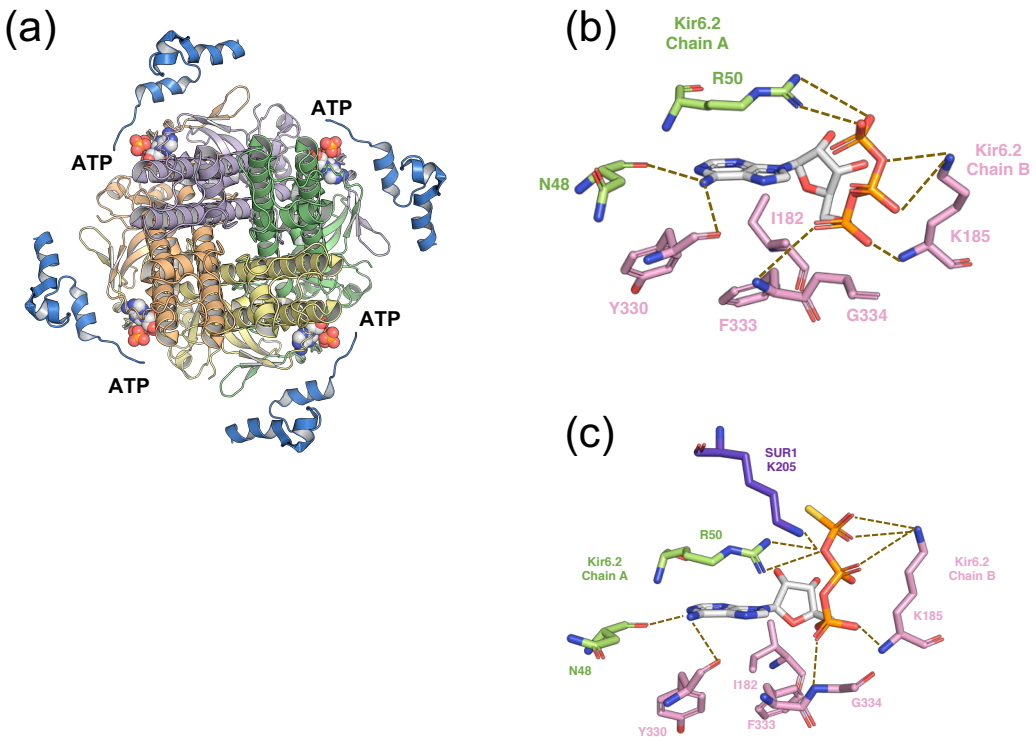


Figure 1.5 The inhibitory ATP-binding site.

(a) Top view of the ATP-sensitive potassium channel (K_{ATP} channel) tetramer showing the four ATP-binding sites on Kir6.2. Each Kir6.2 subunit is shown in a different colour, residues 200–240 in L0 of SUR1 are in blue and ATP is coloured by atom (carbon, white; nitrogen, blue; oxygen, red; phosphorous, orange). Protein Data Bank (PDB) accession number 6BAA.
(b) Close-up of the ATP-binding site, showing residues that participate in ATP binding; these are colour coded according to the Kir6.2 subunit they come from. PDB accession number 6BAA.
(c) A second close-up of the ATP-binding site. As before, the figure shows residues that participate in ATP binding; these are colour coded according to the Kir6.2 subunit they come from. This structure was obtained in the presence of adenosine 5′-(γ-thio) triphosphate (ATPγS). PDB accession number 6JB1.

Figures are taken from Pipatpolkai et. al., 2020. The figures are produced by the author of the thesis.

The ability of Kir6.2 to sense the local intracellular ATP concentration is crucial for its function as a cellular and metabolic sensor. The ATP binding site on Kir6.2 was determined before the release of the structure, using electrophysiology and homology models generated from the C-terminus of Kir3.1. These predicted binding sites were then confirmed by the cryo-EM structure (Figure 1.5a) (Lee *et al.*, 2017; Martin *et al.*, 2017*a*; Wu *et al.*, 2018). The ATP-binding site lies at the interface between N-terminus and the C-terminus close to the phospholipid bilayer interface. N48 and R50 from one subunit stabilise the γ phosphate group of ATP, whilst K185 from an adjacent subunit helps stabilise the β-phosphate (Dabrowski *et al.*, 2004; Martin *et al.*, 2017*a*). The short L0 helix stretch stabilises the geometry of the binding pocket on the CL3 loop from the SUR1 subunit, including K205 stabilization of the γ phosphate (Usher *et al.*, 2020) (Figure 1.5b,c).

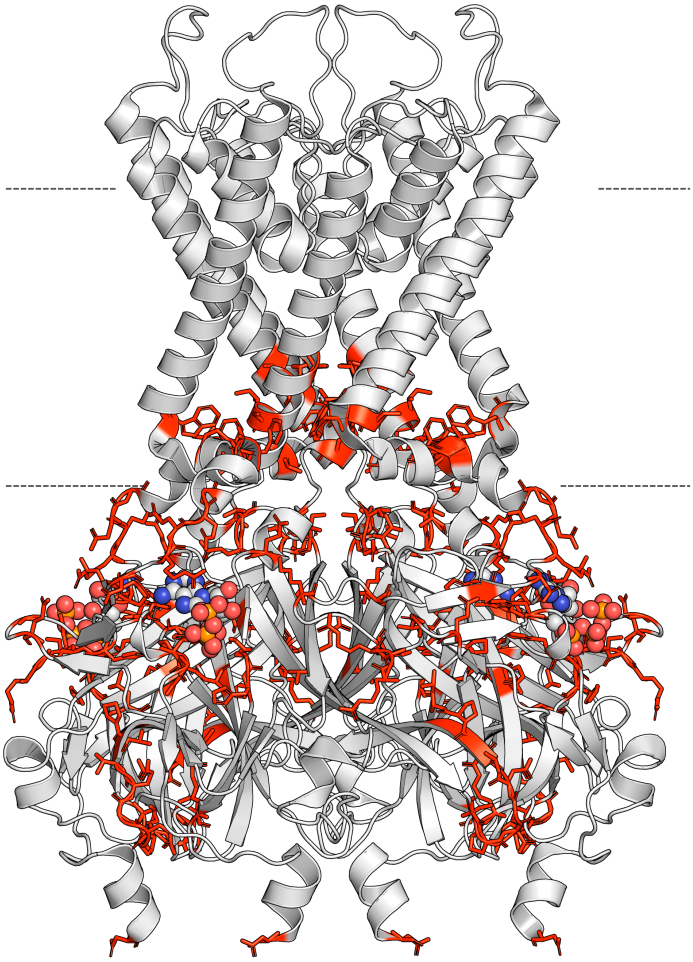


Figure 1.6 Location of Kir6.2 residues altered in neonatal diabetes mellitus.

Kir6.2 tetramer showing residues altered in neonatal diabetes mellitus as red sticks. ATP is shown as spheres coloured by atom (carbon, white; nitrogen, blue; oxygen, red; phosphorous, orange). The dashed lines show the approximate location of the membrane.

This figure is taken from Pipatpolkai et. al., 2020. This figure is produced by the author of the thesis.

Neonatal diabetes mutations in residues at or near the ATP binding site may reduce or completely abolish channel inhibition by ATP (Figure 1.6) (reviewed in Pipatpolkai *et al.*, 2020). In this section, I will discuss a subset of the mutations considered in the review. A mutation hotspot is located at R50, which forms a salt bridge with the γ-phosphate on ATP based on the cryo-EM structure (Lee *et al.*, 2017; Martin *et al.*, 2017*a*; Wu *et al.*, 2018). R50 is associated with three mutations; R50P, R50G and R50Q, which result in two variations of NDM; DEND with R50P (Massa *et al.*, 2005) and R50G (Suzuki *et al.*, 2007; Fraser *et al.*, 2012), and TNDM with R50Q (Shimomura *et al.*, 2006). Both R50P and R50Q cause a reduction in K_{ATP} channel sensitivity to ATP inhibition. By replacing the positive charge (arginine) with a neutral charge amino acid (proline), this is likely going to abolish K_{ATP} channel interaction with ATP. Hence, the channel will fail to close in response to an elevation in an intracellular ATP concentration.

Several other key NDM mutations such as I182V (Gloyn *et al.*, 2005), K185Q (Shimomura *et al.*, 2010) are located on the ATP binding site on the C-terminus of Kir6.2 (Figure 1.6). These residues also form part of the ATP binding pocket based on the cryo-EM structures. An interaction between a positive amine group on K185 and a β-phosphate on ATP is likely required to stabilise an interaction on the interface (Martin *et al.*, 2017*a*).

1.3.4 Activation of Kir and K_{ATP} channel by PIP_2

Many K^+ channels are functionally modulated by specific interactions with lipids from the surrounding membrane. Phosphatidylinositols (PI) such as phosphatidylinositol-4,5-bisphosphates (PIP_2) are glycolipids synthesised based on an attachment of an inositol sugar to PA in the phosphatidic acid pathway (Rajala & Anderson, 2010; Dickson & Hille, 2019). Lipids, in particular PIP_2, were first identified as activators of the ATP-sensitive K^+ (K_{ATP}) channel (Hilgemann & Ball, 1996; Baukrowitz *et al.*, 1998; Shyng & Nichols, 1998). Here, an increase in the level of PIP_2 – by either activation of phospholipase C (PLC) or direct application of PIP_2 – was shown to reduce channel inhibition by ATP (Hilgemann & Ball, 1996; Baukrowitz *et al.*, 1998; Shyng & Nichols, 1998). This could be due to direct binding site competition between ATP and PIP_2 or to allosteric modulation of the ligand-induced opening and closing events (MacGregor *et al.*, 2002; Pipatpolkai *et al.*, 2021).

Later work has shown that PIP_2 activates other Kir channels as well as the K_{ATP} channel. Application of PIP_2 to Kir1, Kir2, Kir3 and Kir6 channels has been shown to increase the size of the current of an excised patch and also elevates their open probability (P_{open}) (Lopes *et al.*, 2002; Haider *et al.*, 2007; Lacin *et al.*, 2017). Interestingly, PIP_2 binding to the Kir channel homologue (KirBac1.1) closes the channel, despite the absence of PIP_2 in bacterial membranes (Enkvetchakul *et al.*, 2005). Mutations at residues predicted to be involved in PIP_2 binding to the K_{ATP} channel produce a reduction in the size of the current in an excised patch (Shyng *et al.*, 2000). Mutations that decrease PIP_2 binding to the Kir6.2 channel also reduce channel burst time duration and

result in a faster *run-down*. Run-down is defined as the time-dependent decline in current in an excised patch (Ribalet *et al.*, 2006). The molecular mechanism of *run-down* has been proposed to be associated with PIP_2 disassociation from its binding site or depletion of the local concentration of PIP_2, which ultimately causes the channel to shut (Suh & Hille, 2008). One method to reduce the rate of *run-down* is to apply a fluoride vanadate and pyrophosphate (FVPP) solution, which inhibits the degradation of PIP_2 by phosphatases (Friedman, 1993). Together, this increases the local PIP_2 concentration in the patch and enhances PIP_2-channel interactions in the excised patch (Huang *et al.*, 1998).

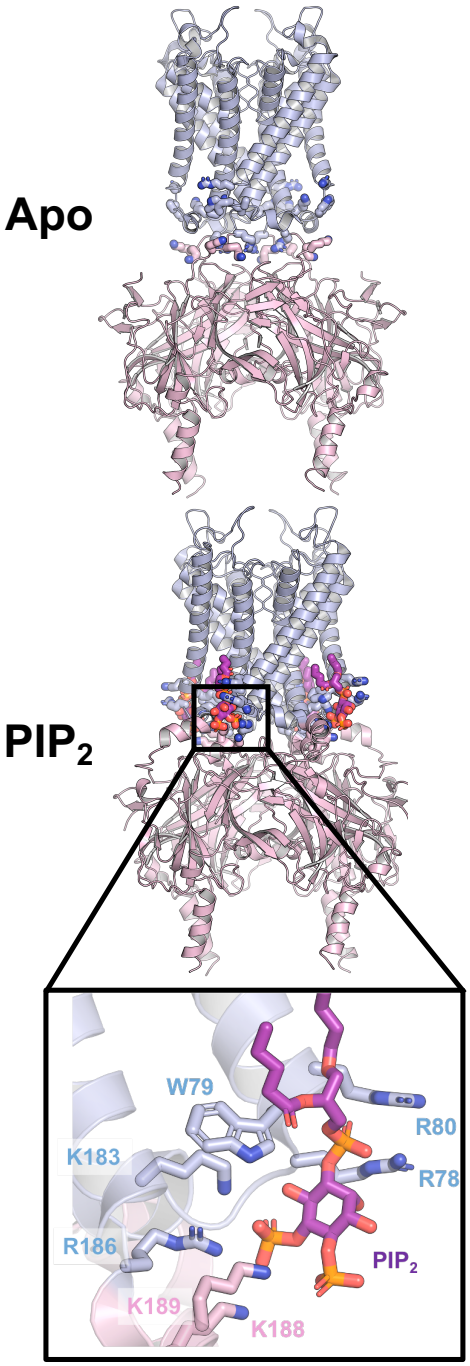


Figure 1.7 Crystal structures of chicken Kir2.2 with and without PIP_2

Atomic structure of the cKir2.2 channel in Apo (PDB ID: 3JYC) (Tao *et al.*, 2009) and PIP_2-bound state (PDB ID: 3SPI) (Hansen *et al.*, 2011). The pore domain of the cKir2.2 is shown in blue and the cytoplasmic domain is shown in pink. PIP_2 is shown as purple stick. All residues within 4 Å of the PIP_2 headgroup are represented as sticks without hydrogen atoms.

This figure is taken from Pipatpolkai et. al., 2021. This figure is produced by the author of the thesis.

Numerous studies attempted to structurally determine the PIP_2 binding site on the Kir channels, including the K_{ATP} channel. In 2011, X-ray structures of Kir2.2 and GIRK (Kir3.2) channels revealed well-defined densities for PIP_2 (Hansen *et al.*, 2011; Whorton & MacKinnon, 2011). These structures highlight the critical amino acid residues coordinating the bound PIP_2 lipid and show how PIP_2 induces conformational changes and drives the opening of the channel gate. The PIP_2 binding site on a Kir channel is predominantly coordinated by basic residues (lysine and arginine), which engage with the phosphate groups (Figure 1.7). These residues are located in both the N-terminal and C-terminal domains of the protein, with the binding site being formed between adjacent subunits. In these structures, both arginine and lysine are predicted to be protonated. The protonation states of arginine and lysine could be predicted further using propKa (Olsson *et al.*, 2011). Otherwise, recent study has demonstrated that free energy perturbation (FEP) could be used to calculate the free energy contribution from the protonation of the basic residues in the PIP_2 binding site. In this thesis, I assumed that all basic residues located at the PIP_2 binding site are protonated (Wu *et al.*, 2019). At the N-terminal end of TM1 of Kir2.2, residue W79 engages with the inositol sugar ring and lipid tails, while R78 and R80 form hydrogen bonds with the 1' phosphate. In the C-terminal domain, K183, R186 and K188 of the C-linker interact with the 5' phosphate and K189 coordinates both 4' and 5' phosphates (Hansen *et al.*, 2011).

Molecular dynamics (MD) simulations enable the unbiased prediction of lipid-binding sites to Kir channels within a membrane environment (Chavent *et*

al., 2016). Due to limits in simulation time and slow diffusion of lipids, coarse-grained (CG) MD simulations are regularly applied to study lipid binding. These simulations usually commence with lipids unbiasedly placed around an ion channel, either as a preformed bilayer or in random orientations to allow membrane self-assembly. As part of the simulation, the lipids and protein are allowed to freely diffuse, with specific lipid-protein interactions recorded as the simulation proceeds (Stansfeld *et al.*, 2009; Duncan *et al.*, 2020). From these simulations, the nature of the binding site can be quantified by several factors, such as residue-lipid contacts, the relative residence time for the lipid within the binding site and membrane deformation around the protein (Duncan *et al.*, 2020). These methods have been shown to excellently agree with structurally determined sites on the chicken Kir2.2 (Duncan *et al.*, 2020). The method has also been used to predict the binding site on human Kir6.2 (Stansfeld *et al.*, 2009).

Conformational changes of Kir channels induced by PIP_2 were proposed by comparing the bound and apo-states of the structures. The binding of PIP_2 to a Kir channel introduces a 6 Å upward conformational change and rotation of the C-terminal domain (Figure 1.7). Similar structural changes have also been observed in the bacterial Kir channel, KirBac, but this closes the channel instead (Clarke *et al.*, 2010). Very recent molecular simulations show that the binding of PIP_2 facilitates opening of the Kir3.2 channel. The work showed that during a 1 μs simulation, F192 moves further away from the same residue in the other subunits. This movement makes the HBC gate large enough for permeations of hydrated K^+ ions. (Bernsteiner *et al.*, 2019)

1.4 Calcium-activated chloride channel (CaCC)

1.4.1 Historical perspective of CaCC

Chloride (Cl^-) channels conduct permeation of Cl^- across the membrane. Chlorine ions are crucial in inhibitory neuronal signalling as the cellular chloride influx usually brings the membrane potential down to a more negative voltage across the membrane. It can also trigger the efflux of chloride ions and membrane depolarisation, depending on the intracellular chloride concentration (Jentsch & Pusch, 2018). Initially, calcium-activated chloride channels (CaCC) were discovered in *Xenopus* oocyte (Miledi, 1982; Barish, 1983). When the sperm fuses with the oocyte, this triggers calcium release inside the cell, which in turn opens CaCC on the oocyte. The efflux of the chloride ion triggers membrane depolarisation and prevents fertilization by other sperm (Wozniak *et al.*, 2018). Approximately 20 years after its discovery, the CaCC channel was cloned from *Xenopus* oocytes and identified as anoctamin-1 or TMEM16A (Yang *et al.*, 2008; Schroeder *et al.*, 2008). These channels are diversely expressed in mammalian cells, including neurones and the epithelial cell of the trachea (Huang *et al.*, 2009).

After the discovery of TMEM16A, the topology of the protein was predicted (Yang *et al.*, 2008). It was initially predicted to have 8 transmembrane helices, and hence the protein was given the name an (for anion) octa (for 8) min-1 (for being the first to be discovered). Another name for this protein family is the TMEM (TransMEMbrane protein) 16 family. There are 10 other proteins in the anoctamin family, and hence, the name is either Ano1-10 or TMEM16A-

K (Yang *et al.*, 2008). Interestingly, some proteins in this family are not ion channels, but are lipid scramblases instead. Some proteins in the TMEM family, such as TMEM16K, can function both as a lipid scramblase and as a chloride channel (Bushell *et al.*, 2019). In this thesis, I have worked with the protein TMEM16A or Ano1. For consistency throughout the thesis, I will use the TMEM16x naming system and use the term TMEM16A.

1.4.2 Structures, function and gating of TMEM16 family protein

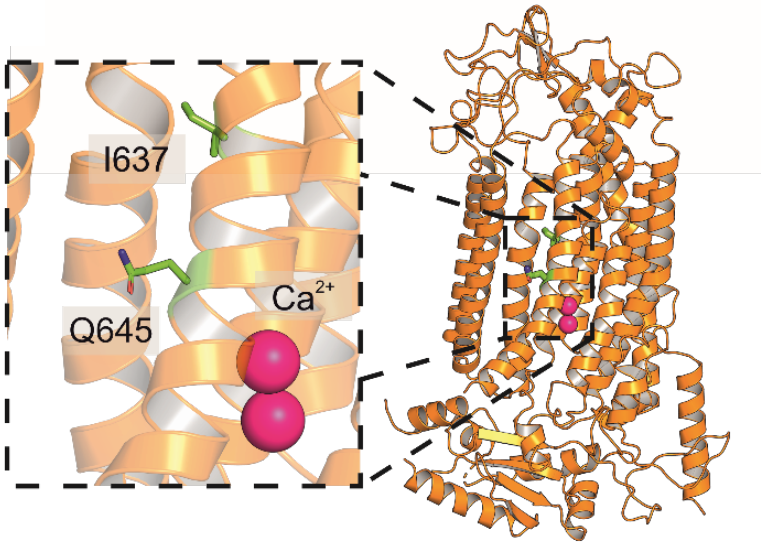


Figure 1.8 Cryo-EM structure of TMEM16A

The cryo-EM structure of TMEM16A with bound Ca^{2+} shown in pink (PDB ID: 5OYB). The left panel represents a zoom-in of transmembrane helix (TM) 6 with two key residues at the gate of the TMEM16A channel : I637 (steric gate) and Q645 (electrostatic gate) in green.

This figure is taken from Dinsdale et. al., 2021. This figure is produced by the author of the thesis.

In 2014, the first TMEM16 family protein structure was obtained by X-ray crystallography from a fungus, *Nectria haematococca* (nhTMEM16) (Brunner *et al.*, 2014). Surprisingly, it was then discovered that the protein does not possess 8 TM helices but 10 helices instead (Brunner *et al.*, 2014). The structure was crystallised as a dimer, similar to previous experimental studies which also suggested that the protein in the TMEM16 family also physiologically functions as a dimer (Tien *et al.*, 2013). Functionally, nhTMEM16 is solely a lipid scramblase, and not a chloride channel (Brunner *et al.*, 2014). In 2017, the

structure of TMEM16A was solved simultaneously by two groups in a calcium-free closed state (Paulino *et al.*, 2017*a*), and a calcium bound closed state (Figure 1.8) (Dang *et al.*, 2017; Paulino *et al.*, 2017*a*). TMEM16A is also activated by intrinsic cellular activators such as PIP_2 (Le *et al.*, 2019*a*). Recent molecular simulation studies have proposed that the TMEM16A channel requires both PIP_2 and calcium for full channel activation (Jia & Chen, 2021). To determine the cryo-EM structure of the open state of the channel, an activating mutation was made in the channel gate (I551A). However, even with an ambiguous density of PIP_2 at its putative binding site, the cavity of the channel is still closed in the presence of calcium (Lam *et al.*, 2021).

The cryo-EM structure of TMEM16A shows that the calcium-binding site is located at the lower portion of the TM6 (Dang *et al.*, 2017; Paulino *et al.*, 2017*a*)(Figure 1.8). The binding of calcium to TM6 is critical to stabilise the conformation of TM6 that promotes channel opening (Peters *et al.*, 2018). Previous work has shown that mutations at the TM6 (I637A, Q645A) promotes channel opening in an absence of calcium (Peters *et al.*, 2018). Interestingly, the predicted path for chloride permeation through TMEM16A is similar to the path for lipid permeation through scramblases such as mammalian TMEM16K (Bushell *et al.*, 2019) or fungal (*Aspergillus fumigatus*) afTMEM16 (Falzone *et al.*, 2019). A recent study has also proposed the possibility that lipids are involved in chloride permeation through the TMEM16F channel, which can function both as an anion channel and a lipid scramblase (Kostritskii & Machtens, 2021).

1.4.3 Permeation and gating of chloride channels

Permeation of chloride ions through an ion channel pore is chemically complex. Some chloride channels, such as the nicotinamide acetylcholine receptor (nAChR), or Cystic fibrosis transmembrane conductance regulator (CFTR), possess positively charged residues such as arginine and lysine at the pore entrance (Corringer *et al.*, 1999; Zhou *et al.*, 2008). These positively charge residues generates an electric field that attracts chloride ions to their binding site in the pore cavity (Stauffer & Karlin, 1994). In TMEM16A a similar effect is achieved through calcium binding to TM6. The presence of calcium in its binding site is proposed to generate an electrostatic field that lures chloride ions to a binding site within the pore. This electrostatic field is called an "electrostatic gate" (Lam & Dutzler, 2018).

Another mechanism that has been commonly shown in multiple chloride channels such as bestrophins or CLC-1 is to use a phenyl group (Dutzler *et al.*, 2002; Kane Dickson *et al.*, 2014). Chloride ions may interact with the edge of the phenyl group, which has a partial positive (δ+) charge (Thomas *et al.*, 1982). The δ+ charge on phenylalanine or tyrosine attracts and drives chloride ion transport down the chloride electrochemical gradient. In some cases, such as in the 5-HT_3 receptor, the edge of an alkyl group (C-H bond) has been proposed to interact with the chloride ion, allowing the ion to move along the hydrophobic amino acid side chains (Lisbjerg *et al.*, 2015; Klesse *et al.*, 2020).

1.5 Aims

In this thesis, I aimed to use molecular dynamics simulations to investigate ion channel gating. The work focuses on two ion channels; the K_{ATP} channel and the TMEM16A channel. The following questions are asked throughout the work in the different chapters of this thesis.

Chapter 3

a) Given that we now have many recent structures of the K_{ATP} channel, can we improve our prediction of the PIP_2 binding site? How does the PIP_2 binding site on the K_{ATP} channel differ from that of other Kir channels?

b) Given the slightly different sequence of the putative PIP_2 binding sites in different Kir channels, how do these affect the channel's PIP_2 affinity?

c) We know that PIP_2 activates the K_{ATP} channel more than PI4P, PI and PC. Can we explain this observation based on molecular simulations and free energy calculations? How does PIP_2 differ from di-C8 in terms of binding free energy?

d) Based on the known repertoire of NDM and CHI mutations, can we identify any mutations which may influence the affinity of PIP_2 binding to the K_{ATP} channel? Can we test our hypotheses derived from MD simulations using coarse-grain MD?

e) Does SUR1 affect the K_{ATP} channel affinity to PIP_2?

Chapter 4

a) Given the PIP_2 binding site predicted by our CG simulations, how do atomistic simulations differ from the CG simulations in Chapter 3?

b) What is the behaviour of ATP in the atomistic simulation in the absence of SUR1?

c) What happens to ATP and PIP_2 molecules and their binding sites when both molecules are bound to the Kir6.2 channel?

d) What is the effect of a gain of function mutation caused by changes in PIP_2 binding on K_{ATP} channel gating by ATP? Do these affect the binding of ATP or the transduction process by which ATP binding leads to channel opening?

Chapter 5

a) How does TMEM16A conduct Cl^- permeation across the membrane?

b) What are the conformational changes in the TMEM16A channel when Ca^{2+} is removed from its binding site? How does the channel gate?

c) How does anthracene-9-carboxylic acid (A9C) activate and inhibit the TMEM16A channel?

d) How do some mutations at the TM6 region of the pore stabilise the open conformation of the channel?

Chapter 2

Methods and Theory

"I was supposed to be a doctor. I realized what I really wanted to do was understand biological systems, and the only way to do this was to do chemistry and physics. Because all of what goes on inside of us is governed by chemistry and physics."

Martin Karplus

Part of the work in this chapters are published in:

Pipatpolkai T., Quetschlich, D. and Stansfeld P.J. **From Bench to Biomolecular simulation: Potassium Channel Modulation by Lipids** *Journal of Molecular Biology*. (2021) https://doi.org/10.1016/j.jmb.2021.167105

2.1 Introduction

The main technique in the thesis is molecular dynamics (MD) simulation, which has been used at both atomistic and coarse-grained (CG) level of representation. MD simulation is a single molecule computational tool which computes changes in the motion of particles over a short timestep, allowing the system to sample conformational dynamics. An increase in protein structure availability due to cryo-electron microscopy and improvements in computational hardware and software has allowed the more complex systems to be studied at longer timescales.

This section aims to introduce scientific concepts behind MD simulation, such as bonded and non-bonded interactions, restraints, barostats, and thermostats. Throughout this thesis, the molecular simulations were carried out by the software GROMACS (Abraham *et al.*, 2015). Simulations of membrane proteins are difficult due to the biphasic environment, with lipids and water. This adds to the size and the complexity of the system and therefore reduces the speed of the calculation. I will look into methods that speed up the simulations, including virtual sites and coarse-grained molecular dynamics simulation.

To simulate a particle, three components are needed. (1) the co-ordinate or the structure of the molecule you wish to simulate. (2) the forcefield, which describes the parameters of each particle (or atom) being simulated. (3) Newton's equation of motion, which is used to calculate interactions between each particle in a box into a distance moved over a short timestep. In this section of a thesis, I will explore these concepts mathematically.

2.2 Newton's equation of motion

In any MD simulation, I treated all atoms as particles attach to springs (bond). This is the primary assumption of classical Newtonian mechanics. First, I consider Newton's equation of motion for a simulation box with N interacting number of atoms.

$$m_i \frac{\partial^2 \boldsymbol{r_i}}{\partial t^2} = \boldsymbol{F_i}, i = 1 \ldots N.$$

where m is the mass of the particle, $\boldsymbol{r_i}$, is the displacement of a single particle i and t represents the time. Thus, the double derivative of the displacement with respect to time is an acceleration ($\Sigma \boldsymbol{F_i} = m_i \boldsymbol{a_i}$) and $\boldsymbol{F_i}$ are the forces exerted on the particle. At a constant temperature and pressure, I can update the co-ordinate of each particle over the short timesteps. The animated motions of co-ordinates over a period of time are called a 'trajectory'.

Next, I consider the energy within the system. As the forces are the negative of the derivatives of a potential energy function, V, in a simulation where the temperature and the pressure were kept constant.

$$\boldsymbol{F_i} = -\frac{\partial V}{\partial \boldsymbol{r_i}}$$

As shown above, an acceleration is a double partial derivative of the displacement, $\boldsymbol{r_i}$. To get the position, I integrate the acceleration over a short timestep. In GROMACS, a leap-frog algorithm is used as a default in order to

integrate this function. The leap-frog algorithm updates the position and the acceleration at the time t = t, and update the velocity at time $t - \frac{1}{2}\Delta t$ as:

$$\boldsymbol{v}\left(t + \frac{1}{2}\Delta t\right) = \boldsymbol{v}\left(t - \frac{1}{2}\Delta t\right) + \frac{\Delta t}{m}\boldsymbol{F}(\boldsymbol{t})$$

$$\boldsymbol{r}(t + \Delta t) = \boldsymbol{r}(t) + \Delta \mathrm{t}\mathbf{v}(\mathrm{t} + \frac{1}{2}\Delta t)$$

This gives a similar result to the Vertlet-scheme algorithm denoted below:

$$\boldsymbol{r}(t + \Delta t) = 2\boldsymbol{r}(t) - r(t - \Delta t) + \frac{1}{m}\boldsymbol{F}(\boldsymbol{t})\Delta t^2 + O(\Delta t^4)$$

Thus, with the leap-frog integrator, the velocity could be updated with the force calculated based on the potential energy function, $V(r)$ at a half-time point. This means that the calculation is symmetrical on both side of time. The symmetrical nature of the algorithm allows that if I reverse the timestep, I will end up at the same time-point as my initial configuration. In addition, the leap frog integrator is also symplectic, which means that energy of the the system is conserved. (unlike the Euler or Runge-Kutta integrator which are not symplectic). The mathematical detail of symplectic integration is beyond the scope of this thesis. The second advantage of leapfrog is the simplicity of the calculation based on the half-time. As there are many terms in the calculation of the potential energy function (will be discussed in the next section), it is important to diverge the computational cost towards computing the force instead of updating the velocity and displacement at a smaller timestep. In this thesis, I am aiming to conduct a simulation over a long period of time, and therefore other integration methods such as truncated Taylor expansion would not be appropriate as the timestep accounting for the calculation would be too small, and thus, would be very computationally costly.

In GROMACS, given the timescale of our simulation, I can only sample a trajectory over time to obtain the nearest local energy minimum from an initial structure, not the global minimum of the whole energy landscape. In the case where an inappropriate initial structure was given, the simulated structure can be trapped at non-representative local energy minimum. Thus, the quality of the initial structure is important to ensure we sample the most physiological state. In our studies, I used the "steepest descent" algorithm to compute for a direction with the greatest negative energy change. The energy minimisation protocol converges when the energy change is negligible, or forces interacting between a pair of particles are below the specified cut-off.

2.3 Forces, interactions and forcefield

In the standard MD simulation, we categorise forces into 2 categories as follows:

a) Non-bonded: Lennard-Jones (van der Waals) and Coulombic interactions. These are calculated based on the radii and the charges of an atom of interest.

b) Bonded: As atoms are bonded together by forces like a spring. I can subdivide these forces into covalent bond-stretching, angle-bending, proper and improper dihedrals.

In this section, I will discuss how each category of forces is parameterised in the simulation.

2.3.1 Non-bonded interaction

GROMACS sums up all pairs of interaction forces $\boldsymbol{F}(r)$ and the potential $V(r)$. Here, I explore different types of interactions as follow:

a) Lennard-Jones (LJ) interaction

The LJ potential between two atoms are:

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

The $C_{ij}^{(12)}$ and $C_{ij}^{(6)}$ are constants depending on the property of an atom. This could be simplified to ϵ_{ij} and σ_{ij} to provide an easier set of constants to work with. The term r_{ij}^{12} represents repulsion between the nuclei of two atoms as they approach one other, whilst r_{ij}^{6} term represents an attractive force between the nuclei and the surrounding electrons. These interactions are arising from the induced dipole interaction between two atoms. The relationship between the distance between two particles r can be represented in the graph in Figure 2.1.

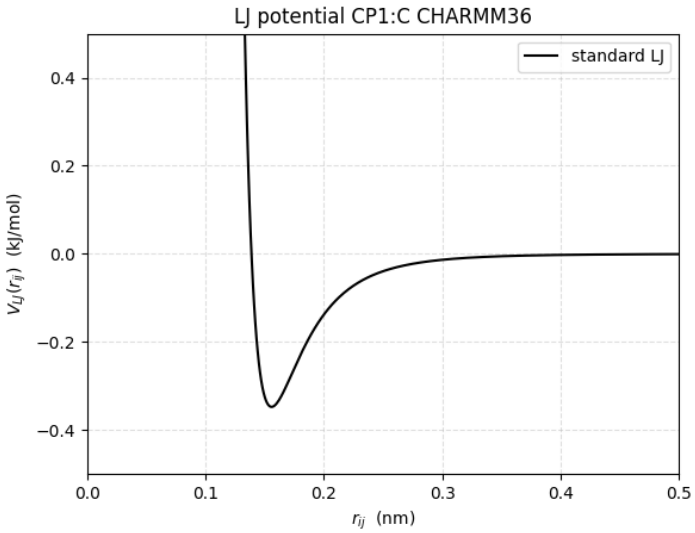


Figure 2.1 Lennard-Jones potential between CP1 and C atom types

A calculated LJ potential between two atom types (CP1 and C) based on CHARMM36 forcefield parameter using equation $V_{LJ}(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\epsilon_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right)$ where ϵ_{ij} = 0.347450500075 and σ_{ij} = 0.138767581228.

From the potential energy, this could be integrated to give forces of

$$\boldsymbol{F_i}(\boldsymbol{r_{ij}}) = \left(12\frac{C_{ij}^{(12)}}{r_{ij}^{13}} - 6\frac{C_{ij}^{(6)}}{r_{ij}^{7}}\right)\frac{\boldsymbol{r_{ij}}}{r_{ij}}$$

In the CHARMM36 forcefield, σ and ϵ term are specified for each pair of interactions.

b) Coulombic interactions

Coulombic interactions between two charged particles are defined as

$$V_{ij} = \frac{1}{4\pi\epsilon_0}\frac{q_i q_j}{\epsilon_r r_{ij}} = k\frac{q_i q_j}{\epsilon_r r_{ij}}$$

where k = 138.935458 kJ mol^{-1} nm e^{-2} which is a product of $\frac{1}{4\pi\epsilon_0}$ where ϵ_0 is the standard dielectric constant. An interaction between like charges lead to a repulsion between two particles, denoted as a positive $V_{ij}(r)$ whereas interaction between two opposite charge lead to an attraction, denoted by a negative $V_{ij}(r)$. The relationship between the energy and the distance between two particles are denoted in Figure 2.2.

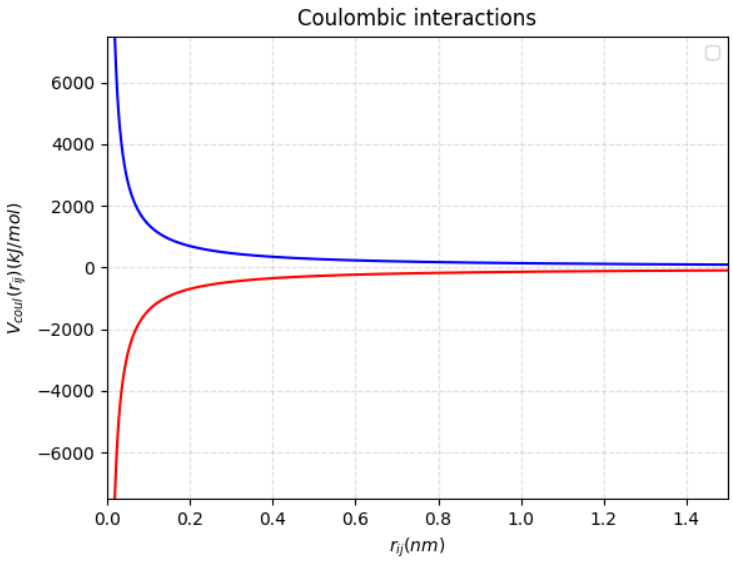


Figure 2.2 Coulombic potential between two particles

A calculated coulombic potential between two particles with the charge of ±1. The blue line represents the potential energy of an interaction between like charge (repulsion) between 0.0 to 1.4 nm calculated based on $V_{ij} = k\frac{q_i q_j}{\epsilon_r r_{ij}}$ where k = 138.935458 kJ mol^{-1} nm e^{-2} and two charges has the same sign. The red line represents the potential energy of an interaction between opposite charge (attraction).

With this, I can derive the force out as:

$$\boldsymbol{F_i}(\boldsymbol{r}_{ij}) = -k\frac{q_i q_j}{\epsilon_r r_{ij}^2}\frac{\boldsymbol{r_{ij}}}{r_{ij}}$$

As shown in the figure, coulombic interaction lasts over a very long distance and hence if all particles in the box are included in the calculation, this makes the process very slow. There are two ways to speed up the process. One is to define an electrostatic cut-off at 1.2 nm, where any interactions beyond 1.2 nm from a particle are not calculated. The other is to use Particle Mesh Ewald (PME) to calculate long-range electrostatics. A full mathematical background of PME will not be discussed. In principle, this method assigns charges to a 3D grid space using cardinal B-spline interpolation, which is then Fourier transformed with a 3D FFT algorithm. This allows the energy term to be calculated, and after the calculation of energy at the grid point, I calculated the force that a single atom experience. It is important to discuss the cut-off distance, and in this case, Verlet cut-off was used instead of just a fixed distance and therefore, the calculation of the PME is modified so that the force at the cut-off is approaching zero.

To optimise the calculation at the short-range interaction to be compatible with PME, the coulombic equation was modified to:

$$V_{ij} = k\frac{q_i q_j}{r_{ij}}\, erfc(\beta r_{ij})$$

β is the parameter which provides a relationship between the direct space and the reciprocal space from Fourier transformation and $erfc(x)$ is the complementary error function.

2.3.2 Bonded interaction

Bonded interactions can be divided into four key terms in the forcefield - bond stretching (fixed distance between two points and energy), bond angle (fixed angle between three points and energy), proper dihedral (fixed dihedral between four points) and the improper dihedral which is a set of forces that ensure that either planarity or chirality is maintained (Figure 2.3a). All of these interactions are held together by a harmonic potential (Figure 2.3b).

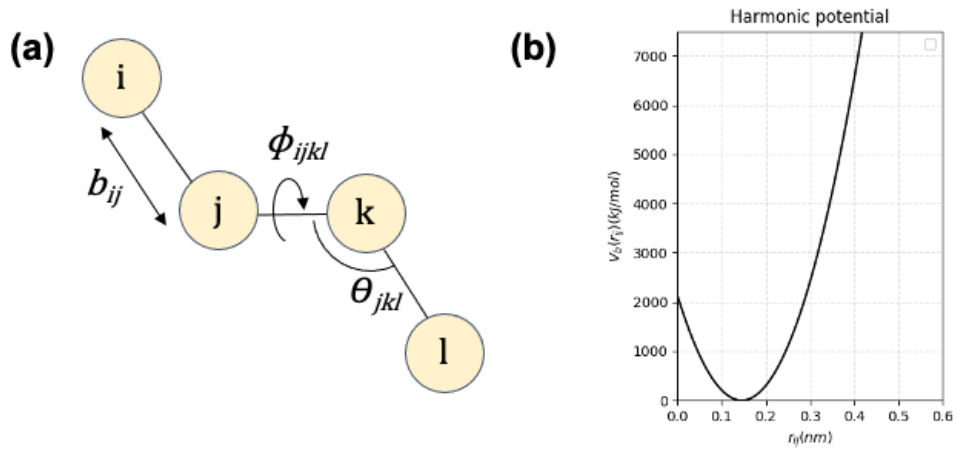


Figure 2.3 Bonded interactions between atoms

(a) Bonded interactions between 4 atoms (i, j, k and l). The bond length between i and j, are denoted as b_{ij} The angle between j, k and l are denoted as θ_{jkl}. The proper dihedral angles are denoted as ϕ_{ijkl}. Improper angles are not shown in the scheme.
(b) A plot representing harmonic potential between two atoms at a fixed distance calculated based on $V_b(r_{ij}) = \frac{1}{2}k_{ij}^b(r_{ij} - b_{ij})^2$. The bond length ($b_{ij}$) is set to 0.1455 nm with the force constant (k_{ij}^b) of 200832 kJ/mol/nm.

a) Bond stretching - two atoms, i and j have an ideal bond distance b_{ij}.

The potential energy $V_b(r_{ij})$ can be calculated as :

$$V_b(r_{ij}) = \frac{1}{2}k_{ij}^b(r_{ij} - b_{ij})^2$$

where r_{ij} is the distance between two particles. Thus, $\left(r_{ij} - b_{ij}\right)^2$ provide the changes in the distance upon bond stretching, k_{ij}^b represents the force constant, which holds the bond together. Thus, the force is simply a differential of this, which is:

$$F_i(r_{ij}) = -k_{ij}^b\left(r_{ij} - b_{ij}\right)\frac{\boldsymbol{r_{ij}}}{r_{ij}}$$

b) Angle - involved three atoms: i, j and k. They have an optimal angle of θ_{ijk}^0 . Thus, the potential energy is:

$$V_a(\theta_{ijk}) = \frac{1}{2}k_{ijk}^{\theta}\left(\theta_{ijk} - \theta_{ijk}^0\right)^2$$

The force calculation is, however, less simple. It may be calculated from the differential of $V_a(\theta_{ijk})$ by splitting three vectors into three terms, i, j and k, allowing the following to be obtained:

$$\boldsymbol{F}_i = -\frac{dV_a(\theta_{ijk})}{d\boldsymbol{r_i}}$$

$$\boldsymbol{F}_k = -\frac{dV_a(\theta_{ijk})}{d\boldsymbol{r_k}}$$

and by solving two vectors, this gives us:

$$\boldsymbol{F_j} = -\boldsymbol{F_i} - \boldsymbol{F_k}$$

and thus:

$$\theta_{ijk} = \arccos\frac{\boldsymbol{r}_{ij} \cdot \boldsymbol{r}_{kj}}{r_{ij} \cdot r_{kj}}$$

c) Proper dihedral is periodic with the angle ϕ is the angle between two planes made of ijk and jkl planes where 0 defines when i and l are on the same side. The potential between the two dihedrals is defined as

$$V_d(\phi_{ijkl}) = k_\phi(1 + \cos(n\phi + \phi_s))$$

Given that ϕ_s is the defined dihedral and k_ϕ is the force constant.

d) Improper dihedral is defined as harmonic potential and thus, discontinuous. The harmonic restraint of the dihedral is defined as:

$$V_{id}(\xi_{ijkl}) = \frac{1}{2}k_\xi(\xi_{ijkl} - \xi_0)^2$$

where ξ_0 represents the defined optimal dihedral and k_ξ is defined as a force constant to maintain the harmonic restraint.

Thus, I can sum all the energy of bonded interactions as:

$$V_{bond} = \frac{1}{2}k_{ij}^b(r_{ij} - b_{ij})^2 + \frac{1}{2}k_{ijk}^\theta(\theta_{ijk} - \theta_{ijk}^0)^2 + \frac{1}{2}k_\xi(\xi_{ijkl} - \xi_0)^2 + k_\phi(1 + \cos(n\phi + \phi_s))$$

2.3.3 Application of restraints

Throughout this thesis, I restrained the position of the protein using two types of restraints. The first was a harmonic position restraint, where I restrained the position of the xyz co-ordinates of the molecule as shown in Figure 2.3. The energy results from breaking the restraint is defined as:

$$V_{pr}\left(\boldsymbol{r}_i\right) = \frac{1}{2} k_{pr} \left|\boldsymbol{r}_i - \mathbf{R}_i\right|^2$$

and therefore, forces in a single dimension can be defined as:

$$F_{pr}\left(\boldsymbol{r}_i\right) = -k_{pr}\left(\boldsymbol{r}_i - \mathbf{R}_i\right)$$

Thus, the minimum is only fixed at one point with the harmonic position restraint.

In the free energy calculations, I applied a flat bottom restraint to allow particle overlap at closer distance. The flat-bottom restraints provide a wider range of energy well, allowing the lipid to remain in its position after the transformation from one lipid to PC (Pipatpolkai *et al.*, 2020*a*). Thus, the lipid would not experience any forces of restraint until it diffuses away at distance r_{i} (Figure 2.4).

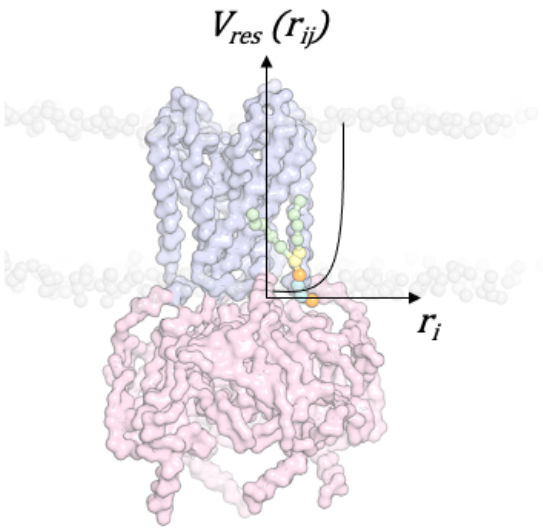


Figure 2.4 Application of flat bottom restraint on protein

The backbone of the Kir6.2 protein is represented in a coloured surface. The pore domain of Kir6.2 is shown in light blue. The cytoplasmic domain of Kir6.2 is shown in light pink. The PIP_2 molecule is shown as coloured sphere. The PIP_2 molecule is restrained with the flat bottom restrain shown as a graph representing a function between the energy ($V_{res}\,(\boldsymbol{r}_i)$) and the distance between protein and lipids (r_i). The PIP_2 headgroup is held at a maximum of 8 Å between the backbone of the PIP_2 binding residues before it experiences the restraint force at 1000 kJ/mol/nm^2. The graph represents the relationship between the distance between protein and lipids (r_i) and is overlaid on top of the Kir6.2.

2.4 Maintaining the simulations

To simulate particles in a box, I need to ensure that both the temperature and the pressure are maintained. I will also discuss how periodic boundary conditions (PBC) are applied to ensure that the molecules would be freely diffused. In this section, I will discuss these ideas in a greater depth.

2.4.1 Periodic boundary conditions

Imagine you are trying to simulate the system in a solvated box with water and lipids, and a molecule that is freely diffusible. At a certain point in the simulation, the molecule will approach the face of the simulation box as the simulation space is finite. This finite space may introduce the potential artefact causes by the interaction between the molecules of interest and the edge of the box. Applying PBC allows the molecules to be simulated in an infinite space, as it would diffuse to the other side of the box instead (Figure 2.5). For example, if the protein moves to the left side of the box, it will reappear to the right side and vice versa. The distance between the protein and the edge of the box is kept to 20 Å for minimum-image convention to prevent an interaction between the protein with itself.

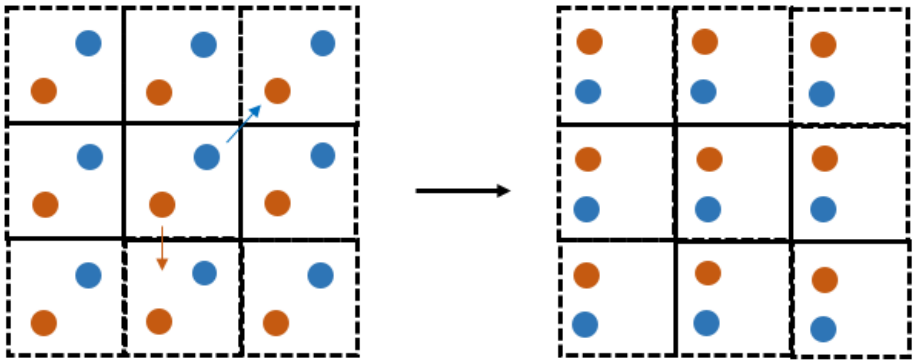


Figure 2.5 Periodic boundary condition

Diagram show periodic boundary condition of two particles (orange and blue). The particle moves by the direction shown in the arrow of its corresponding colour to the position on the right panel. Once the particle moves from its initial position, it reappears on the other side of the box as shown on the right.

2.4.2 Thermostat

In the MD simulation, the temperature of the simulation is calculated based on the total kinetic energy in the simulation system box space. Given that most of my simulations apply either an NPT (number of molecules, pressure and temperature are constant) or an NVT (number of molecules, box volume and temperature are constant) ensemble, I apply a thermostat so that the temperature of the system is maintained. As the temperature (and the energy) of the system fluctuates, the thermostats will also ensure that the fluctuation of the temperature remains at a sensible size. In this thesis, I applied the velocity rescaling (v-rescale) thermostat. This thermostat is an extension of the previously commonly used thermostat (Berendsen thermostat), where random forces are added to maintain the distribution of the kinetic energy in the system (Lemak & Balabaev, 1994). However, the v-rescale thermostat has added consideration of relaxation time into the quantity of force added to the system, allowing the overall dynamics of the system to not be influenced by the thermostat itself (Bussi *et al.*, 2007).

2.4.3 Barostat

During the NPT ensemble, the pressure inside the box needs to be kept constant. This is achieved through fluctuations in the box vectors. Throughout this thesis, pressure coupling was applied semi-isotopically (applying pressure coupling in the xy and z directions independently) to the simulation. To ensure that the pressure is maintained at 1 atmosphere, two algorithms were used: Berendsen during equilibration, and Parrinello-Rahman during the production run (Parrinello & Rahman, 1981; Berendsen *et al.*, 1984).

a) Berendsen barostat: Unlike the thermostat, Berendsen barostat does not alter the velocity of the particle as the motion of the particle is affected by the pressure (Berendsen *et al.*, 1984). Instead, Berendsen alters the co-ordinate of the box vector, the compressibility and the coupling time, allowing the pressure inside the box to be relaxed. This control of the pressure allows Berendsen to yields a correct “average” pressure, but not the correct NPT ensemble as the volume inside the box does not oscillate, and hence, fail to yield the correct distribution of the box volume (Rizzi *et al.*, 2020). The box volume does not fluctuate as the pressure coupling is modelled under the first-order differential equation. The accounting of the change in the volume fluctuation in the box is crucial as we are aiming to calculate the Gibb’s free energy, not the Helmholtz’s free energy. In principle, we would expect the change in the box volume causes by the binding of the ligand. By not yielding the correct distribution of the box volume, this may cause an inaccuracy in the calculation of the binding free energy of the ligand to the protein. However, in

the coarse-grain simulation system, the difference in the Gibb's free energy affected by the change in volume should be negligible.

b) Parrinello-Rahman barostat: To put this in a simple terms, the box vector is represented as a matrix and the motions of all atoms in the box are calculated with an additional term (Parrinello & Rahman, 1981). The barostat calculates the strength of the pressure coupling based on the defined approximate isothermal compressibility and the pressure-time constant, allowing constant pressure to be reached in the NPT ensemble with the Gibb's free-energy in the system being conserved. If the initial pressure is very far away from the desirable pressure, it is common that the fluctuation in the box vectors will be too large, causing the simulation to crash. Thus, it is a general practice to equilibrate the pressure in the system using Berendsen so that the desired pressure is almost reached, and then using Parrinello-Rahman barostat in the production run to ensure that the free energy in the simulation is conserved throughout the simulation.

2.5 Speeding up the simulations

Throughout this thesis, I have applied two tools to increase the speed of the simulation to gain access to a greater simulation timestep, which means that I can sample over a longer simulation timescale. In this section, I discuss how these tools help us to sample a longer timescale of a simulation.

2.5.1 Virtual sites

Virtual sites are extensions to atomistic forcefields for molecular dynamics simulation. The virtual site used in this work involves the addition of a zero-charge dummy atom. When the virtual site is applied, the mass of three hydrogen atoms become zero and become redistributed between the non-hydrogen atoms (carbon and nitrogen) and the dummy atom generated (Figure 2.6). In both NH_3^+ and CH_3 functional groups, two dummy atoms are anchored at the two centres-of-mass between four atoms with a constraint applied to it (Olesen *et al.*, 2018). As we aim to increases an integration timestep from 2 fs to 4 fs, or indeed 5 fs, all bonds are constraints using LINCS algorithm to prevent harmonic oscillation around the C-H bond. The distance and the angle constraint are also provided between the dummy atoms (MC1, MC2) and the heavy atom (such as C or N). If the C-H bond rotation violates the constraints, LINCS algorithm will resets the bond length and the angle to the set values.

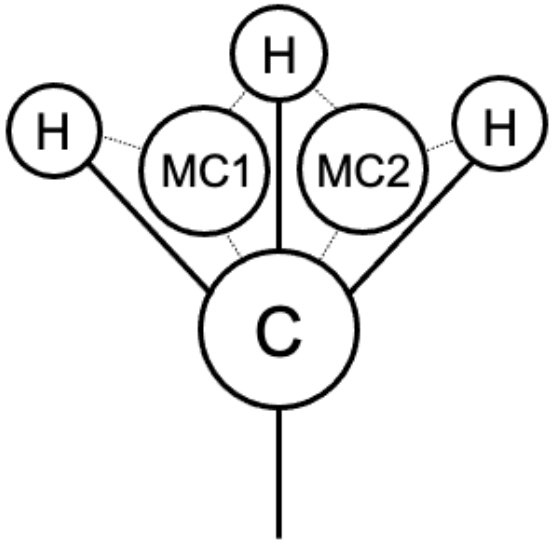


Figure 2.6 Virtual sites

Standard CH_3 bond is represented as thick line. All atoms contributed in the virtual site is shown as circle. To generate a virtual site, two dummy particles (MC1 and MC2) are placed between the C-H bond with restraint (dashed line) to allow slower C-H bond rotation and mass distribution. The position of H and C atoms are restrained to the two dummy particles. All atomistic simulations with the virtual sites were run with all bonds restrained to precent rotation of the C-H at 4 fs timestep.

2.5.2 Coarse-grained molecular dynamics simulation

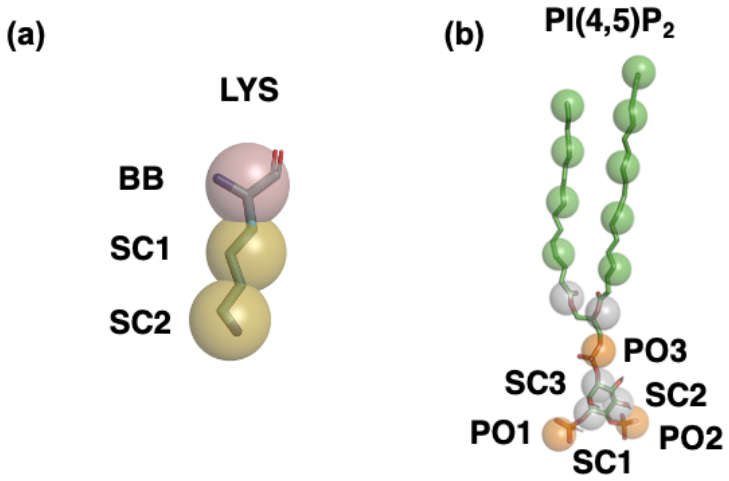


Figure 2.7 Coarse-grained representation

Coarse-grained representation of (a) amino acid such as lysine and (b) lipid such as PIP_2. Three to four non-H atoms are converted to a single particle. For example, N, CA, CO and O atoms on the amino acid backbone are converted to a single BB particle. In the naming system, BB stands for backbone, SC for side-chain and PO for phosphates.

Coarse-grained (CG) molecular dynamics simulation was first applied as a tool to increase the simulation timestep when simulating a soft material (Padding & Briels, 2003). The technique involved combining three to five non-hydrogen atoms into a particle (or a bead), with a set of assigned properties for bonded and non-bonded interactions in the forcefield (Figure 2.7). Throughout this thesis, all simulations are carried out using MARTINI2.2 forcefield (Marrink *et al.*, 2007). In this forcefield, in the bonded interaction term, the bond length and the bond angle are maintained with softer harmonic restraints. In the non-bonded interaction, the electrostatic interactions are commonly calculated with a cut-off of 1.2 nm, instead of using PME (Marrink *et al.*, 2004). These interactions are defined empirically or based on the atomistic forcefield. An

elastic force (or an elastic network) is defined between two particles, within 0.8 nm, to maintain the secondary and tertiary structure (Bahar *et al.*, 1997; Tama *et al.*, 2002). Overall, this decreases bond oscillations, allowing the timestep to be increased to 20 fs. Although this technique increases the simulation timestep, the use of elastic networks prevents protein conformational change (Marrink *et al.*, 2007). One major benefit of CG simulations is the ability to identify lipid-binding sites on membrane proteins, and to quantify their strength of binding (Corey *et al.*, 2019*a*).

2.6 Quantifying the strength of an interaction: Free energy calculations

When one thinks about understanding protein-ligand interaction, it is important to consider two questions. (1) where does the ligand bind? (2) what is the strength of binding? The former may be addressed using unbiased molecular dynamics simulations as discussed above. However, this does not answer the question regarding the strength of the interactions. In an experimental study, this is commonly quoted as the dissociation constant (K_d), which is related to Gibbs free energy (ΔG). Quantifying the strength of an interaction between membrane proteins and their ligands is experimentally difficult due to the low expression yield, and the poor solubility of the ligand in a biphasic environment. Hence, this provides a window for MD simulation studies. There are several MD simulation methods to calculate the binding free energies between protein and the ligand. In this section, I will discuss how ΔG could be calculated using MD simulation.

2.6.1 Free energy perturbation (FEP)

Free energy perturbation is sometimes referred to as "alchemical transformation" (Chodera *et al.*, 2011). This is simply because one molecule (or a certain part of a molecule) is transformed into others using this calculation. The general concept is that free energy differences are computed between the two states as one is transformed to another (Aldeghi *et al.*, 2018*a*). The free

energy in a single NPT ensemble is calculated based on statistical mechanics as:

$$A_j = -k_B T \ln Q_j$$

where A_j represents the Helmholtz free energy at the state j, k_B represents the Boltzmann constant, T for temperature in Kelvin and Q_j represents a canonical partition function of state j. The partition function (Q_j) is simply a probability density function representing the distribution of energy in each state. This is defined as:

$$Q_j = \int_{\Gamma_j} e^{-\frac{U_j(\boldsymbol{q})}{k_B T}} \, d\boldsymbol{q}$$

where $U_j(\boldsymbol{q})$ represents the potential energy in the function of co-ordinates with the position $\boldsymbol{q}$ and Γ_j represents the phase space sample of the momentum $\boldsymbol{p}$**.** Thus, the free energy **difference** between two states, i and j, could be calculated by:

$$\Delta A_{ij} = (-k_B T \ln Q_j) - (-k_B T \ln Q_j) = -k_B T \ln \frac{Q_j}{Q_i}$$

where ΔA_{ij} represents the free energy difference between states j and state i.

To put this in a qualitative term, the system is allowed to sample between state i and state j. These two states are referred to as the phase space (Γ_j), which must overlap between the two states. The ratio of probabilities is then calculated between the two-phase spaces. It is important to implement these transformations in small steps so that the phase space overlap well, leading to an accurate calculation of the binding free energy.

These transformations are performed in small steps. The electrostatics is first perturbed from state i to j over λ windows. This is then followed by the perturbation of the Lennard-Jones over the next series of λ windows. During the alchemical transformation, the exchange of a given particle (with mass and charge) for a dummy particle (with the same mass but with zero charge and Lennard-Jones interactions), may allow two particles to occupy the same space ($r_{ij} = 0$). This is overcome by applying a soft-core potential to prevent particle overlap (Klimovich *et al.*, 2015). This is a modification of a Lennard-Jones equation defined as:

$$U(\lambda, r) = 4\epsilon\lambda^n \left[\left(\alpha(1-\lambda)^m + \frac{r^6}{\sigma^6} \right)^{-2} - \left(\alpha(1-\lambda)^m + \frac{r^6}{\sigma^6} \right)^{-1} \right]$$

where $U(\lambda, r)$ defines a potential energy function over the chemical state, λ and radii, r. Three constants, α, m, n are set at 0.5, 1 and 1 respectively. As you may observe, this modification of the Lennard's Jones equation has added the λ term to the calculation, allowing the force to be rescaled at a λ window where the force reach infinity as two particles overlapped (Figure 2.8). This results in a finite amount of energy when the two radii eventually overlap and prevent the simulation to explode.

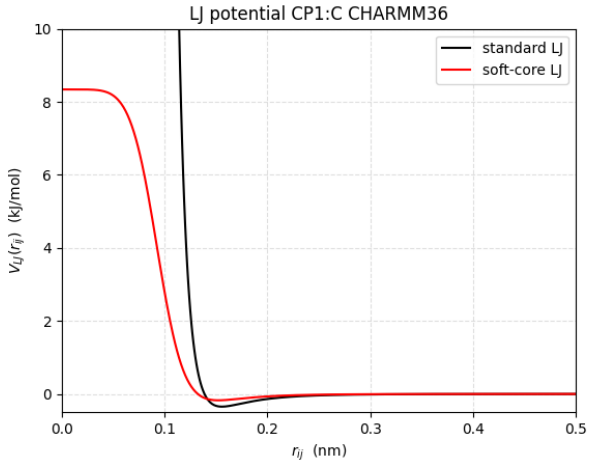


Figure 2.8 Lennard-Jones potential between CP1 and C atom types with soft-core potentials

Black - A calculated LJ potential between two atom types (CP1 and C) at a varied distance based on CHARMM36 forcefield parameter using equation $V_{LJ}(r_{ij}) = 4\epsilon_{ij}\left(\left(\frac{\epsilon_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right)$ where ϵ_{ij} = 0.347450500075 and σ_{ij} = 0.138767581228.
Red - A calculated LJ potential between two atom types (CP1 and C) at a varied distance based on CHARMM36 forcefield parameter with a modified soft-core potential using equation $V_{LJ}(r_{ij}) = 4\epsilon\lambda^{n}\left[\left(\alpha(1-\lambda)^{m} + \frac{r^6}{\sigma^6}\right)^{-2} - \left(\alpha(1-\lambda)^{m} + \frac{r^6}{\sigma^6}\right)^{-1}\right]$ where α = 0.5, m = 1, n = 1, ϵ_{ij} = 0.347450500075, σ_{ij} = 0.138767581228 and λ = 0.5.

To calculate the free energy differences, the Multistate Bennett Acceptance Ratio (MBAR) was used to calculate the free energy difference between states i and j. This method has shown to have the lowest variance amongst all other methods used for free energy calculations. It allows the data from all states to be assessed and hence, the energy of every state is sampled and calculated. This is an extension of Bennett's Acceptance Ratio (BAR) where a single partition function is modified to

$$Q_j = \int_{\Gamma_j} \alpha_j(\boldsymbol{q}) e^{-\frac{U_j(\boldsymbol{q})}{k_B T}} \, d\boldsymbol{q}$$

where $\alpha_j(\boldsymbol{q})$ is calculated from variational calculus to minimise the variance of the free energy calculated across all states. In all states, enough sampling is required to ensure the normal distribution of each state is achieved. Once sufficiently sampled, the average energy calculated will be constant, and hence the simulation has converged. Failing to reach convergence in the free energy calculation leads to inaccuracies.

In practice, FEP is very difficult when charged species are involved due to an imbalance in the calculation of the non-bonded electrostatic interaction using PME (Domański *et al.*, 2017; König *et al.*, 2018). This could be counteracted by a different approach, such as perturbing an ion to couple the charge difference so that the net charge in the simulation system remains at zero (Dixit & Chipot, 2001). This causes atomistic FEP with charged species to be very difficult to compute. In this thesis, coarse-grained FEP is applied with an electrostatic cut-off which helps with convergence issues (Kastenholz & Hünenberger, 2006).

Free energy perturbation has been a widely used method to answer both chemical and biological problems. Initial work involved the calculation of the solvation free energy of organic molecules (Jorgensen *et al.*, 1988; Jorgensen, 1989). The method has then been extended to several branches: perturbation of functional groups in a drug molecule to aid structural-based drug design (Wang *et al.*, 2015*b*); perturbation of amino acid side chains to assess the

thermal stability or the contribution of a specific side chain to a binding site (Seeliger & de Groot, 2010; Gapsys *et al.*, 2015); perturbation of a water molecule to a vacuum space to assess the binding free energy of water molecules in the binding site (Aldeghi *et al.*, 2018*b*); and most recently, perturbing one lipid species to another to calculate lipid-binding free energy in a coarse-grained system (Corey *et al.*, 2019*b*).

2.6.2 Potential of Mean Force (PMF) calculation

PMF calculations sample the free energy of lipid binding between bound and unbound states, usually along a linear path, called a collective variable (CV) (Figure 2.9). The collective variable is selected so that the path of motion is the rate-limiting step i.e. the slowest step of the motion. This selection of CV allows the energy landscape of lipid accessing the binding site to be sampled thoroughly. By sampling with umbrella potentials, lipids may be simulated at fixed, incremental distances from the protein binding site (Corey *et al.*, 2019*b*). This gives rise to a calculated one-dimensional lipid binding free energy profile from bound to unbound, whilst also providing a total free energy difference between the bound and unbound states.

The first attempt of PMF calculations to calculate protein-lipid affinity was for cardiolipin binding sites on the cytochrome c oxidase and complex IV (Arnarez *et al.*, 2013*a*, 2013*b*). Since this initial approach, the method has been applied to study lipid binding affinity to receptors, transporters and ion channels (Hedger *et al.*, 2016*a*, 2016*b*; Domański *et al.*, 2017; Corey *et al.*, 2019*b*; Wang *et al.*, 2020). The method has also shown to be useful when comparing differences in the affinities for a range of lipid types and provides details of their 1D energy landscape.

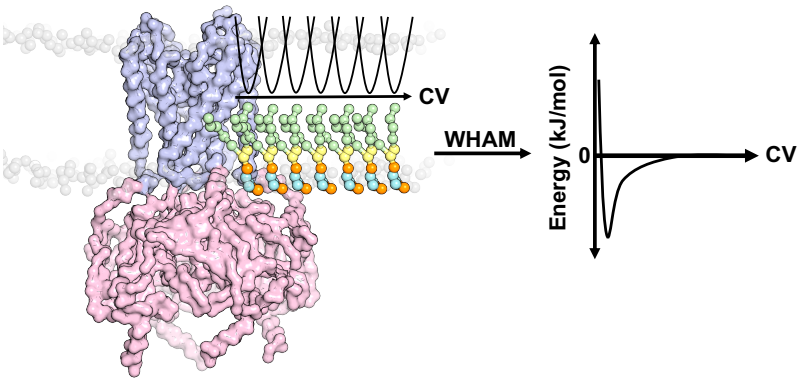


Figure 2.9 Potential of mean force calculation

Potential of mean force (PMF) calculation. PIP_2 lipids (acyl tail - green, glycerol - yellow, phosphates - orange, inositol ring - cyan) are sampled every 0.5 A with a harmonic potential of a force constant at 1000 kJ/nm^2/mol along the collective variable (CV) parallel to the bilayer. Each PIP_2 represents a sampling position. The protein is displayed only in coarse-grained backbone where the transmembrane region is shown in blue and cytoplasmic region is shown in pink. PIP_2 binding free energy to the protein is then calculated using weighted histogram analysis method (WHAM) based on the energies it samples along the CV (right).

To calculate the binding free energy, simulations generally start from the ligand-bound state. This may be obtained from an experimental structure or an unbiased simulation. A CV is generated using steered MD simulation to exert forces in a certain dimension on the ligand. This may be visualised as the ligand is pulled along the CV. At each point along the CV, the ligand is restrained at that position using a harmonic restraint and the external forces of interaction are sampled. The shape of the restraint is an umbrella and hence, this defines the name of this sampling method as "umbrella sampling". Once energy within each point is sampled, the energy values are combined using the Weighted Histogram Analysis Method (WHAM). The energy distribution at each point is then used to construct the energy landscape along the collective variable axis. Error estimates of these calculations represent the fluctuation of energy at a certain point along the CV.

To put this in a more qualitative way, at each point, the energy is sampled in a normal distribution shape. At each point the energy distribution is sampled, with an overlap required between adjacent windows. If a greater force constant is imposed, the narrower the region is sampled and therefore, the distance between two sampling points will need to be closer to each other. Failure to obtain an overlapping distribution leads to inaccuracies in the PMF calculation.

Chapter 3

Phosphoinositides interaction with Kir channels gating

“パタモン.... しんか~…アンジェモン”
“Patamon … digivolved to … Angemon”

Digimon Adventure Episode 13 (1999)

The work in this chapter is published in:

Pipatpolkai, T., Corey R.A., Proks P., Ashcroft F.M. and Stansfeld P.J. **Evaluating Inositol phospholipid interactions with Inward Rectifier Potassium Channels and their role in Disease**. *Communications Chemistry* **3,** 147 (2020). https://doi.org/10.1038/s42004-020-00391-0

3.1 Introduction

In this chapter, I characterised the regulation of ion channels by phospholipids. One of the most well documented examples of phospholipid regulation of ion channels is that of inwardly rectifying K channel (Kir channels). Kir channels are activated by phosphoinositides, in particular phosphatidylinositol-4,5-bisphosphate (PIP_2) (Suh & Hille, 2008). For a eukaryotic Kir channel to be activated, PIP_2 needs to bind to its respective binding site, and trigger a protein conformational transition from a closed to an open state of a channel. I term the first event “binding” and the second events “transduction”.

Understanding phospholipid binding and how this is translated into activation of Kir channels has been a tremendous challenge in the field for several reasons. The usual way to investigate Kir channel activation by an agent, would be to apply it directly to a membrane patch and record channel opening and closing using electrophysiological methods (Huang *et al.*, 1998). However, PIP_2 presents a particular challenge because it is very difficult to calculate how much phosphoinositide is incorporated in the bilayer. This is because PIP_2 concentration in the patch declines after excision due to its hydrolysis by phosphatase. This is thought to underlie the “rundown” of the channel. To prevent the rundown, fluorovanadate and pyrophosphate (FVPP), a phosphatase inhibitor, could be applied to the patch. The rate of rundown varies between patch to patch due to their different experimental condition.

Direct application of PIP_2 is difficult as the alkyl tail of PIP_2 is hydrophobic, and this makes the molecules accumulate in the membrane. Thus, the concentration of PIP_2 in the perfusion system is not going to be the same as the concentration in the membrane. A more soluble form of lipid (such as diC8:PIP_2) could also be used (Rohács *et al.*, 2003). However, the concentration of diC8:PIP_2 in the membrane is not the same as the local concentration around the channel. Several methods have been developed to reduce PIP_2 concentration in the membrane such as the used of voltage sensitive phosphatases (VSPs) or phospholipase C (PLC) (Hilgemann & Ball, 1996; Rjasanow *et al.*, 2015). However, neither of the methods provide a quantitative assessment to the PIP_2 concentration (Schulze *et al.*, 2003; Ribalet *et al.*, 2006). Quantitatively, neomycin, can be used to screen negative charges from PIP_2 and preventing its interaction with the Kir channel. Neomycin has been previously used to characterise PIP_2 binding sites on Kir channels (Haider *et al.*, 2007). However, the correlation between neomycin concentration and how much charges are screened out are not known. In this chapter, I collaborated with Dr Peter Proks to use neomycin to assess an effect of neonatal diabetes mutation on the effect of neonatal diabetes mutation (E179K) on K_{ATP} channel activation by PIP_2.

Because of the problems mentioned above it is very difficult to separate lipid binding events from transduction. However, recent advances in structural biology and computational power, enable us to approach this problem using computational methods. The focus of this chapter is the development of novel

computational approach to assess the relative phosphoinositide binding affinities of different lipids to different types of Kir channels.

Different Kir channels exhibit variable affinities for different phosphoinositides and different magnitudes of channel activation (Fan & Makielski, 1997; Huang *et al.*, 1998; Rohács *et al.*, 2003). One hypothesis is that the extent of channel activation is correlated with the number of positively charged amino acids that reside at the PIP_2 binding site (Suh & Hille, 2008). However, there is not a good correlation between the number of charged residues and the gating behaviour of Kir channels. Interestingly, despite the highly conserved nature of the PIP_2 binding site, the binding of PIP_2 activates eukaryotic Kir channels, but inhibits the bacterial Kir channel (KirBac) (Enkvetchakul *et al.*, 2005).

Recent advances in cryo-electron microscopy (cryo-EM) have enabled several high-resolution structures of the K_{ATP} channel complex, both Kir6.2/SUR1 and Kir6.1/SUR2B (Lee *et al.*, 2017; Li *et al.*, 2017; Martin *et al.*, 2017*b*, 2017*a*; Wu *et al.*, 2018; Sung *et al.*, 2021). In this thesis, I only considered the pancreatic K_{ATP} channel which comprise of Kir6.2/SUR1. Earlier studies had shown that SUR1 enhances the channel sensitivity to ATP inhibition (Tucker *et al.*, 1997; Usher *et al.*, 2020). Given that the ATP and PIP_2 binding sites lie near each other, I investigated how SUR1 may influences K_{ATP} channel activation by PIP_2 using both experimental and computational approaches.

Naturally occurring mutations in the Kir6.2 subunit that led to insulin secretory disorders are commonly found at the ATP-binding site, and have been shown to alter channel inhibition by ATP. My work interested in the mutations near the PIP_2 binding site which are associated with the clinical symptoms. I was interested in exploring how mutations in the putative PIP_2 binding site might similarly lead to neonatal diabetes (NDM; e.g. E179K/A) and congenital hyperinsulinism (CHI; e.g. K67N) (Reimann *et al.*, 2003; Haider *et al.*, 2007; De Franco *et al.*, 2020). A previous study had shown that the K67N mutation does not alter channel surface expression but reduced channel activation in response to inhibition of cell metabolism (Reimann *et al.*, 2003). The mechanism of how the E179K/A and K67N mutations affect channel activity is currently unclear.

Based on the prevalent numbers of crystal structures available, Kir channels form an attractive target for applying a computational approach to compare binding affinity between phosphoinositides and to assess the impact of mutations on binding. Coarse-grained (CG) molecular dynamics (MD) simulations have previously been used to identify lipid-binding sites on ion channels (Marrink *et al.*, 2007; Monticelli *et al.*, 2008; Hedger & Sansom, 2016), as well as to predict the affinity of interactions (Corey *et al.*, 2019*b*). These simulations were commonly approached at a coarse-grained detail to reduce the simulation time. Despite a great reduction in the atomic resolution, the predicted binding site on the Kir channel were in good agreement with previous experimental studies and the existing crystal structure (Stansfeld *et al.*, 2009; Duncan *et al.*, 2020).

For many years, the application of atomistic free energy perturbation (FEP) methods have been successfully applied to determine small molecule, lipid, and drug binding affinities (Wang *et al.*, 2015*b*) as well as to study the impact of amino acid side chain mutations on binding (Gapsys *et al.*, 2015; Gapsys & de Groot, 2017). These approaches were initially done at the atomic resolution. However, free energy calculation of atomistic simulation converges very slowly and alternatively, using an unconverging free energy landscape lead to an inaccurate prediction of binding free energy (Beutler *et al.*, 1994; Domański *et al.*, 2017; Meng *et al.*, 2018; König *et al.*, 2018). However, a lower resolution approach using a coarse-grained protocol (CG-FEP) to assess protein-lipid binding free energies allows protein-lipid binding affinities to be calculated using more rapidly. This approach agrees well with other free energy calculation methods such as potential of mean force calculation (PMF) and well-tempered metadynamics (WTMetaD) (Corey *et al.*, 2019*b*).

In this chapter, I used CG-FEP to compare the relative binding free energies between different phospholipids and the human Kir6.2 channel (hKir6.2), capturing the full thermodynamic cycle for the transition of PIP_2 to PC, either directly or via the intermediates phosphatidylinositol-4-phosphate (PI4P) and phosphatidylinositol (PI), and thereby reporting on the affinity of each interaction (Corey *et al.*, 2019*b*). I extended the methodology to investigate the functional effects of the lipid-associated neonatal diabetes mellitus (NDM) mutations (E179K/A) and a congenital hyperinsulinism (CHI) mutation (K67N) in hKir6.2 (Huopio *et al.*, 2002; Flanagan *et al.*, 2007). Based on the predicted binding site for PIP_2, I calculated that these residues interact

directly with PIP_2 in the membrane and that this interaction is perturbed by their mutation. This therefore provides a biochemical and structural explanation for the different clinical phenotypes.

The work in this chapter includes computational analyses by the author and electrophysiology experiments done by Dr Peter Proks. These assess both the affinity for PIP_2, and channel activation by PIP_2, respectively. I also extend the computational methodology to assess the binding free energy differences between a range of hKir channels (hKir1.1, hKir2.2, and hKir3.2) and other phosphoinositide lipid species such as phosphatidylinositol-4-phosphate (PI4P) and phosphatidylinositol (PI). Together, my CG-FEP approach enables us to describe the affinity of membrane proteins with a range of different lipids, as well as examining how biologically important mutations affect these interactions.

3.2 Methods

3.2.1 Molecular modelling

Modeller9v16 (Webb & Sali, 2016) was used to add the missing loops and amino acid residues to the cryo-EM structure of the Kir6.2 channel (PDB entry: 6BAA) (Martin *et al.*, 2017*a*) and to generate a human Kir6.2 model based on residues 32-352. In both the cryo-EM structure and my model there are 32 amino acids missing at the N-terminus and 39 at the C-terminus. Modeller9v16 was also used to generate a model of the human K_{ATP} channel octamer (hKir6.2 tetramer + four SUR1 - PDB entry: 6BAA) and the models of the hKir6.2 mutant channels. The modeller script used to generate the hKir6.2 model is available on https://github.com/WillPPK/Thesis. Each model was compared to its initial template structure to ensure that the modelling had not demonstrably altered the original secondary structure or the rotation of the amino acid sidechains (overall RMSD of all protein atoms < 1.0 Å). All protonation states are calculated based on Swiss-Model, resulting in all basic residues (arginine and lysine) near PIP_2 binding site being protonated and has a positive charge. This is a primary assumption in this thesis. The model shows a MolProbity score of 1.09 with 95.9% sequence identity. Models of the other hKir channels were generated using Swiss-Model (Waterhouse *et al.*, 2018), with human Kir1.1 and Kir2.2 based on 3SPH and human Kir3.2 based on 3SYC. The MolProbity score of the hKir1.1, hKir2.2 and hKir3.2 models are 1.70, 1.43 and 1.51 respectively. It is important to note that hKir1.1 model generated has 48.17% sequence identity to the cKir2.2 structure. On the other

hand, hKir2.2 and hKir3.2 has a much higher sequence identity to its template (89.52% and 100%). The chicken Kir2.2 structure with bound diC8-PIP_2 (PDB entry: 3SPI) (Hansen *et al.*, 2011) was aligned with all hKir channels to provide the co-ordinate of the PIP_2 headgroup in the PIP_2 binding site. All structural models, Ramachandran plot, sequence alignments and coarse-grained structures with PIP_2 bound for CG-MD and CG-FEP calculation are available on https://github.com/WillPPK/Thesis.

3.2.2 Coarse-grained (CG) Simulations

All protein structures were converted to their CG representation and embedded in a PC (1-palmitoyl-e-oleoyl-sn-glycero-3-phosphocholine) bilayer using the self-assembly MemProtMD protocol (Stansfeld *et al.*, 2015; Newport *et al.*, 2018) with the MARTINI v2 biomolecular forcefield (Marrink *et al.*, 2007). This approach orients the structure of the transmembrane protein parallel to the z-axis using MEMEMBED (Nugent & Jones, 2013). The protein is then placed in a periodic box at minimum distance of 30 Å from the edge of the box in both x and y directions, and with a z dimension of 80 Å. The structure is then converted to a CG representation with *martinize.py* with an application of an elastic network with a force constant of 1,000 kJ/mol/nm^2 between backbone beads within 0.5-0.9 nm to maintain their secondary and tertiary structure. The PC lipid is then added to the periodic box, allowing them to assemble freely around the protein. The z-dimension of the box is then extended so that the minimum distance between the protein and the face of the box is 30 Å apart, and then flooded with the coarse-grain water particles, Na^+ and Cl^- ions to a final concentration of 0.15 M to neutralize the system. The total number of the

molecules in the setup is described in Table 3.1. A temperature of 323 K was maintained with V-rescale temperature coupling (Bussi *et al.*, 2007), while 1 atm pressure was controlled using semi-isotropic Parrinello-Rahman pressure coupling (Parrinello & Rahman, 1981). Systems were energy minimised using the steepest descents algorithm and equilibrated for 5 ns with 1,000 kJ/mol/nm^2 position restraints on backbone beads, prior to 1 μs production. All simulations, root mean square deviation (RMSD) calculations and distance analyses were carried out using GROMACSv2018 (Abraham *et al.*, 2015) and all structural alignments and docking were carried out using PyMOL.(Schrodinger LLC, 2015)

Unbiased simulations

Protein	PIP_2	POPC	Water	Na	Cl	Box size	Length
hKir1.1	1	495	14196	314	330	134x134x149 Å^3	1μs x5
hKir2.2	1	538	16324	378	353	140x140x155 Å^3	1μs x5
hKir3.2	1	463	12451	321	285	131x131x141 Å^3	1μs x5
hKir6.2	1	508	13989	328	308	135x137x145 Å^3	1μs x5

Table 3.1 Number of molecules in the simulation system.

The number of PIP_2, POPC, Water, Na and Cl, the box size, simulation length and number of repeats used in the unbiased simulation.

Protein	Perturbations	λ Windows	Length	Replicates
hKir1.1 (with flat-bottom restraint)	PIP_2 → PI4P	21	300 ns	3
	PI4P → PI	21	300 ns	3
	PI → PC	21	300 ns	3
	PIP_2 → PC	21	300 ns	3
hKir2.2 (with flat-bottom restraint)	PIP_2 → PI4P	21	300 ns	3
	PI4P → PI	21	300 ns	3
	PI → PC	21	300 ns	3
	PIP_2 → PC	21	300 ns	3
hKir3.2 (with flat-bottom restraint)	PIP_2 → PI4P	21	300 ns	3
	PI4P → PI	21	300 ns	3
	PI → PC	21	300 ns	3
	PIP_2 → PC	21	300 ns	3
hKir6.2 (with flat-bottom restraint)	PIP_2 → PI4P	21	300 ns	3
	PI4P → PI	21	300 ns	3
	PI → PC	21	300 ns	3
	PIP_2 → PC	21	300 ns	3
	PIP_2 → diC8 PIP_2	21	300 ns	3
hKir6.2 - 90% PC + 10% PS (with flat-bottom restraint)	PIP_2 → PC	21	300 ns	3
hKir6.2 (without flat-bottom restraint)	PIP_2 → PI4P	21	300 ns	3
	PI4P → PI	21	300 ns	3
	PI → PC	21	300 ns	3
	PIP_2 → PC	21	300 ns	3
	PIP_2 → diC8 PIP_2	21	300 ns	3
K_{ATP} channel	PIP_2 → PC	21	300 ns	3
Bulk PC	PIP_2 → PI4P	21	300 ns	3
	PI4P → PI	21	300 ns	3
	PI → PC	21	300 ns	3
	PIP_2 → PC	21	300 ns	3
hKir6.2 with PIP_2	E179K → WT	21	300 ns	3
	WT → E179A	21	300 ns	3
	WT → K67N	21	300 ns	3
	WT → C166S	21	300 ns	3
hKir6.2 without PIP_2	E179K → WT	21	300 ns	3
	WT → E179A	21	300 ns	3
	WT → K67N	21	300 ns	3
	WT → C166S	21	300 ns	3

Table 3.2 Number of repeats in free energy perturbation calculation

The number of simulations windows, length and repeats used in every FEP calculations and their conditions.

3.2.3 Free energy perturbation (FEP) calculation of PIP lipids

The hKir6.2 tetramer with one bound PIP_2, obtained after equilibration, was used as the initial co-ordinates for the majority of the FEP calculations. Here, I calculate a relative binding free energy ($\Delta\Delta G$) by converting from one lipid type (such as PIP_2) to a series of other phospholipids (such as PIP_2, PI4P, PI and PC) along a reaction coordinate in a chemical space denoted λ. As is standard for FEP calculations, separate transformations were performed with either the lipid bound to the channel or in the bulk membrane.

I applied FEP to hKir6.2, hKir3.2, hKir2.2 and hKir1.1 and the following pairs of inositol lipids: (PIP_2 and PI4P), (PI4P and PI), (PI and PC) and (PIP_2 and PC). For these, specific phosphate and inositol sugar particles were transformed into a dummy particle with no interaction properties, in a stepwise process. Coulombic (charge interactions) and Lennard-Jones (Van der Waals interactions) were turned off separately, with a soft-core parameter used for the Lennard-Jones interactions. The coulombic interactions were perturbed linearly (λ = 0, 0.1, 0.2 … 0.9, 1.0) in the first 10 simulation windows, with the van der Waals interactions perturbed linearly (λ = 0, 0.1, 0.2 …… 0.9, 1.0) in the last 10 simulation windows with the soft-core parameters of α = 0.5 and σ = 0.3. Each simulation window was energy minimised and equilibrated as described above, before three production runs were carried out for 300 ns with randomised initial velocities, using a leap-frog stochastic dynamics integrator. A flat-bottom distance restraint between the PO4 phosphate group and protein backbone beads at 6 Å radii from the PIP molecule was applied using Plumed

(1000 kJ/mol/nm^2, 8 Å cut-off) (Tribello *et al.*, 2014). The flat-bottom restraint prevents the bound lipid from drifting away from its binding pocket and increases the accuracy of the calculation. As lipid molecule rarely experience the force of interaction from the flat-bottom restraint, I assume that the energetic cost from applying the flat-bottom restraint to the system is negligible. The free energy pathways were constructed using the *alchemical-analysis* software package (Klimovich *et al.*, 2015), where the energies are calculated based on the 300 ns of the data for a good convergence (Corey *et al.*, 2019*b*). Analyses was run using the Multistate Bennett Acceptance Ratio (MBAR) (Fajer *et al.*, 2009). All values are reported are reported as mean ± 95% confidence interval. All simulations were carried out using GROMACS v2018 (Abraham *et al.*, 2015). The details of all simulations set-up for FEP calculation is shown in Table 3.2.

3.2.4 Free energy perturbation (FEP) between amino acid residues

As before, the PIP_2-bound equilibrated hKir6.2 system was used as the initial co-ordinates for the FEP calculations. When assessing the influence of NDM or CHI mutations, I calculated changes in the relative binding free energy ($\Delta\Delta G$) by alchemically transforming the wild-type amino acid residue to its mutant counterpart. This was performed using a change in chemical space denoted as λ and applying the previously described protocol. The series of transformations carried out are shown in Table 3.3. Additional simulations were run by performing the same perturbation of the protein in a bulk POPC membrane in the absence of PIP_2.

Mutation	Direction of Perturbation	From		To	
		SC1	SC2	SC1	SC2
K67N	WT → K67N	C3	Qd	P5	Dum
C166S	WT → C166S	C5	-	P1	-
E179K	E179K → WT	C3	Qd	Qa	Dum
E179A	WT → E179A	Qa	-	Dum	-

Table 3.3 Direction of free energy perturbation of amino acid residues

The direction of the transformation and the particle types for free energy perturbation analysis. The particles are suggested based on MARTINI2.2 forcefield. Dum is a dummy particle with a mass of 72 and no bonded and non-bonded interactions.

3.2.5 Potential of Mean Force Calculation (PMF)

PMF calculations were set up by building the hKir6.2 in the PC bilayer, and a single POPC lipid was pulled from the binding site using steered-MD, where the collective variables (CV) are the distance between the lipid headgroup and the centre of mass of the protein. The vector of the CV is determined as the slowest step as PC is getting pulled out of the PIP_2 binding site and the POPC headgroup is pulled at 0.001 nm/ps at 1000 kJ/mol. The initial position for PC was modified from an initial PIP_2 co-ordinate. The simulations were calculated along the CV at 0.2 Å interval for optimal histogram overlap, with a 1000 kJ/mol/nm^2 umbrella potential applied to restrain the position of the lipid along the CV. Positional restraints of 100 kJ/mol/nm^{-2} were applied to the protein backbone to prevent rotation of the protein in the bilayer. For each window, the simulations were run for 500 ns, which was sufficient to see convergence. In fact, this set of simulation is oversampled and would converge with half of the sampling (0.5 Å interval). The 1D energy profile was generated using weighted histogram analysis method (WHAM) using the *gmx wham* tool (200 rounds Bayesian Bootstrap) (Hub *et al.*, 2010).

3.3 Results

3.3.1 PIP_2 binding conformation to the hKir6.2 channel

First, I modelled the putative PIP_2 binding pocket of the Kir6.2 channel based on the recent cryo-EM structure of the rat K_{ATP} channel in complex with ATP and glibenclamide (PDB ID: 6BAA) (Martin *et al.*, 2017*a*) using classical coarse-grained molecular dynamics simulations (CG-MD). The Kir6.2 tetramer of this structure is similar to other high-resolution K_{ATP} channel structures published more recently (Lee *et al.*, 2017; Li *et al.*, 2017; Martin *et al.*, 2017*b*, 2017*a*, 2019; Wu *et al.*, 2018; Ding *et al.*, 2019). The initial position of the PIP_2 molecule was obtained from the chicken Kir2.2 channel:diC8-PIP_2 complex (Hansen *et al.*, 2011). After structural alignment of hKir6.2 with chicken Kir2.2, one of the bound diC8-PIP_2 molecules was extracted, converted to CG and the resultant hKir6.2-PIP_2 complex was built into a PC membrane and simulated for 1 μs (n = 5) using CG lipid self-assembly in order to find energy minima (Marrink *et al.*, 2007; Monticelli *et al.*, 2008; Stansfeld *et al.*, 2015). I defined residues that were within a 6 Å radius of the whole PIP_2 molecule for >75 % of the simulation time as proximal residues (Figure 3.1a). I found that PIP_2 binds in the vicinity of both the N- and C-termini of the hKir6.2 channel, including ^{67}KWP^{69}, on the N-terminus, and the residues between 170 and 179 on the C-terminus. These residues are mostly basic, which allows them to interact with the negative charged phosphates on the inositol ring of the PIP_2 headgroup. E179 is the only negatively charged amino acid to be within a 6 Å cut-off range of the lipid (Figure 3.2).

I assessed the stability of PIP_2 in its binding site using Root Mean Square Deviation (RMSD) analysis over the 1 μs simulations (Figure 3.1b). The data show that the position of the PIP_2 diverges very little in 1 μs (RMSD = *ca.* 0.8 Å). This is also confirmed by analysing the distance between the PIP_2 molecule and two amino acids near the PIP_2 binding site, K67 and E179 (Figures 3.1c, 3.1d). I found that the minimum distance between K67 or E179 and the PIP_2 head group is approximately 5 Å (Figures 3.1c-d) and ~7 Å between lipid and protein backbone (Figure 3.3). Therefore, I hypothesise that mutations to these residues may affect the binding free energy of PIP_2 to the channel.

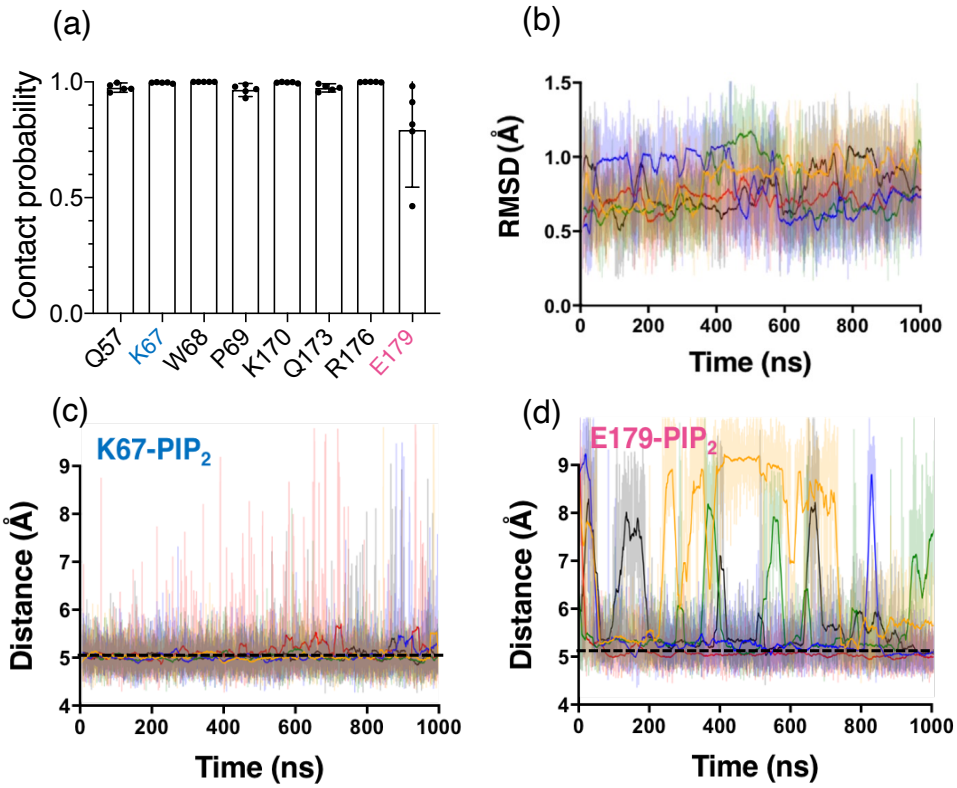


Figure 3.1 PIP_2 binding dynamics on hKir6.2

(a) PIP_2 contact analysis showing the fraction of time that the indicated residues are in 6 Å proximity to the PIP_2 molecule (contact probability). Only residues with a >75% contact probability are shown. Data from 5 repeats of 1 µs simulations. Error bar represents 95% confidence interval around the mean.
(b) Root mean square deviation (RMSD) analysis of the PIP_2 molecule headgroup when bound to the hKir6.2 tetramer. The different colours indicate the individual repeats of the simulations (n=5). The darker lines show the running average for each simulation.
(c) Calculated minimum distance between K67 and the PIP_2 headgroup during five 1 µs simulation. The different colours indicate the individual repeats of the simulation. The darker line shows a running average. A dashed black line denotes the distance cutoff used to denote a contact in panel (a).
(d) As in panel (c) but for E179 and the PIP_2 headgroup.

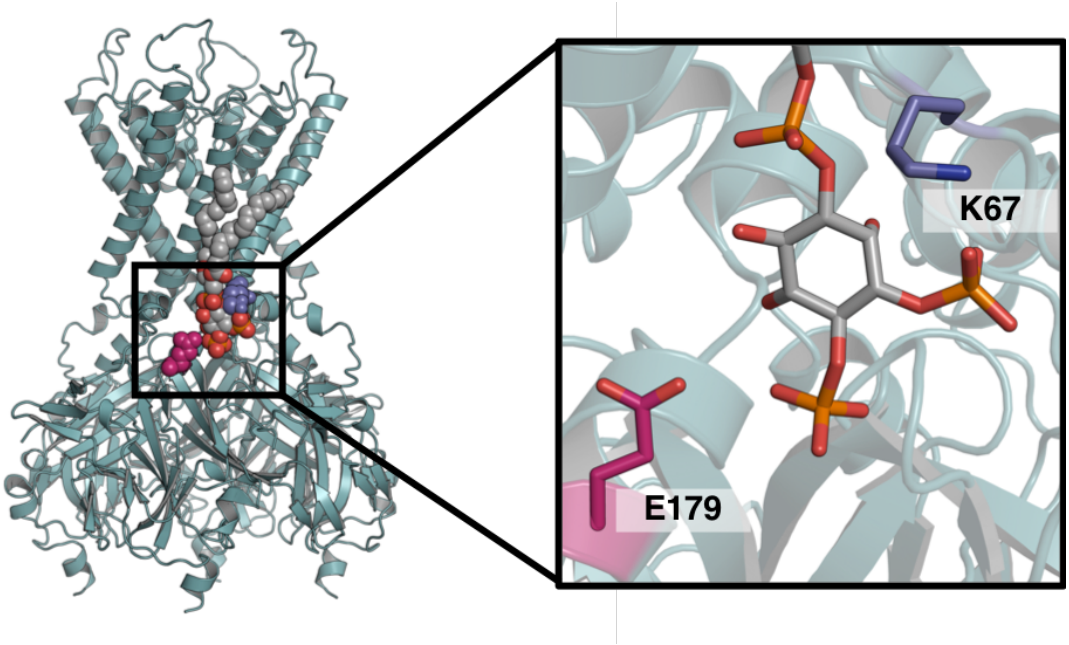


Figure 3.2 A PIP_2 binding site on hKir6.2

PIP_2 binding site on the hKir6.2 tetramer (green) showing PIP_2 (grey with CPK colours), E179 (magenta) and K67 (blue). All hydrogens are hidden in this schematic representation. However, in the simulations, K67 is protonated, E179 is deprotonated and all phosphates on the PIP_2 headgroup are deprotonated. The PIP_2 carries -5 charge throughout the CG simulation.

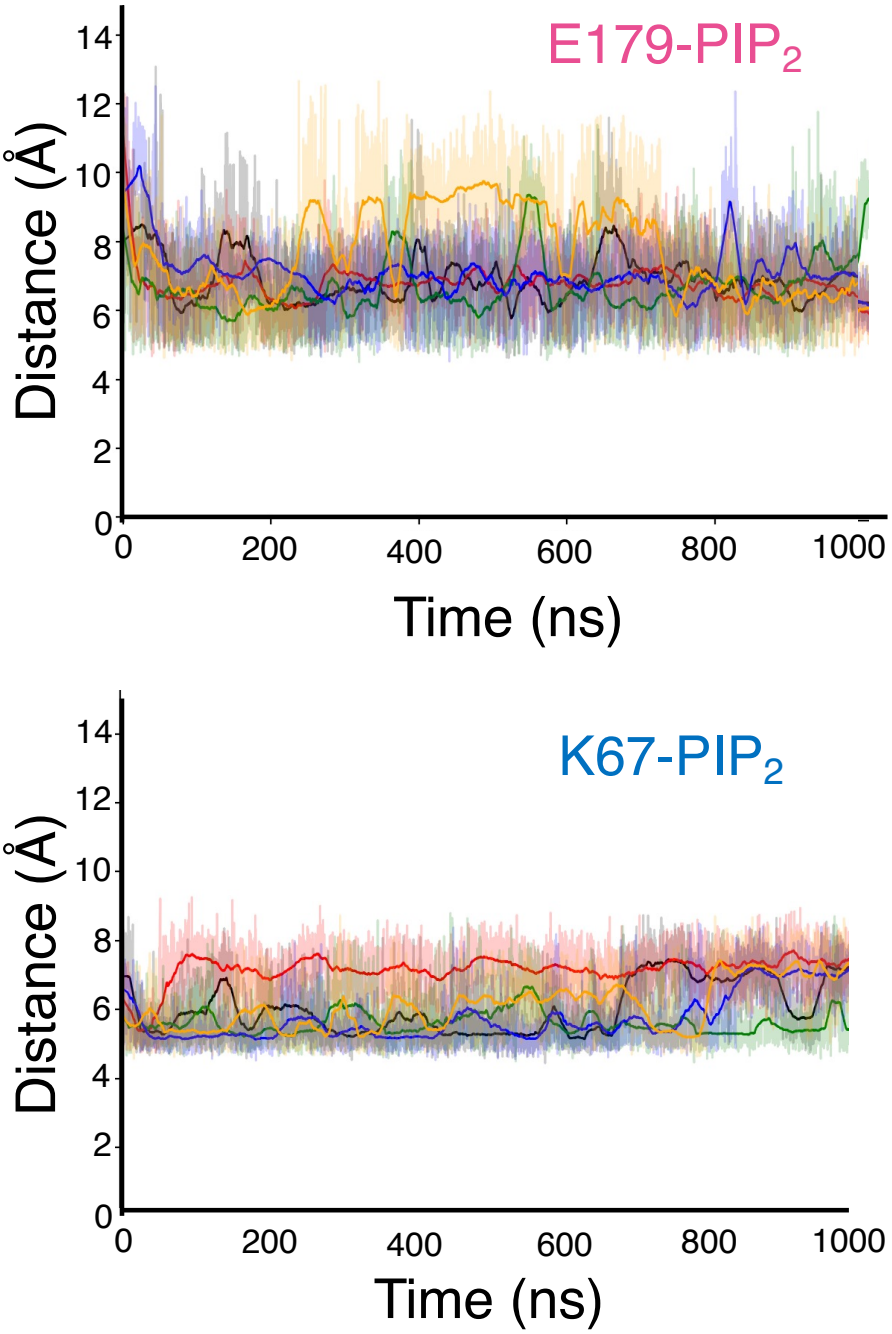


Figure 3.3 Distance between protein backbone and the PIP_2 headgroup

Calculated distance between the PIP_2 headgroup and the backbone of either E179 or K67, over five 1 μs simulation. The different colours indicate the individual repeats of the simulation. The darker lines show the running average for each simulation.

3.3.2 Step-wise perturbation of the PIP molecule bound to the hKir6.2 channel

I then assessed the contribution that each of the different PIP headgroup moieties (i.e. each phosphate group and the inositol ring) make to the free energy of binding to the closed hKir6.2 tetramer using CG-FEP. For this, I iteratively perturbed single beads to transform from one phospholipid (such as PIP_2) into another (such as PC). This enabled us to calculate the binding free energy difference ($\Delta\Delta G$) of the two different phospholipids to hKir6.2, embedded in a PC bilayer (Figure 3.4a). For simplicity, all lipids have both palmitoyl and oleoyl alkyl chains. The energies were computed using Multistate Bennett Acceptance Ratio (MBAR) (Fajer *et al.*, 2009), with convergence seen within *ca.* 200 ns per window (Figures. 3.5a-b, 3.6a-b, 3.7a-b and 3.8a-b).

To prevent the molecule from leaving its binding site, I applied a flat-bottom distance restraint between the protein and lipid using Plumed (Bonomi *et al.*, 2009; Tribello *et al.*, 2014) (Figures. 3.5a-c and 3.6a-c). This was mostly applicable to calculations in which the lipid was transformed to PC. In the cases where the lipids remain bound at the binding site, applying a flat-bottom restraint makes no difference to the binding free energy and its convergence. (Figures 3.7c and 3.8c). This procedure also reduced the errors between simulation replicas. As lipid infrequently experiences the restraint, thus, I assumed that the restraint would have negligible effect on the binding energies (Figures 3.5d and 3.6d). Thus, I did not factor the energetic from the flat-bottom restraint to the calculation. Alternative approach such as running multiple repeats of shorter simulations may provide a greater sampling of the PIP_2 binding site, without having to apply a restraint to the protein in the binding pocket.

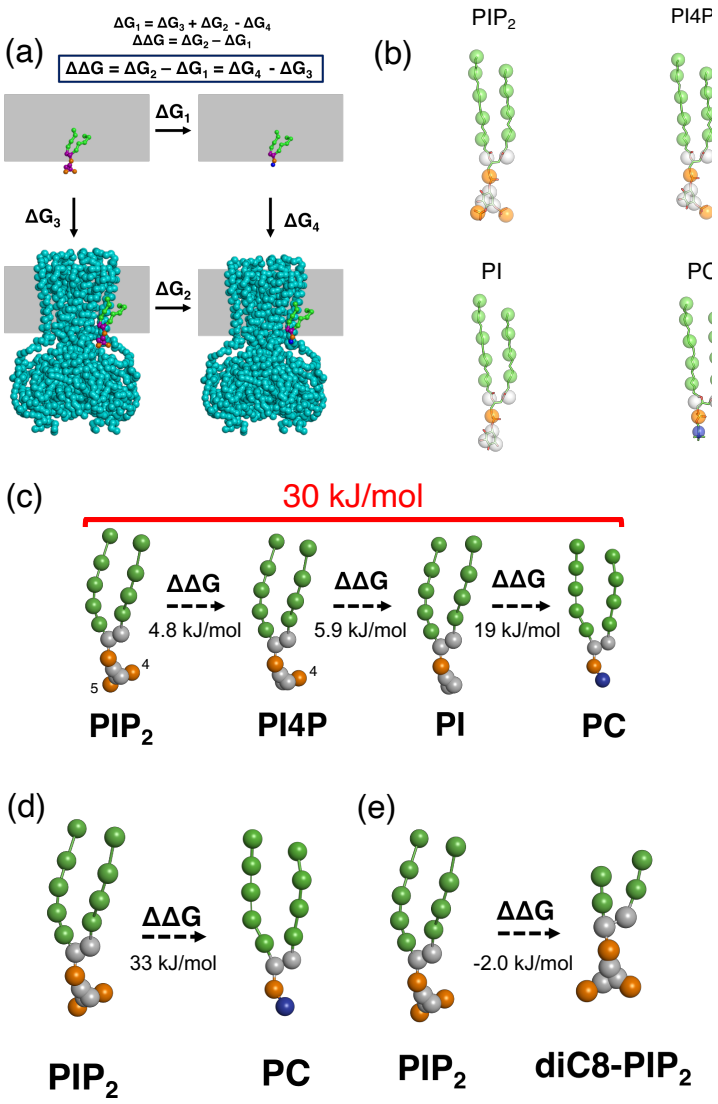


Figure 3.4 The free energy calculation of an individual phosphate group and fatty acid chains on a hKir6.2 tetramer

(a) Thermodynamic cycle used for the relative binding free energy calculations. The perturbation of the PIP_2 headgroup (purple) was calculated in both the channel-bound state (ΔG_2) and unbound state (grey rectangle (ΔG_1)).
(b) Coarse-grain to atomistic mapping of the phosphoinositides PIP_2, PI4P, PI and PC.
(c) Change in binding free energy ($\Delta\Delta G$) when individual phosphate groups are perturbed: i.e. from PIP_2 to PI4P, from PI4P to PI and from PI to PC (values in black). The sum of these free energy changes (i.e. from PIP_2 to PC) is given in red. Values are rounded to two significant figures.
(d) Change in binding free energy when PIP_2 is perturbed to PC. Values are rounded to two significant figures.
(e) Change in binding free energy when PIP_2 is perturbed to PIP_2-diC8. Values are rounded to two significant figures.

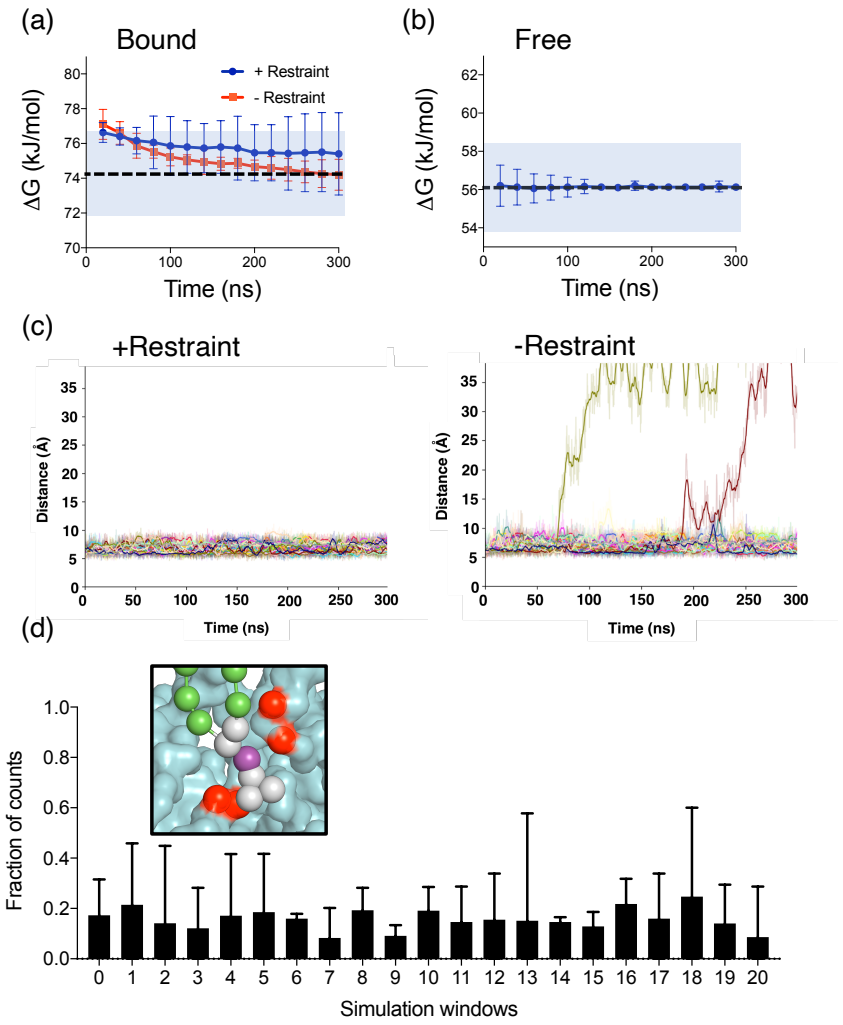


Figure 3.5 Convergence analysis of PI to PC free energy perturbation.

(a) Convergence of the free energy calculation during PI to PC perturbation in the presence of hKir6.2. The analysis was carried out with (blue) and without (red) the flat-bottom restraint. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(b) Convergence of the free energy calculation during PI to PC perturbation in the bulk PC bilayer in the absence of hKir6.2. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(c) The distance between the phosphate headgroup (PO4 particle) and the average position of four amino acid residues (red), in the presence and absence of a flat-bottom restraint. These four residues were chosen as they are 6 Å away from the lipid headgroup. Different colours represent the simulations in the different alchemical states (λ windows) of PI to PC transformation.
(d) Fraction of counts where the PO4 particle experiences the flat-bottom restraint in each simulation window. Inset: A flat bottom restraint was imposed between the PO4 particle (purple) and the protein backbone (red). The error bars are representing 95% confidence interval around the mean.

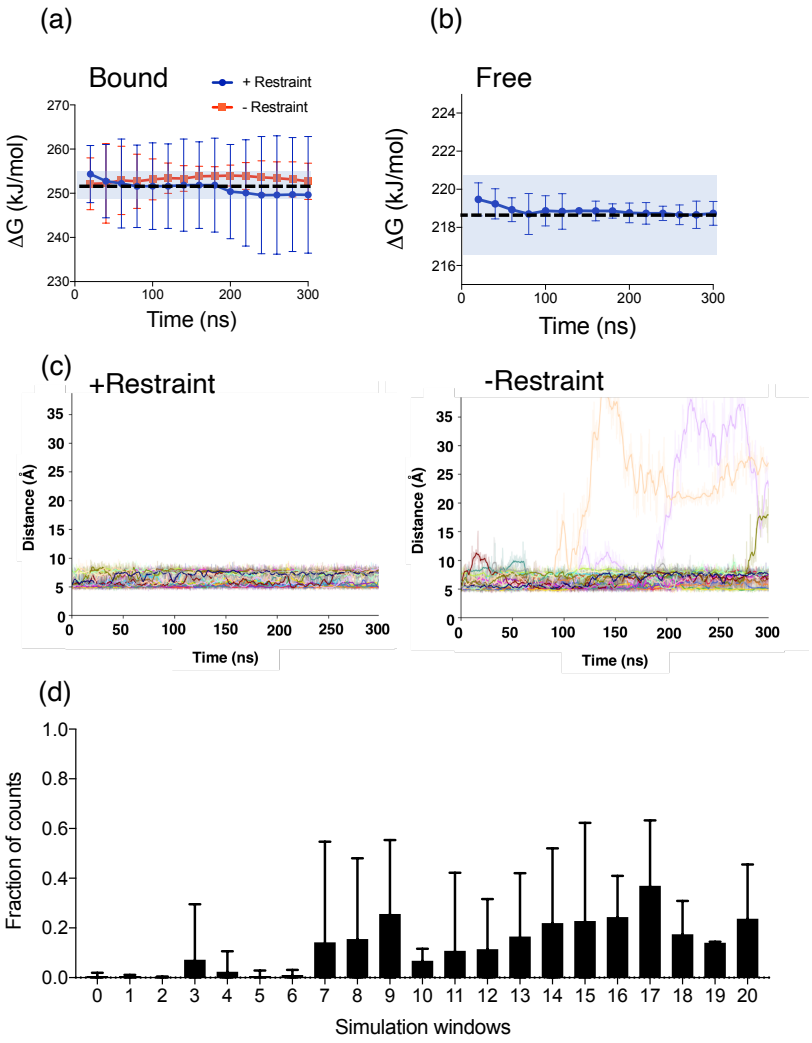


Figure 3.6 Convergence analysis of PIP_2 to PC free energy perturbation.

(a) Convergence of the free energy calculation during PIP_2 to PC perturbation in the presence of hKir6.2. The analysis was carried out with (blue) and without (orange) the flat-bottom restraint. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(b) Convergence of the free energy calculation during PIP_2 to PC perturbation in the bulk PC bilayer. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(c) The distance between the phosphate headgroup (PO4 particle) and the centre position between the backbone of 4 amino acid residues (red) in the presence and absence of a flat-bottom restraint. These residues were chosen as they are 6 Å away from the lipid headgroup. Different colours represent the simulations in the different alchemical states (λ windows) of PIP_2 to PC transformation.
(d) Fraction of counts where the PO4 particle experiences the flat-bottom restraint in each simulation windows. The error bars are representing 95% confidence interval around the mean.

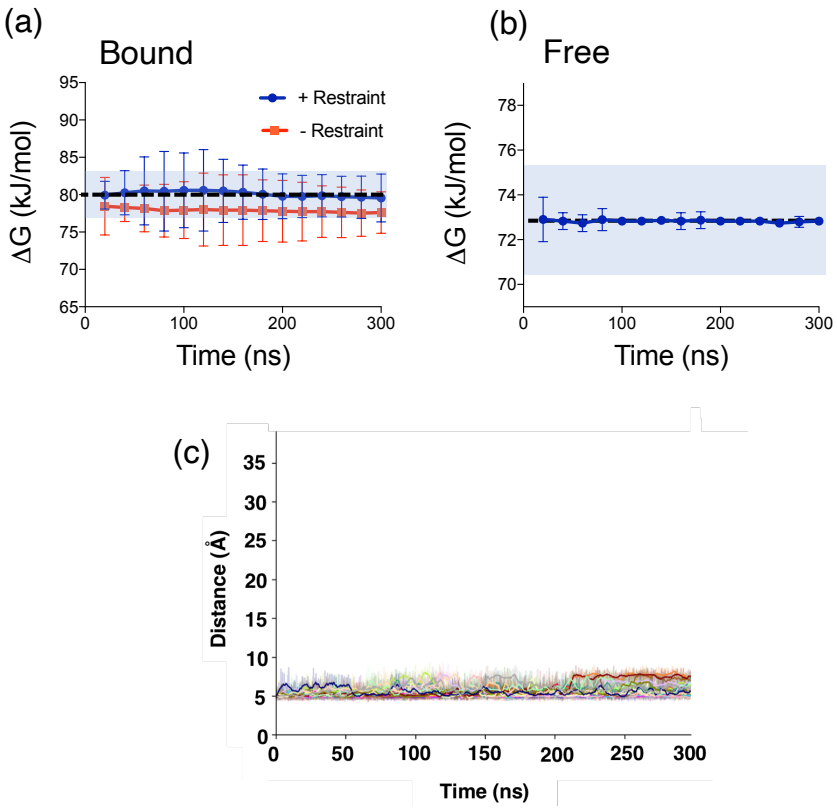


Figure 3.7 Convergence analysis of PIP_2 to PI4P free energy perturbation.

(a) Convergence of the free energy calculation during PIP_2 to PI4P perturbation in the presence of hKir6.2. The analysis was carried out with (blue) and without (orange) the flat-bottom restraint. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(b) Convergence of the free energy calculation during PIP_2 to PI4P perturbation in the bulk PC bilayer. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(c) The distance between the phosphate headgroup (PO4 particle) and the flat-bottom restraint. Different colours represent the individual simulations in the different alchemical states (λ windows) of the PIP_2 to PI4P transformation.

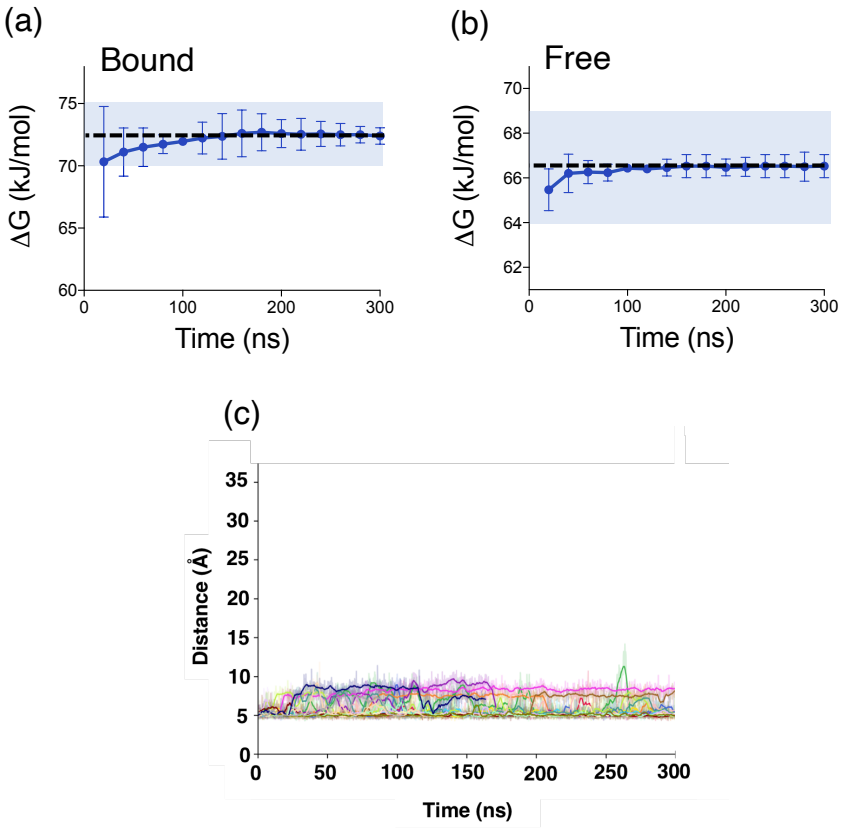
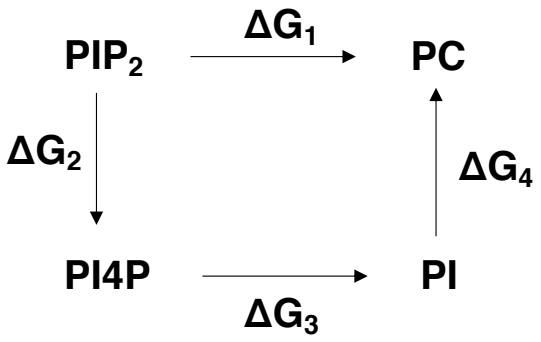


Figure 3.8 Convergence analysis of PI4P to PI free energy perturbation.

(a) Convergence of the free energy calculation during PI4P to PI perturbation in the presence of hKir6.2. The analysis was carried out with (blue) and without (orange) the flat-bottom restraint. The blue shaded region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.
(b) Convergence of the free energy calculation during PI4P to PI perturbation in the bulk PC bilayer. The blue shaded region illustrates the thermal fluctuation of the system, i.e,. k_BT. The error bars are representing 95% confidence interval around the mean.
(c) The distance between the phosphate headgroup (PO4 particle) and the flat-bottom restraint. Different colours represent the simulations in the distinct alchemical states (λ windows) of the PI4P to PI transformation.

Transformation from PIP_2 to PI4P showed a very small relative free energy change (Figure 3.4b and Table 3.4). This suggested that a phosphate group at either the 5' position does not make a substantial contribution to PIP_2 binding (Figure 3.4b and Table 3.4). However, I observed large free energy changes when the last phosphate group at the 4' position and the inositol ring were perturbed (Figure 3.4b and Table 3.4). Summation of each individual moiety from PIP_2 to PC gives a binding free energy of 30 ± 3.7 kJ/mol, which is remarkably similar to the 33 ± 13 kJ/mol I obtain for direct perturbation of PIP_2 to PC (Figure 3.4c, Table 3.4, and Figure 3.9). This suggests that the approach I use is valid for both single and multi-step free energy calculations.



$$\Delta G_1 = \Delta G_2 + \Delta G_3 + \Delta G_4$$
$$\Delta G_1 = 5 + 6 + 19$$
$$\Delta G_1 = 30 \text{ kJ mol}^{-1}$$

$$\Delta G_{1(calc)} = 31 \text{ kJ mol}^{-1}$$

Figure 3.9 A complete thermodynamic cycle in PIP_2 stepwise perturbation

The thermodynamic cycle representing relative binding free energy between phosphoinositide lipids (PIP_2, PI4P and PI) and PC. The calculation representing the sum of the small steps is equivalent to the direct perturbation.

Transformation		ΔΔG
From	To	(kJ/mol)
PIP_2	PI4P	4.8 ± 2.7
PI4P	PI	5.9 ± 0.62
PI	PC	19 ± 2.4
PIP_2	PC	33 ± 13
PIP_2	PIP_2-diC8	-2.0 ± 1.8

Table 3.4 Calculations of the ΔΔG on an individual phosphate group and the fatty acid chains on a hKir6.2 tetramer

Change in binding free energy (ΔΔG) when individual phosphate groups are perturbed: i.e. from PIP_2 to PI4P, from PI4P to PI and from PI to PC, when PIP_2 is perturbed to PC and when PIP_2 is perturbed to PIP_2-diC8. Values are rounded to the nearest 2 significant figures (n=3). Errors are quoted with 95% confidence interval.

A previous study demonstrated that crosstalk between different anionic lipids can affect the affinities of each lipid for the Kir2.2 channel (Duncan *et al.*, 2020). Figure 3.10 shows that the presence of 10% anionic lipid (PS) in the lower leaflet of the bilayer in my simulations does not affect the overall ΔΔG for PIP_2 binding to Kir6.2.

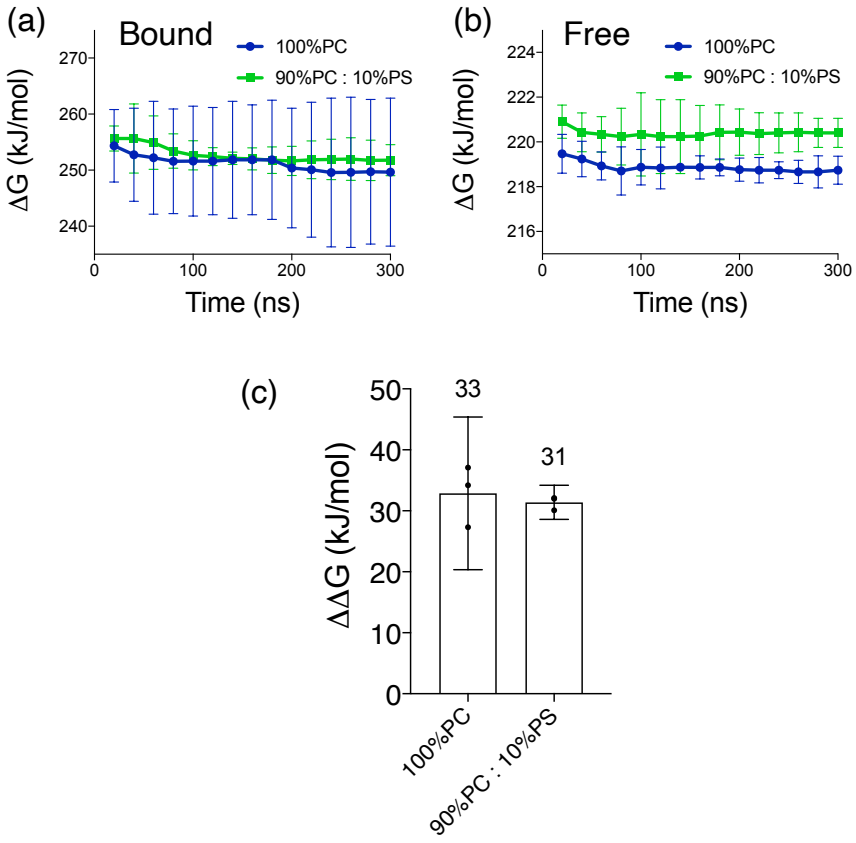


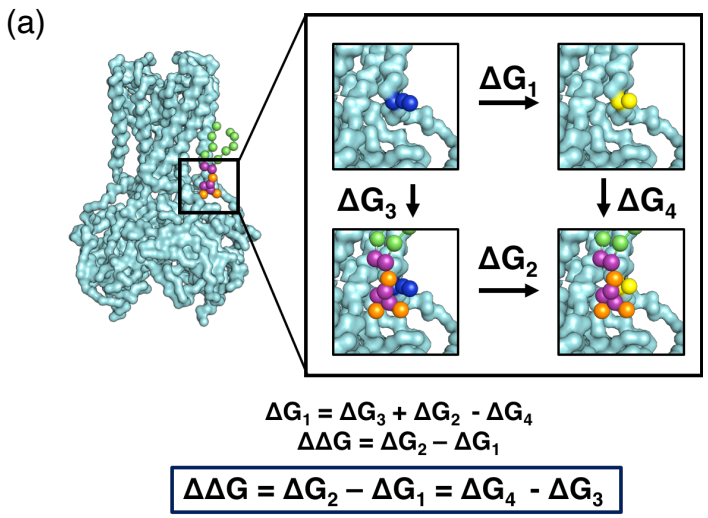
Figure 3.10 Convergence analysis of PIP_2 to PC free energy perturbation in anionic lipid environment.

(a) Convergence of the free energy calculation during PIP_2 to PC perturbation in the presence of hKir6.2. The analysis was carried out in 100% PC bilayer (blue) or mixed bilayer containing 10% PS (green). The error bars are representing 95% confidence interval around the mean.
(b) Convergence of the free energy calculation during PIP_2 to PC perturbation in the bulk PC bilayer. The analysis was carried out in 100% PC bilayer (blue) or mixed bilayer containing 10% PS (green). The error bars are representing 95% confidence interval around the mean.
(c) Binding free energy between PIP_2 and Kir6.2 in 100% PC bilayer or in 10% PS. Values are rounded to the nearest 2 significant figures. Error bars represent the 95% confidence interval (n=3)

Previous electrophysiological and crystallographic studies have commonly used the soluble eight-carbon atom phosphatidylinositol, diC8-PIP_2, to study channel activation (Haider *et al.*, 2007; Hansen *et al.*, 2011; Whorton & MacKinnon, 2011). Therefore, I investigated the effect of the length of the acyl chain on PIP_2 binding affinity. I found that truncation of the acyl chain from either 4 or 5 particles (i.e., palmitoyl and oleoyl) to 2 particles (equivalent to 8 carbon atoms) had no effect on the calculated PIP_2 affinity (Figure 3.4d). This suggests that PIP_2-diC8 is indeed an effective substitute for investigating the impact of PIP_2 binding in electrophysiological and structural studies.

3.3.3 Relative binding free energy calculations for PIP_2 interactions with both wild-type and mutant hKir6.2

Based on the closed state model of mouse Kir6.2 (PDB: 6BAA), I generated three structural models of hKir6.2 with disease-associated mutations; K67N which causes CHI, and E179A and E179K which cause neonatal diabetes (Huopio *et al.*, 2002; Flanagan *et al.*, 2007). As these residues are near the PIP_2 binding site (Figure 3.1e), I hypothesised that mutation to these residues would modulate PIP_2 affinity and thereby affect the channel open probability (P_{open}). An increase in PIP_2 binding affinity should correlate with an increase in channel activation and thus also a reduced inhibition by ATP.



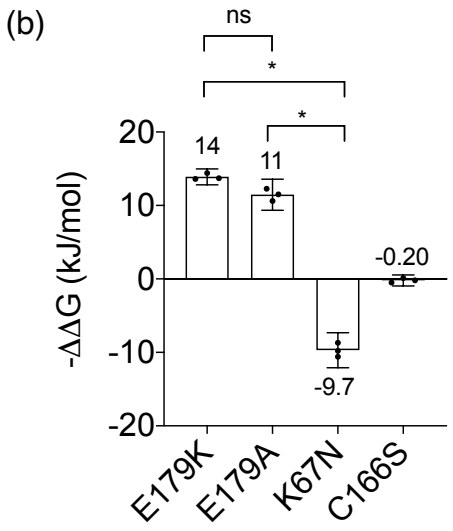


Figure 3.11 Free energy calculations of the disease associated Kir6.2 mutations.

(a) Schematic diagram showing the free energy calculation. An amino acid residue - lysine (blue sphere) is transformed into glutamine (yellow sphere) between two states (PIP_2 bound and free). The surface of the protein backbone is shown as cyan. Inositol group on PIP_2 are shown in purple. The phosphate particles are shown in orange. The alkyl tails are shown in green.
(b) PIP_2 binding free energy differences between E179K, E179A, K67N and C166S mutations comparing to WT during amino acids CG-GEP calculated based on the schematic diagram in (a). The error bars are representing 95% confidence interval around the mean. * $P < 0.05$ (Student's t-test).

I next performed calculations in which I mutated hKir6.2 residue in the presence and absence of PIP_2 (Figure 3.11a). This allows us to calculate the relative changes in the PIP_2 binding free energy between the wild-type and mutant channels. For the highest energy residue substitution (E179K), I observed convergence of the free energy calculations within 50 ns per window (Figure 3.12). The data show an increase in PIP_2 binding energy, and hence an increased affinity, with the E179K transformation. I also observed an increase in binding free energy with the E179A mutation (Figure 3.11b and Table 3.5). Conversely, I observed a reduction in PIP_2 binding affinity with the K67N transformation (Figure 3.11b and Table 3.5). This quantitatively confirms that both the E179K and E179A mutations increase PIP_2 channel affinity, whereas the K67N mutation decreases channel affinity. This agrees with the patient phenotypes: E179K causes NDM, due to an increase in channel activity, whereas K67N causes CHI, which is associated with a reduction in channel activity. This supports the in *vitro* experimental results conducted by Dr Peter Proks (Pipatpolkai *et al.*, 2020*a*).

Mutation	-ΔΔG (kJ/mol)
E179K (K→E)	14 ± 1.1
E179A	11 ± 2.2
K67N	-9.7 ± 2.4
C166S	-0.20 ± 0.75

Table 3.5 Free energy calculations using disease associated hKir6.2 mutations.

The energetic cost of making the residue mutation based on the schematic diagram (Figure 3.11b). ΔΔG are calculated as the binding free energy difference between mutants and WT. The positive ΔΔG indicates an increase in binding free energy. All values are quoted to 2 significant figures with t-statistic.

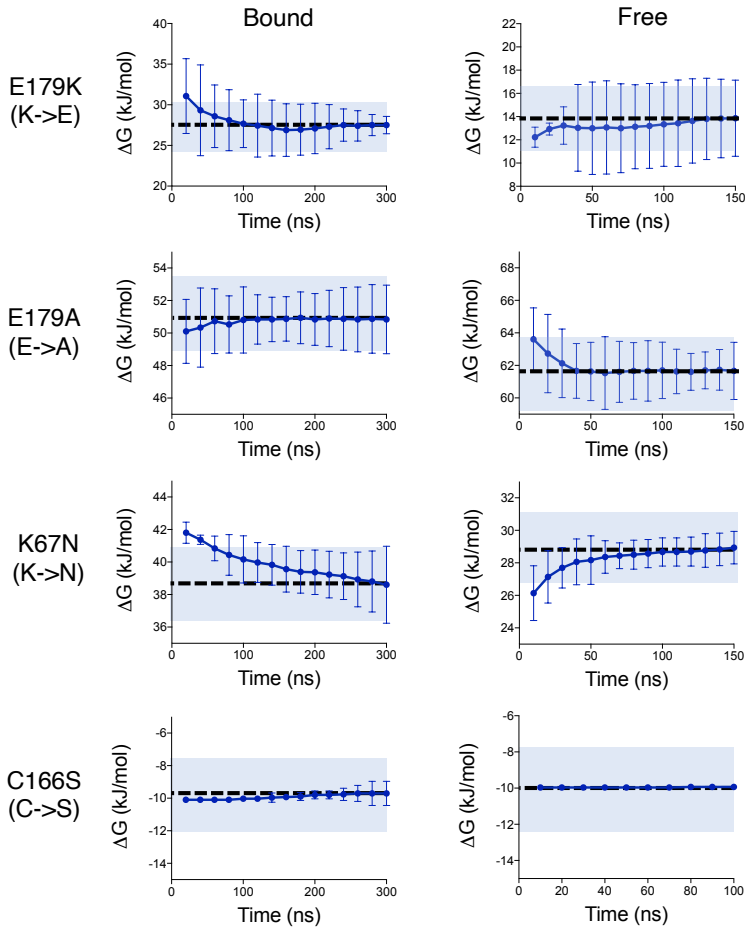


Figure 3.12 Convergence analysis of E179K (K→E), E179A, K67N and C166S free energy perturbation.

Convergence of the free energy calculation during lysine to glutamate perturbation for residue 179 (E179K), glutamate to alanine perturbation for residue 179 (E179A), lysine to asparagine perturbation for residue 67 (K67N) and cysteine to serine perturbation for position 166 (C166S) of hKir6.2 in (a) the presence and (b) the absence of PIP_2. The shaded blue region illustrates the thermal fluctuation of the system, i.e., k_BT. The error bars are representing 95% confidence interval around the mean.

As a control, I investigated an NDM mutation, C166S, which is distant from the PIP_2 binding site. I expected that this mutation would have no impact on the PIP_2 binding affinity, despite having an influence on the channel opening probability (Trapp *et al.*, 1998). When perturbing this site using single residue FEP, I see no change in the PIP_2 affinity of the channel, and therefore, unlike the E179A/K mutations, the C166S mutant does not appear to increase channel opening probability by increasing PIP_2 affinity (Table 3.5). Rather, it must exert its effect by a different mechanism. For example, it might affect the open state stability of the channel directly.

3.3.4 Assessment of the PIP_2 activity dependency on SUR1 subunit

We collaborated with Dr Peter Proks to experimentally assessed the importance of the SUR1 subunit for PIP_2 binding affinity and activation. Previous studies suggested that the presence of SUR1 enhances the P_{open} of Kir6.2ΔC (Tucker *et al.*, 1997; Baukrowitz *et al.*, 1998). However, the contribution of PIP_2 to this modulation and the relationship between SUR1 and PIP_2 sensitivity remains unclear. To address this issue, I calculated the relative binding free energy of hKir6.2 and PIP_2 in both the presence and absence of the SUR1 subunit. I showed that addition of SUR1 only marginally increased the PIP_2 binding free energy (Figure 3.13a, Table 3.6). Therefore, this result suggests that SUR1 makes only a minor contribution to PIP_2 affinity even though it is only approximately 8 Å away from PIP_2 headgroup. This agrees with experiments conducted by Dr Peter Proks, which assessed the neomycin sensitivity of hKir6.2 in the presence and absence of SUR1 on *Xenopus* oocyte. The ΔC construct is used to allows Kir6.2 to traffic to the plasma membrane in an absence of SUR1. His works showed that the presence of SUR1 does not significantly alters channel neomycin sensitivity. (Kir6.2/SUR1: IC_{50} = 81 μM, h = 0.96, Kir6.2ΔC, IC_{50} = 54 μM, h = 0.61) (Figure 3.13b). This is in qualitative agreement with the lack of a change in the free energy of PIP_2 binding obtained from my FEP calculations.

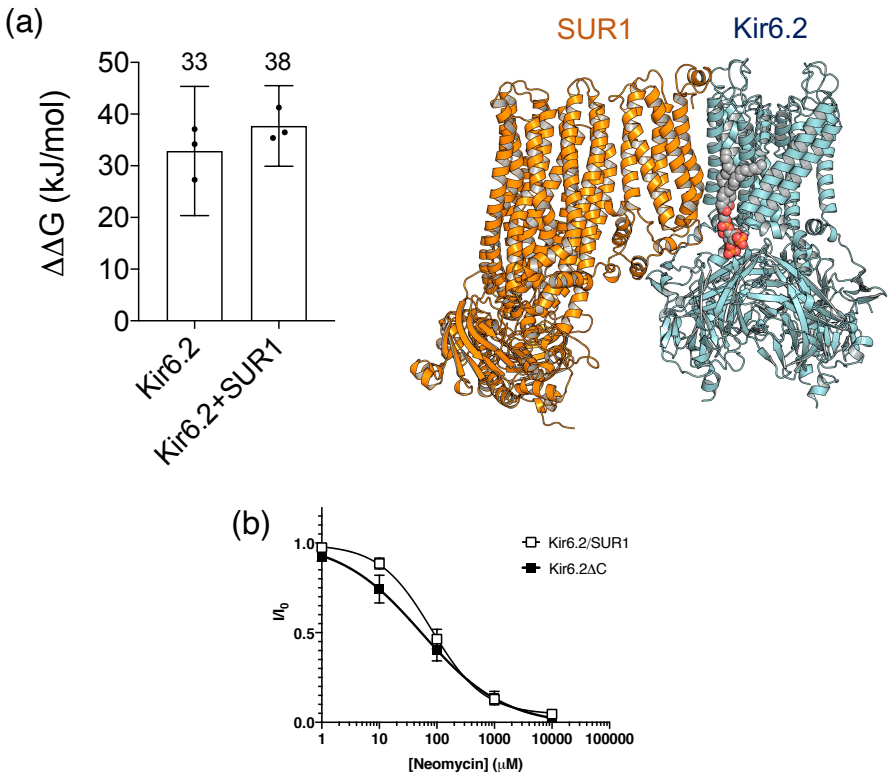


Figure 3.13 Effect of SUR1 on PIP_2 affinity and activation.

(a) (Left) Binding free energy between PIP_2 and Kir6.2 ± SUR1. Values are rounded to the nearest 2 significant figures. Error bars represent the 95% confidence interval (n=3) (Right) The PIP_2 (grey) binding sites between hKir6.2 and SUR1. Kir6.2 are shown in blue and SUR1 are shown in orange.
(b) Mean relationship between the neomycin concentration and the K_{ATP} current (I), expressed relative to the current in the absence of neomycin (I_0), for Kir6.2 co-expressed with SUR1 (open squares, n=5) or hKir6.2ΔC expressed without SUR1 (filled squares, n=5). The ΔC construct removes the ER retention signal from the C-terminus of the protein. This allows the Kir6.2 channel to be trafficked to the plasma membrane in the absence of SUR1.

Protein	ΔΔG (kJ/mol)
hKir6.2	33 ± 13
hKir6.2 + SUR1	38 ± 7.8

Table 3.6 Free energy calculations using PIP_2 to PC transformation

Binding free energy changes between wild-type and mutated hKir6.2 channels and hKir6.2 channel with SUR1 as I perturbed PIP_2 to PC. Values are rounded to two significant figures. The number represents the mean and 95% confidence interval (n=3).

3.3.5 PIP_2 binding affinity to other human inward rectifying potassium (hKir) channels

I next assessed the binding of PIP_2 to other hKir channels by calculating the binding free energy of PIP_2 perturbation to PC for the human Kir1.1, Kir2.2 and Kir3.2 (hKir1.1, hKir2.2 and hKir3.2) channels using the thermodynamic cycle described in Figure 3.3a. The electrophysiological behaviour of these channels on PIP_2 binding is well characterised (Rohács *et al.*, 2003). I perturbed PIP_2 to PC, the dominant phospholipid species in the eukaryotic plasma membrane. Note that a potential of mean force (PMF) calculation shows the binding energy of PC to hKir6.2 is 0 ± 2 kJ/mol (Figure 3.14). Due to the absence of human Kir1.1 and human Kir2.2 structures in the PIP_2 bound conformation, I generated hKir molecular models of both, as described in the Methods. The PIP_2 binding sites and interacting residues on these proteins are all highly conserved (Figures 3.15, 3.16).

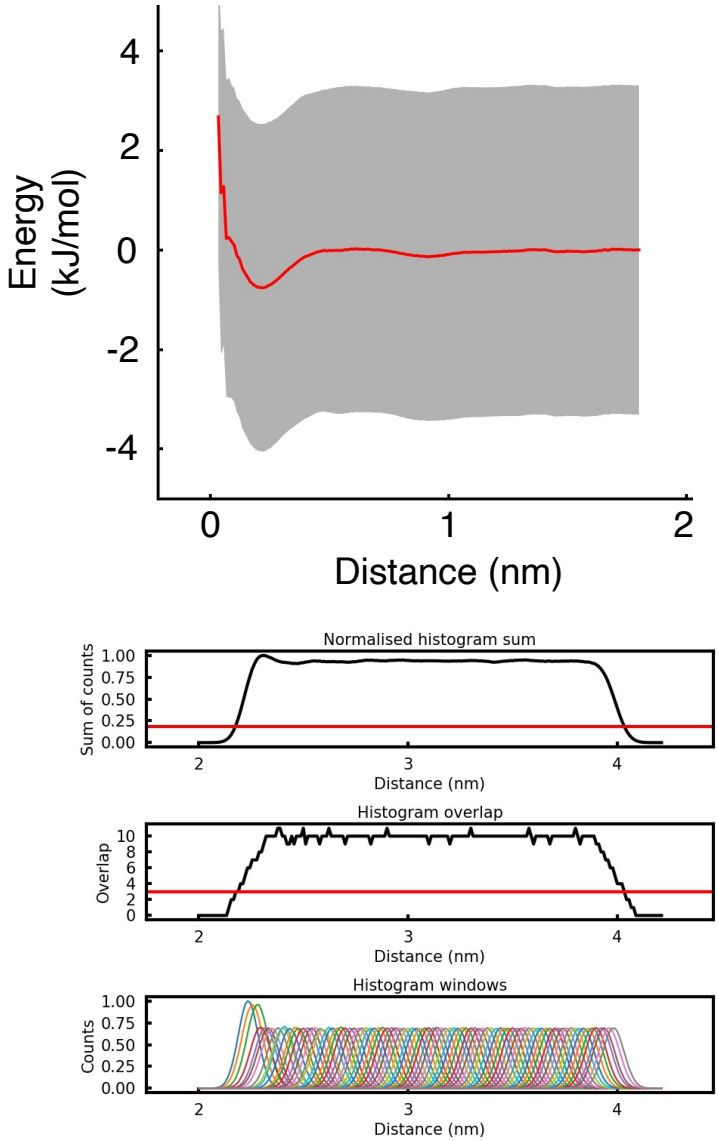


Figure 3.14 PMF calculation and convergence analysis for PC binding to hKir6.2

The 1D energy landscape for PC binding to hKir6.2. The PC headgroup is steered away from the PIP_2 binding site along the collective variable. The top graph represents the energy being sampled within each umbrella sampling window at different distance as PC headgroup is being steered away. This yield an energy of *ca*. 0 kJ/mol. The sum of the total counts along the reaction co-ordinate between Kir6.2 and PC extracted from a histogram, which is marked on the second panel. A line at 0.2 marks sufficient sampling. The total counts are normalised to 1.0. The count of overlap histograms at each point along the reaction coordinate from the data above where only histograms which overlap at >10% of the total histogram height are counted. A red line marks the optimal overlap threshold of 3 based on Corey *et al.* (2019). The histogram windows are extracted from WHAM analysis. Each colour of the histogram represents a single simulation window. Here, the window is clearly oversampled and thus, similar calculation could be achieved with 0.5 Å spacing instead of 0.2 Å.

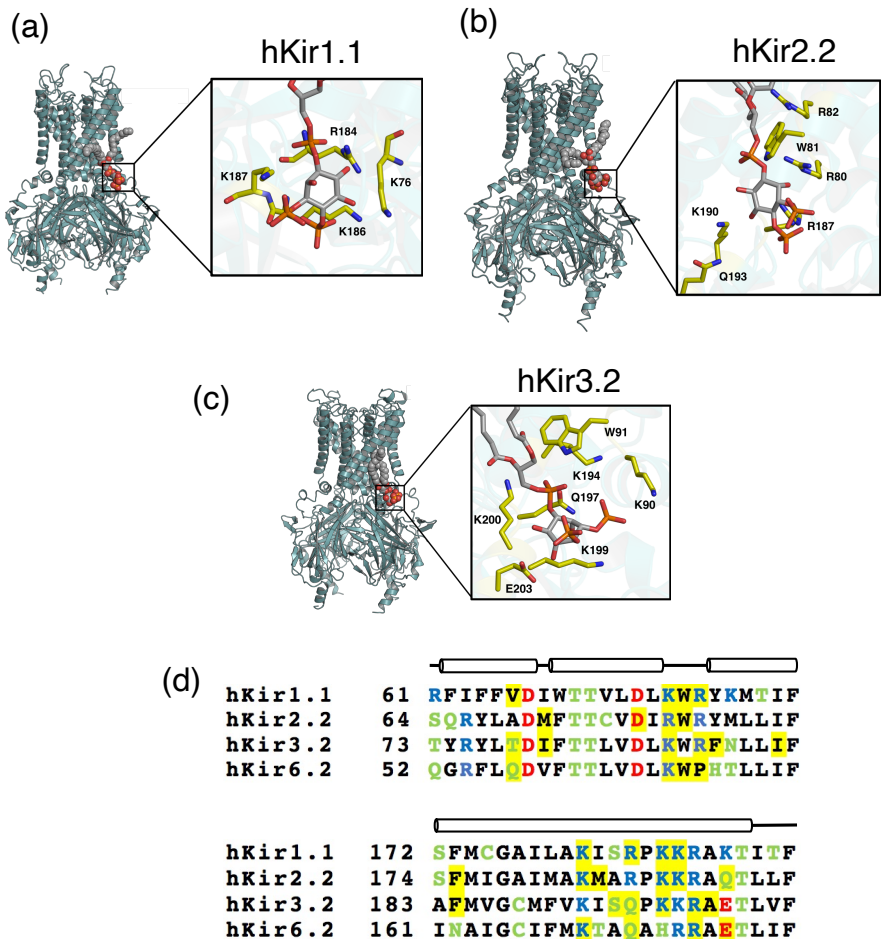


Figure 3.15 PIP_2 binding configuration in different hKir channels.

Models of (a) hKir1.1, (b) hKir2.2 and (c) hKir3.2 channels in the PIP_2 bound conformations after 1 μs of CG simulation and conversion back to an atomistic description. Insets: Carbons of key PIP_2 binding residues are highlighted in yellow, with the rest of PIP_2 shown in CPK colours.

(d) Sequence alignment between hKir1.1, hKir2.2, hKir3.2 and hKir6.2 channels in the region where the contacts are conserved between more than two channels. Highlighted in yellow are residues that contact PIP_2 for more than 70% of the 1 μs simulations (n=5). The cylinders represent α-helices in the secondary structure and the lines represent either a disordered region or a kink within the α helix. Acid residues (Asp, Glu) are shown in red, basic residues (Lys, Arg) are shown in blue, Polar residues (Ser, Thr, Cys, Gln, Asn) are shown in green. Other residues are shown in black.

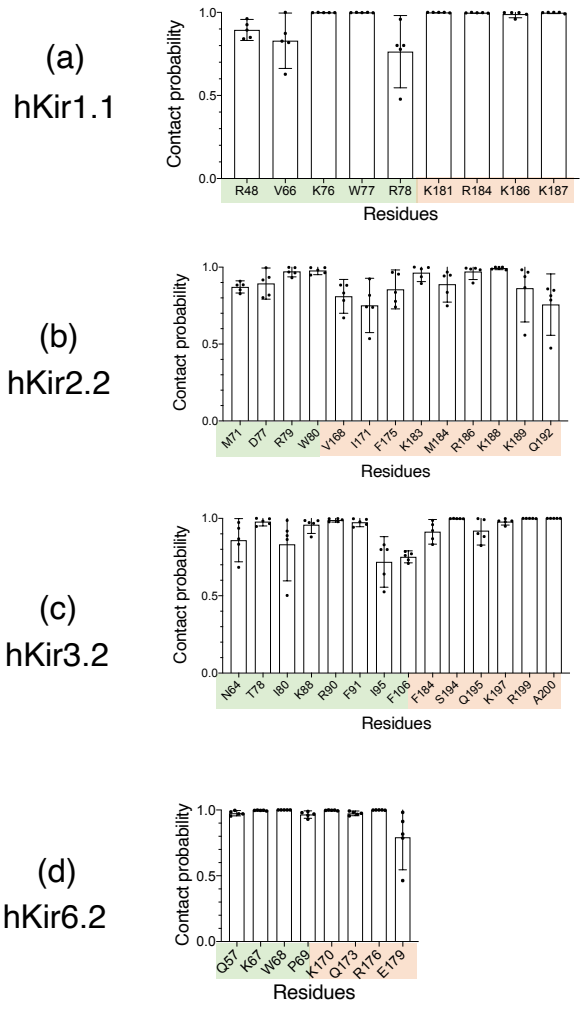


Figure 3.16 A PIP_2 binding site on hKir channels

PIP_2 contact analysis showing residues which make contact >75% of the time during five 1 μs simulations for (a) hKir1.1, (b) hKir2.2, (c) hKir3.2 or (d) hKir6.2 channels. Residues in the N-terminal domain are highlighted in green and in the C-terminal domain are coloured in orange. The error bars are representing 95% confidence interval around the mean.

Our data reveal that the binding free energy of PIP_2 to hKir1.1, hKir2.2 and hKir3.2 channels is higher than that to hKir6.2 channels (Figures 3.17a and Table 3.7). In my study, the PIP_2 binding free energy value for hKir2.2 is similar to that previously recorded for the chicken Kir2.2 channel, which was reported as *ca.* -46 kJ/mol both using FEP and PMF (Domański *et al.*, 2017; Corey *et al.*, 2019*b*). The higher affinity of hKir3.2 was rather unexpected, as hKir3.2 also has a glutamate in the equivalent position to E179 in hKir6.2 (denoted as E201) (Figure 3.14d). However, the minimum distance between the PIP_2 headgroup and hKir3.2-E201 (*ca.* = 8.5 Å) is greater than that of than that of hKir6.2-E179 (*ca.* = 5 Å) in my molecular models.

Protein	**ΔΔG (kJ/mol)**
hKir1.1	53 ± 4.9
hKir2.2	41 ± 4.8
hKir3.2	47 ± 3.0
hKir6.2	33 ± 13

Table 3.7 Free energy calculations on different hKir channels

Binding free energy changes between different hKir channels as I perturbed PIP_2 to PC. Values are rounded to the 2 significant figures with 95% confidence interval (n=3).

To account for the free energy difference between PIP_2 and PC to the different hKir channels, I examined the contribution of the individual phosphate and inositol groups of PIP_2 in the binding to each of the channels, as described above (Figure 3.4a). First, I showed that the binding energy contributed by the 5' phosphate is higher in hKir1.1 than the other three channels (PIP_2>PI4P; Figure 3.17b and Table 3.8). In addition, I observed that 4' phosphate and the inositol ring contribution is stronger in all other channels relative to hKir6.2 (PI4P>PI; Figure 3.17b and Table 3.8). Overall, this accounts for most of the energy differences as I perturbed PIP_2 to PC in hKir1.1, hKir2.2 and hKir3.2. For both single- and multi-step approaches, hKir6.2 exhibited the lowest binding free energy for PIP_2.

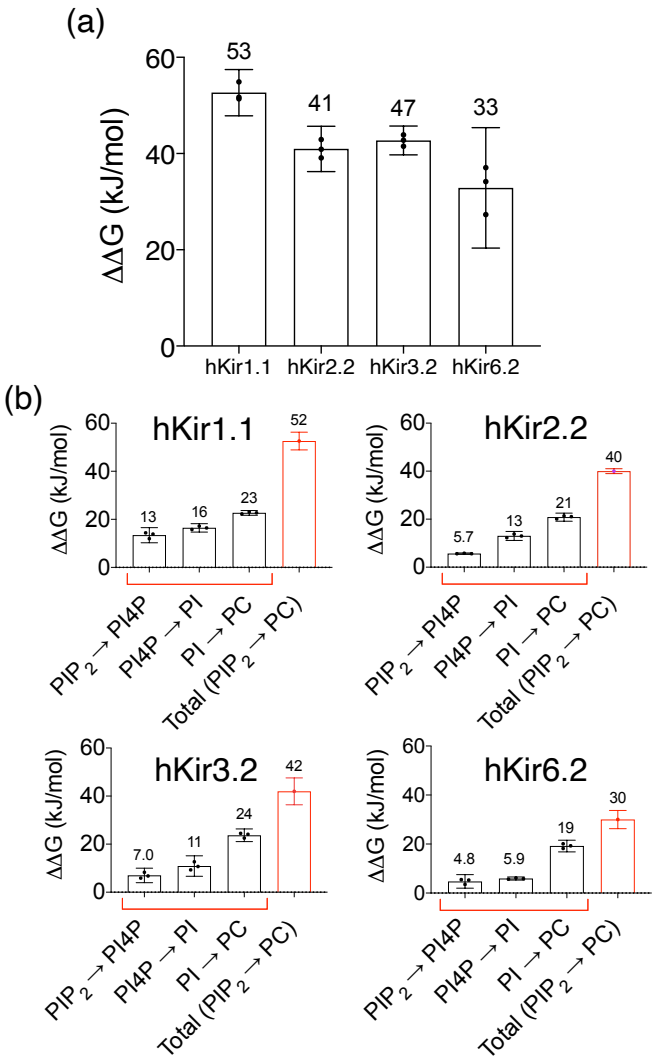


Figure 3.17 Free energy calculations for different hKir channels

(a) Binding free energy changes between different hKir channels as PIP_2 is perturbed to PC. (n=3). The error bars are representing 95% confidence interval around the mean.
(b) Binding free energy changes between PIP_2 and hKir1.1, hKir2.2, hKir3.2 or hKir6.2 as each PIP_2 phosphate group is sequentially perturbed: from PIP_2 to PI4P, then to PI and finally to PC (black). The sum of the free energy change from PIP_2 to PC is shown in red. Values are rounded to the nearest two significant figures. Error bars represent the 95% confidence interval (n=3).

Protein	Transformation		ΔΔG (kJ/mol)
	From	To	
hKir1.1	PIP_2	PI4P	13 ± 3.1
	PI4P	PI	16 ± 1.8
	PI	PC	23 ± 1.0
	Sum		**52 ± 3.7**
hKir2.2	PIP_2	PI4P	5.7 ± 0.30
	PI4P	PI	13 ± 1.8
	PI	PC	21 ± 1.7
	Sum		**40 ± 2.5**
hKir3.2	PIP_2	PI4P	7.0 ± 3.0
	PI4P	PI	11 ± 4.3
	PI	PC	24 ± 2.6
	Sum		**42 ± 5.9**
hKir6.2	PIP_2	PI4P	4.8 ± 2.7
	PI4P	PI	5.9 ± 0.62
	PI	PC	19 ± 2.4
	Sum		**30 ± 3.7**

Table 3.8 Calculations of the ΔΔG on an individual phosphate group on hKir channels.

Change in binding free energy (ΔΔG) when individual phosphate groups are perturbed (i.e., from PIP_2 to PI4P, from PI4P to PI and from PI to PC. The sum of these values is shown in bold. Values are rounded to the nearest 2 significant figures (n=3). Errors are quoted with 95% confidence interval.

3.3.6 Convergence analysis and method development

In my simulations, I showed that the calculation converged at 300 ns (n=3). However, given that my simulations were calculated at CG-level, I have attempted different approach at running more repeats and calculated the error based on *t*-statistic to decrease the computational cost of the future projects. Using similar approach to previous literature, I showed that the binding free energy of hKir2.2 can be reliable calculated using 12 ns simulation at 15 repeats (95% CI is lower than 1 k_BT) (Figure 3.18) (Corey *et al.*, 2021). This approach is simply possible due to a much smoother energy landscape of the CG simulation. Thus, only up to 12 ns is required for a decent distribution convergence of a simulation based on the t-statistic. This yielded a total of 180x2=360 ns total simulation time instead of 300x3x2 = 1800 ns. These could also be improved using non-equilibrium FEP, which is currently being developed to reduce the simulation cost even further in the future studies.

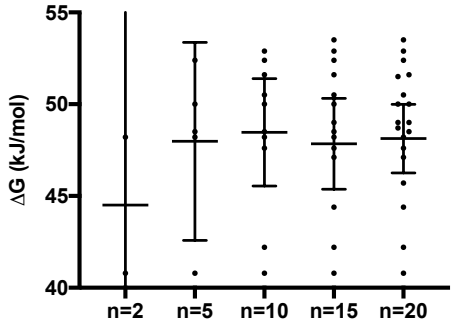


Figure 3.18 PIP_2 binding free energy on hKir2.2 at different repeats

ΔG binding free energy of PIP_2 binding to hKir2.2 channel calculated in different numbers of repeats. The calculation involve transformation of PIP_2 to PC on the hKir2.2 PIP_2 binding site where the energy is expected to be ~46 kJ/mol. The error bars show 95% confident interval based on *t-statistic.* The data suggested that 10-15 repeats of the short FEP calculation with t-statistic is optimal to determine the PIP_2 binding free energy.

3.4 Discussion

Our results showed that coarse-grained simulations could be used to predict the binding free energy of a ligand at a lesser computational cost than atomistic simulations. This enabled us to predict the binding free energy of PIP_2 to Kir6.2 and to other Kir channels. My data revealed that PIP_2 binds to Kir6.2 with a higher affinity than PI4P and with a much higher affinity than PI. I also found that SUR1 has no significant effect on PIP_2 binding. Mutations of residues in the PIP_2 binding site associated with neonatal diabetes led to increased PIP_2 binding, whereas those that reduced binding were associated with CHI. Comparison of phosphoinositide binding to other Kir channels showed changes in affinity that can account for the functional data.

Molecular modelling

I modelled the PIP_2 binding site of the hKir1.1, hKir2.2, hKir3.2 and hKir6.2 channels and predicted their PIP_2 binding affinity using CG-MD. The models of the hKir.1.1 and hKir2.2 were generated based on the PIP_2 bound chicken Kir2.2 (cKir2.2) crystal structure. The structure of hKir3.2 was based on the crystal structure of PIP_2 bound mouse Kir3.2. The hKir6.2 model was based on the cryo-EM structure of the ATP and glibenclamide bound K_{ATP} channel in the absence of SUR1. This is an improvement to the existing model of Kir6.2 which was generated based on cKir2.2 and KirBac channels (Stansfeld *et al.*, 2009). In all structures and models, the diameter of the lower

bundle-crossing gate is narrower than the size of the hydrated K^+ ion, which suggests that this channel is closed, despite bound PIP_2. Thus, it may be an intermediate state that occurs between PIP_2 binding and channel opening. Despite this caveat, I demonstrated that the residues in binding site obtained from my models and CG-simulations are involved in Kir channels activation by PIP_2 based on previous electrophysiological studies (Zhang *et al.*, 1999; Shyng *et al.*, 2000; Soom *et al.*, 2001; Haider *et al.*, 2007).

Calculation of binding free energy on different inositides and its caveats

To determine the contribution of the inositol ring and the alkyl group to the binding of PIP_2, I built my CG-free energy calculation based on my previous approach for comparing the binding free energy between two lipid species to a given site on a membrane protein (Corey *et al.*, 2019*b*). I showed that this approach enables us to complete a thermodynamic cycle where the sum of the individual perturbation steps (i.e., PIP_2>PI4P>PI>PC) is equivalent to the single-step transformation from PIP_2 to PC. The results of my free-energy calculations are within the range of values calculated from CG-simulation which are commonly observed for PIP_2 interactions with other membrane proteins (Hedger *et al.*, 2016*a*; Domański *et al.*, 2017; Naughton *et al.*, 2018; Corey *et al.*, 2019*b*). My simulations also showed that the affinity of the channel for diC8:PIP_2 is not significantly different from that of PIP_2. This argues that diC8:PIP_2 would be a suitable substitute for PIP_2, if the concentration of diC8:PIP_2 applied to the membrane is the same as the concentration of PIP_2 at the channel. My calculations also showed that removal of the inositol group

reduces the PIP_2 binding activity more than the removal of the phosphate groups. However, electrophysiological studies have shown that PI is unable to activate the K_{ATP} channel (Fan & Makielski, 1999; Rohács *et al.*, 2003). Thus, this suggests that PIP_2 binding may be driven by the inositol ring, but that the interaction of the phosphate group with the protein is required for transduction.

One concern is an inability of the CG forcefield to distinguish between similar inositol lipids in the free energy calculations. This application of CG-FEP has demonstrated that the method can show differences between PIP_2, PI4P, PI and PC at the accuracy of at least 5 kJ/mol (~1.5 k_BT). These were all computed at the coarse-grain level, which may lead to inaccurate calculations due to a lesser resolution of the model and the modification of the entropic term in the forcefield parameter. However, analysis in atomistic detail has never been previously shown to converge to a reliable value due to variability in long-range electrostatic interactions (König *et al.*, 2018). Given the lack of experimental data and the impossibility of performing atomistic FEP calculations, I am left with CG methodologies. My work complements earlier free energy calculations within the field of CG methodology, which showed that the CG forcefield could distinguish between PIP_2 and PIP_3 (Naughton *et al.*, 2018). Overall, this illustrates the robustness of the application and demonstrates its power as a relatively cheap and effective *in silico* approach for comparing lipid-binding free energies to a membrane protein of interest. Nevertheless, although there is good agreement between the single- and multi-step approaches, I cannot exclude the possibility that the individual particle contributions may be either over- or under-estimated due to the CG approach.

Calculation of binding free energy on different Kir channels

To assess the binding affinity between PIP_2 and different Kir channels, I calculated the binding free energy between PC and Kir6.2 channel. My PMF calculation predicted that the binding free energy of PC to the Kir channel is ~0 kJ/mol. Thus, using CG-FEP calculation, a transformation from PIP_2 to PC can be approximately taken as the absolute binding free energy of the PIP_2 to the Kir channels. My simulation showed that PIP_2 affinity rank for Kir channels was hKir1.1 > hKir3.2 ~ hKir2.2 > hKir6.2. This ranking disagrees with the extent of channel activation by PIP_2 based on the experimental studies, which gave a rank order of Kir2 > Kir1 > Kir3 (Rohács *et al.*, 2003). The primary assumption in our model is that I assumed that all basic residues contributed to a PIP_2 binding site are all protonated with a +1 charge and the PIP_2 headgroup is fully deprotonated (-5 charge). By estimating the net charge on the surface, my calculations correlate well with the net charge of the site (total number of the basic residues subtracted by the total number of acidic residues comprising the PIP_2 binding site) (8, 4 and 2 for hKir1.1, hKir2.2/hKir3.2 and hKir6.2, respectively). Thus, the results from my simulation could simply allow me to hypothesise that the binding of PIP_2 to the Kir channel is mainly driven by electrostatic interactions, under the assumption that all basic residues are protonated. The protonation states of the basic residues and the PIP_2 headgroup will need to be further investigated either using constant pH simulation, propKa or atomistic free-energy perturbation to calculate the energetic contribution of a proton in the binding site.

The disagreements between my simulations and the previous experimental studies may be due to the following reasons. First, my CG representation may not be detailed enough to capture the subtle differences between the PIP_2 binding pockets of different Kir channels. Second, the energy required for the channel to transit from closed to open states may vary between different Kir channels. Thus, the correlation between PIP_2 binding and channel activation will be affected by this energy difference. Last, because of its hydrophobicity, the actual concentration of PIP_2 that was seen by the channel in the electrophysiological experiments is unknown.

Clinical mutations and PIP_2

I demonstrated that amino acid mutations could also be investigated using CG-FEP, which allowed us to probe the effect of a given amino acid's substitution on lipid binding. My method extends the traditional atomistic FEP mutation approach which has been demonstrated to be much more difficult to converge (Gapsys *et al.*, 2015). My CG-calculations are more appropriate for charge calculations which are very difficult to capture in atomistic. However, my method would fail to calculate free energy differences where the side chain perturbation is more subtle, such as from leucine to isoleucine. My work showed that the naturally occurring NDM mutation (E179K/A) increases the affinity of hKir6.2 for PIP_2. This would be expected to enhance KATP channel activity and reduce its inhibition by ATP, thus leading to less insulin secretion when blood glucose rises. This mechanism is novel as most NDM mutations have previously been found to alter ATP inhibition by impairing ATP binding or

increasing the open state probability(Pipatpolkai *et al.*, 2020*b*). On the other hand, a mutation that disrupts PIP_2 binding, such as K67N, is associated with CHI. Thus, this identifies a potential new mechanism of action for a CHI mutation.

Conclusion

The application of CG-FEP is potentially a valuable tool for the relatively high-throughput analysis of multiple disease-causing mutations related to lipid binding to ion channels. I have validated this method using PIP_2-Kir channel interactions. However, numerous ion channels are regulated by PIP_2. Recent advances in cryo-EM have provided us with a multitude of ion channel structures with PIP_2 bound, such as KCNQ1 and KCNQ4. Providing the site of PIP_2 binding is known, I can calculate its relative binding affinity and therefore identify how natural mutations can cause human disease. For example, mutations at the PIP_2 binding site of the KCNQ1 channel are associated with long QT syndrome (Sun & MacKinnon, 2020). The next step in this study is to understand the dynamic interplay of PIP_2 with other ligands (such as ATP) and how defects in this process may be associated with disease.

Chapter 4

Competition between PIP_2 and ATP on the Kir6.2 channel gating

"My advice to everyone is to find a friend. It makes science much more fun and you can do so much more."

Frances Mary Ashcroft (2018)

The work in this chapter is published in:

Pipatpolkai, T., Usher S*., Vedovato N*. Corey R.A., Ashcroft F.M., and Stansfeld P.J., **The dynamic interplay of PIP_2 and ATP in the regulation of the K_{ATP} channel.** bioRxiv.(2021) https://doi.org/10.1101/2021.05.06.442933

4.1 Introduction

Introduction

In chapter 3, I used CG-MD to predict the PIP_2 binding site and its affinity for different phosphoinositides. I also examined the effects of mutations in the PIP_2 binding site on the affinity of PIP_2 binding. In this chapter, I am interested in the dynamic interplay between PIP_2 and ATP when both are present in their respective binding sites. In this chapter, I have used atomistic simulations to observe conformational transitions when both PIP_2 and ATP are bound to the K_{ATP} channel.

Previous literatures on ATP and PIP_2 competition

The mechanism by which PIP_2 reduces ATP inhibition of the channel has long been debated. Previous studies have shown that the application of PIP_2 markedly increases the channel P_{open} and reduces the channel sensitivity to ATP (Fan & Makielski, 1999). Because an increase in channel open probability is associated with a reduction in ATP inhibition (Trapp *et al.*, 1998; Enkvetchakul *et al.*, 2000), it is possible that at least part of the effect of PIP_2 is mediated via changes in P_{open}. However, it has also been argued that PIP_2 may have an additional effect on the channel ATP sensitivity independent of P_{open} (Enkvetchakul *et al.*, 2000). Both molecules carry similar negatively charged phosphate groups, and previous studies have proposed that PIP_2 competes with ATP for the same binding site on the C-terminus of the protein. (MacGregor

et al., 2002). However, a comparison of recent structural studies (with bound ATP) and docking and MD simulations (with PIP_2) suggested that ATP and PIP_2 have separate binding pockets (Haider *et al.*, 2005, 2007). Nevertheless, they share a single common residue, K39, which lies in the N-terminus of Kir6.2.

PIP_2 binding mutations

Mutations at the PIP_2 binding site influence both K_{ATP} channel activation by PIP_2 and inhibition by ATP (Shyng *et al.*, 2000; Haider *et al.*, 2007). The previous chapter focused on two neonatal diabetes mutations (E179K/A) located near the putative PIP_2 binding site. The work from the previous chapter also predicted that the E179K mutation increases PIP_2 affinity. This chapter addresses the mechanisms by which neonatal diabetes mutations near the PIP_2 binding site (such as E179K/A and K39R) influence channel inhibition by ATP. I collaborated with Dr Natascia Vedovato and Samuel Usher on electrophysiological and patch-clamp fluorometry experiments to determine the validity the predictions described by the MD simulation data.

Experimental approaches

My collaborator used two approaches to understand the effect of the mutations on the K_{ATP} channel inhibition, patch-clamp recordings of K_{ATP} currents and Förster resonance energy transfer FRET to assess ATP binding. Together, we investigated the effect of mutations at two residues near the PIP_2 binding site (E179A/K, K39R/E/A) on K_{ATP} channel current inhibition by ATP.

We also investigated whether the observed change in K_{ATP} current inhibition by ATP caused by these mutations was due to the enhanced activation by PIP_2. And if so, is this caused by changes in PIP_2 binding or the way in which binding is transduced into changes in P_{open}. We first measured ATP inhibition of K_{ATP} currents using standard patch-clamp methods. We then measured ATP binding using a novel FRET assay (Puljung *et al.*, 2019; Usher *et al.*, 2020). This involved insertion of a fluorescent amino acid 6-propionyl-2-(N,N-dimethyl)aminonaphthalene (ANAP) into Kir6.2 (Kir6.2*), at a position close to the ATP-binding site (W311). We then used a fluorescent ATP analogue TNP-ATP whose absorption spectrum overlaps with the emission spectrum of ANAP. This means that when TNP-ATP binds it is close enough to ANAP to permit FRET.

FRET occurs when two fluorophores are in proximity (~20-30 Å) and the emission spectra of one fluorophore, in this case ANAP, overlaps with the absorption spectra of the other (TNP-ATP). This results in a reduction in the fluorescence emission (quenching) of ANAP and an increase in fluorescence of TNP-ATP. By simultaneously recording K_{ATP} currents and FRET, we can calculate the dissociation constant for ATP binding (K_d^{app} or EC_{50}) and the half inhibitory concentration (IC_{50}) for ATP. One caveat with this method is that TNP-ATP is not identical to ATP. The sugar group is bulkier and the molecule has a greater net charge (of -6 compared with -4 for ATP). The previous study have shown that ANAP substitution at W311 reduces ATP binding to the K_{ATP} channel (Usher *et al.*, 2020). We assume that TNP-ATP binding to the W311

mutant channel will show similar changes when residues K39 and E179 are mutated as the ATP response of the wild-type channel. I hypothesised that when the channel has a higher affinity to PIP_2, this will reduce channel inhibition and binding to TNP-ATP.

Computational approach

Computationally, I used atomistic MD simulations (AT-MD) to determine the dynamics of K39 when both ATP and PIP_2 occupy their respective binding sites. My simulations suggest that K39 coordinates with both ATP and PIP_2 but with a stronger preference for PIP_2 when both ligands are present. I also extended my work to include hydrogen bond analysis of K39R and wild-type channels.

4.2 Methods

4.2.1 Coarse-Grained system preparation

The CG simulation system containing human Kir6.2 (Kir6.2) model without SUR1 but with PIP_2 bound was generated as described in the last chapter (PDB entry: 6BAA) (Martin *et al.*, 2017*a*) All simulations were carried out using GROMACS-2019.4 (Abraham *et al.*, 2015).

4.2.2 Atomistic simulation set up

The coarse-grained simulation system (hKir6.2, lipids - PIP_2 and POPC, ions and water) was converted to atomistic using the CG2AT pipeline (Stansfeld & Sansom, 2011). The K39R mutant model was generated using PyMOL with the minimum initial stearic clashes (Schrodinger LLC, 2015). The exact positions of the ATP molecules were taken from the cryo-EM structure (PDB entry: 6BAA) and placed in the Kir6.2 ATP binding sites. The initial position of the hydrogen atoms was added using PyMOL. All simulations were carried out using CHARMM36 biomolecular forcefield with the virtual sites on the proteins and lipids CH_3 and NH_3^+ groups (Huang & Mackerell, 2013; Olesen *et al.*, 2018), allowing the integration timestep of 4 fs in the production run. The force field for the ATP molecule was derived from the CHARMM-GUI (Jo *et al.*, 2008; Kim *et al.*, 2017). In this study, four different conditions were set (Apo, ATP-bound, PIP_2-bound, both ATP and PIP_2 bound), and simulations were carried out for 3 repeats.

The systems were energy minimised using the steepest descents algorithm with non-hydrogen atoms restrained at 1000 kJ/mol/nm^2. This was followed by a 5 ns equilibration for the system where the C_α backbone on the protein and the atoms on the ATP molecules were restrained with 1000 kJ/mol/nm^2 with 4 fs timesteps. A temperature of 310 K was maintained with V-rescale temperature coupling, while 1 atm pressure was controlled using semi-isotropic Parrinello-Rahman pressure coupling (Parrinello & Rahman, 1981). The simulation was then further equilibrated with only C_α restraint on the protein for another 15 ns in similar conditions. Then a 380 ns production run was implemented, in which the first 80 ns of the simulations were discarded as equilibration. All simulations, H-bond calculations and distance calculations were calculated using GROMACS-2019.4 (Abraham *et al.*, 2015). I used MDAnalysis packages to calculate pairwise RMSD to define changes in protein dynamics across the trajectory (Theobald, 2005; Beckstein *et al.*, 2009; Michaud-Agrawal *et al.*, 2011; Gowers *et al.*, 2016). Kir6.2 is a symmetrical homo-tetramer; I assumed that all sites behave identically during the short simulation timescale (300 ns).

4.3 Results

4.3.1 PIP_2 and ATP binding site on Kir6.2

In the previous chapter, I used coarse-grained simulations to predict the PIP_2 binding site. In this chapter, I refined my predicted PIP_2 binding site on Kir6.2 in the absence of SUR1 using atomistic molecular dynamic simulations. I also aim to explore the relationship between the ATP and PIP_2 binding sites of Kir6.2 in the absence of SUR1. There is no structure of Kir6.2, in either a PIP_2 bound or an ATP-bound conformation, in which SUR1 is not present. Thus, I built two atomistic simulation systems, Kir6.2 with ATP and Kir6.2 with PIP_2 and simulated these independently for 380 ns. To ensure that the initial protein structure was stable after converting to an atomistic system, I used pairwise root mean square deviation (RMSD) analysis over the last 300 ns of the simulation on the C_α atoms of Kir6.2 as a measure of the stability of protein tertiary structure. I observed that the C_α RMSD never deviated more than 4 Å across the trajectory in all simulation set-ups. Interestingly, the simulations with either PIP_2 and/or ATP bound to the channel showed slightly less C_α rearrangement than the apo state. This suggested that the 3D structure of the protein is highly stable throughout my simulation and further stabilised by the ligand (Figure 4.1).

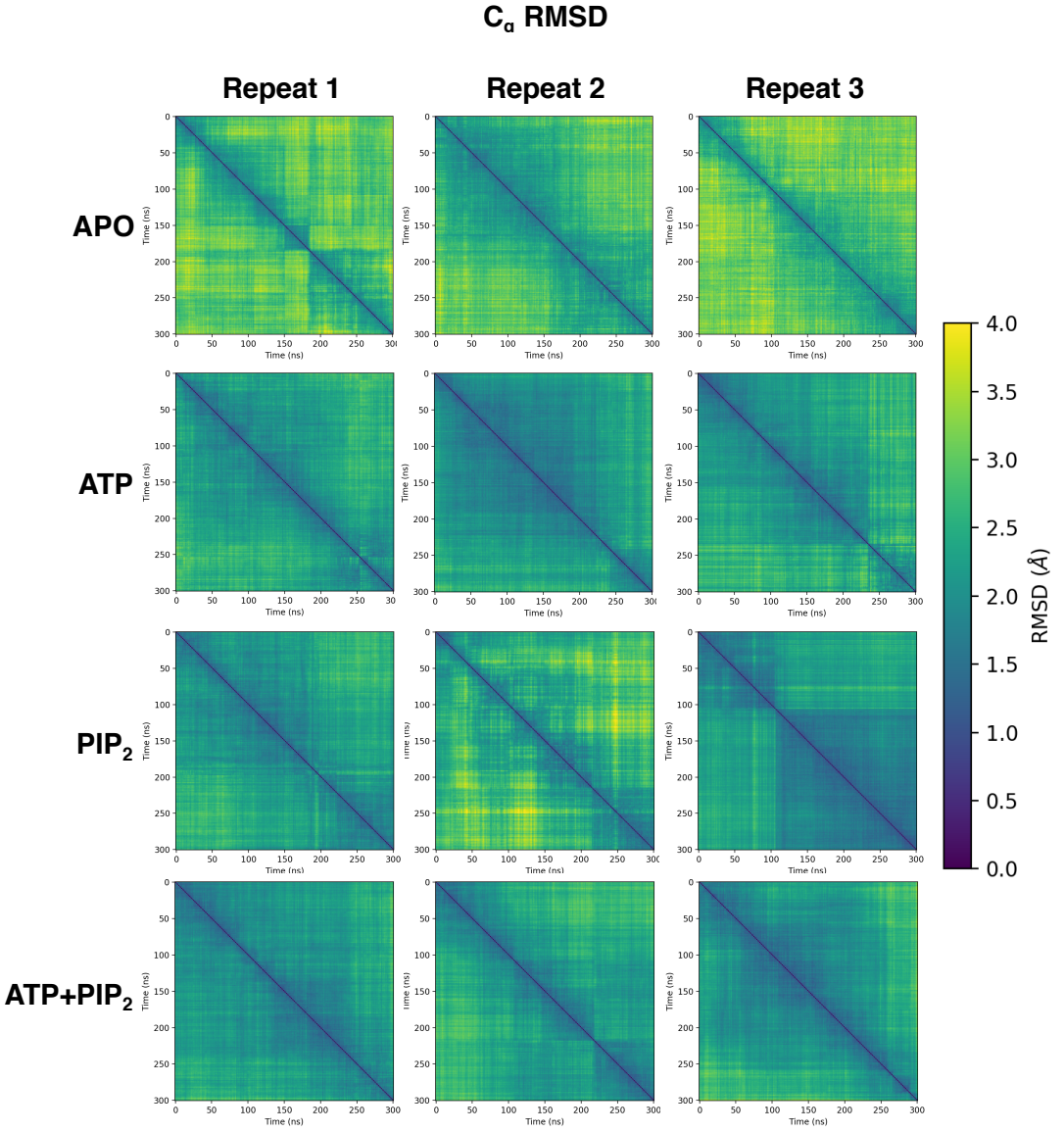


Figure 4.1 Overall protein dynamics in different states of the simulation.

Pairwise Root Mean Square Deviation (RMSD) analysis of the C_α atoms on Kir6.2 in the different ligand-bound states, measured over the final 300 ns simulation trajectory (n=3). The colour bar indicates the RMSD value ranging from 0 to 4 Å.

To investigate the effect of PIP_2 and ATP binding on the local geometry of their binding sites, I selected amino acid residues within 4 Å of the ligand for > 40% of the time and defined them as contacting residues. Between the different Kir6.2 subunits, I found no significant difference in the contact that these residues made to either ATP or PIP_2 in the final 300 ns of the simulation (Figure 4.2). Thus, in all subsequent analyses, I defined each Kir6.2 subunit as a separate calculation to increase the sampling of an ATP molecule in the binding pocket. As there are 4 subunits and three separate simulations were run, this yielded a total of 12 data points per contacting residue.

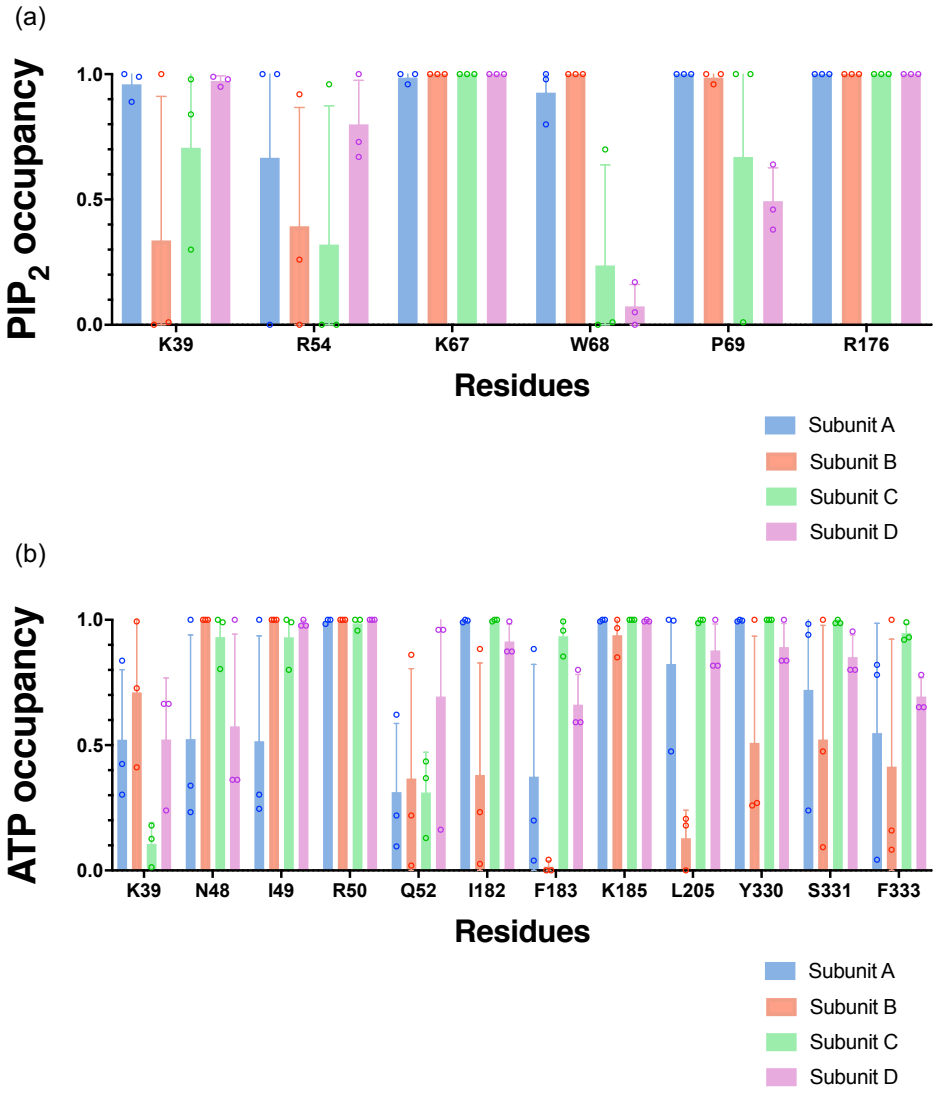


Figure 4.2 PIP_2 and ATP binding sites in four Kir6.2 subunits

The fraction of time that the indicated residues are in < 4 Å proximity to (a) PIP_2 or (b) ATP during the final 300 ns of the simulations (defined as "PIP_2 or ATP occupancy"). Only residues where the mean occupancy is greater than 40% are plotted. Three simulations were run, and the data collected from each subunit of the tetramer are shown in different colours. The error bar indicates the standard deviation.

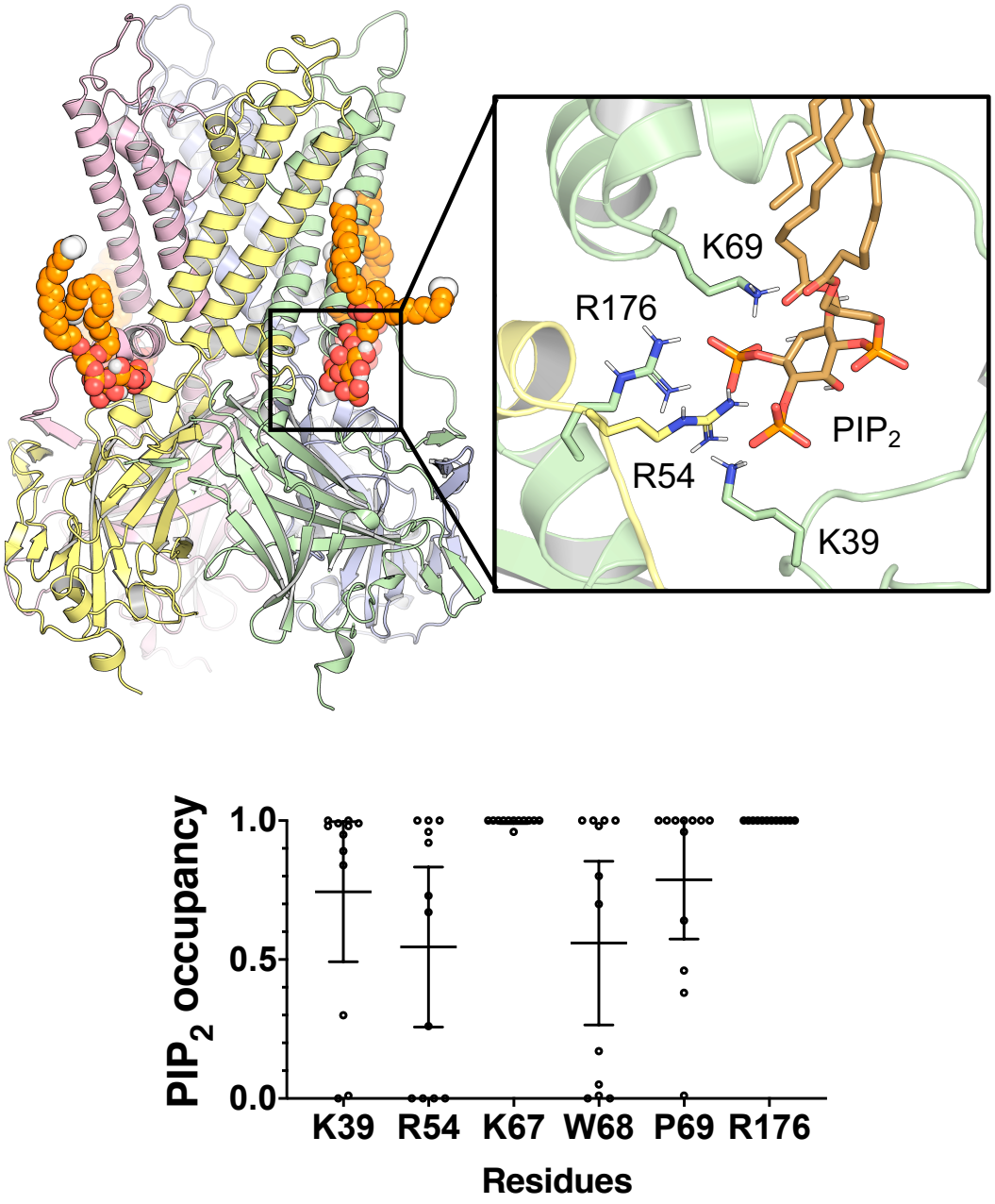


Figure 4.3 PIP_2 binding sites

Above. The structure of the Kir6.2 tetramer shown with four PIP_2 molecules in their binding sites. Inset: Interactions between PIP_2 (bronze) and basic residues in two chains of Kir6.2 (yellow and green).
Below. The fraction of time that the indicated residues are in < 4Å proximity to PIP_2 during the final 300 ns of the simulations (defined as "PIP_2 occupancy"). Only residues where the mean occupancy is greater than 40% are plotted. Three simulations were run, and each subunit of the tetramer is treated as an individual data point (which makes a total of 12 data points). The error bar indicates the 95% confidence interval around the mean.

In the PIP_2 binding site, I found that R54 and K67 from one chain and R176 from an adjacent chain coordinated with the 4' phosphate of PIP_2 and that K39 coordinates the 5' phosphate (Figure 4.3). Other uncharged residues (W68 and P69) that lie at the membrane-water interface also make > 40% contact with PIP_2, but for less time than K67 and R176. It is interesting to note that the basic residues, K67 and R176 are making stronger contact than K39 and R54. Overall, this suggested than an electrostatic interaction is crucial for the PIP_2 binding to Kir6.2. This agrees with my CG model from the previous chapter which suggested that Kir6.2 binds to the 1' and 4' phosphates are more tightly than the 5' phosphate.

By evaluating the root mean square fluctuation (RMSF) of all contact residues, I showed that the binding of PIP_2 mildly reduced the dynamics of all side chains which make contact to PIP_2. The reduction is only statistically significant at residues K67 and R176 (Figure 4.4a). Interestingly, only the difference at K67 showed a decrease of ~1 Å or more. Mutations at all PIP_2-contacting residues have been previously shown to interfere with channel gating and make the channel insensitive to ATP inhibition (Männikkö *et al.*, 2011).

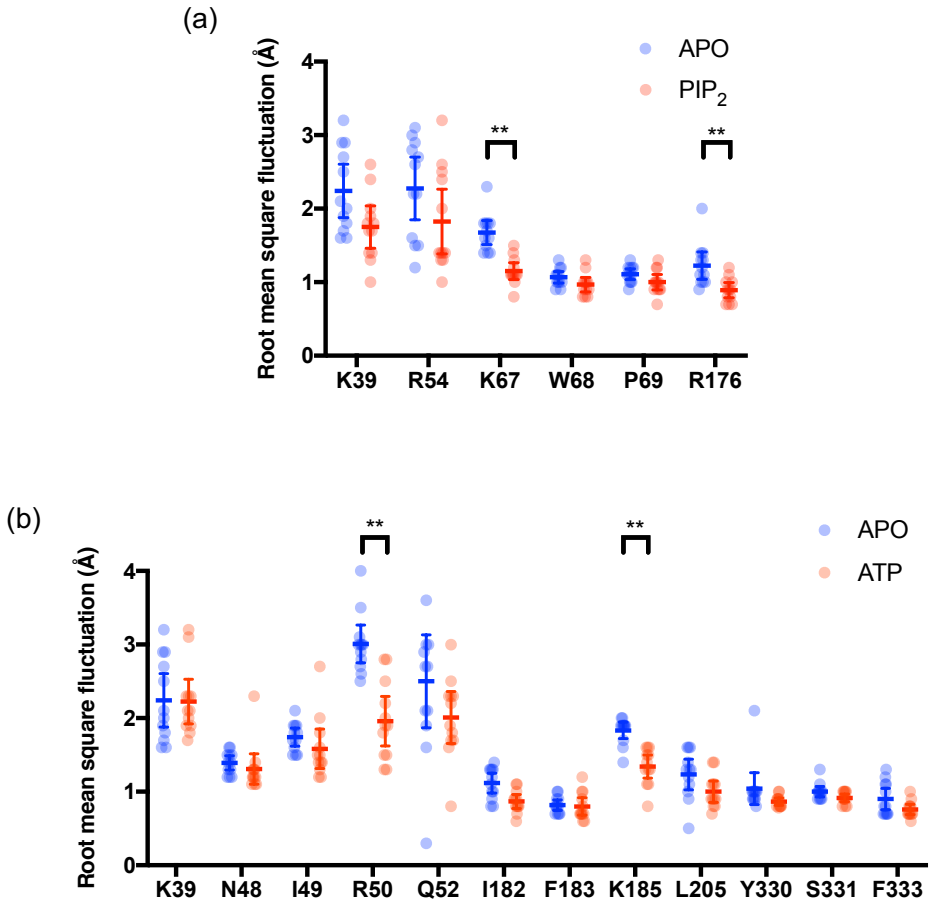


Figure 4.4 The dynamics of individual side chains in the simulation

(a) Root mean square fluctuation (RMSF) analysis of the residues on the Kir6.2 tetramer, which contact the PIP_2 molecule in the absence (blue) and the presence (red) of PIP_2. Three simulations were run, and each subunit of the tetramer is treated as an individual data point (which makes a total of 12 data points). The error bar indicates the 95% confidence interval around the mean. ** $P < 0.01$ (Student's t-test).

(b) Root mean square fluctuation (RMSF) analysis of the residues on the Kir6.2 tetramer, which contact the ATP molecule in the absence (blue) and the presence (red) of ATP. Three simulations were run, and each subunit of the tetramer is treated as an individual data point (which makes a total of 12 data points). The error bar indicates the 95% confidence interval around the mean. ** $P < 0.01$ (Student's t-test).

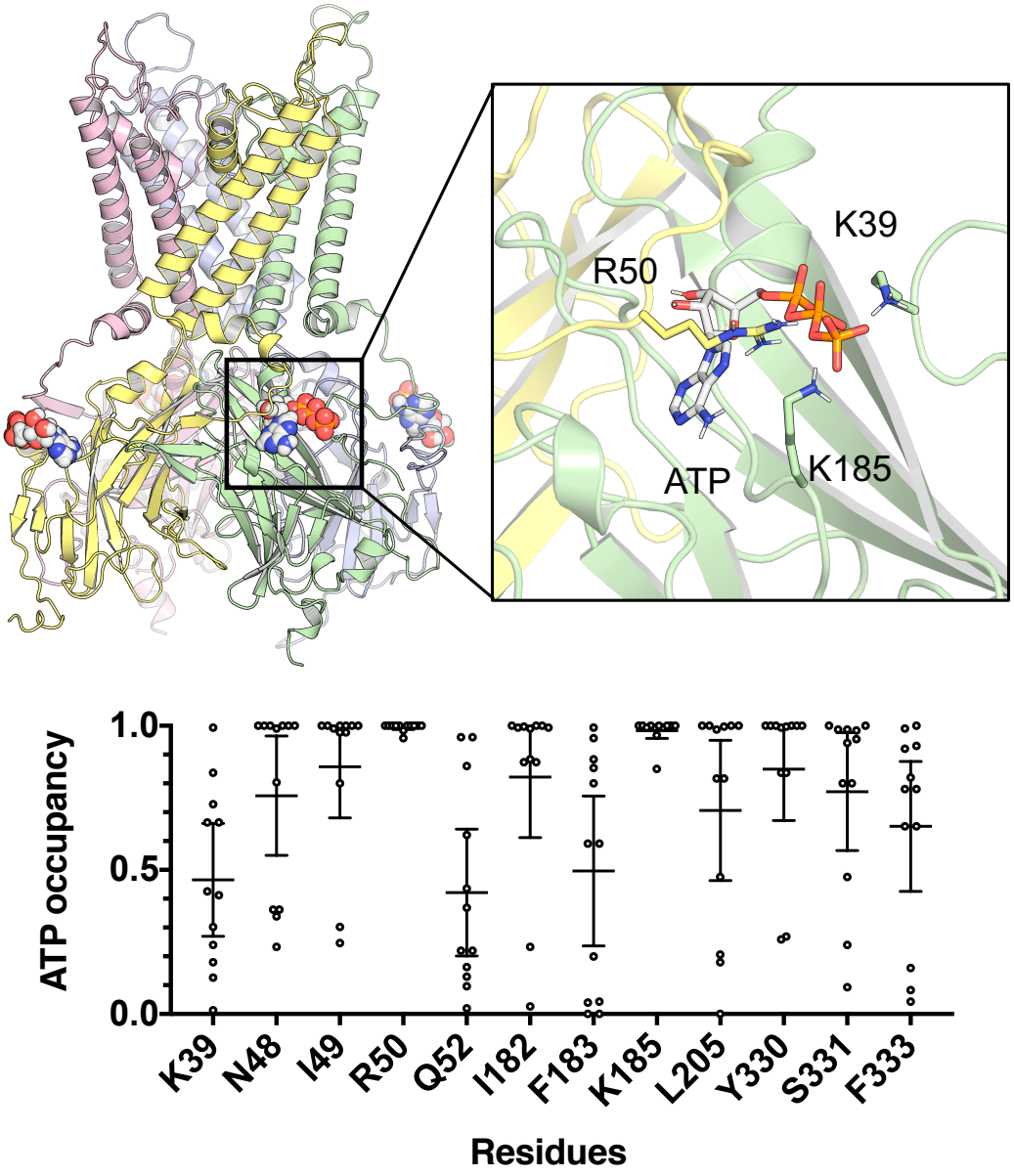


Figure 4.5 ATP binding sites

Above, The structure of the Kir6.2 tetramer shown with four ATP molecules in their binding site. Inset: Interactions between ATP (CPK) and basic residues in two chains of Kir6.2 (yellow and green) after a 380 ns simulation.
Below. The fraction of time that the indicated residues are in < 4 Å proximity to ATP during the final 300 ns of the simulations (defined as “ATP occupancy”). Only residues where the mean occupancy is greater than 40% are plotted. Three simulations were run, and each subunit of the tetramer is treated as an individual data point (which makes a total of 12 data points). The error bar indicates the 95% confidence interval around the mean.

In the ATP binding site, I found that R50 coordinates both the β and the γ phosphate, K39 coordinates the γ phosphate, and K185 coordinates the α and the β phosphate (Figure 4.5). Both R50 and K185 dynamics are stabilised when ATP binds to the channel (Figure 4.4b). These findings agree with the previous studies in which the ATP-binding residues in Kir6.2 were mapped using site-directed mutagenesis (Tucker *et al.*, 1998; Dabrowski *et al.*, 2004). Interestingly, the sidechain of K39, which is *ca.* 7 Å from the ATP molecule in the cryo-EM structure, moved towards ATP and made contact in some simulations. Interestingly, the K39 on every subunits of Kir6.2 apart from the subunit C makes strong contact with ATP. All residues that contact ATP in the cryo-EM structure (N48, I49, Q52, I182, F183, L205, Y330, S331, F333) also make contact in my simulations. These residues are crucial for ATP inhibition of K_{ATP} currents as shown by electrophysiological studies. (Tucker *et al.*, 1998). Mutations of residues that make contact with ATP are also associated with neonatal diabetes (Pipatpolkai *et al.*, 2020*b*).

4.3.2 The competition between PIP_2 and ATP for K39 coordination

Earlier studies have shown that PIP_2 reduces K_{ATP} channel inhibition by ATP (Fan & Makielski, 1999; Shyng *et al.*, 2000). However, it was unclear if PIP_2 competes directly for the ATP binding site or if the transduction caused by PIP_2 interferes with ATP-dependent gating. Comparison of the cryo-EM structure of the ATP binding site and the predicted PIP_2 binding site suggest they lie *ca.* 25 Å from one another (Stansfeld *et al.*, 2009; Martin *et al.*, 2017*a*). In my studies, K39 contacts with both ATP and PIP_2 in independent simulations, but the position of the amine head group is different (Figures 4.3, 4.5).

I next explored the dynamics of K39 when both ligands occupied their respective binding sites (Figure 4.6a). I calculated the distance between ATP or PIP_2 and the amine group of K39 for 12 independent simulations (4 subunits x 3 repeats). I observed that the position of the amine head group of K39 oscillated between ATP and PIP_2 during the early part of the simulation in 7 repeats of the simulations, in all of which the headgroup ended up favouring PIP_2 at the end of the simulation (Figures 4.6b, 4.7). In the 5 other repeats, the K39 side chain was stably bound to the PIP_2 during the equilibration process and never sampled any interactions with ATP in the last 300 ns of the simulation (Figure 4.7). In the last 300 ns of the simulation, I found that a statistically significant decrease ($P < 0.01$) in ATP contacts only occurred at residue K39 when PIP_2 was present (Figure 4.6c). Thus, I hypothesised that K39 might change its coordination from ATP to PIP_2 when both molecules are bound.

I calculated the change in ATP occupancy of other residues that also contact ATP, in both absence and presence of PIP_2. However, no other residues, apart from K39, showed significant differences in their contact probability between ATP bound and ATP+PIP_2 bound state (Figure 4.6c). In fact, I observed a non-significant more stable interaction between ATP and M48, I49, Q52, I182, F183, L205, Y330 and F333. These amino acid side chains do not carry a positive charge. This may suggest that when PIP_2 is present, aromatic, and polar amino acid sidechains undergo a small-scale rearrangement to compensate for the loss of K39 from the ATP binding site.

In the final 300 ns of the simulation, I showed that K39 co-ordinates with PIP_2 for 75% of the simulation time in both the presence and the absence of ATP (Figure 4.6d). The presence of ATP also mildly, but not significantly reduced the contact that PIP_2 made with Kir6.2 at residues K67, W68 and R176. This may suggest that the presence of ATP disrupts an electrostatic interaction between PIP_2 and Kir6.2.

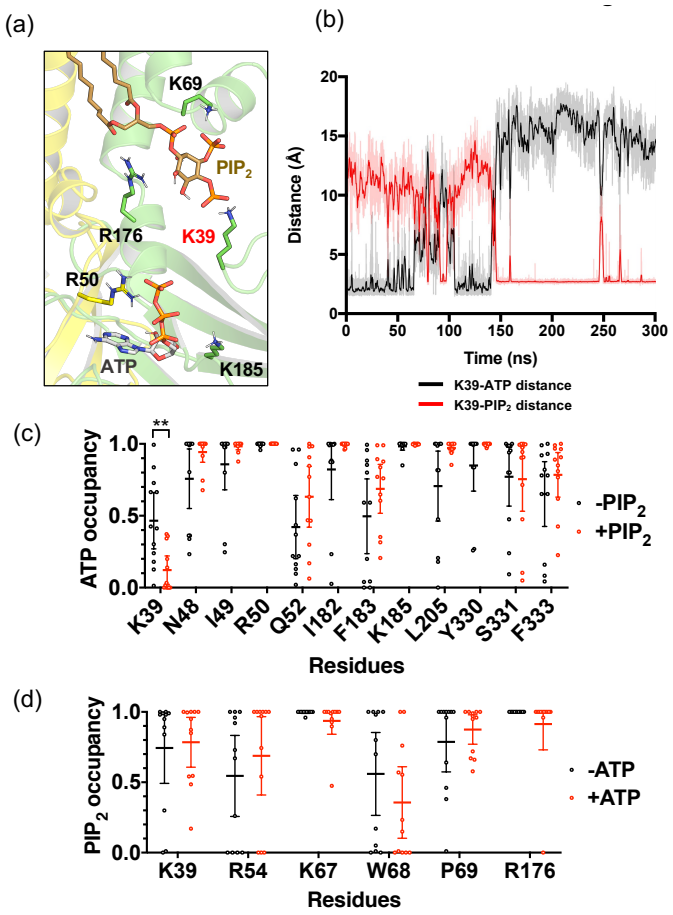


Figure 4.6 Changes in the ATP and PIP_2 binding sites when both ligands are present

(a) Interactions between Kir6.2 (two chains are indicated in yellow and green), ATP (CMYK) and the PIP_2 (bronze) headgroup. Only the basic residues of the protein are shown.
(b) A representative trace showing the distance calculation between K39 and ATP (black) and K39 and PIP_2 (red) across a single 300 ns trajectory. The darker lines show running averages of every 1 ns of the simulation.
(c) The fraction of time that the residues are in < 4 Å proximity to ATP during the final 300 ns of the simulations (defined as "ATP occupancy") in the absence (black) and presence (red) of PIP_2. Only residues where the mean occupancy is greater than 40% are plotted. Three simulations were run, and each subunit of the tetramer is treated as an individual data point (which makes a total of 12 data points). The error bar indicates the 95% confidence interval around the mean. The asterisk indicates the significant difference at $P < 0.01$ using the Student's t-test.
(d) The fraction of time that the residues are in < 4Å proximity to PIP_2 during the final 300 ns of the simulations (defined as "PIP_2 occupancy") in the absence (black) and presence (red) of ATP. Only residues where the mean occupancy is greater than 0.4 are plotted. Three simulations were run, and each subunit of the tetramer is treated as an individual data point (which makes a total of 12 data points). The error bar indicates the 95% confidence interval around the mean.

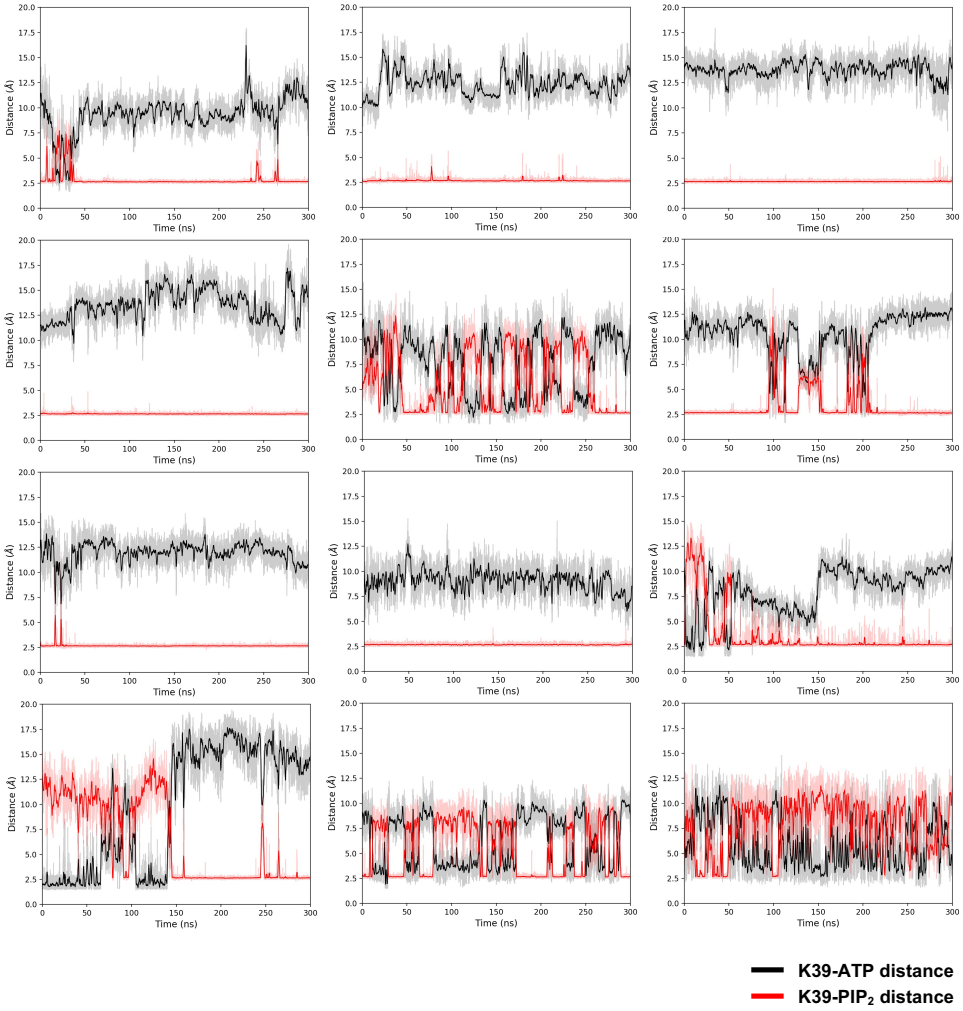


Figure 4.7 The distance between K39, ATP and PIP_2

Distance calculation of the ATP (black) and PIP_2 (red) and the K39 when both ATP and PIP_2 are present, measured over the final 300 ns simulation trajectory. The darker lines show running averages of every 1 ns of the simulation. Four vertical panels show behaviour across four subunits and three horizontal panel represents three independent simulation repeats.

Together, I propose that the salt bridges between K39 and ATP are broken in the presence of the PIP_2, causing the amine group on K39 to swing towards the PIP_2 headgroup. This agrees with previous functional studies in which a different mutation at the same residue (K39A) decreases both channel inhibition by ATP (Tucker *et al.*, 1998), and also reduces channel activation by PIP_2 (Cukras *et al.*, 2002).

4.3.3 Neonatal diabetes mutation (K39R)

A mutation at K39 (K39R) is associated with transient neonatal diabetes (Zhang *et al.*, 2015). The K39R mutation does not alter the charge (as both lysine and arginine are positively charged), but arginine has a guanidium group instead of an amine. I simulated Kir6.2 containing the K39R substitution and compared the contacts between the arginine guanidium group and PIP_2 or ATP. In simulations with ATP alone, residue 39 in the mutant channel spent more time contacting ATP than the wild-type channel (Figure 4.8). This suggested that K39R mutation may favour ATP binding when PIP_2 is not present in its binding site.

In simulations with PIP_2 alone, residue 39 in both the wild-type and the mutant channel spent almost all its time coordinating with PIP_2. When both ATP and PIP_2 were present, I found that the contact probability of both residue 39 and K39R with the PIP_2 headgroup was not significantly different in either the absence or the presence of ATP (Figures 4.8a, 4.8b) From this, one might conclude that K39R mutation does not affect the channel preference for PIP_2 in either the presence or the absence of ATP. As K39 lies at the interface between Kir6.2 and SUR1, the K39R mutation might stabilise an interaction with SUR1 which may lead to an activatory effect of the channel. The SUR1 residue which is in closest proximity to K39 in Kir6.2, is F198, which lies *ca.* 12 Å away – but may of course move closer in a different state.

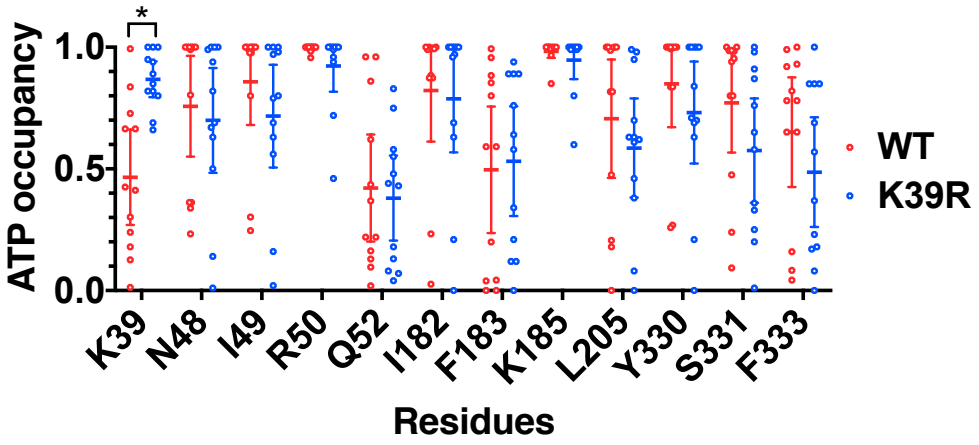


Figure 4.8 Changes in the ATP occupancy caused by the K39R mutation.

The fraction of time that the residues are in < 4Å proximity to ATP during the final 300 ns of the simulations (defined as "ATP occupancy") in the WT (red) or the K39R mutant channel (blue). Only the residues at which the occupancy are greater than 40% are plotted. There are three simulations, and each subunit of a tetramer is being treated as an individual data point (a total of n = 12). The error bar indicates the 95% confidence interval around the mean. The asterisk indicates the significant difference at $P < 0.01$ using the Student's t-test.

I hypothesise that the K39R mutation may enhance the binding strength of PIP_2 and K39. This may be expected to activate the channel at a lower PIP_2 concentration. By increasing the open probability, PIP_2 will reduce the ability if ATP to inhibit the channel. To quantify the strength of the interaction between PIP_2 and K39 or K39R, I carried out a hydrogen bond (H-bond) analysis to determine the number of H-bonds formed between the PIP_2 headgroup and the side chain of residue 39. I assumed that all H-bonds has the same binding free energy regardless of their distance and their partial charge. Note that arginine can form up to 5 H-bonds spread over 3 amine groups, whereas lysine can only form a maximum of 3 bonds from a single amine. I observed that the guanidium group on the arginine forms two hydrogen bonds with the PIP_2 headgroup, whereas the lysine amine group forms only a single hydrogen bond. In both cases, the hydrogen bonds are formed with the 5' phosphate on the PIP_2 inositol headgroup both in the presence and the absence of the ATP (Figures 4.6c, 4.6d). Thus, the K39R mutation enhances the strength of the interaction of residue 39 with PIP_2. I postulate this leads to reduced channel inhibition by ATP and impairs insulin secretion leading to neonatal diabetes.

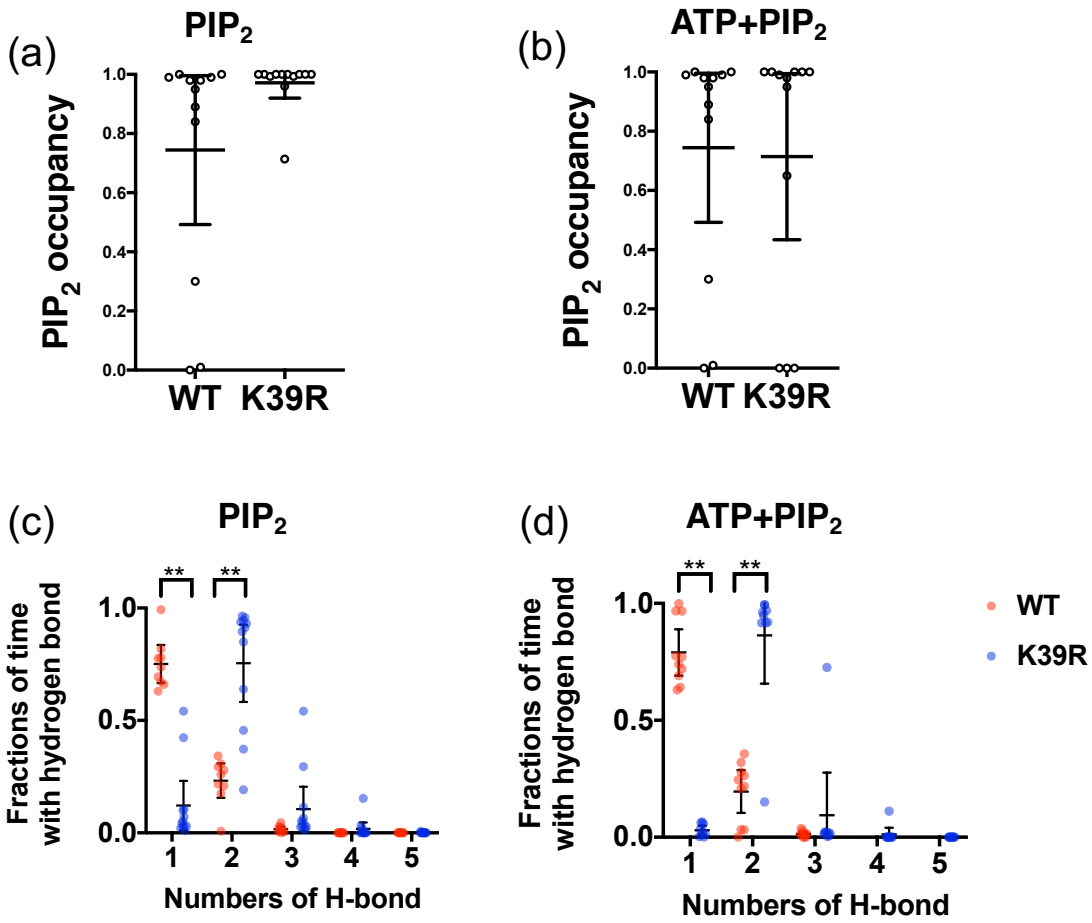


Figure 4.9 The K39R mutation changes the PIP_2 binding configuration

(a,b) The fraction of time that the K39 (or K39R) are in < 4 Å proximity to PIP_2 during the final 300 ns of the simulations (defined as "PIP_2 occupancy") in the absence (a) and presence (b) of ATP. There are three simulations, and each subunit of a tetramer is being treated as an individual data point (a total of n = 12). The error bar indicates the 95% confidence interval around the mean.

(c,d) Hydrogen bond analysis showing the fraction of time when K39 (red) or K39R (blue) forms a different number of H-bonds with the PIP_2 headgroup during the final 300 ns of the simulations in the absence (c) and presence (d) of ATP. There are three simulations, and each subunit of a tetramer is being treated as an individual data point (a total of n = 12). The error bar indicates the 95% confidence interval around the mean. ** P < 0.01 (Student's t-test).

4.3.4 Supporting experimental data

I collaborated with Dr Natascia Vedovato and Samuel Usher on experimental studies designed to test the predictions of my MD simulations. The data, figures and figure legends in this section are all provided by my collaborators.

We first investigated the effect of K39 mutations (K39R/E/A) on inhibition of the K_{ATP} current by ATP on HEK293T cells in patch clamp studies using giant inside-out membrane patches (Dr Natascia Vedovato). We aimed to test two hypotheses as follows:

a) My computational study predicted that the K39R mutant would have a higher affinity to PIP_2. We assumed that PIP_2 is present in an excised patch (and is only slowly lost following patch excision). Therefore, when a solution of ATP is applied to the inside of the patch, this represents the situation in which both ATP and PIP_2 are bound to the channel and is similar to my simulation. The simulations predict the K39R mutation would have a stronger interaction with PIP_2 than ATP, and thus that the channel would be inhibited less by ATP.

b) We further hypothesised that both the K39A and K39E mutations reduce both the ATP and PIP_2 affinity of the channel. This is simply because the positive charge on K39, which is crucial for ATP and PIP_2 co-ordination, is now a neutral (A) or a negative (E) charge. Thus, as both ATP and PIP_2 bind less to the channel, we expected a reduction in K_{ATP} channel inhibition by ATP.

First, we used a Kir6.2 construct in which GFP was fused to the C-terminus of the protein. This allows Kir6.2 to traffic to the cell surface in the absence of SUR1. Previous studies have shown that the GFP-tag modification does not alter the channel inhibition by ATP (Schwappach *et al.*, 2000). We found that the K39R mutation significantly reduced the ATP sensitivity of the channel: the IC_{50} for the wild-type channel was 208±36 μM (n=3) compared with 384±19 μM (n=3) for K39R. Unfortunately, we could not conduct an experiment with the K39A or K39E mutants, because the macroscopic currents were too small to measure accurately, or the channels were not expressed.

To increase the amplitude of the macroscopic currents, we co-transfected Kir6.2 with SUR1 to generate octameric K_{ATP} channels, which increased the current amplitude approximately 10-fold. We showed that both K39E and K39R mutants significantly reduce the ATP sensitivity of the K_{ATP} channel (Figure 4.10). Thus, the results for the K39R mutation agree well with my computational predictions. However, the data for the K39A channel deviates from my hypothesis, as we found no significant reduction in channel inhibition by ATP. This contrasts with a previous experimental study (Tucker *et al.*, 1998) which displayed a mild reduction in inhibition by 0.1 and 1 mM ATP. I would have expected that substitution of a positively charged lysine (K) at the ATP binding site with an alanine (A) would reduce channel sensitivity to ATP significantly.

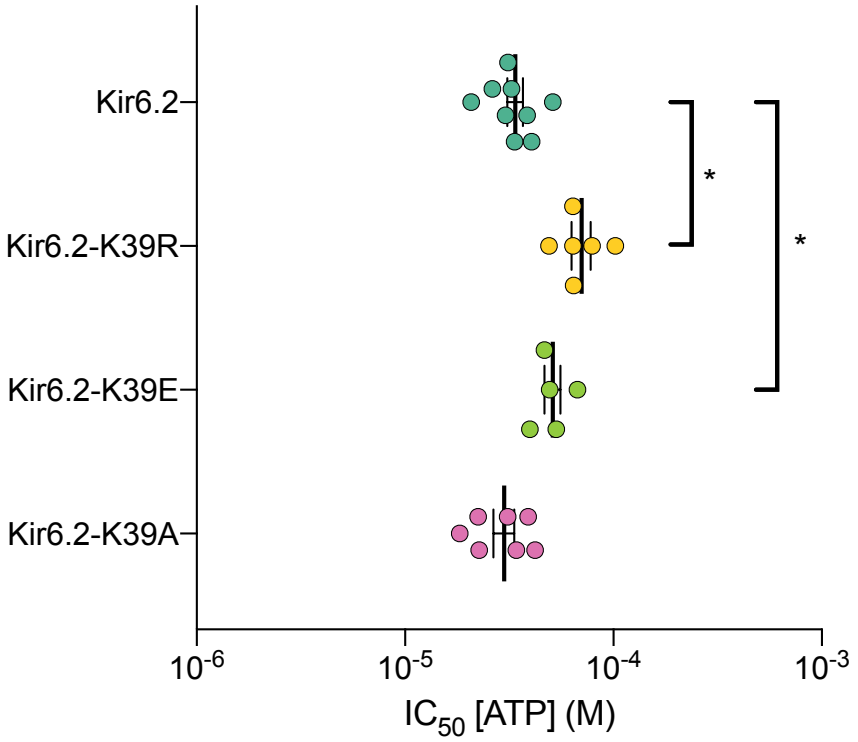


Figure 4.10 Effect of ATP on K_{ATP} current.

IC_{50}s from ATP current inhibition on wild-type and different K39 mutants derived from Hill fits to individual experiments are plotted as coloured data points. The means and 95% confidence intervals for each construct are overlaid as black error bars. Each construct is labelled with GFP at the C-terminus and co-expressed with SUR1.

I then proposed two further hypotheses to explain the effect of K39R mutation as follows:

a) the K39R mutation favours PIP_2 co-ordination. This reduces the binding of ATP to the K_{ATP} channel.

b) the K39R mutation stabilises the open state of the channel via an increase in PIP_2 affinity. This indirectly reduces ATP inhibition, as ATP stabilises the closed state of the channel.

We therefore attempted to measure the binding of ATP to the K_{ATP} channel, using the FRET assay described in the Introduction of this chapter. In this chapter, the ANAP-substituted Kir6.2 construct is denoted as Kir6.2*. We attempted to assess the macroscopic currents of the Kir6.2* channel without SUR1. However, the macroscopic currents were too small and thus all Kir6.2* constructs were co-transfected with SUR1. Thus, this set-up differs from the simulations which were not done in the presence of SUR1. A further difference is that the experiments used TNP-ATP instead of ATP (as TNP-ATP is fluorescent and can be used in FRET studies).

We compared the effect of the K39 mutations on Kir6.2* current inhibition by ATP. We found that the Kir6.2*-K39R currents displayed a reduction in ATP sensitivity (Figure 4.11a-c), whilst the K39A/E constructs had no significant difference in ATP sensitivity from wild-type (Figure 4.11d). This agrees well with the previous experiment carried out on the wild-type channel (without the W311 mutation).

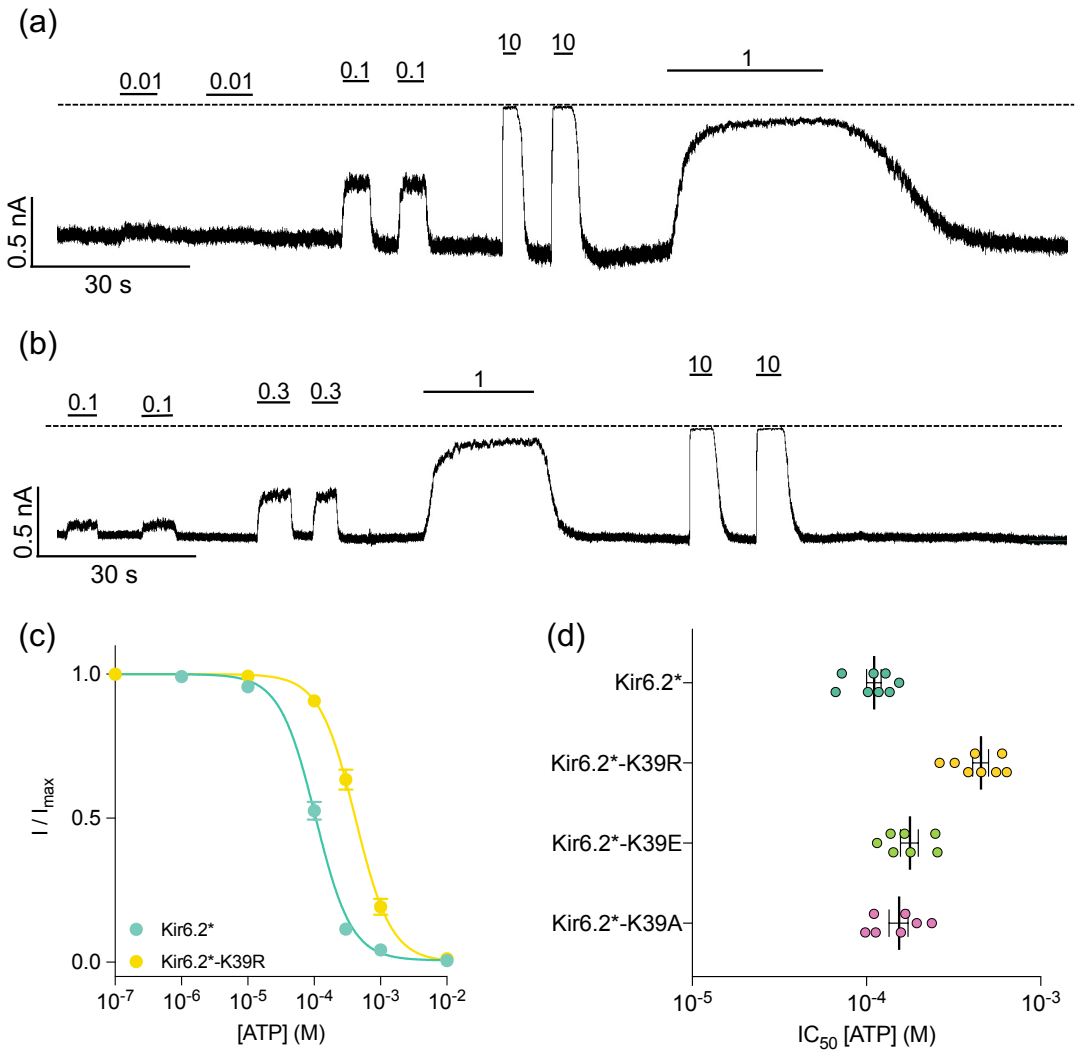


Figure 4.11 Effect of ATP on Kir6.2* and Kir6.2*-K39R current

(a,b) Representative traces for Kir6.2*(A), and Kir6.2*-K39R (b). Top solid bars mark the different ATP concentrations applied (in mM). The dashed line indicates the zero-current level.
(c) Concentration-response relationships for ATP inhibition of Kir6.2* and Kir6.2*-K39R currents (I/I_{max}), fitted to the Hill equation. IC_{50}: 105 µM for Kir6.2*, 417 µM for Kir6.2*-K39R, fitted to the Hill equation, h=1.6 and I_{max} fixed to 1.
(d) IC_{50}s from ATP current inhibition derived from Hill fits individual experiments are plotted as coloured data points. The means and 95% confidence intervals for each construct are overlaid as black error bars.

Figure 4.12a shows that TNP-ATP binds and inhibits the K_{ATP} channel with a lower IC_{50} than ATP, as previously reported (Usher *et al.*, 2020). When TNP-ATP are within 20-30 Å of the ANAP (i.e., TNP-ATP binds to the K_{ATP} channel), FRET between ANAP and TNP-ATP occurs. This binding quenches the fluorescence emitted by the ANAP at 470 nm (Figure 4.12b), which was used to measure binding. Combining data for TNP-ATP binding and current inhibition, we can calculate the binding constant (K_d^{app} or EC_{50}) and the half inhibitory concentration (IC_{50}) for TNP- ATP. (Figure 4.12e,f; EC_{50} = 94.4±14.0 μM; IC_{50} = 7.06±1.79 μM.; n=11)

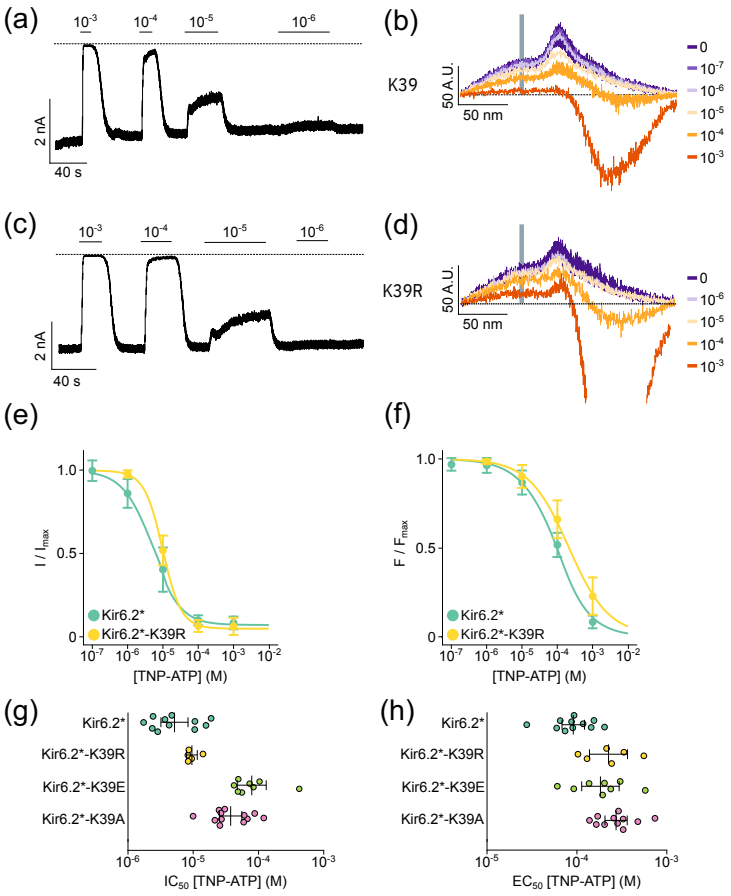


Figure 4.12 Effect of TNP-ATP on K_{ATP} current

(a,c) Representative current for Kir6.2* and (c) the K39R mutant channel. Top solid bars indicate application of TNP-ATP, with concentrations given in M. Dashed line, indicates the zero-current level.
(b,d) Fluorescence traces for (b) Kir6.2* and (d) K39R mutant channel. The grey bar indicates the fluorescence peak corresponding to ANAP. The colours indicate the application of TNP-ATP, with concentrations given in M. The dashed line, indicates the zero-fluorescence level with respect to the background.
(e,f) Concentration-response relationships for (e) TNP-ATP current inhibition (I/I_{max}) and (f) fluorescence quenching (F/F_{max}) of Kir6.2* (Kir6.2*, green) and Kir6.2*-K39R (Kir6.2*-K39R, yellow). Data points are the mean, and error bars are the 95% confidence intervals. Smooth lines are fitted to the Hill equation.
(g) IC_{50}s from TNP-ATP current inhibition (h) and EC_{50}s from TNP-ATP fluorescence quenching
Data in (g) and (h) are derived from Hill fits individual experiments are plotted as coloured data points. The means and 95% confidence intervals for each construct are overlaid as black error bars. Each construct is labelled with GFP at the C-terminus, ANAP at residue W311 and co-expressed with SUR1.

We investigated the effect of three K39 mutations (K39R/E/A) on the channel sensitivity to TNP-ATP. Interestingly, the K39R mutant showed no significant difference in either TNP-ATP binding or current inhibition compared to the wild-type channel (Figures 4.12c-f). However, both K39E and K39A mutants exhibited a reduction in the channel binding and inhibition by TNP-ATP (Figures 4.12g-h). This is opposite to what was found for ATP.

To address the basis of this inconsistency between ATP and TNP-ATP, we aligned the structures of Kir6.2 bound with ATP and the TNP-ATP molecule in the binding site (Figure 4.13). My alignment showed that the position of the TNP moiety on the TNP-ATP overlaps with the position of the K39 sidechain (Figure 4.13). Thus, there is a steric clash, and it is possible that the TNP-ATP molecule may displace K39 from the position found in the solved structures. However, the charge on TNP-ATP is more negative than that on ATP (-4 on ATP, -6 on TNP-ATP). It is possible, however, that this does not occur in practice or that the larger TNP-ATP molecule distorts the binding pocket such that K39 interacts with its γ-phosphate better than that of ATP. This might favour K39 co-ordination with TNP-ATP more than ATP and explain why it is a more potent inhibitor of the channel.

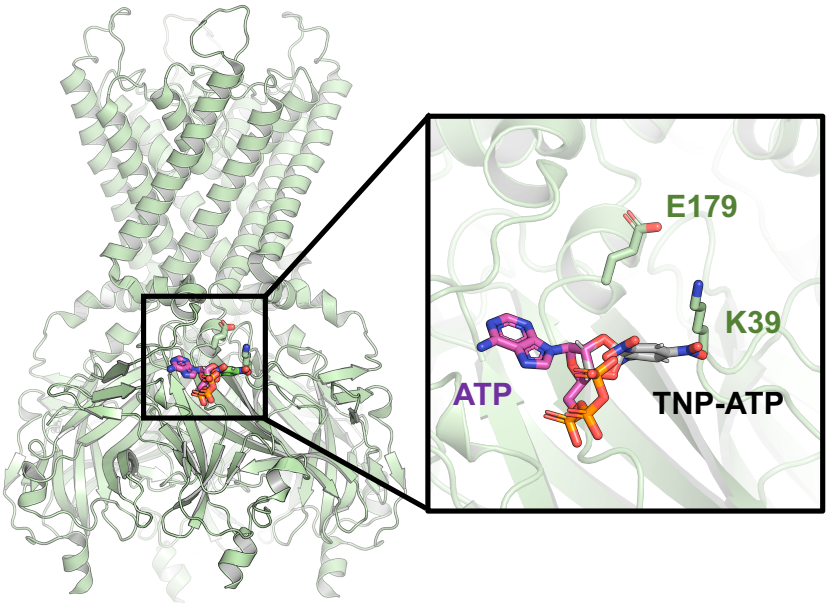


Figure 4.13 Alignment of TNP-ATP and ATP on the Kir6.2 structure.

Alignment of TNP-ATP to ATP on the Kir6.2 structure. TNP-ATP is represented in grey, ATP in purple. K39 and E179 side chains are represented as sticks in green. This alignment suggested that the TNP moiety of TNP-ATP overlaps with the side chain of K39 (green)

It seems clear from the previous section we cannot adequately resolve why mutations at residue 39 seem to have different effects on TNP-ATP binding, ATP inhibition and my computational predictions. Thus, I next explored the effects of PIP_2 on ATP inhibition on another set of mutations (E179K/A) at residue K179, which is predicted to lie within the PIP_2 binding site but not interact directly with ATP. Both mutations increase PIP_2 affinity in my CG-FEP study.

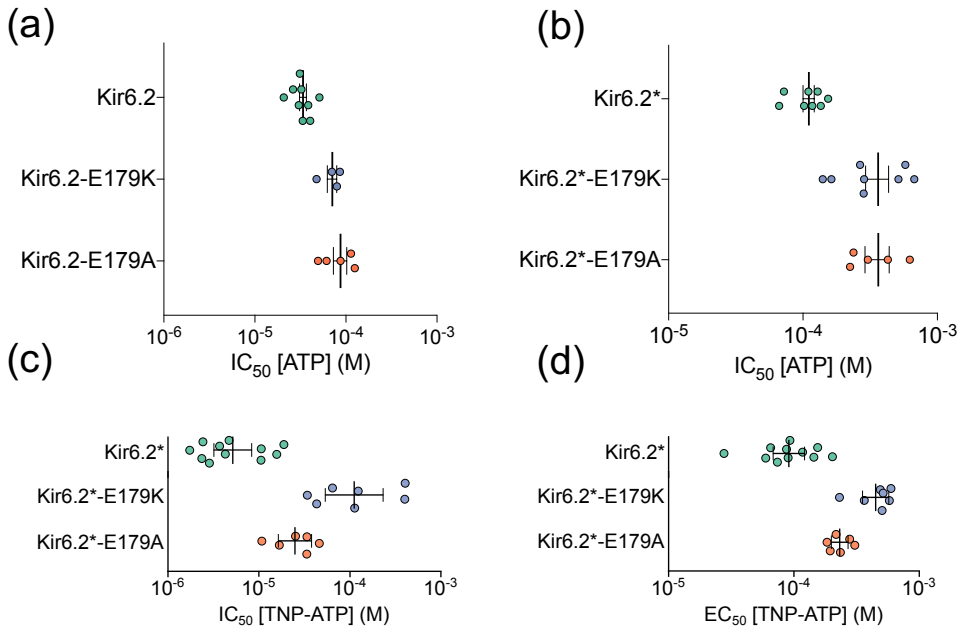


Figure 4.14. Effect of E179 mutations on ATP and TNP-ATP inhibition on K_{ATP} current.

(a,b) IC_{50}s from ATP current inhibition in the K_{ATP} channel (a) without ANAP substituted at W311 or (b) with ANAP substituted at W311.
(c) IC_{50}s from TNP-ATP current inhibition in the K_{ATP} channel with ANAP substituted at W311
(d) EC_{50}s from TNP-ATP fluorescence quenching
All data derived from Hill fits individual experiments are plotted as coloured data points. The means and 95% confidence intervals for each construct are overlaid as black error bars. Each construct is labelled with GFP at the C-terminus and co-expressed with SUR1

We investigated the effect of the E179K/A mutations on the ATP sensitivity on the wild-type K_{ATP} channel (Kir6.2/SUR1) and on Kir6.2* co-expressed with SUR1 (Figure 4.14a, b). In both cases, both E179K and E179A mutant channels displayed a reduction in current inhibition by ATP. We also observed a decrease in TNP-ATP binding and inhibition in both Kir6.2* E179A/K constructs (Figure 4.14c, d). Together, this supports the idea that a mutation that enhances PIP_2 binding can lower the affinity of the K_{ATP} channel to ATP.

4.4 Discussion

Our atomistic MD simulations suggested that a single critical residue, K39, forms part of both the ATP and PIP_2 binding sites. When both ligands are present, K39 has a stronger preference for coordination with PIP_2 than with ATP. This reduces the channel affinity for ATP as K39 is no longer available to coordinate with the ATP γ phosphate (Figure 4.15). No other residue (in either site) alters its contact probability in the presence of the other ligand. My experimental study shows that the K39R mutant exhibits a significant decrease in current inhibition by ATP, validating my prediction from the simulation. A similar reduction was found for K39E but not for K39A. However, a decrease in ATP sensitivity has been previously reported for the K39A mutation (Tucker *et al.*, 1998; Cukras *et al.*, 2002).

Experimental studies

To explain the complex observations from our experimental studies, I asked three questions:

a) Why is TNP-ATP a more potent inhibitor of the wild-type channel than ATP?
b) How do all K39 mutations (K39R/E/A) decrease TNP-ATP binding to the K_{ATP} channel?
c) Why does the K39R mutation reduce TNP-ATP binding to the channel but not reduce channel inhibition?

First question: Why is TNP-ATP a more potent inhibitor than ATP? We assume that PIP_2 is present in the patch when it is excised and that it remains bound to the channel. We estimated the distance between the molecules and K39: TNP-ATP being the closest due to steric bulk, PIP_2 based on the CG-MD simulation, and ATP based on the cryo-EM structure. Thus, although my simulations show that residue K39 (which is positively charged) prefers PIP_2 over ATP, I hypothesise that it would be more likely to interact with TNP-ATP than PIP_2. If this is the case, it will have two effects: (i) It will enhance TNP-ATP binding to K39 (ii) it will reduce PIP_2 binding and thereby activation of the K_{ATP} channel. Both these effects will mean that channel inhibition by TNP-ATP will be more potent than ATP, which is what was found experimentally.

Second question: How do all K39 mutations (K39R/E/A) decrease TNP-ATP binding to the K_{ATP} channel? As all three mutations reduce binding to the TNP-ATP to the channel, it is simplest to hypothesise that all 3 mutations may destabilise the TNP-ATP binding site. The K → E mutation creates an electrostatic repulsion as the charge on the surface of the binding site becomes more negative. The K → A mutation removes the positive charge on the TNP-ATP binding site, and the K → R mutation may sterically hinder the geometry of the TNP-ATP binding site.

Third question: Why does K39R reduce TNP-ATP binding to the channel but not reduce TNP-ATP inhibition? For an ion channel to gate from open to closed, TNP-ATP needs to bind to the channel, and produce a conformational change than is then somehow translated into closure of the pore. The simplest explanation of our findings is that the K39R mutation reduces ATP binding, as the data show, but that this is offset by a second process that simultaneously decreases channel activation by a similar amount. For example, the K39R mutation might cause an increase in transduction in the TNP-ATP inhibition pathway. Alternatively, it might allosterically destabilise PIP_2 binding.

All these ideas are highly speculative and needed to be tested further both computationally and experimentally.

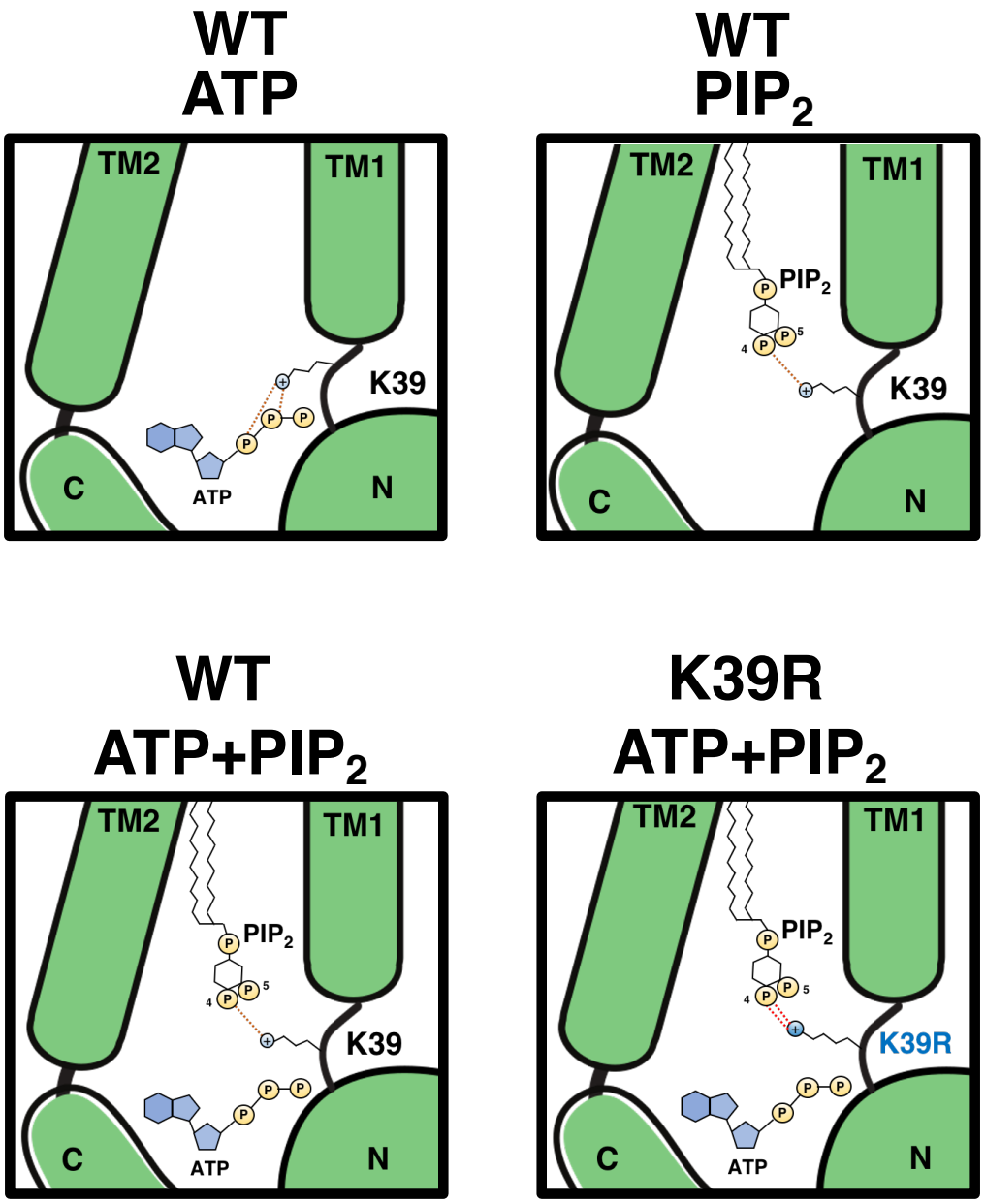


Figure 4.15 Schematic representation of the interaction between Kir6.2, ATP and PIP_2

Interaction between ATP (silver sphere), PIP_2 (hexagon with silver phosphate group – P) and Kir6.2 (green). A positive charge on K39 is denoted in yellow (or red in the K39R mutant). The orange dashed line represents hydrogen bonds between K39 and the ligand. The arrow represents the change in motion of K39.

Significance of neonatal diabetes mutations

I investigated how mutations at two residues (K39R, E179K/A) may cause neonatal diabetes. Currently, it is challenging to calculate the change in Gibb's free energy when I mutate acidic or basic residues. Thus, I used the H-bond dynamics to infer the binding energy. I found that K39R mutation increases the number of H-bonds between the side chain of residue 39 and the PIP_2 headgroup (Figure 4.15). However, an increase in the number of hydrogen bonds only quantifies the change in the enthalpy of interaction, not the binding free energy. The current challenge in computational biophysics is how to precisely calculate the enthalpy, entropy and the free energy change as charged amino acid residues are being perturbed. With the assumption that the entropy change in the binding process is negligible, I may potentially conclude that E179A/K and K39R mutations enhance PIP_2 affinity and reduce channel inhibition by ATP.

Caveats, assumptions, and future works

The major caveat in my study is that my system does not include SUR1. A previous study has shown that SUR1 is involved in the binding of ATP at the binding site on Kir6.2 (Martin *et al.*, 2017*a*). Mutation of K205 to glutamate (E) in SUR1 reduces both TNP-ATP binding to the K_{ATP} channel and channel inhibition (Usher *et al.*, 2020). A previous study has also shown that the IC_{50} of channel inhibition by ATP decreases by ~10-fold in the absence of SUR1 (Tucker *et al.*, 1997). This strongly suggests that SUR1 is involved in the ATP binding site and gating of the K_{ATP} channel. How ATP binds to the K_{ATP} channel is, however, different to PIP_2 binding. My prediction from the previous chapter suggested that SUR1 has no significant effect on the K_{ATP} channel affinity for PIP_2. The work from Dr Peter Proks also showed that the K_{ATP} channel expressed without SUR1 (ΔC construct) shows no significant change in its sensitivity to neomycin (Pipatpolkai *et al.*, 2020*a*).

There are several reasons why I excluded SUR1 from my simulations. Firstly, SUR1 is a large protein (1581 amino acids – 108 x 82 x 118 Å in size). This adds complexity to my simulation system. To include SUR1 in my simulation system, the computational cost would be four times more expensive. A possible approach might be to truncate SUR1 to TMD0 plus the first helix of the transmembrane segment of TMD1 to reduce the computational cost, but this is not desirable as it may alter the effect of the local dynamic of the L0 region (ATP binding site) and the global dynamics of the protein. Longer simulations of the octameric K_{ATP} channel should be achievable very soon in

the future thanks to rapid developments in computational hardware and software (Abraham *et al.*, 2015). The second reason is that the SUR1 is quite distant from K39 (~ 11 Å edge to edge distance based on PDB entry: 6BAA) (Martin *et al.*, 2017*a*). Thus, it should not interfere with the conformation of the K39. The last reason is that SUR1 coordinates only with the γ phosphate on the ATPγS molecule in several recent structures (PDB ID: 6JB1) (Wu *et al.*, 2018). Therefore, SUR1 may only help to stabilise the conformation of ATP but might not interfere with how ATP interacts with the side chain of K39. The simulation with the full K_{ATP} octameric complex is strongly needed to validate this statement.

I have also made several assumptions in the MD simulations. Those assumptions are discussed below as follows:

a) The structure of the K_{ATP} channel used in the MD simulation was solved in the absence of PIP_2, an activator of the K_{ATP} channel. Instead, it was solved in the closed state in the presence of ATP and glibenclamide, both of which are K_{ATP} channel inhibitors. There was a proposition that PIP_2 molecules may be bound to the solved K_{ATP} channel structure (Martin *et al.*, 2017*a*). However, the electron density map at the putative PIP_2 binding site is ambiguous. This means that the structure used in my simulations may mimic these possible situations; either the case where PIP_2 is spontaneously added to the closed state ATP-bound K_{ATP} channel; or the case where both PIP_2 and ATP are initially bound to the closed state K_{ATP} channel.

b) The simulations were run using virtual sites to reduce the computational cost. As previously discussed, virtual site simulation placed two dummy atoms between the CH_3 and NH_3, allowing the mass of the hydrogen atoms to be repartitioned into carbon and nitrogen atoms (Hopkins *et al.*, 2015). This method has been commonly used for lipids and proteins to reduce computational costs (Olesen *et al.*, 2018). However, there has been no attempt to use virtual site in the simulation for ATP molecule, only for GDP (Sun *et al.*, 2018). The virtual site should not affect the nature of the binding site in my simulations which are primarily driven by the electrostatic charge.

c) I used hydrogen bond analysis to infer the change in the enthalpy change affected by the K39R mutation. That is very different from the binding free energy, and it does not account for the entropic term ($\Delta G = \Delta H - T\Delta S$). I assumed that the entropic term would not be significant. It is also important to note that I assumed that all hydrogen bonds have an equal amount of binding free energy. This further assumes that the number of hydrogen bonds formed and broken between side chains and the surrounded water molecule is equal. Thus, I assume that the possible rearrangement of water molecules/side chains affected by the mutation is likely to be negligible.

In the previous chapter, I discussed that perturbing between one charged species to another is very difficult to achieve accurately under a reasonable simulation timescale. This difficulty also extends to a non-neutral charge side-chain perturbation such as aspartate to glutamate or lysine to arginine. Several methods have been developed to speed up the calculation,

such as non-equilibrium free energy perturbation (non-equilibrium FEP) (Gapsys *et al.*, 2021). These additional simulations should be attempted in the future to validate the effect of hydrogen bonding on the enthalpy and the binding free energy.

There are also several assumptions in the supporting experimental studies. I have discussed the conflict between the data on binding and inhibition by TNP-ATP found for the K39R mutant channel in the Results section of this chapter. On my simulation front, the natural next step would be to simulate the system with ANAP incorporated into the protein at W311 and with TNP-ATP instead of ATP. My primary hypothesis would be that incorporation of ANAP at W311 would not affect the dynamics of the K_{ATP} channel within my short simulation timescale. I predict that a simulation with TNP-ATP would be different to the simulation with ATP. I would expect the binding dynamics between two ligands, the ligand binding geometry, and the conformational change in the binding pocket to be very different. There is no cryo-EM structure with TNP-ATP in the binding site, so its initial binding pose would need to be obtained from molecular docking or extrapolation based on the ATP binding site, as demonstrated in Figure 4.13.

Parameterisation of TNP-ATP will also be a considerable challenge as electrons can be delocalised around the trinitrophenol group. This makes charge assignments on a single atom a tough challenge for classical MD simulation. The system would need to be simulated in a non-classical way to overcome these difficulties, either using a polarisable forcefield or quantum

mechanics-molecular mechanics (QM-MM) simulations. Both have exceptionally high computational costs but should be achievable shortly.

Summary

My work is consistent with PIP_2 regulating the function of Kir6.2 both by enhancing the P_{open} of the channel and by allosterically modulating the binding of ATP. In addition, I have demonstrated a mechanism by which mutations at residue 39 can lead to neonatal diabetes. My simulations have predicted that both K39R and E179K mutations enhance PIP_2 binding, and this may lead to a reduced sensitivity of Kir6.2 to ATP inhibition.

Chapter 5

Modulation of TMEM16A channel gating by Ca^{2+} and A9C

"Okay... I have 4 hours to get the impossible manuscript, Spitemmoskis doesn't open until 11:30, how am I going to get the steak?"

Andy Sachs – The Devils wears Prada (2006)

The work in this chapter is published in:

Dinsdale R.L., **Pipatpolkai T.***, Agostinelli E.*, Russell A.J., Stansfeld P.J. and Tammaro P. **The outer pore gate controls the pharmacology of the TMEM16A channel**. (2021). *Proceedings of the National Academy of Science* (2021). https://doi.org/10.1073/pnas.2023572118

5.1 Introduction

In the last two chapters, I investigated the binding of PIP_2 to the K_{ATP} channel and its competition with ATP adjacent to the PIP_2 binding site. One limitation of this study was not being able to observe a gating transition from a closed to an open state of the K_{ATP} channel.

As a proof of the principle that unbiased MD simulation can be used to capture the gating process of an ion channel, I changed my target protein to TMEM16A, the Ca^{2+} activated Cl^- channel (CaCC). In this chapter, I generated an open state model for the TMEM16A channel. Using atomistic MD simulations, I observed gating transitions from open to closed state. I also observed a Cl^- binding site and a test compound, anthracine-9-carboxylate (A9C) binding site within the channel.

TMEM16A is an endogenously expressed CaCC and is involved in many cellular processes such as controlling muscle tone and Cl^- ion homeostasis (Pedemonte & Galietta, 2014). Channel opening is influenced by different ligands, such as Ca^{2+} ions and/or PIP_2 (Yang *et al.*, 2008; Le *et al.*, 2019*a*). Recent structures of TMEM16A were solved with and without Ca^{2+} bound in its binding site for both wild-type channels and those with the I551A mutation (residue number based on acTMEM16A variant) (Paulino *et al.*, 2017*a*; Lam *et al.*, 2021). However, none of the structures solved has a wide enough cavity for water and Cl^- permeation, suggesting that all structures solved are all in the closed state (Rao *et al.*, 2019; Lam *et al.*, 2021).

When Ca^{2+} binds to its binding site on the lower portion of the transmembrane helix 6 (TM6), it triggers a conformational change where the lower half of TM6 swings outward with the hinge point at G640 (Peters *et al.*, 2018). This conformational change widens the cavity of the channel and allows hydrated Cl^- ions to permeate through the pore. This chapter uses the term "steric gate" to address this portion of the TMEM16A channel. TM6 mutations such as Q645A and I637A stabilise the opening of the channel in the absence of Ca^{2+} (Peters *et al.*, 2018). As Ca^{2+} binding site is negatively charge, the negative charge environment repels Cl^- and prevents its permeation (Lam & Dutzler, 2018). This repulsion is named an "electrostatic gate". The binding of Ca^{2+} at its binding site screened the negative charge which then permits Cl^- permeation. These gates are highly conserved amongst all TMEM16 family proteins, including those which function as a lipid scramblase such as TMEM16K and TMEM16F (Bushell *et al.*, 2019; Le *et al.*, 2019*b*; Ye *et al.*, 2019).

Despite the vast amount of knowledge on the Ca^{2+} activation of the TMEM16A channel, there is no structure with a wide enough cavity at the upper half of TM6 that indicates the channel is in an open state. Simulations and electrophysiological studies have led to the proposal of a second gate in the TMEM16A channel.

TMEM16A is highly expressed on the apical membrane and on the arteries. Thus, this diversity of expression makes TMEM16A recently become a novel therapeutic target for a range of diseases. Thus, both activations and

inhibitions of TMEM16A could have therapeutic value (Kunzelmann *et al.*, 2019). For example, TMEM16A activators may provide potential novel treatments for cystic fibrosis (CF) (Danahay *et al.*, 2020). CF is a genetic disease caused by a loss-of-function mutation in CFTR, an ABC transporter Cl^- channel (Groman *et al.*, 2002). During CF, defective Cl^- transport through CFTR channels in the apical membrane of airways epithelial cells impairs water movement to the surface of mucus. This defects in water movement leads to an accumulation of dense mucus in the airways, which is the hallmark feature of CF. Thus, activation of TMEM16A in the apical membrane of the epithelial cells may correct for loss of Cl- transport via CFTR channels and thus favour mucus fluidification.

TMEM16A inhibitors may aid treatments of conditions such as pulmonary and systemic hypertension. In pulmonary hypertension patients, TMEM16A are overexpressed in the pulmonary arteries and increases Cl^- current. An increase in the Cl^- current is implicated in the proliferation of the (Wang *et al.*, 2015*a*). Furthermore, TMEM16A depolarisation in pulmonary artery smooth muscle cell cells lead to arterial constriction which exacerbate the pulmonary HT (Papp *et al.*, 2019) . Thus, TMEM16A inhibition may aid smooth muscle relaxation and reduction in the proliferation of the pulmonary artery smooth muscle cell during pulmonary HT (Liang *et al.*, 2009; Sun *et al.*, 2012).

Recently, both TMEM16A activators and inhibitors have been developed. Recent study shows that an activator, 3,4,5-Trimethoxy-N-(2-methoxyethyl)-N-(4-phenyl-2-thiazolyl)-benzamide, potentiate CaCC current in

airway epithelial cell line (Namkung *et al.*, 2011). In addition, several pore inhibitors have been designed. These include 2-(4′-chloro-2′-methylphenoxy)-*N*-[(2″-methoxyphenyl)methylideneamino]-acetamide (Ani9) (Seo *et al.*, 2016) and N-((4-methoxy)-2-naphthyl)-5-nitroanthranilic acid (MONNA) (Oh *et al.*, 2013). However, the selectivity of these compounds is not fully defined. In this chapter, I have focussed on anthracene-9-carboxylic acid (A9C) (Figure 5.1). A9C is a small molecule that produced a complex, biphasic effect on the channel. A9C is both blocks the pore and allosterically increases activation of the TMEM16A channel by Ca^{2+} and voltage (Ta *et al.*, 2016). Thus, understanding the action of A9C could provide insight towards activatory and inhibitory mechanisms of the channel. In this chapter, I provide simulations that, together with whole-cell patch-clamp and the development of ultra-fast perfusion system performed by Dr Ria Dinsdale and Dr Emillio Agostinelli from the Tammaro group, provide new mechanistic insights into how A9C regulates the TMEM16A channel.

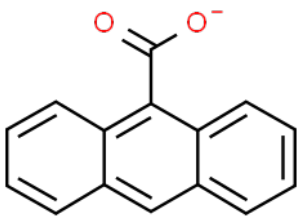


Figure 5.1 The structure of anthracine-9-carboxylate

The structure of the antracine-9-carboxylate (A9C) drawn in the skeletal formula. In our simulations, the A9C is parameterised in a deprotonated state with one negative charge on one of the oxygen atoms. This parametrisation is generally acceptable as the molecule is symmetrical.

Several computational methods have been developed to assess whether the channel is opened or closed. The standard approach calculates the pore radii and whether it is wide enough for permeation of a hydrated ion (Smart *et al.*, 1996). However, this approach does not account for the hydrophobicity of the amino acid residues at the gate and may lead to an over/underprediction of the wetness of the pore. Later software (such as Channel Annotation Packages - CHAP) has incorporated calculations of the free energy for water permeation and the relative pore hydrophobicity (Rao *et al.*, 2019; Klesse *et al.*, 2019). This development enables a more heuristic assessment of whether the pore poses a barrier to water. However, the bottleneck of both methods is the difficulty in the assessment when the pore is asymmetrical. This poses a challenge for the analysis of the TMEM16A channel as its pore is not only asymmetrical, but lipids may also form the pore lining. To simplify the difficulty of the analysis, in this chapter, I simply assessed the water permeation event using “VolMap”, a software that calculates the occupancy of water molecules as they permeate through the pore (Humphrey *et al.*, 1996).

In this chapter, I used atomistic MD simulation to determine the conformational changes underlying the gating process of the TMEM16A channel. I identified a gate at the outer pore of the TMEM16A channel. This gate closes when Ca^{2+} is removed from its binding site both in the WT channel and the I637A and Q645A mutant channels. We assessed the path of Cl^- permeation through the pore. I also identified the putative A9C binding site, which can only be accessed from the extracellular solution.

5.2 Methods

5.2.1 Structural modelling

All molecular modelling was performed using Swiss-Model (Waterhouse *et al.*, 2018) and PyMOL (Schrodinger LLC, 2015). The mouse TMEM16A in a Ca^{2+} free state was modelled based on the PDB entry: 5OYG (Paulino *et al.*, 2017*a*), while the Ca^{2+} bound state was modelled using the PDB entry: 5OYB (Paulino *et al.*, 2017*a*). The Ca^{2+} bound state of the channel (PDB entry: 5OYB) is not fully open at the outer vestibule (the pore size is smaller than 4 Å and also shown in Rao *et al.*, 2019), especially when compared to the Ca^{2+} bound scramblase afTMEM16 or TMEM16K structures (Bushell *et al.*, 2019; Falzone *et al.*, 2019). I modelled an open vestibule for the TMEM16A channel based on TMEM16K (PDB entry: 5OC9) and compared it with the model of the Ca^{2+} bound state. This produced a TMEM16A model with a fully open conduit across the membrane that I, therefore, tested with molecular simulation for water and ion permeation properties. The simulated structure of a PIP_2 bound TMEM16A conductive state channel was kindly provided by the Chen lab (Jia & Chen, 2021). Structures of the open-state TMEM16A channels, and the co-ordinate of the PIP_2 bound to the open-state TMEM16A channel are available on https://github.com/WillPPK/Thesis.

5.2.2 Coarse-grained MD preparation for atomistic simulation

The simulation systems were initially set up using coarse-grained MD (CG-MD) simulations. All structures were converted to CG representation and embedded in a palmitoyl oleoyl phosphatidylcholine (POPC) bilayer and solvated in water and 0.15 M NaCl using *insane* (Wassenaar *et al.*, 2015). All simulations were carried out with the Martini2.3 biomolecular forcefield (Monticelli *et al.*, 2008; de Jong *et al.*, 2013). The tertiary and quaternary structures of the protein were maintained through the application of an elastic network with a force constant of 1,000 kJ/mol/nm^2 between CG backbone particles within 0.5-0.9 nm. Systems were energy minimised using the steepest descents algorithm for 5000 steps. The system was then equilibrated for 500 ns. During the equilibration process, a temperature of 323 K was maintained using V-rescale temperature coupling throughout the simulation. The pressure of 1 atm was maintained using semi-isotropic Parrinello-Rahman pressure coupling (Parrinello & Rahman, 1981; Bussi *et al.*, 2007). All simulations were carried out using GROMACS-2020.1 (Van Der Spoel *et al.*, 2005).

5.2.3 All-atom MD simulations and water accessibility inside the pore

The final time-point of the simulated CG systems was converted to all-atom using CG2AT (Vickery & Stansfeld, 2021), with the original protein structure aligned with the coordinates of the protein from the CG simulation. All simulations were carried out using the CHARMM36 biomolecular forcefield (Huang & Mackerell, 2013). Virtual sites were applied to the CH_3 and NH_3^+ groups of the proteins and lipids, allowing an integration time-step of 4 fs in the

800 ns production runs (Olesen *et al.*, 2018). Simulations were energy minimised using the steepest descents algorithm, with simulations performed at a temperature of 310 K. In all simulations, distances were measured using GROMACS-2020.1. The systems were equilibrated for 20 ns with position restraints imposed on the C_α backbone of the protein. Data were collected from 800 ns simulations (3 repeats) with and without Ca^{2+} present in its binding site. To calculate water accessibility in the pore, I used the last 100 ns of the simulation trajectory to calculate an occupancy of the water molecule (3 Å radii) using VolMap packages in VMD1.9.4 every 1 ns frame (Humphrey *et al.*, 1996). All images were processed using PyMOLv2.1(Schrodinger LLC, 2015).

5.2.4 Identification of Cl^- and A9C binding sites in the pore

In my 800 ns simulations with bound Ca^{2+}, I identified a central binding site for Cl^-. However, I could not observe a complete ion translocation from one side of the membrane to the other. This appears to be due to Cl^- preferentially binding to a central site during the simulations, in agreement with a previous study (Jia & Chen, 2021). The central Cl^- binding site was identified using the VolMap tool in VMD based on the occupancy of the Cl^- in the pore over the simulation trajectory (Humphrey *et al.*, 1996). To observe a full ion permeation path, Cl^- was placed on both the outer and the inner sides of the pore and simulated for 50 ns with 3 repeats to determine the paths towards the central binding site. The systems were energy minimised using the steepest descents followed by a 5 ns equilibration for the system, where the C_α backbone on the protein and the Cl^- was restrained with 1000 kJ/mol/nm^2 with a 4 fs time-step.

From these simulations, I tracked the motions of the ions towards the central binding site.

To identify the A9C binding site, an A9C molecule was placed at the outer mouth of the pore using AutoDock Vina (Trott & Olson, 2010*a*). The A9C was deprotonated, and the partial charges were assigned using Maestro (Schrödinger Release 2021-1: Maestro, Schrödinger, LLC, New York, NY, 2021). To place A9C at the outer mouth with no steric clashes, I centred the grid space on the extracellular half of the Ca^{2+} bound open-state of the TMEM16A channel with a search space of 76x76x88 $Å^3$ for 20 iterations. This initial binding site for A9C, on the extracellular side, was used for the subsequent simulations to identify if A9C could permeate further into the pore. An initial position of A9C in the open state TMEM16A channel is available for download on https://github.com/WillPPK/Thesis. The simulation parameters for A9C were generated using CHARMM-GUI (Kim *et al.*, 2017). The systems were energy minimised before a 20 ns equilibration during which the C_α backbone on the protein and all atoms of the A9C molecules were restrained by 1000 $kJmol^{-1}nm^{-2}$. Data were collected from 200 ns simulations (3 repeats), an adequate duration for A9C to adopt a stable binding pose in the pore.

5.3 Results

5.3.1 Molecular modelling and dynamics offer insight into the structural rearrangement of the outer pore region.

I used molecular modelling and molecular dynamics (MD) to gain insights into the Ca^{2+} dependent conformational rearrangements of the outer pore. As there is not yet an experimentally determined structure of a fully open-state of TMEM16A, I used the structure of TMEM16K (PDB entry: 5OC9) to build a model of TMEM16A in the open conformation (Figure 5.2a) (Paulino *et al.*, 2017*a*; Bushell *et al.*, 2019). The model clearly exhibits three gates, an opening outer pore of the channel, an opening at the lower portion of the TM6, and the Ca^{2+} bound in their respective binding site. The model shows high α-helical and β-strand content, with the Molprobity score of 1.96 (Figure 5.2b).

The coordinates of the TMEM16A channel with Ca^{2+} bound (PDB entry: 5OYB) were structurally aligned with the TMEM16K structure, for comparison (Figures 5.3a-b). The primary difference between the two structures is the position of the TM4 helix. We modelled the coordinates of TM4 from the TMEM16K structure within the Ca^{2+} bound TMEM16A structure. The modelled "open-state" TMEM16A channel (orange) has a wider outer pore than the previously solved Ca^{2+} bound TMEM16A structures (Figures 5.3a-b). I also compared our model to the close-state structure without Ca^{2+} (Figure 5.3c). The structure generally resembled each other in the most region, apart from TM4 and TM6. The structure without Ca^{2+} shows a narrower TM4-TM6 distance and a kink at the end of TM6. Our model also has a wider pore than an I551A mutant channel previously proposed to open in the absence of Ca^{2+} (Lam *et al.*, 2021) (Figure 5.3d).

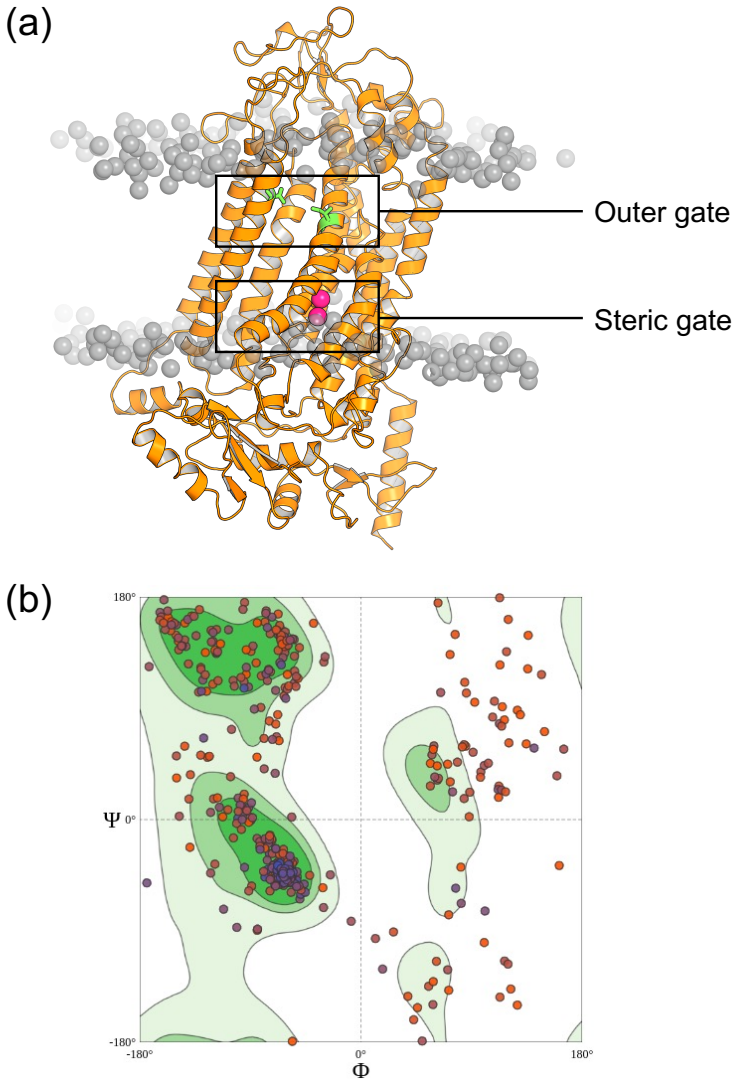


Figure 5.2 Validation of an open-state TMEM16A model

(a) The structure of an open-state TMEM16A channel with calcium remains bound to the binding site. The protein is embedded in the POPC bilayer with 0.15 M NaCl. The POPC alkyl tail, water and ions in the simulation system are not shown for clarity. The outer gate is located on the top half of the protein and is denoted by two key residues, V539 and I636, represented as a green stick. The location of the steric gate is at the lower half of the protein, denoted by a kink in the TM6 region.
(b) Ramachandran plot representing the distribution of the ϕ and ψ angles of the peptide bond in the TMEM16A model. Each dot represents the dihedrals around a peptide bond. Most region of the modelled structures retained their secondary structures as most ϕ and ψ angles are in the α-helical or β-strand region.

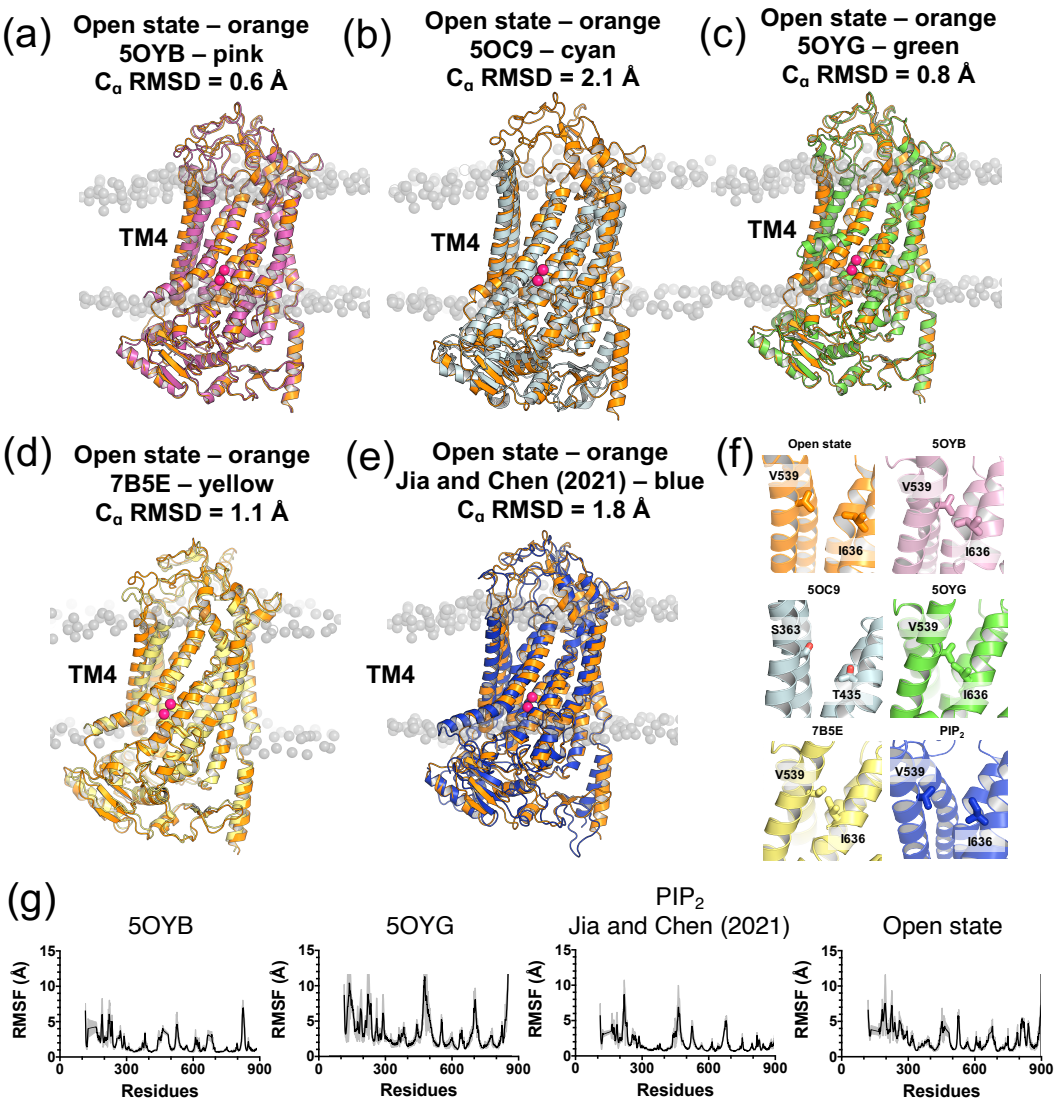


Figure 5.3 Structural alignment between the open-state model and the existing structures

The hybrid model (orange) is aligned with (a) 5OYB Ca^{2+} bound TMEM16A cryo-EM structure (pink), (b) 5OC9 Ca^{2+} bound TMEM16K structure (cyan), (c) 5OYG Ca^{2+} free TMEM16A cryo-EM structure (green), (d) 7B5C Ca^{2+} bound TMEM16A mutant cryo-EM structure (yellow) and (e) the model generated from spontaneous opening of the channel caused by PIP_2 (blue). (f) Expanded view at the pore of the channel to show the constriction at V539 and I636 in the different cryo-EM structures and models, as indicated. (g) C_α RMSF and secondary structure retention of the cryo-EM structures and the existing models. Calculated residue RMSF of the first 200 ns simulation of the cryo-EM structures with Ca^{2+} bound (5OYB), Ca^{2+} free (5OYG), PIP_2 bound model and our hybrid model. The secondary structure analysis was sampled every 1 ns. The shaded region indicates the standard deviation around the mean of three repeats.

To validate this open-state model, I considered three criteria: (i) the structural similarity to the existing cryo-EM structures/models using static comparison of C_α root mean square deviation (RMSD) analysis; (ii) the stability of the local secondary and tertiary structures using C_α root mean square fluctuation (RMSF) analysis through 200 ns MD simulation and (iii) conduction of water molecules and Cl^- ions through the pore.

Figure 5.3 provides a structural alignment of our open state TMEM16A model with the deposited TMEM16x structures and the recent simulated structure of the PIP_2-bound open TMEM16A channel. During the course of my thesis project, an independent research group simulated the structure of TMEM16A with both Ca^{2+} and PIP_2 bound (Jia & Chen, 2021). Structural alignment analysis demonstrated that our model closely resembles these structures and almost completely overlaps with the PIP_2-bound open TMEM16A channel model, with C_α RMSD of 1.8 Å (Figures 5.3e-f). Our open-state model also showed comparable stability with the TMEM16A wild-type structures and PIP_2-bound model over three 200 ns simulations (Figure 5.3g).

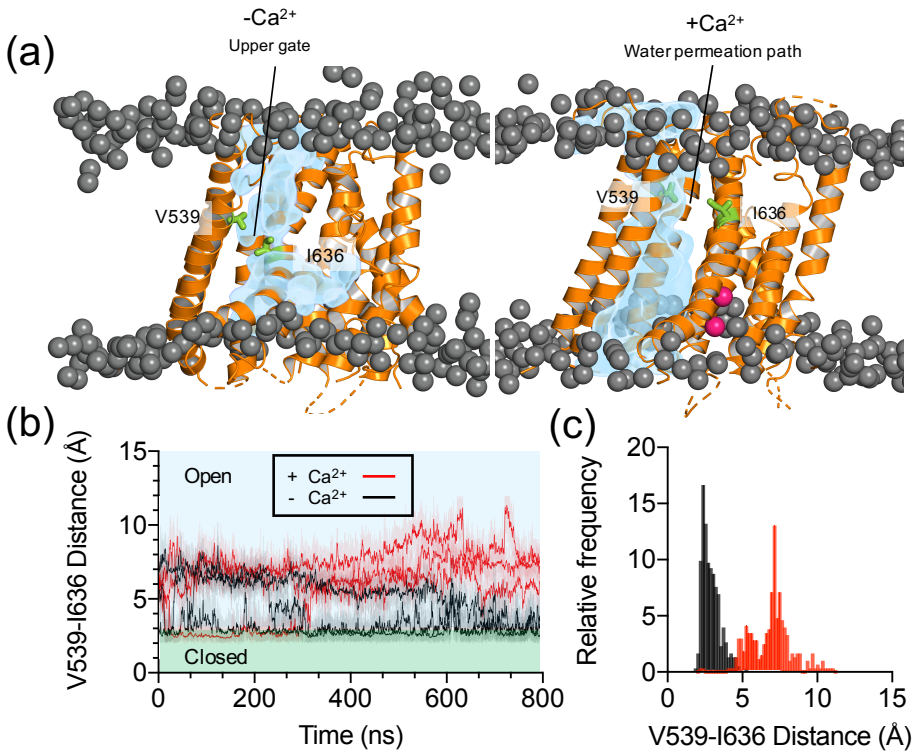


Figure 5.4 Occupancy of water molecules inside the TMEM16A pore

(a) Representative calculations of the average occupancy of water molecules inside the pore of the TMEM16A channel over the last 100 ns of an 800 ns simulation. The occupancy of the water molecules is shown in blue. The phosphate headgroups are shown as the dark grey spheres, and the protein is represented as a cartoon in orange, the relative positions of V539 and I636 in the modelled Ca^{2+} bound open state and the Ca^{2+} free closed state are shown in green (PDB entry: 5OYG). Ca^{2+} are shown in pink.
(b) Calculated minimum distance between V539 and I636 in the Ca^{2+} bound (red) and the Ca^{2+} free (black) state across 800 ns simulations (n=3). The lighter shaded line shows the distance calculated every 10 ps, and the darker line shows running averages over 1 ns.
(c) A histogram showing the V539-I636 distance in the last 100 ns of the simulation. Two conditions, Ca^{2+} bound (red) and the Ca^{2+} free (black) state, were sampled with the bin size of 0.17 Å and sampled every 1 ns (n=3).

To test whether our model was conductive to solvent, I calculated the occupancy of the water molecules within the pore in the last 100 ns of the 800 ns simulation trajectory (Figure 5.4a). The open-state model showed water molecules within the outer pore region. I studied the effect of Ca^{2+} on the open pore in 800 ns simulations (n=3); as Ca^{2+} was removed, the pore narrowed, especially in proximity of V539 (TM4) and I636 (TM6) (Figure 5.4). These residues are in the vicinity of a gate proposed in recent experimental studies from the Dutzler group (Lam & Dutzler, 2021; Lam *et al.*, 2021). My simulations suggested that by removing the Ca^{2+} from its binding site, the conformation of the outer pore converges towards the closed-state Ca^{2+} free structure (5OYG). Together, my simulations indicate that Ca^{2+} ions are required to retain a stable opening of the outer pore of the TMEM16A channel.

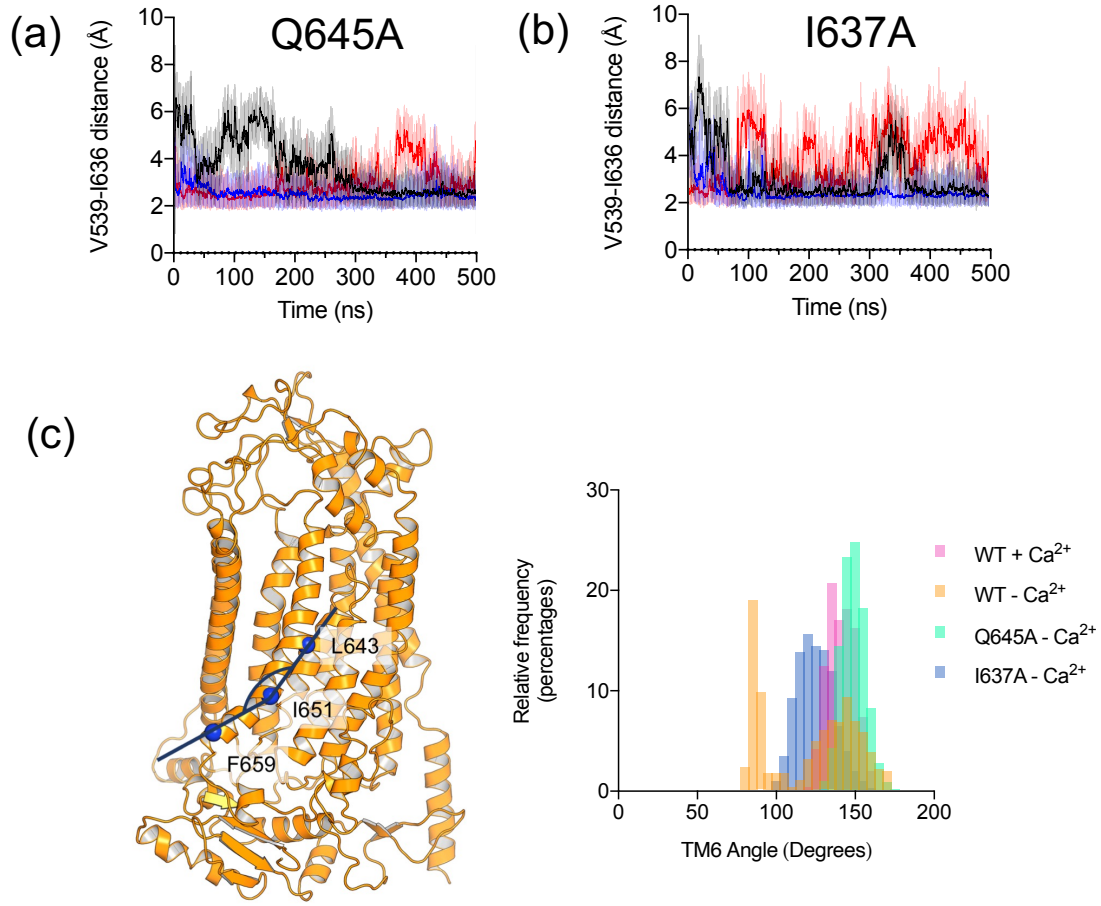


Figure 5.5 Effect of mutations on the outer gate and steric gate conformation

(a-b) The minimum distance between V539 and I636 in Q645A (a) and I637A (b) mutants simulations in the absence of Ca^{2+}. The lighter shaded line shows the distance calculated every 10 ps, and the darker line shows running averages over 1 ns. The different colours indicate the three independent repeats.
(c) Left panel: Representative structure of the open-state model, showing the locations of L643, I651 and F659 on TM6. Right panel: Histogram of the angle distributions between the Cα atoms of L643-I651-F659 on TM6, representing the kinking of the TM helix over 500 ns of simulations of wild-type (orange), Q645A (green), and I637A (blue) channels in the absence of Ca^{2+}. The comparison for the wild-type model in the presence of Ca^{2+} is shown in purple. Histograms were sampled with the bin size of 5 degrees and sampled every 1 ns (n=3).

I investigated how Q645A and I636A mutants promote channel opening in the absence of Ca^{2+}. First, I observed that the outer pore of both I637A and Q645A mutants are closed when Ca^{2+} is removed (Figures 5.5a-b). This suggested that Ca^{2+} is required for a stable opening of the outer pore of the channel. However, simulations of the Q645A and I637A mutant channels do not show the same degree of transmembrane helix kinking in the absence of Ca^{2+} over 500 ns simulations (n=3) (Figure 5.5). This suggested that the Q645A and I637A stabilise the opening of the steric gate and favour the open state of the channel. Thus, I propose that the steric gate and the outer gate function independently; however, both are being opened by Ca^{2+}.

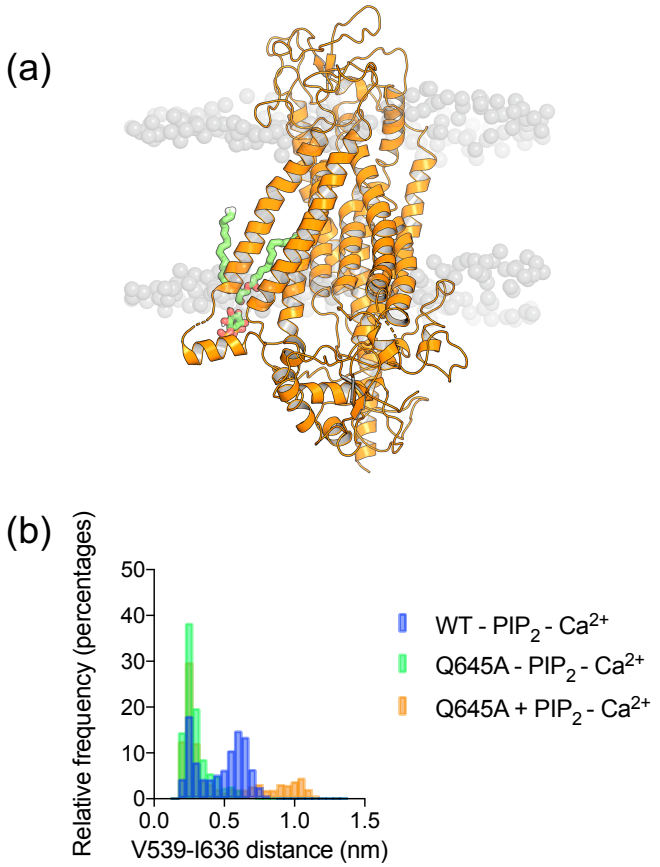


Figure 5.6 PIP_2 does not stabilise the outer gate opening

(a) The position of PIP_2 in the TMEM16A channel. The co-ordinate of PIP_2 molecule is taken from Jia and Chen (2021). The PIP_2 molecule is docked to the PIP_2 binding site after aligning my open state model to their open state model. 500 ns simulations were run to analyse an effect of the PIP_2 binding to outer gate opening (n=3)
(c) Histogram of the V539-I636 distance distributions over 500 ns of simulations of wild-type with Ca^{2+} (blue), Q645A without PIP_2 and without Ca^{2+} (green), and Q645A with PIP_2 and with Ca^{2+} (orange). Histograms were sampled with the bin size of 0.5 Å and sampled every 1 ns (n=3).

One of the potential candidates that might contribute to an opening of the outer gate of the TMEM16A mutant channel in the absence of Ca^{2+} is PIP_2. To investigate the role of PIP_2 in the opening of the outer gate of TMEM16A, I aligned the position of PIP_2 based on the structure of the PIP_2 bound open-state model to the Q645A mutant channel (Figure 5.6a). The Ca^{2+} was then

removed from the binding site and the system was then simulated for 500 ns (n=3). Here, I calculated the distance between V539 and I636, similar to our previous analysis of an opening of the outer gate (Figure 5.6b). The simulation of the Q645A mutant with PIP_2 in the binding site displays a slightly greater frequency of an opening event at the outer gate comparing to the mutant channel without PIP_2. Though, the extent of the opening of the outer gate is to a much lesser extent, comparing to the WT channel in the presence of Ca^{2+}. This result is very preliminary and suggest that PIP_2 might stabilise an opening configuration of the open gate. Though, other factors must be contributing to an opening of the outer pore so that Q645A mutant is remains conductive even in an absence of calcium.

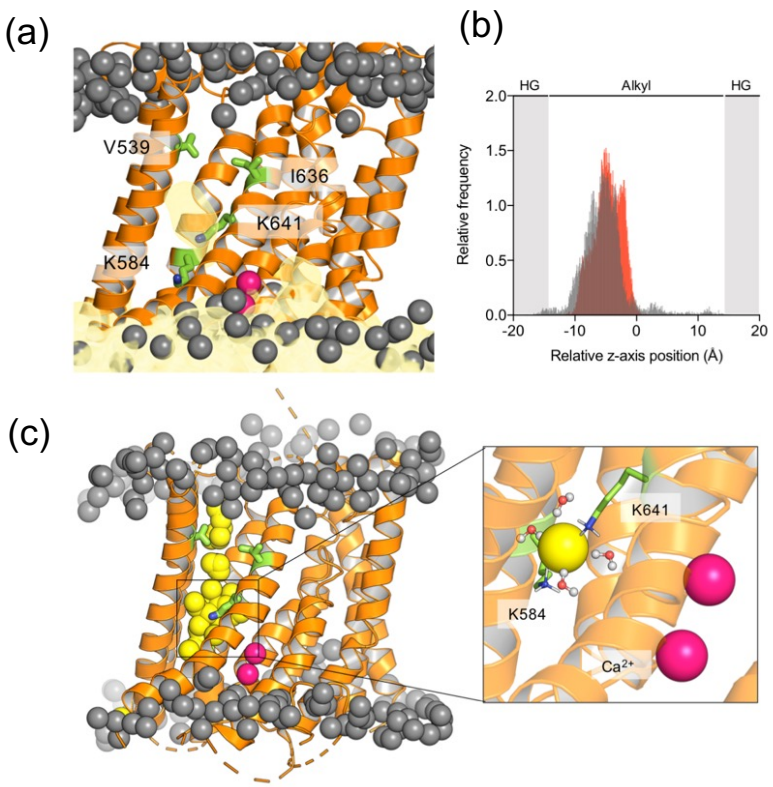


Figure 5.7 Cl^- binding site in the TMEM16A open-state model

(a) Calculations of the average density of Cl^- inside the pore of the TMEM16A open-state model over an 800 ns trajectory using VolMap. The occupancy of Cl^- is shown in yellow. The phosphate headgroups are shown in the dark grey sphere. The protein is represented as a cartoon in orange, the positions of V539, I636, K584 and K641 are shown in green and Ca^{2+} are shown in pink.
(b) Histogram of the position of Cl^- relative to the membrane bilayer within the model of the open TMEM16A channel across 50 ns simulations (n=3) with a bin size of 0.08 Å and sampling every 10 ps. The starting position of the Cl^- ion was either on the extracellular (grey) or intracellular (red) side of the membrane. In either case, Cl^- reached the same central binding site during the simulation. The regions shaded in grey denoted the positions of the phosphate headgroups.
(c) Left panel: Overlay of snapshots of different Cl^- positions as it permeates through the pore. Right panel: Representative position of a Cl^- (yellow) at the end of a 50 ns of simulation. The residues that make contact with Cl^- are shown in green, and the Ca^{2+} within the TMEM16A binding sites are shown in pink.

Given that my simulations show water permeation, I aim to assess Cl^- permeation through the Ca^{2+} bound open-state model. I identified a central binding site coordinated by two lysine residues (K584 and K641) and, on average 4 water molecules (Figure 5.7a). This binding site is consistent with published experimental studies (Paulino *et al.*, 2017*a*). The single-channel recording shows that the conductance of Cl^- is 1-3 pS, so I would then expect the time for the ion to find its binding site to be approximately 350 − 1000 ns (Manoury *et al.*, 2010). In my 800 ns simulations, this site was only apparent when the ion permeated from the intracellular side of the membrane. To hasten the time required for the simulation, I conducted an additional set of simulations whereby the ion was placed at either the intracellular or extracellular mouth, with the ion released to sample its optimal binding pose inside the pore. This computational set-up could be viewed as an increase in a local Cl^- concentration nearer the pore region, facilitating ion movement to the central cavity. The Cl^- ion remains stable in the binding site throughout 50 ns. The Cl^- binding site is identical to the one observed in one out of three 800 ns simulations. However, I did not observe any full permeation of the Cl^- ion, despite an ion remaining bound for almost 300 ns. Nevertheless, although permeation through the channel was not observed, I can track an ion permeation path through the membrane (Figure 5.7b). The prevalence of this Cl^- interaction site within my simulations suggests a highly favourable binding site for the conduction of Cl^-, with coulombic knock-on required to promote ion permeation.

5.3.2 A9C binding site

I next aimed to identify the binding site for A9C in the pore of the TMEM16A channel. I used AutoDock Vina to position A9C into the outer pore of the open-state TMEM16A model (Trott & Olson, 2010*b*). In the presence of Ca^{2+}, A9C reached a binding site in the first 100 ns and remained at this site for the entire duration (200 ns) of the simulation (Figure 5.8a). In a simulation where A9C was applied intracellularly side, there was no entry into the pore, in close agreement with a previous study, where A9C was shown not to act from the intracellular side of the membrane during patch-clamp experiments (Ta *et al.*, 2016) (Figure 5.9).

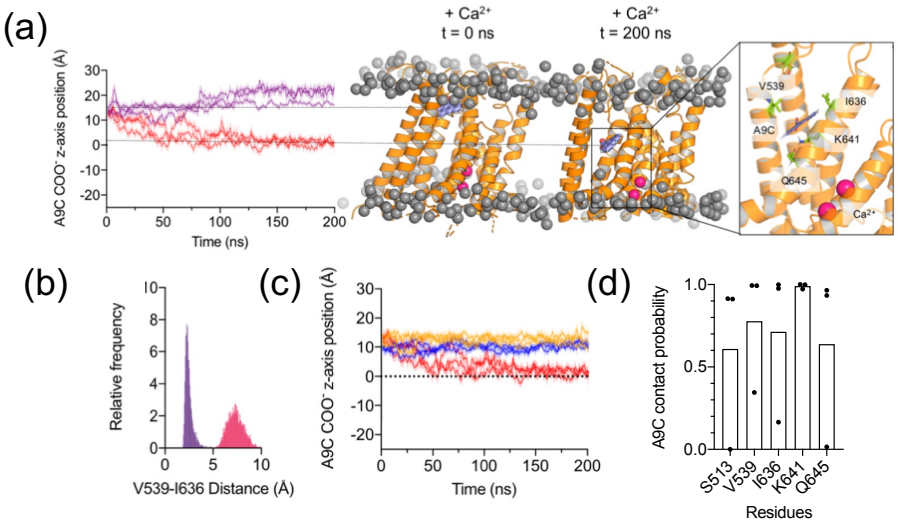


Figure 5.8 Position of the A9C molecule relative to the membrane in the TMEM16A open-state model with the presence and absence of Ca^{2+}

(a) The calculated z-axis position of an A9C molecule, where the centre of the bilayer is set as 0. The positive direction marks the movement of the A9C towards the extracellular side of the pore, and the negative marks the movement towards the cytoplasmic side. The simulations were run in the presence (red) or the absence of Ca^{2+} (purple). An initial position of both simulations is marked with the purple line. The red line shows the position of an A9C molecule relative to the membrane bilayer as it moves downward in the Ca^{2+} bound TMEM16A channel during simulations. The lighter shaded line shows the distance calculated every 10 ps, and the darker line shows running averages over 1 ns (n=3). Subpanel: A representative position of an A9C molecule (blue) after 200 ns of simulation. The residues with a contact probability greater than 50% are shown in green.
(b) A histogram showing V539-I636 distance in the Ca^{2+} bound (red) and the Ca^{2+} free (purple) state during the last 100 ns of the simulation with the bin size = 0.05 Å and sampled every 10 ps (n=3).
(c) The calculated z-axis position of an A9C molecule. The simulations were run in the presence of Ca^{2+} in the hybrid model (red), 5OYB structure (yellow) and 5OYG (blue)
(d) Contact probability across 3 repeats in the last 100 ns of the 200 ns simulation trajectory of residues shown in (a). All individual repeats are shown as individual data points with the mean as a bar.

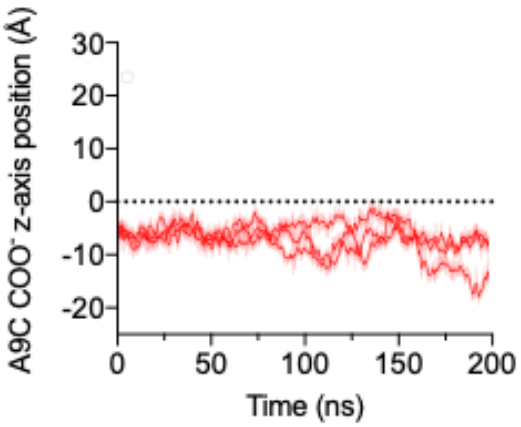


Figure 5.9 Inability of intracellular A9C to reach a binding site within the pore of the open TMEM16A model

Z-axis position of A9C within the pore of the open TMEM16A model generated in this study. The simulations were run in the presence of intracellular Ca^{2+}. A9C was applied to the intracellular side. In three repeats of 200 ns simulation, the A9C did not translocate to and reach the putative binding site.

In the absence of Ca^{2+}, outer gate closure (between V539 and I636) prevented A9C from reaching its putative binding site (Figures 5.8a-b). My simulations also demonstrated that A9C could not reach its binding site in any cryo-EM structures, which all present a closed outer pore (Figure 5.8c). The residues forming the putative A9C binding site were identified as the residues within a 4 Å distance from the A9C molecule for at least half of the simulation time. These residues include S513, V539, I636, K641, and Q645. The identical residues were identified when the simulation was performed on the PIP_2-bound TMEM16A open state model produced by Jia and Chen, 2021 (Figure 5.10). The carboxyl group of A9C coordinated with the K641 residue, while its aromatic system interacted with the hydrophobic side chains of I636 and V539 (Figure 5.8d). Overall, this suggested that a Ca^{2+} dependent outer pore

rearrangement crucially enables A9C to reach its binding site, consistent with the electrophysiology data. These experimental data will be discussed in the next section.

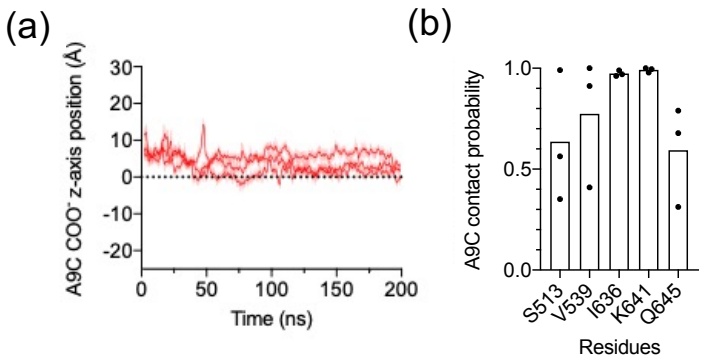


Figure 5.10 Entry of extracellular A9C into the pore of the PIP_2-bound TMEM16A model

(a) z-axis position of A9C within the pore of the PIP_2-bound model of the open TMEM16A channel described in Jia and Chen (2021). The simulations were run in the presence of intracellular Ca^{2+}. A9C was applied extracellularly.
(b) Contact probability across 3 repeats in the last 100 ns of the 200 ns simulation trajectory. All individual repeats are shown as individual data points with the mean as a bar.

5.3.3 Experimental studies

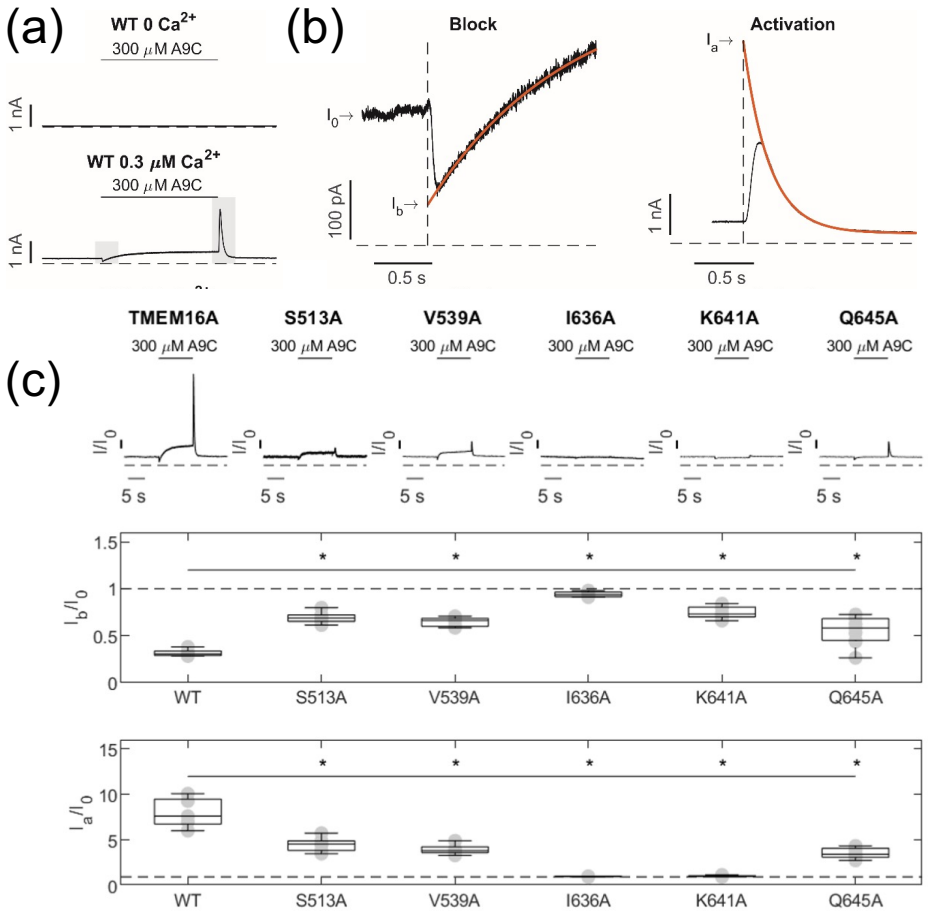


Figure 5.11 Separation of inhibiting and activating effects of A9C on TMEM16A and the A9C binding site

(a) Whole-cell currents recorded from HEK-293T cells expressing WT TMEM16A, channels, as indicated. V_m was +70 mV and $[Ca^{2+}]_i$ was 0 or 0.3 μM, as indicated. Extracellular A9C (300 μM) was applied as indicated by the horizontal bar via a fast-perfusion system ('concentration jump'). Dashed horizontal lines represent the zero-current level. The shaded grey bars indicate regions of the recording that were expanded in (b).
(b) The continuous red traces represent single exponential fits used to back-extrapolate the currents to obtain I_b (left panel) and I_a (right panel).
(c) (Top) Whole-cell currents recorded from HEK-293T cells expressing TMEM16A, TMEM16A-K641A, TMEM16A-V539A, TMEM16A-I636A, TMEM16A-Q645A or TMEM16A-S513A channels, as indicated. Extracellular A9C (300 μM) was applied as indicated by the horizontal bar via a fast-perfusion system ('concentration jump'). Dashed horizontal lines indicate the zero-current level. (Middle) Mean current inhibition (I_b/I_0) or (Bottom) current activation (I_a/I_0) plotted as box plot for each channel type, as indicated. The number of experiments was 5-9 in each case. *P<0.05, compared to TMEM16A.

I collaborated with Dr Ria Dinsdale and Dr Emilio Agostinelli to provide molecular insights into their electrophysiological studies. The collaborators provide the figures and figure legends discussed in this section.

We were interested in the biphasic effect (activation and inhibition) of A9C on the TMEM16A current. All experimental studies discussed in this chapter were done in the whole-cell configuration with A9C applied to the cell (extracellular application) via a computer-controlled fast perfusion system ("concentration-jump" experiments) (Figure 5.11a) (Ta *et al.*, 2016). The concertation-jump protocol is described in detailed in Ta *et al.*, 2016. Briefly, the method enables the quantification of (i) the fast A9C block by assessing the current amplitude at the start of the A9C application (I_b) (ii) the slower allosteric activation produced by A9C on the channel. This allosteric activation is measured as a large current increase immediately after the A9C is removed (I_a). This method has enabled the study of the effect of the single point mutation at the putative A9C binding site on both channel block and activation. Specifically, the extent of block is expressed as I_b/I_0, where I_0 is the current amplitude in the absence of A9C and activation was expressed as I_a/I_0 (Figure 5.11b)

To validate the putative A9C binding site obtained from my simulation studies, we mutated residues at the binding site to alanine and assessed changes in channel activation and inhibition by A9C (Figure 5.11c). We showed a significant abolishment of both blockage (I_b/I_0) and activation by A9C for all mutations at the A9C binding site (I_a/I_0). These results are consistent with the proposition that these residues may form part of the A9C binding site.

5.4 Discussion

An open state model of the TMEM16A channel was generated using the open state of the TMEM16K scramblase and the TMEM16A closed state channel as templates. This model allowed the path of Cl^- permeation to be predicted using MD simulation. My finding highlights novel and unanticipated conformational rearrangement at the outer pore which affects pharmacological property of the channel. I also observed key conformational changes underlying the transition from the open state to the closed state of the channel as a result of the simulation.

The generation of the novel open state model allowed us to investigate the first putative binding site for A9C, a biphasic drug (Ta *et al.*, 2016). The prediction obtained from MD simulations agrees well with experimental data from our collaborator. This allows us to validate the binding site of the A9C, and hence, potentially postulate an inhibitory and activatory mechanism for A9C on the TMEM16A channel.

Gating of the TMEM16 family protein

The current gating model of the TMEM16A channel involves two gates: a steric gate at the lower part of TM6 and an 'electrostatic gate' that is open when Ca^{2+} is present in its binding site, as described in the introduction of this chapter (Lam & Dutzler, 2018; Peters *et al.*, 2018). My work postulates three

roles of the Ca^{2+} on TMEM16A channel; an opening of the steric gate, an opening of the electrostatic gate and the dilation of the outer pore (Figure 5.12).

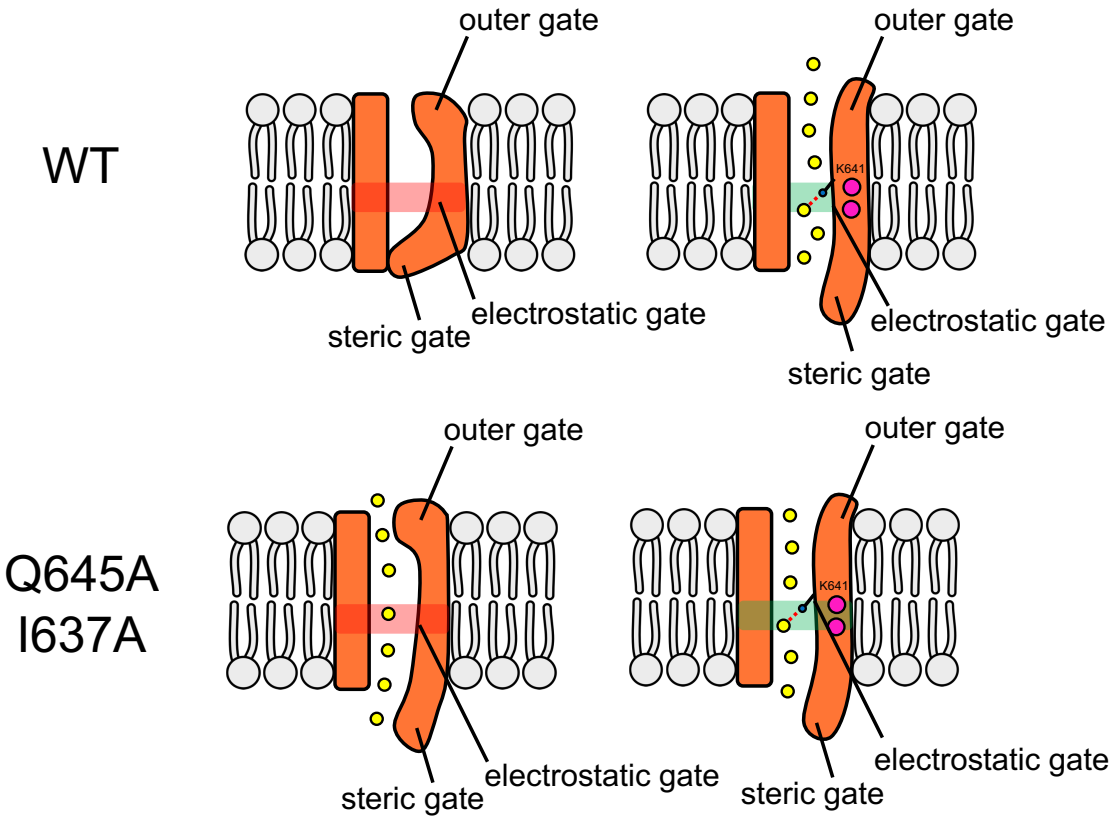


Figure 5.12 Schematic representation of TMEM16A and mutant channels activation by Ca^{2+}

Left – Wild-type TMEM16A channel (top) and mutant TMEM16A channel (bottom) are shown in orange. The steric gate of the WT channel is closed in an absence of calcium. The outer pore of both WT and mutant channels are not dilated in the absence of calcium. The electrostatic gate is closed in both WT and mutant channels (red highlight). Cl^- are shown as yellow circle. Mutant TMEM16A channels has the steric gate opened in an absence of calcium.

Right – Ca^{2+} activates TMEM16A channels by opening both steric gate and dilating the outer pore. Ca^{2+} are shown as pink circle. The electrostatic gate is opened (green highlight). Cl^- are shown as yellow circle. Residue K641 is shown with the blue sphere to show its co-ordination with chloride ion.

The work in this chapter provides evidence of a novel gate constituted by the hydrophobic residues between TM4 and TM6, in particular V539 and I636. Other hydrophobic residues on TM4 (such as I546 and I547) and TM6 (such as I637) contribute to this gate. Following my studies, but prior to their publication, several papers were published showing the same findings (Paulino *et al.*, 2017*a*; Jia & Chen, 2021; Lam *et al.*, 2021). I showed that Ca^{2+} binding to the channel led to dilation of the outer pore at residues V539 and I636. This mechanism is similar to that seen in TMEM16 scramblases such as *Nectria haematococca* TMEM16 (nhTMEM16) (Kalienkova *et al.*, 2019). Lipid transport in nhTMEM16 requires the opening of the outer region of the permeation pathway involving residues homologous to V539 and I636 in TMEM16A (nhTMEM16 T333 and Y439) (Khelashvili *et al.*, 2019). Our data suggest that TMEM16 channels and scramblases may share a common Ca^{2+} dependent mechanism that widens the outer pore. The notion that a single point pore mutation confers scramblase activity on TMEM16A is consistent with the idea that the lipid and ion permeation pathways in TMEM16x proteins may share a similar overall geometry.

A9C interaction at the pore

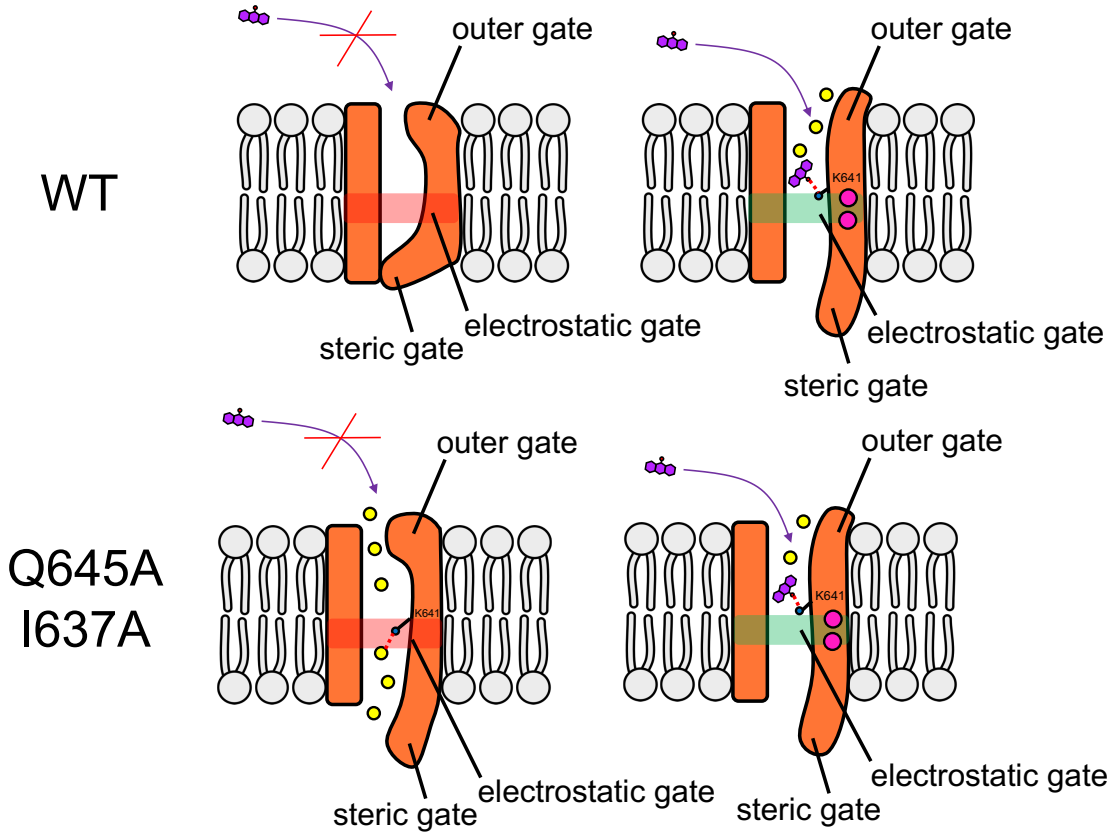


Figure 5.13 Schematic representation of A9C action in WT and mutant channels

Left – Wild-type TMEM16A channel (top) and mutant TMEM16A channel (bottom) are shown in orange. The steric gate of the WT channel is closed in an absence of calcium. The outer pore of both WT and mutant channels are not dilated in the absence of calcium. The electrostatic gate is closed in both WT and mutant channels (red highlight). Cl^- are shown as yellow circle. Mutant TMEM16A channels has the steric gate opened in an absence of calcium. A9C is shown in purple. The outer pore is too narrow for A9C to access its binding site. Residue K641 is shown with the blue sphere to show its co-ordination with chloride ion.

Right – Ca^{2+} activates TMEM16A channels by opening both steric gate and dilating the outer pore. Ca^{2+} are shown as pink circle. The electrostatic gate is opened (green highlight). Cl^- are shown as yellow circle. The outer pore is wide enough for A9C to access its binding site. Residue K641 is shown with the blue sphere to show its co-ordination with chloride ion. The bottom panel shows how K641 occluded the chloride binding site as A9C competes for the K641 sidechain.

When we consider the permeation of Cl^-, we only have to concern about the electrostatic gate and the steric gate. However, this is not the case for A9C. For A9C to access the binding site, the outer pore is needed to be dilated by calcium to allow its access to the binding site. My simulation showed that A9C penetrates the channel when the outer pore is dilated by Ca^{2+} (Figure 5.13) and reaches a region composed of five amino acids (S513, V539, I636, K641, and Q645). While separating the effects of mutations on gating and binding poses challenges, the combined use of electrophysiology and MD simulations suggest that this region is a *bona fide* binding site for A9C.

Further postulations

One of the caveats underlying our model is that our mutant open state model (I637A, Q645A) fails to be Cl^- conductive when Ca^{2+} is absent. Our model predicted that the outer gate is closed in the absence of Ca^{2+}. But then, if the outer gate is closed, how does the Cl^- permeate in the absence of Ca^{2+}? I postulated that the outer gate might be opened to a different extent (Figure 5.10). This means that the mutant channel may have a wide enough cavity for Cl^- to permeate in the absence of Ca^{2+} but not wide enough for A9C to gain access to its binding site. The recent structure of the I551A mutant shows that even with this gating mutation, the outer gate of the TMEM16A channel is closed (Lam & Dutzler, 2021; Lam *et al.*, 2021). A previous independent simulation study has shown that the outer gate of the channel is only open when both PIP_2 and Ca^{2+} are present. I attempted simulation with PIP_2. However, I also failed to capture the intermediate state Cl^- conductive of the channel in the

absence of Ca^{2+}. I have also attempted (in collaboration with Dr Emillio Agostinelli) to use a voltage-sensitive phosphatase (DrVSP) to remove PIP_2 from the membrane. However, we failed to abolish the Cl^- current when DrVSP was activated. Our observation, therefore, suggested that other factors may aid the opening of the TMEM16A channel, in addition to PIP_2. These factors may need to be included in the simulation and/or the cryo-EM structure to maintain the conductive Cl^- state of the mutant channel in the absence of Ca^{2+}.

The second postulation is a putative mechanism for inhibition by A9C. My work shows that A9C is negatively charge and shares a binding site with the putative Cl^- binding site at K641 (Paulino *et al.*, 2017*a*, 2017*b*). The A9C binding site is only accessible when Ca^{2+} is present, where the outer pore is more dilate. I proposed that A9C blocks the Cl^- ion permeation as it competes for the putative Cl^- binding site. Thus, it acts almost like a “cork in a bottle” and blocks the chloride efflux in the steady-state current. This action is similar to other Cl^- channel blockers, such as Ani9 (Seo *et al.*, 2016).

As A9C is washed out of the binding site, we observed a net activatory effect of the channel. One possible hypothesis to this is that A9C stabilises the open state and extend the burst time of the channel. As A9C is removed, the channel needs a short period of time to transits from open to close. This transition event disrupts the fast gated (flickering) of the channel. Together, this shift the net effect between opening and closing and yield a total activatory effect.

Conclusion

To conclude, I propose a Cl^- permeation pathway for the TMEM16A channel and postulate potential mechanisms for activation and inhibition by A9C. Using long atomistic MD simulation, I successfully demonstrated a transition of the channel from an open to a closed state when Ca^{2+} is removed. I have also identified an outer gate to the TMEM16A channel, which is highly conserved amongst other TMEM16 family proteins. The outer gate forms a binding site for the A9C drug. An identification of the A9C binding site aids further therapeutics drug design to the TMEM16A channel, which is currently the clinical and pharmacological interest. Together, the work in this chapter has demonstrated the potential of MD simulations, in conjunction with the pharmacological and electrophysiological studies, to determine the molecular details of the gating mechanisms of the TMEM16A channel.

Chapter 6

Discussion and conclusion

“No good sittin’ worryin’ abou’ it. What’s comin’ will come, an’ we’ll meet it when it does.”

Rubeus Hagrid – Chapter 37 Harry Potter and the Goblet of Fire

In this thesis, I used MD simulations to understand the gating process of two ion channels, K_{ATP} and TMEM16A. This chapter aims to summarise key findings in each chapter and address questions asked in the aim of this thesis.

6.1 Coarse-grained studies of protein-lipid interactions

There are now structures of Kir2.2, Kir3.2 and K_{ATP} channels (Hansen *et al.*, 2011; Whorton & MacKinnon, 2011; Lee *et al.*, 2017; Martin *et al.*, 2017*a*). Previously, the PIP_2 binding site on the K_{ATP} channel was predicted using homology modelling based on the Kir3.1-KirBac1.3 chimaera (Stansfeld *et al.*, 2009). Eight years later, the structure of the K_{ATP} channel with ATP and glibenclamide was solved (Li *et al.*, 2017; Martin *et al.*, 2017*a*). Using the PIP_2 binding geometry of chicken Kir2.2 as a starting point, we used CG-MD to predict the PIP_2 binding site on the K_{ATP} channel (Hansen *et al.*, 2011). My predicted PIP_2 binding site on the K_{ATP} channel agrees with the binding site predicted based on the unbiased CG-MD simulation in the previous study. Mutations at the predicted PIP_2 binding site on the K_{ATP} channel resulted in a reduction in channel neomycin sensitivity and the size of the current after the excision patch (Shyng *et al.*, 2000; Haider *et al.*, 2007).

In this thesis, I generated a structure of Kir1.1 based on chicken Kir2.2. Given the rapid improvement in structural prediction tools such as AlphaFold2 (Jumper *et al.*, 2021), we could use a different method to build the homology model for other Kir channels such as Kir1, Kir4 and Kir5.

To address the question regarding the binding free energy of PIP_2 to Kir channels, I have developed a “coarse-grained free energy perturbation” method to compare binding free energies between different lipid species. I also extended the method to calculate the influence of mutations in the K_{ATP} channel on the PIP_2 binding free energy. This approach has also been further used for in *silico* alanine scanning to assess the contribution of each residue in the cardiolipin binding site of all *E. coli* membrane protein structures (Corey *et al.*, 2021).

A second, alternative, approach to the equilibrium free energy calculation is to use non-equilibrium FEP. This method has been attempted multiple times with several preliminary data (Dr Mariana Bunoro, Dr Robin Corey, *personal communications*). Their further development of the method uses a reduced computational cost, with the results obtained from their method agreeing well with my approach. The reduced cost of the method means that this approach could be applied widely to a range of proteins, for example, quantitative assessments of all putative PIP_2 binding proteins in the database.

The CG-FEP approach developed in chapter 3 has shown that the PIP_2 binding affinity is correlated with the number of the basic residues found in the binding site. These electrostatic interactions are easily captured in CG simulations at a lower cost.

Despite several caveats within the CG-FEP methodology, the method works well with the K_{ATP} channel. The predictions from our simulations agree well with the experimental data on several points, as follow:

a) The simulations predicted that SUR1 does not affect the K_{ATP} channel affinity for PIP_2. This agrees with the experimental data, which shows that SUR1 does not affect channel sensitivity to neomycin.

b) The simulations predicted that PI4P, PI and PC bind to the K_{ATP} channel less than PIP_2. The lower free energy calculated for these phosphoinositides agrees with the previous literature, which shows that a higher concentration of PI4P and PI are required for the K_{ATP} channel activation.

c) Our simulations predicted that E179K increases the K_{ATP} channel affinity to PIP_2. The calculated binding free energy agrees with our experimental data, which show that E179K increases the open probability of both the K_{ATP} channel and the ΔC construct of the Kir6.2 channel. The mutation also decreases the K_{ATP} and the Kir6.2ΔC channel sensitivity to neomycin.

6.2 Gating process of the K_{ATP} channel

After the CG simulation, I refined the PIP_2 binding site on the Kir6.2 channel using atomistic simulation. The SUR1 subunits were excluded to reduce the computational cost. We also refined the ATP-binding site on Kir6.2. We were then interested in the interplay between ATP and PIP_2 in the binding site. The binding site predicted by the atomistic simulations agrees well with the binding site predicted from the CG simulation as expected.

The limitation of the work in chapter 4 is the relatively short simulation time (300 ns). Because of our short simulations, we can only investigate competitions between ATP and PIP_2 on the surface of Kir6.2. We cannot see how the binding of ATP and/or PIP_2 induce conformational change at the cytoplasmic or the bundle crossing gate. The difficulty in tracking the conformational change at the bundle crossing gate was also rectified in the recent computational study by the other group (Bründl *et al.*, 2021). However, it was clear from our simulation that K39 swings upward to co-ordinate with PIP_2. Our simulation predicted that residue K39 would prefer to co-ordinate to PIP_2 when both ATP and PIP_2 are present. This provides additional information to our previous understanding of the competition between ATP and PIP_2. Thus, a K39 mutation that enhances PIP_2 binding to the K_{ATP} channel should decrease ATP binding to the K_{ATP} channel. As a consequence, the mutation would be expected to also reduce channel inhibition by ATP.

A criticism of the work in chapter 4 is that none of our simulations included SUR1. Previous studies have shown that SUR1 is critical in the gating

process of the K_{ATP} channel. Our experimental work was also conducted in the presence of SUR1. Simulations that include SUR1 should be conducted when the computational power is sufficient.

6.3 Gating process of the TMEM16A channel.

My work also addressed the gating of TMEM16A, Ca^{2+} activated Cl^- channel. I have identified an "outer gate" at the top of the Cl^- permeation path using unbiased atomistic MD simulation. The opening of this outer gate requires Ca^{2+} binding at the lower part of TM6 of the channel. Using MD simulation, we showed that mutations in the lower portion of TM6 (I637A, Q645A) stabilise the opening of the steric (lower) gate of the channel. My simulations are consistent with whole cell patch-clamp recordings demonstrating that these mutations result in channels open in the absence of Ca^{2+} (Peters *et al.*, 2018). Opening of the outer gate allows A9C to access its binding site. My simulations identified an A9C binding site in the centre of the pore. This binding site was further validated by single point mutagenesis and electrophysiology experiments by the Tammaro group.

In chapter 5, I also asked how A9C blocks the pore. I showed that A9C competes with Cl^- for K641, effectively preventing passage of Cl^- through the pore (i.e. "cork in a bottle" mechanism). As noted in chapter 5, A9C has a complex, biphasic action on the TMEM16A channel. The complex biphasic action of A9C is the simultaneous open channel block, and also an allosteric activation of the channel. The mechanism of activation was not fully elucidated by my study. However, I hypothesised that the binding of the A9C to the pore

stabilises the open state of the channel. Indeed, as the binding of the ligand is favourable (releasing Gibb's free energy), the free energy release from ligand binding may stabilise the open state conformation of TMEM16A channel. This proposition might be tested computationally by examining the stabilisation of the open state when A9C is removed, and the energy landscape of conformational transition of TMEM16A from open to closed using umbrella sampling, or well-tempered metadynamics simulation.

The key limitation in this study is the lack of the structure of the open state TMEM16A channel. However, our open state model, built based on TMEM16K lipid scramblase, shows conductivity to Cl^-, water permeation and no lipid scramblase activity. Thus, this makes a very interesting point in the evolutionary origin of the TMEM16 family protein, which may function as a Cl- channel and/or lipid scramblase. A recent study also shows that PC mediated the transport of Cl^- on the nhTMEM16F protein (Kostritskii & Machtens, 2021). More detailed review on the evolution of TMEM16A were discussed in chapter 5 and in the very recent review (Kalienkova *et al.*, 2021). Our open state channel also closely resembles the structures generated from the simulations done by another research group. It will be interesting to see whether our model resembles an open state structure solved in the future.

6.4 Conclusions and future directions of MD simulation

So, what's next in the field of molecular dynamics simulation and ion channels? My interest focuses on ligand binding and the associated conformational change of ion channels. The recent development of AlphaFold2 in 2021 allows many structures to be predicted. However, these structures are only predicted as static structures in a single conformation. The next step for model development is to use machine learning to understand the concept behinds protein conformational changes. These concepts can then be used to build a potential platform to predict multiple conformations of a given protein's structure.

Obtaining protein structures in multiple conformations still does not explain the path of protein conformational change from one state to another. Using MD simulation to access multiple conformations of the protein is currently our technical challenge in the field (Figure 6.1). Several machine learning and sampling methods have been performed to cluster multiple protein conformations, assessing their path of conformational change along the energy landscape. Understanding the path and energy landscape underlying protein conformational change is crucial to identify amino acid residues involved in the process. Locating these residues and linking them to their functions provides a better understanding of ion channels gating, enzyme mechanisms and their relationship with multiple disease-causing mutations for a more efficient pharmacological targeting.

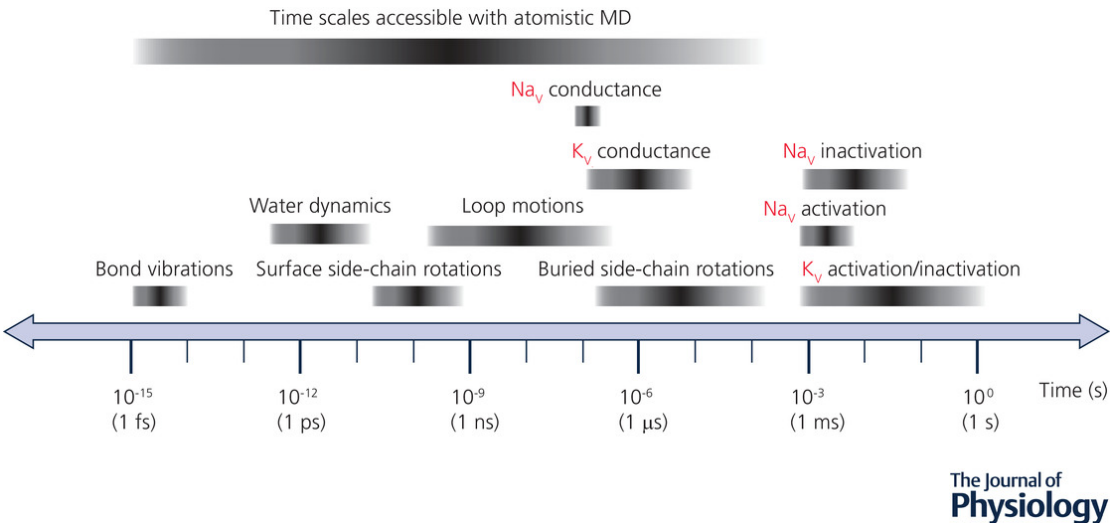


Figure 6.1 Timescale accessible with MD simulation related to ion channels function

Timescale accessible using MD simulation ranges from bond vibrations (fs) to conformational changes and side-chain rotations (sub µs timescales).

Taken from DeMarco *et al.*, 2019

Recent studies used unbiased CG-MD simulation to identify multiple lipid-binding sites (Duncan *et al.*, 2020; Ansell *et al.*, 2021). The method also quantifies the strength of the protein-lipid interaction. The simplicity and the usability of the method are ideal for integration into the cryo-EM software pipeline. The software may potentially identify lipid-binding site when the density is unclear, almost straight after structural assignment of the cryo-EM density map. This method would facilitate the identification of multiple unclear potential lipid densities on the cryo-EM electron density map. For a single lipid-binding site, coupling this method with the CG-FEP approach in this thesis would allow a rapid assessment of the lipid-binding affinity within the binding site.

To conclude, my thesis has used a range of methods in molecular dynamics simulation to study ligand binding and the associated conformational changes in K_{ATP} and TMEM16A channels. We highlighted the role of the mutations on these channels and their association with the genetic disease using MD simulation. Integrating knowledge from the clinic with MD simulations may provide a broader horizon to study channelopathies in the future.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B & Lindah E (2015). Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2,** 19–25.

Aguilar-Bryan L et al. (1995). Cloning of the beta cell high-affinity sulfonylurea receptor: a regulator of insulin secretion. *Neuron* **18,** 423 LP – 426.

Alberts B (2015). *Molecular biology of the cell*.

Aldeghi M, Bluck JP & Biggin PC (2018*a*). Absolute Alchemical Free Energy Calculations for Ligand Binding: A Beginner's Guide. *Methods Mol Biol* **1762,** 199.

Aldeghi M, Ross GA, Bodkin MJ, Essex JW, Knapp S & Biggin PC (2018*b*). Large-scale analysis of water stability in bromodomain binding pockets with grand canonical Monte Carlo. *Commun Chem* **1,** 19.

Ansell TB, Curran L, Horrell MR, Pipatpolkai T, Letham SC, Song W, Siebold C, Stansfeld PJ, Sansom MSP & Corey RA (2021). Relative Affinities of Protein-Cholesterol Interactions from Equilibrium Molecular Dynamics Simulations. *bioRxiv*2021.06.02.446704.

Arnarez C, Marrink SJ & Periole X (2013*a*). Identification of Cardiolipin Binding Sites on Cytochrome c Oxidase at the Entrance of Proton Channels. *Sci Rep* **3,** 1263.

Arnarez C, Mazat J-P, Elezgaray J, Marrink S-J & Periole X (2013*b*). Evidence for Cardiolipin Binding Sites on the Membrane-Exposed Surface of the Cytochrome Bc_1. *J Am Chem Soc* **135,** 3112.

Ashcroft FM (2005). ATP-sensitive potassium channelopathies: Focus on

insulin secretion. *J Clin Invest* **115,** 2047–2058.

Ashcroft FM (2007). ATP-sensitive K^+ channels and disease: from molecule to malady. *Am J Physiol Metab* **293,** E880–E889.

Ashcroft FM, Harrison DE & Ashcroft SJH (1984). Glucose induces closure of single potassium channels in isolated rat pancreatic β-cells. *Nature* **312,** 446–448.

Ashcroft FM & Rorsman P (2012). Diabetes mellitus and the β cell: The last ten years. *Cell* **148,** 1160–1171.

Ashcroft SJH & Ashcroft FM (1992). The sulfonylurea receptor. *BBA - Mol Cell Res* **1175,** 45–59.

Bahar I, Atilgan AR & Erman B (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. *Fold Des* **2,** 173–181.

Barish ME (1983). A transient calcium-dependent chloride current in the immature Xenopus oocyte. *J Physiol* **342,** 309–325.

Baukrowitz T, Schulte U, Oliver D, Herlitze S, Krauter T, Tucker SJ, Ruppersberg JP & Fakler B (1998). PIP_2 and PIP as determinants for ATP inhibition of K_{ATP} channels. *Science (80)* **282,** 1141–1144.

Beckstein O, Denning EJ, Perilla JR & Woolf TB (2009). Zipping and Unzipping of Adenylate Kinase: Atomistic Insights into the Ensemble of Open↔Closed Transitions. *J Mol Biol* **394,** 160–176.

Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A & Haak JR (1984). Molecular dynamics with coupling to an external bath. *J Chem Phys* **81,** 3684–3690.

Bernsteiner H, Zangerl-Plessl E-M, Chen X & Stary-Weinzinger A (2019).

Conduction through a narrow inward-rectifier K+ channel pore. *J Gen Physiol* **151,** 1231–1246.

Beutler TC, Mark AE, van Schaik RC, Gerber PR & Van Gunsteren WF (1994). Avoiding Singularities and Numerical Instabilities in Free Energy Calculations Based on Molecular Simulations. *Chem Phys Lett* **222,** 529.

Boesze-Battaglia K & Schimmel R (1997). Cell membrane lipid composition and distribution: implications for cell function and lessons learned from photoreceptors and platelets. *J Exp Biol* **200,** 2927–2936.

Bond PJ & Sansom MSP (2007). Bilayer deformation by the Kv channel voltage sensor domain revealed by self-assembly simulations. *Proc Natl Acad Sci U S A* **104,** 2631–2636.

Bonomi M, Branduardi D, Bussi G, Camilloni C, Provasi D, Raiteri P, Donadio D, Marinelli F, Pietrucci F, Broglia RA & Parrinello M (2009). PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics. *Comput Phys Commun* **180,** 1961.

Bowie JU (2011). Membrane protein folding: how important are hydrogen bonds? *Curr Opin Struct Biol* **21,** 42–49.

Bründl M, Pellikan S & Stary-Weinzinger A (2021). Simulating PIP(2)-Induced Gating Transitions in Kir6.2 Channels. *Front Mol Biosci* **8,** 711975.

Brunner JD, Lim NK, Schenck S, Duerst A & Dutzler R (2014). X-ray structure of a calcium-activated TMEM16 lipid scramblase. *Nature* **516,** 207–212.

Bushell SR et al. (2019). The structural basis of lipid scrambling and inactivation in the endoplasmic reticulum scramblase TMEM16K. *Nat Commun* **10,** 3956.

Bussi G, Donadio D & Parrinello M (2007). Canonical sampling through velocity

rescaling. *J Chem Phys* **126,** 14101.

Casares D, Escribá P V & Rosselló CA (2019). Membrane Lipid Composition: Effect on Membrane and Organelle Structure, Function and Compartmentalization and Therapeutic Avenues. *Int J Mol Sci* **20,** 2167.

Chavent M, Duncan AL & Sansom MS (2016). Molecular dynamics simulations of membrane proteins and their interactions: from nanoscale to mesoscale. *Curr Opin Struct Biol* **40,** 8–16.

Chiu SW, Jakobsson E & Scott HL (2001). Combined Monte Carlo and molecular dynamics simulation of hydrated dipalmitoyl–phosphatidylcholine–cholesterol lipid bilayers. *J Chem Phys* **114,** 5435–5443.

Chodera JD, Mobley DL, Shirts MR, Dixon RW, Branson K & Pande VS (2011). Alchemical free energy methods for drug discovery: progress and challenges. *Curr Opin Struct Biol* **21,** 150–160.

Clarke OB, Caputo AT, Hill AP, Vandenberg JI, Smith BJ & Gulbis JM (2010). Domain Reorientation and Rotation of an Intracellular Assembly Regulate Conduction in Kir Potassium Channels. *Cell* **141,** 1018–1029.

Corey RA, Song W, Duncan AL, Ansell TB, Sansom MSP & Stansfeld PJ (2021). Identification and assessment of cardiolipin interactions with E. coli inner membrane proteins. *bioRxiv*2021.03.19.436130.

Corey RA, Stansfeld PJ & Sansom MSP (2019*a*). The energetics of protein–lipid interactions as viewed by molecular simulations. *Biochem Soc Trans* **48,** 25–37.

Corey RA, Vickery ON, Sansom MSP & Stansfeld PJ (2019*b*). Insights into

Membrane Protein-Lipid Interactions from Free Energy Calculations. *J Chem Theory Comput* **15,** 5727–5736.

Corringer P-J, Bertrand S, Galzi J-L, Devillers-Thiéry A, Changeux J-P & Bertrand D (1999). Mutational Analysis of the Charge Selectivity Filter of the α7 Nicotinic Acetylcholine Receptor. *Neuron* **22,** 831–843.

Cukras CA, Jeliazkova I & Nichols CG (2002). The role of NH_2-terminal positive charges in the activity of inward rectifier K_{ATP} channels. *J Gen Physiol* **120,** 437–446.

Dabrowski M, Tarasov A & Ashcroft FM (2004). Mapping the architecture of the ATP-binding site of the KATPchannel subunit Kir6.2. *J Physiol* **557,** 347–354.

Danahay HL, Lilley S, Fox R, Charlton H, Sabater J, Button B, McCarthy C, Collingwood SP & Gosling M (2020). TMEM16A Potentiation: A Novel Therapeutic Approach for the Treatment of Cystic Fibrosis. *Am J Respir Crit Care Med* **201,** 946–954.

Dang S, Feng S, Tien J, Peters CJ, Bulkley D, Lolicato M, Zhao J, Zuberbühler K, Ye W, Qi L, Chen T, Craik CS, Jan YN, Minor DL, Cheng Y & Jan LY (2017). Cryo-EM structures of the TMEM16A calcium-activated chloride channel. *Nature* **552,** 426–429.

DeMarco KR, Bekker S & Vorobyov I (2019). Challenges and advances in atomistic simulations of potassium and sodium ion channel gating and permeation. *J Physiol* **597,** 679–698.

Dickson EJ & Hille B (2019). Understanding phosphoinositides: rare, dynamic, and essential membrane phospholipids. *Biochem J* **476,** 1–23.

Ding D, Wang M, Wu J-X, Kang Y & Chen L (2019). The Structural Basis for

the Binding of Repaglinide to the Pancreatic K_{ATP} Channel. *Cell Rep* **27,** 1848-1857.e4.

Dixit SB & Chipot C (2001). Can Absolute Free Energies of Association Be Estimated from Molecular Mechanical Simulations? The Biotin−Streptavidin System Revisited. *J Phys Chem A* **105,** 9795–9799.

Domański J, Hedger G, Best RB, Stansfeld PJ & Sansom MSP (2017). Convergence and Sampling in Determining Free Energy Landscapes for Membrane Protein Association. *J Phys Chem B* **121,** 3364–3375.

Doyle DA, Cabral JM, Pfuetzner RA, Kuo A, Gulbis JM, Cohen SL, Chait BT & MacKinnon R (1998). The Structure of the Potassium Channel: Molecular Basis of K^+ Conduction and Selectivity. *Science (80)* **280,** 69 LP – 77.

Duncan AL, Corey RA & Sansom MSP (2020). Defining how multiple lipid species interact with inward rectifier potassium (Kir2) channels. *Proc Natl Acad Sci* **117,** 7803 LP – 7813.

Duncan AL, Song W & Sansom MSP (2019). Lipid-Dependent Regulation of Ion Channels and G Protein–Coupled Receptors: Insights from Structures and Simulations. *Annu Rev Pharmacol Toxicol*; DOI: 10.1146/annurev-pharmtox-010919-023411.

Dutzler R, Campbell EB, Cadene M, Chait BT & MacKinnon R (2002). X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. *Nature* **415,** 287–294.

Enkvetchakul D, Jeliazkova I & Nichols CG (2005). Direct modulation of Kir channel gating by membrane phosphatidylinositol 4,5-bisphosphate. *J Biol Chem* **280,** 35785–35788.

Enkvetchakul D, Loussouarn G, Makhina E, Shyng SL & Nichols CG (2000).

The kinetic and physical basis of K(ATP) channel gating: toward a unified molecular understanding. *Biophys J* **78,** 2334–2348.

Fajer M, Swift R V & McCammon JA (2009). Using multistate free energy techniques to improve the efficiency of replica exchange accelerated molecular dynamics. *J Comput Chem* **30,** 1719–1725.

Falzone ME, Rheinberger J, Lee B-C, Peyear T, Sasset L, Raczkowski AM, Eng ET, Di Lorenzo A, Andersen OS, Nimigean CM & Accardi A (2019). Structural basis of Ca^{2+}-dependent activation and lipid transport by a TMEM16 scramblase ed. Csanády L & Aldrich R. *Elife* **8,** e43229.

Fan Z & Makielski JC (1997). Anionic Phospholipids Activate ATP-sensitive Potassium Channels. *J Biol Chem* **272,** 5388–5395.

Fan Z & Makielski JC (1999). Phosphoinositides decrease ATP sensitivity of the cardiac ATP-sensitive K(+) channel. A molecular probe for the mechanism of ATP-sensitive inhibition. *J Gen Physiol* **114,** 251–269.

Flanagan SE, Patch A-M, Mackay DJG, Edghill EL, Gloyn AL, Robinson D, Shield JPH, Temple K, Ellard S & Hattersley AT (2007). Mutations in ATP-sensitive K^+ channel genes cause transient neonatal diabetes and permanent diabetes in childhood or adulthood. *Diabetes* **56,** 1930–1937.

De Franco E et al. (2020). Update of variants identified in the pancreatic β-cell KATP channel genes KCNJ11 and ABCC8 in individuals with congenital hyperinsulinism and diabetes. *Hum Mutat* **41,** 884–905.

Fraser CS, Rubio-Cabezas O, Littlechild JA, Ellard S, Hattersley AT & Flanagan SE (2012). Amino acid properties may be useful in predicting clinical outcome in patients with Kir6.2 neonatal diabetes. *Eur J Endocrinol* **167,** 417–421.

Friedman ZY (1993). Tamoxifen and vanadate synergize in causing accumulation of polyphosphoinositides in GH4C1 membranes. *J Pharmacol Exp Ther* **267,** 617–623.

Gapsys V & de Groot BL (2017). pmx Webserver: A User Friendly Interface for Alchemistry. *J Chem Inf Model* **57,** 109–114.

Gapsys V, Michielssens S, Seeliger D & de Groot BL (2015). pmx: Automated protein structure and topology generation for alchemical perturbations. *J Comput Chem* **36,** 348–354.

Gapsys V, Yildirim A, Aldeghi M, Khalak Y, van der Spoel D & de Groot BL (2021). Accurate absolute free energies for ligand–protein binding based on non-equilibrium approaches. *Commun Chem* **4,** 61.

Gloyn AL, Reimann F, Girard C, Edghill EL, Proks P, Pearson ER, Temple IK, Mackay DJG, Shield JPH, Freedenberg D, Noyes K, Ellard S, Ashcroft FM, Gribble FM & Hattersley AT (2005). Relapsing diabetes can result from moderately activating mutations in KCNJ11. *Hum Mol Genet* **14,** 925–934.

Gowers RJ, Linke M, Barnoud J, Reddy TJE, Melo MN, Seyler SL, Domański J, Dotson DL, Buchoux S, Kenney IM & Beckstein O (2016). MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations. *Proc 15th Python Sci Conf*98–105.

Groman JD, Meyer ME, Wilmott RW, Zeitlin PL & Cutting GR (2002). Variant cystic fibrosis phenotypes in the absence of CFTR mutations. *N Engl J Med* **347,** 401–407.

Haider S, Antcliff JF, Proks P, Sansom MSP & Ashcroft FM (2005). Focus on Kir6.2: A key component of the ATP-sensitive potassium channel. *J Mol Cell Cardiol* **38,** 927–936.

Haider S, Tarasov AI, Craig TJ, Sansom MSP & Ashcroft FM (2007). Identification of the PIP2-binding site on Kir6.2 by molecular modelling and functional analysis. *EMBO J* **26,** 3749–3759.

Hannesschlaeger C, Horner A & Pohl P (2019). Intrinsic Membrane Permeability to Small Molecules. *Chem Rev* **119,** 5922–5953.

Hansen SB, Tao X & MacKinnon R (2011). Structural basis of PIP2 activation of the classical inward rectifier K^+ channel Kir2.2. *Nature* **477,** 495–498.

Harayama T & Riezman H (2018). Understanding the diversity of membrane lipid composition. *Nat Rev Mol Cell Biol* **19,** 281–296.

Hattersley AT & Ashcroft FM (2005). Activating mutations in Kir6.2 and neonatal diabetes: new clinical syndromes, new scientific insights, and new therapy. *Diabetes* **54,** 2503–2513.

Hedger G, Rouse SL, Domański J, Chavent M, Koldsø H & Sansom MSP (2016*a*). Lipid-Loving ANTs: Molecular Simulations of Cardiolipin Interactions and the Organization of the Adenine Nucleotide Translocase in Model Mitochondrial Membranes. *Biochemistry* **55,** 6238.

Hedger G & Sansom MSP (2016). Lipid interaction sites on channels, transporters and receptors: Recent insights from molecular dynamics simulations. *Biochim Biophys Acta* **1858,** 2390–2400.

Hedger G, Shorthouse D, Koldsø H & Sansom MSP (2016*b*). Free Energy Landscape of Lipid Interactions with Regulatory Binding Sites on the Transmembrane Domain of the EGF Receptor. *J Phys Chem B* **120,** 8154.

Hilgemann DW & Ball R (1996). Regulation of Cardiac Na^+/Ca^{2+} Exchange and K_{ATP} Potassium Channels by PIP_2 *Science (80)* **273,** 956 LP – 959.

Hille B (1978). Ionic channels in excitable membranes. Current problems and

biophysical approaches. *Biophys J* **22,** 283–294.

Hopkins CW, Le Grand S, Walker RC & Roitberg AE (2015). Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J Chem Theory Comput* **11,** 1864–1874.

Huang C-L, Feng S & Hilgemann DW (1998). Direct activation of inward rectifier potassium channels by PIP_2 and its stabilization by Gβγ. *Nature* **391,** 803–806.

Huang F, Rock JR, Harfe BD, Cheng T, Huang X, Jan YN & Jan LY (2009). Studies on expression and function of the TMEM16A calcium-activated chloride channel. *Proc Natl Acad Sci* **106,** 21413 LP – 21418.

Huang J & Mackerell AD (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem* **34,** 2135–2145.

Hub JS, De Groot BL & Van Der Spoel D (2010). g_wham-A Free Weighted Histogram Analysis Implementation Including Robust Error and Autocorrelation Estimates. *J Chem Theory Comput* **6,** 3713.

Humphrey W, Dalke A & Schulten K (1996). VMD: Visual molecular dynamics. *J Mol Graph* **14,** 33–38.

Huopio H, Jääskeläinen J, Komulainen J, Miettinen R, Kärkkäinen P, Laakso M, Tapanainen P, Voutilainen R & Otonkoski T (2002). Acute Insulin Response Tests for the Differential Diagnosis of Congenital Hyperinsulinism. *J Clin Endocrinol Metab* **87,** 4502–4507.

Inagaki N, Gonoi T, Clement IV JP, Namba N, Inazawa J, Gonzalez G, Aguilar-Bryan L, Seino S & Bryan J (1995). Reconstitution of IKATP: An inward rectifier subunit plus the sulfonylurea receptor. *Science (80)*; DOI:

10.1126/science.270.5239.1166.

Jentsch TJ & Pusch M (2018). CLC Chloride Channels and Transporters: Structure, Function, Physiology, and Disease. *Physiol Rev* **98,** 1493–1590.

Jia Z & Chen J (2021). Specific PIP2 binding promotes calcium activation of TMEM16A chloride channels. *Commun Biol* **4,** 259.

Jo S, Kim T, Iyer VG & Im W (2008). CHARMM-GUI: A web-based graphical user interface for CHARMM. *J Comput Chem* **29,** 1859–1865.

de Jong DH, Singh G, Bennett WFD, Arnarez C, Wassenaar TA, Schäfer L V, Periole X, Tieleman DP & Marrink SJ (2013). Improved Parameters for the Martini Coarse-Grained Protein Force Field. *J Chem Theory Comput* **9,** 687–697.

Jorgensen WL (1989). Free energy calculations: a breakthrough for modeling organic chemistry in solution. *Acc Chem Res* **22,** 184–189.

Jorgensen WL, Buckner JK, Boudon S & Tirado-Rives J (1988). Efficient computation of absolute free energies of binding by computer simulations. Application to the methane dimer in water. *J Chem Phys*; DOI: 10.1063/1.454895.

Jumper J et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*; DOI: 10.1038/s41586-021-03819-2.

Kalienkova V, Clerico Mosina V, Bryner L, Oostergetel GT, Dutzler R & Paulino C (2019). Stepwise activation mechanism of the scramblase nhTMEM16 revealed by cryo-EM ed. Aldrich R, Swartz KJ, Ballesteros A, Meyerson J & Long SB. *Elife* **8,** e44364.

Kalienkova V, Clerico Mosina V & Paulino C (2021). The Groovy TMEM16 Family: Molecular Mechanisms of Lipid Scrambling and Ion Conduction. *J*

Mol Biol **433,** 166941.

Kane Dickson V, Pedi L & Long SB (2014). Structure and insights into the function of a Ca(2+)-activated Cl(-) channel. *Nature* **516,** 213–218.

Kastenholz MA & Hünenberger PH (2006). Computation of methodology-independent ionic solvation free energies from molecular simulations. I. The electrostatic potential in molecular liquids. *J Chem Phys* **124,** 124106.

Khelashvili G, Falzone ME, Cheng X, Lee B-C, Accardi A & Weinstein H (2019). Dynamic modulation of the lipid translocation groove generates a conductive ion channel in Ca2+-bound nhTMEM16. *Nat Commun* **10,** 4972.

Kim S, Lee J, Jo S, Brooks III CL, Lee HS & Im W (2017). CHARMM-GUI ligand reader and modeler for CHARMM force field generation of small molecules. *J Comput Chem* **38,** 1879–1886.

Klesse G, Rao S, Sansom MSP & Tucker SJ (2019). CHAP: A Versatile Tool for the Structural and Functional Annotation of Ion Channel Pores. *J Mol Biol* **431,** 3353–3365.

Klesse G, Rao S, Tucker SJ & Sansom MSP (2020). Induced Polarization in Molecular Dynamics Simulations of the 5-HT3 Receptor Channel. *J Am Chem Soc* **142,** 9415–9427.

Klimovich P V, Shirts MR & Mobley DL (2015). Guidelines for the analysis of free energy calculations. *J Comput Aided Mol Des* **29,** 397–411.

König G, Brooks BR, Thiel W & York DM (2018). On the convergence of multi-scale free energy simulations. *Mol Simul* **44,** 1062–1081.

Kostritskii AY & Machtens J-P (2021). Molecular mechanisms of ion conduction and ion selectivity in TMEM16 lipid scramblases. *Nat Commun* **12,** 2826.

Kramer W, Oekonomopulos R, Pünter J & Summ HD (1988). Direct photoaffinity labeling of the putative sulfonylurea receptor in rat beta-cell tumor membranes by [3H]glibenclamide. *FEBS Lett* **229,** 355–359.

Kunzelmann K, Ousingsawat J, Cabrita I, Doušová T, Bähr A, Janda M, Schreiber R & Benedetto R (2019). TMEM16A in Cystic Fibrosis: Activating or Inhibiting? . *Front Pharmacol* **10,** 3.

Lacin E, Aryal P, Glaaser IW, Bodhinathan K, Tsai E, Marsh N, Tucker SJ, Sansom MSP & Slesinger PA (2017). Dynamic role of the tether helix in PIP_2-dependent gating of a G protein–gated potassium channel. *J Gen Physiol* **149,** 799–811.

Lam AKM & Dutzler R (2018). Calcium-dependent electrostatic control of anion access to the pore of the calcium-activated chloride channel TMEM16A ed. Chanda B & Aldrich R. *Elife* **7,** e39122.

Lam AKM & Dutzler R (2021). Mechanism of pore opening in the calcium-activated chloride channel TMEM16A. *Nat Commun* **12,** 786.

Lam AKM, Rheinberger J, Paulino C & Dutzler R (2021). Gating the pore of the calcium-activated chloride channel TMEM16A. *Nat Commun* **12,** 785.

Le SC, Jia Z, Chen J & Yang H (2019*a*). Molecular basis of PIP_2-dependent regulation of the Ca2+-activated chloride channel TMEM16A. *Nat Commun* **10,** 3769.

Le T, Jia Z, Le SC, Zhang Y, Chen J & Yang H (2019*b*). An inner activation gate controls TMEM16F phospholipid scrambling. *Nat Commun* **10,** 1846.

Lee KPK, Chen J & Mackinnon R (2017). Molecular structure of human katp in complex with ATP and ADP. *Elife*; DOI: 10.7554/eLife.32481.

Lemak AS & Balabaev NK (1994). On The Berendsen Thermostat. *Mol Simul*

13, 177–187.

Li N, Wu JX, Ding D, Cheng J, Gao N & Chen L (2017). Structure of a Pancreatic ATP-Sensitive Potassium Channel. *Cell* **168,** 101-110.e10.

Liang W, Ray JB, He JZ, Backx PH & Ward ME (2009). Regulation of proliferation and membrane potential by chloride currents in rat pulmonary artery smooth muscle cells. *Hypertens (Dallas, Tex 1979)* **54,** 286–293.

Lisbjerg M, Valkenier H, Jessen BM, Al-Kerdi H, Davis AP & Pittelkow M (2015). Biotin[6]uril Esters: Chloride-Selective Transmembrane Anion Carriers Employing C—H···Anion Interactions. *J Am Chem Soc* **137,** 4948–4951.

Lopatin AN, Makhina EN & Nichols CG (1994). Potassium channel block by cytoplasmic polyamines as the mechanism of intrinsic rectification. *Nature* **372,** 366–369.

Lopes CMB, Zhang H, Rohacs T, Jin T, Yang J & Logothetis DE (2002). Alterations in Conserved Kir Channel-PIP_2 Interactions Underlie Channelopathies. *Neuron* **34,** 933–944.

Lynch CI, Rao S & Sansom MSP (2020). Water in Nanopores and Biological Channels: A Molecular Simulation Perspective. *Chem Rev* **120,** 10298–10335.

MacGregor GG, Dong K, Vanoye CG, Tang L, Giebisch G & Hebert SC (2002). Nucleotides and phospholipids compete for binding to the C terminus of K_{ATP} channels. *Proc Natl Acad Sci* **99,** 2726 LP – 2731.

Männikkö R, Stansfeld PJ, Ashcroft AS, Hattersley AT, Sansom MSP, Ellard S & Ashcroft FM (2011). A conserved tryptophan at the membrane-water interface acts as a gatekeeper for Kir6.2/SUR1 channels and causes

neonatal diabetes when mutated. *J Physiol* **589,** 3071–3083.

Manoury B, Tamuleviciute A & Tammaro P (2010). TMEM16A/anoctamin 1 protein mediates calcium-activated chloride currents in pulmonary arterial smooth muscle cells. *J Physiol* **588,** 2305–2314.

Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP & de Vries AH (2007). The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations. *J Phys Chem B* **111,** 7812–7824.

Marrink SJ, de Vries AH & Mark AE (2004). Coarse Grained Model for Semiquantitative Lipid Simulations. *J Phys Chem B* **108,** 750–760.

Martin GM, Kandasamy B, DiMaio F, Yoshioka C & Shyng S-L (2017*a*). Anti-diabetic drug binding site in a mammalian K_{ATP} channel revealed by Cryo-EM ed. Swartz KJ. *Elife* **6,** e31054.

Martin GM, Sung MW, Yang Z, Innes LM, Kandasamy B, David LL, Yoshioka C & Shyng S-L (2019). Mechanism of pharmacochaperoning in a mammalian KATP channel revealed by cryo-EM ed. Aldrich R, Yellen G, Moiseenkova-Bell VY, Nichols CG & Agar J. *Elife* **8,** e46417.

Martin GM, Yoshioka C, Rex EA, Fay JF, Xie Q, Whorton MR, Chen JZ & Shyng SL (2017*b*). Cryo-EM structure of the ATP-sensitive potassium channel illuminates mechanisms of assembly and gating. *Elife*; DOI: 10.7554/eLife.24149.

Massa O, Iafusco D, D'Amato E, Gloyn AL, Hattersley AT, Pasquino B, Tonini G, Dammacco F, Zanette G, Meschi F, Porzio O, Bottazzo G, Crino A, Lorini R, Cerutti F, Vanelli M & Barbetti F (2005). KCNJ11 activating mutations in Italian patients with permanent neonatal diabetes. *Hum Mutat* **25,** 22–27.

Matsuda H, Saigusa A & Irisawa H (1987). Ohmic conductance through the inwardly rectifying K channel and blocking by internal Mg^{2+}. *Nature* **325,** 156–159.

van Meer G, Voelker DR & Feigenson GW (2008). Membrane lipids: where they are and how they behave. *Nat Rev Mol Cell Biol* **9,** 112–124.

Meng X-Y, Kang S & Zhou R (2018). Molecular mechanism of phosphoinositides' specificity for the inwardly rectifying potassium channel Kir2.2. *Chem Sci* **9,** 8352–8362.

Michaud-Agrawal N, Denning EJ, Woolf TB & Beckstein O (2011). MDAnalysis: A toolkit for the analysis of molecular dynamics simulations. *J Comput Chem* **32,** 2319–2327.

Mikhailov M V, Campbell JD, de Wet H, Shimomura K, Zadek B, Collins RF, Sansom MS, Ford RC & Ashcroft FM (2005). 3-D structural and functional characterization of the purified K_{ATP} channel complex Kir6.2-SUR1. *Embo J* **24,** 4166–4175.

Miledi R (1982). A calcium-dependent transient outward current in Xenopus laevis oocytes. *Proc R Soc London Ser B Biol Sci* **215,** 491–497.

Monticelli L, Kandasamy SK, Periole X, Larson RG, Tieleman DP & Marrink S-J (2008). The MARTINI Coarse-Grained Force Field: Extension to Proteins. *J Chem Theory Comput* **4,** 819.

Nagle JF & Tristram-Nagle S (2000). Structure of lipid bilayers. *Biochim Biophys Acta - Rev Biomembr* **1469,** 159–195.

Namkung W, Yao Z, Finkbeiner WE & Verkman AS (2011). Small-molecule activators of TMEM16A, a calcium-activated chloride channel, stimulate epithelial chloride secretion and intestinal contraction. *FASEB J Off Publ*

Fed Am Soc Exp Biol **25,** 4048–4062.

Naughton FB, Kalli AC & Sansom MSP (2018). Modes of Interaction of Pleckstrin Homology Domains with Membranes: Toward a Computational Biochemistry of Membrane Recognition. *J Mol Biol* **430,** 372–388.

Nessa A, Aziz QH, Thomas AM, Harmer SC, Tinker A & Hussain K (2015). Molecular mechanisms of congenital hyperinsulinism due to autosomal dominant mutations in ABCC8. *Hum Mol Genet* **24,** 5142–5153.

Newport TD, Sansom MSP & Stansfeld PJ (2018). The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic Acids Res* **47,** D390–D397.

Nugent T & Jones DT (2013). Membrane protein orientation and refinement using a knowledge-based statistical potential. *BMC Bioinformatics* **14,** 276.

Oh S-J, Hwang SJ, Jung J, Yu K, Kim J, Choi JY, Hartzell HC, Roh EJ & Lee CJ (2013). MONNA, a potent and selective blocker for transmembrane protein with unknown function 16/anoctamin-1. *Mol Pharmacol* **84,** 726–735.

Olesen K, Awasthi N, Bruhn DS, Pezeshkian W & Khandelia H (2018). Faster Simulations with a 5 fs Time Step for Lipids in the CHARMM Force Field. *J Chem Theory Comput*; DOI: 10.1021/acs.jctc.8b00267.

Olsson MHM, Søndergaard CR, Rostkowski M & Jensen JH (2011). PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pKa Predictions. *J Chem Theory Comput* **7,** 525–537.

Ortiz D & Bryan J (2015). Neonatal Diabetes and Congenital Hyperinsulinism Caused by Mutations in ABCC8/SUR1 are Associated with Altered and Opposite Affinities for ATP and ADP . *Front Endocrinol* **6,** 48.

Overington JP, Al-Lazikani B & Hopkins AL (2006). How many drug targets are there? *Nat Rev Drug Discov* **5,** 993–996.

Padding JT & Briels WJ (2003). Coarse-grained molecular dynamics simulations of polymer melts in transient and steady shear flow. *J Chem Phys* **118,** 10276–10286.

Papp R et al. (2019). Targeting TMEM16A to reverse vasoconstriction and remodelling in idiopathic pulmonary arterial hypertension. *Eur Respir J* **53,** 1800965.

Parrinello M & Rahman A (1981). Polymorphic transitions in single crystals: A new molecular dynamics method. *J Appl Phys* **52,** 7182–7190.

Paulino C, Kalienkova V, Lam AKM, Neldner Y & Dutzler R (2017*a*). Activation mechanism of the calcium-activated chloride channel TMEM16A revealed by cryo-EM. *Nature* **552,** 421–425.

Paulino C, Neldner Y, Lam AKM, Kalienkova V, Brunner JD, Schenck S & Dutzler R (2017*b*). Structural basis for anion conduction in the calcium-activated chloride channel TMEM16A ed. Chanda B. *Elife* **6,** e26232.

Pedemonte N & Galietta LJ V (2014). Structure and Function of TMEM16 Proteins (Anoctamins). *Physiol Rev* **94,** 419–459.

Peters CJ, Gilchrist JM, Tien J, Bethel NP, Qi L, Chen T, Wang L, Jan YN, Grabe M & Jan LY (2018). The Sixth Transmembrane Segment Is a Major Gating Component of the TMEM16A Calcium-Activated Chloride Channel. *Neuron* **97,** 1063-1077.e4.

Petersen MC & Shulman GI (2018). Mechanisms of Insulin Action and Insulin Resistance. *Physiol Rev* **98,** 2133–2223.

Pipatpolkai T, Corey RA, Proks P, Ashcroft FM & Stansfeld PJ (2020*a*).

Evaluating inositol phospholipid interactions with inward rectifier potassium channels and characterising their role in disease. *Commun Chem* **3,** 147.

Pipatpolkai T, Usher S, Stansfeld PJ & Ashcroft FM (2020*b*). New insights into K_{ATP} channel gene mutations and neonatal diabetes mellitus. *Nat Rev Endocrinol*; DOI: 10.1038/s41574-020-0351-y.

Pipatpolkai T, Usher SG, Vedovato N, Ashcroft FM & Stansfeld PJ (2021). The dynamic interplay of PIP_2 and ATP in the regulation of the K_{ATP} channel. *bioRxiv*2021.05.06.442933.

Popot JL & Engelman DM (2000). Helical membrane protein folding, stability, and evolution. *Annu Rev Biochem* **69,** 881–922.

Proks P, Capener CE, Jones P & Ashcroft FM (2001). Mutations within the P-loop of Kir6.2 modulate the intraburst kinetics of the ATP-sensitive potassium channel. *J Gen Physiol* **118,** 341–353.

Proks P, Girard C, Haider S, Gloyn AL, Hattersley AT, Sansom MSP & Ashcroft FM (2005). A gating mutation at the internal mouth of the Kir6.2 pore is associated with DEND syndrome. *EMBO Rep* **6,** 470 LP – 475.

Puljung M, Vedovato N, Usher S & Ashcroft F (2019). Activation mechanism of ATP-sensitive K+ channels explored with real-time nucleotide binding ed. Kuriyan J, Swartz KJ, Swartz KJ, Shyng S-L & Goldschen-Ohm MP. *Elife* **8,** e41103.

Puljung MC (2018). Cryo-electron microscopy structures and progress toward a dynamic understanding of K_{ATP} channels. *J Gen Physiol*; DOI: 10.1085/jgp.201711978.

Rajala RVS & Anderson RE (2010). Focus on molecules: phosphatidylinositol-4,5-bisphosphate (PIP_2). *Exp Eye Res* **91,** 324–325.

Rao S, Klesse G, Stansfeld PJ, Tucker SJ & Sansom MSP (2019). A heuristic derived from analysis of the ion channel structural proteome permits the rapid identification of hydrophobic gates. *Proc Natl Acad Sci* **116,** 13989 LP – 13995.

Reimann F, Huopio H, Dabrowski M, Proks P, Gribble FM, Laakso M, Otonkoski T & Ashcroft FM (2003). Characterisation of new K_{ATP}-channel mutations associated with congenital hyperinsulinism in the Finnish population. *Diabetologia* **46,** 241–249.

Ribalet B, John SA, Xie L-H & Weiss JN (2006). ATP-sensitive K^+ channels: regulation of bursting by the sulphonylurea receptor, PIP_2 and regions of Kir6.2. *J Physiol* **571,** 303–317.

Rizzi A, Jensen T, Slochower DR, Aldeghi M, Gapsys V, Ntekoumes D, Bosisio S, Papadourakis M, Henriksen NM, de Groot BL, Cournia Z, Dickson A, Michel J, Gilson MK, Shirts MR, Mobley DL & Chodera JD (2020). The SAMPL6 SAMPLing challenge: assessing the reliability and efficiency of binding free energy calculations. *J Comput Aided Mol Des* **34,** 601–633.

Rjasanow A, Leitner MG, Thallmair V, Halaszovich CR & Oliver D (2015). Ion channel regulation by phosphoinositides analyzed with VSPs-$PI(4,5)P_2$ affinity, phosphoinositide selectivity, and $PI(4,5)P_2$ pool accessibility. *Front Pharmacol* **6,** 127.

Rohács T, Lopes CMB, Jin T, Ramdya PP, Molnár Z & Logothetis DE (2003). Specificity of activation by phosphoinositides determines lipid regulation of Kir channels. *Proc Natl Acad Sci* **100,** 745 LP – 750.

Sakura H, Ammälä C, Smith PA, Gribble FM & Ashcroft FM (1995). Cloning and functional expression of the cDNA encoding a novel ATP-sensitive

potassium channel subunit expressed in pancreatic beta-cells, brain, heart and skeletal muscle. *FEBS Lett*; DOI: 10.1016/0014-5793(95)01369-5.

Sakura H, Wat N, Horton V, Millns H, Turner RC & Ashcroft FM (1996). Sequence variations in the human Kir6.2 gene, a subunit of the beta-cell ATP-sensitive K-channel: no association with NIDDM in white caucasian subjects or evidence of abnormal function when expressed in vitro. *Diabetologia* **39,** 1233–1236.

Schrodinger LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.

Schroeder BC, Cheng T, Jan YN & Jan LY (2008). Expression Cloning of TMEM16A as a Calcium-Activated Chloride Channel Subunit. *Cell* **134,** 1019–1029.

Schulze D, Krauter T, Fritzenschaft H, Soom M & Baukrowitz T (2003). Phosphatidylinositol 4,5-bisphosphate (PIP_2) modulation of ATP and pH sensitivity in Kir channels. A tale of an active and a silent PIP_2 site in the N terminus. *J Biol Chem* **278,** 10500–10505.

Schwappach B, Zerangue N, Jan YN & Jan LY (2000). Molecular Basis for K_{ATP} Assembly: Transmembrane Interactions Mediate Association of a K^+ Channel with an ABC Transporter. *Neuron* **26,** 155–167.

Seeliger D & de Groot BL (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des* **24,** 417–422.

Seo Y, Lee HK, Park J, Jeon D, Jo S, Jo M & Namkung W (2016). Ani9, A Novel Potent Small-Molecule ANO1 Inhibitor with Negligible Effect on ANO2. *PLoS One* **11,** e0155771.

Shimomura K, Flanagan SE, Zadek B, Lethby M, Zubceuic L, Girard CAJ, Petz O, Mannikko R, Kapoor RR, Hussain K, Skae M, Clayton P, Hattersley A,

Ellard S & Ashcroft FM (2009). Adjacent mutations in the gating loop of Kir6.2 produce neonatal diabetes and hyperinsulinism. *EMBO Mol Med* **1,** 166–177.

Shimomura K, Girard CAJ, Proks P, Nazim J, Lippiat JD, Cerutti F, Lorini R, Ellard S, Hattersley AT, Barbetti F & Ashcroft FM (2006). Mutations at the same residue (R50) of Kir6.2 (KCNJ11) that cause neonatal diabetes produce different functional effects. *Diabetes* **55,** 1705–1712.

Shimomura K, Horster F, de Wet H, Flanagan SE, Ellard S, Hattersley AT, Wolf NI, Ashcroft F & Ebinger F (2007). A novel mutation causing DEND syndrome: a treatable channelopathy of pancreas and brain. *Neurology* **69,** 1342–1349.

Shimomura K, De Nanclares GP, Foutinou C, Caimari M, Castaño L & Ashcroft FM (2010). The first clinical case of a mutation at residue K185 of Kir6.2 (KCNJ11): A major ATP-binding residue. *Diabet Med* **27,** 225–229.

Shyng S-L, Cukras CA, Harwood J & Nichols CG (2000). Structural Determinants of PIP_2 Regulation of Inward Rectifier K_{ATP} Channels. *J Gen Physiol* **116,** 599 LP – 608.

Shyng SL & Nichols CG (1998). Membrane phospholipid control of nucleotide sensitivity of KATP channels. *Science (80)* **282,** 1138–1141.

Smart OS, Neduvelil JG, Wang X, Wallace BA & Sansom MSP (1996). HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *J Mol Graph* **14,** 354–360.

Song W, Duncan AL & Sansom MSP (2021). Modulation of adenosine A2a receptor oligomerization by receptor activation and PIP_2 interactions. *Structure*; DOI: https://doi.org/10.1016/j.str.2021.06.015.

Soom M, Schönherr R, Kubo Y, Kirsch C, Klinger R & Heinemann SH (2001). Multiple PIP2 binding sites in Kir2.1 inwardly rectifying potassium channels. *FEBS Lett* **490,** 49–53.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE & Berendsen HJC (2005). GROMACS: Fast, flexible, and free. *J Comput Chem* **26,** 1701–1718.

Stansfeld PJ, Goose JE, Caffrey M, Carpenter EP, Parker JL, Newstead S & Sansom MSP (2015). MemProtMD: Automated Insertion of Membrane Protein Structures into Explicit Lipid Membranes. *Structure* **23,** 1350–1361.

Stansfeld PJ, Hopkinson R, Ashcroft FM & Sansom MSP (2009). PIP2-binding site in Kir channels: Definition by multiscale biomolecular simulations. *Biochemistry* **48,** 10926–10933.

Stansfeld PJ & Sansom MSP (2011). From coarse grained to atomistic: A serial multiscale approach to membrane protein simulations. *J Chem Theory Comput* **7,** 1157–1166.

Stauffer DA & Karlin A (1994). Electrostatic potential of the acetylcholine binding sites in the nicotinic receptor probed by reactions of binding-site cysteines with charged methanethiosulfonates. *Biochemistry* **33,** 6840–6849.

Sturgess N, Cook D, Ashford M. & Hales CN (1985). THE SULPHONYLUREA RECEPTOR MAY BE AN ATP-SENSITIVE POTASSIUM CHANNEL. *Lancet* **326,** 474–475.

Suh B-C & Hille B (2008). PIP_2 is a necessary cofactor for ion channel function: how and why? *Annu Rev Biophys* **37,** 175–195.

Sun H, Xia Y, Paudel O, Yang X-R & Sham JSK (2012). Chronic hypoxia-induced upregulation of Ca^{2+}-activated Cl^- channel in pulmonary arterial myocytes: a mechanism contributing to enhanced vasoreactivity. *J Physiol* **590,** 3507–3521.

Sun J & MacKinnon R (2020). Structural Basis of Human KCNQ1 Modulation and Gating. *Cell* **180,** 340-347.e9.

Sun X, Singh S, Blumer KJ & Bowman GR (2018). Simulation of spontaneous G protein activation reveals a new intermediate driving GDP unbinding ed. Faraldo-Gómez JD & Kuriyan J. *Elife* **7,** e38465.

Sung MW, Yang Z, Patton BL, Mostofian B, Russo J, Zuckerman DM & Shyng S-L (2021). Vascular K_{ATP} channel structural dynamics reveal regulatory mechanism by Mg-nucleotides. *bioRxiv*2021.05.15.444267.

Suzuki S, Makita Y, Mukai T, Matsuo K, Ueda O & Fujieda K (2007). Molecular basis of neonatal diabetes in Japanese patients. *J Clin Endocrinol Metab* **92,** 3979–3985.

Ta CM, Adomaviciene A, Rorsman NJG, Garnett H & Tammaro P (2016). Mechanism of allosteric activation of TMEM16A/ANO1 channels by a commonly used chloride channel blocker. *Br J Pharmacol* **173,** 511–528.

Tama F, Wriggers W & Brooks CL 3rd (2002). Exploring global distortions of biological macromolecules and assemblies from low-resolution structural information and elastic network theory. *J Mol Biol* **321,** 297–305.

Tammaro P, Flanagan SE, Zadek B, Srinivasan S, Woodhead H, Hameed S, Klimes I, Hattersley AT, Ellard S & Ashcroft FM (2008). A Kir6.2 mutation causing severe functional effects in vitro produces neonatal diabetes without the expected neurological complications. *Diabetologia* **51,** 802–

810.

Tao X, Avalos JL, Chen J & MacKinnon R (2009). Crystal structure of the eukaryotic strong inward-rectifier K^+ channel Kir2.2 at 3.1 A resolution. *Science* **326,** 1668–1674.

Theobald DL (2005). Rapid calculation of RMSDs using a quaternion-based characteristic polynomial. *Acta Crystallogr Sect A* **61,** 478–480.

Thomas KA, Smith GM, Thomas TB & Feldmann RJ (1982). Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc Natl Acad Sci U S A* **79,** 4843–4847.

Thorens B (2015). GLUT2, glucose sensing and glucose homeostasis. *Diabetologia* **58,** 221–232.

Tien J, Lee HY, Minor DL, Jan YN & Jan LY (2013). Identification of a dimerization domain in the TMEM16A calcium-activated chloride channel (CaCC). *Proc Natl Acad Sci* **110,** 6352 LP – 6357.

Trapp S, Proks P, Tucker SJ & Ashcroft FM (1998). Molecular Analysis of ATP-sensitive K Channel Gating and Implications for Channel Inhibition by ATP. *J Gen Physiol* **112,** 333 LP – 349.

Tribello GA, Bonomi M, Branduardi D, Camilloni C & Bussi G (2014). PLUMED 2: New feathers for an old bird. *Comput Phys Commun* **185,** 604–613.

Trott O & Olson AJ (2010*a*). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. *J Comput Chem* **31,** 455–461.

Trott O & Olson AJ (2010*b*). AutoDock Vina. *J Comput Chem* **31,** 445–461.

Tucker SJ, Gribble FM, Proks P, Trapp S, Ryder TJ, Haug T, Reimann F &

Ashcroft FM (1998). Molecular determinants of K_{ATP} channel inhibition by ATP. *EMBO J* **17,** 3290–3296.

Tucker SJ, Gribble FM, Zhao C, Trapp S & Ashcroft FM (1997). Truncation of Kir6.2 produces ATP-sensitive K^+ channels in the absence of the sulphonylurea receptor. *Nature* **387,** 179–183.

Usher SG, Ashcroft FM & Puljung MC (2020). Nucleotide inhibition of the pancreatic ATP-sensitive K^+ channel explored with patch-clamp fluorometry ed. Yellen G & Aldrich RW. *Elife* **9,** e52775.

Vickery ON & Stansfeld PJ (2021). CG2AT2: An Enhanced Fragment-based approach for Serial Multi-scale Molecular Dynamics simulations. *bioRxiv*2021.03.25.437005.

Wang B, Li C, Huai R & Qu Z (2015*a*). Overexpression of ANO1/TMEM16A, an arterial Ca^{2+}-activated Cl^- channel, contributes to spontaneous hypertension. *J Mol Cell Cardiol* **82,** 22–32.

Wang L et al. (2015*b*). Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. *J Am Chem Soc* **137,** 2695–2703.

Wang L, Bose PS & Sigworth FJ (2006). Using cryo-EM to measure the dipole potential of a lipid membrane. *Proc Natl Acad Sci* **103,** 18528 LP – 18533.

Wang Q, Corey RA, Hedger G, Aryal P, Grieben M, Nasrallah C, Baronina A, Pike ACW, Shi J, Carpenter EP & Sansom MSP (2020). Lipid Interactions of a Ciliary Membrane TRP Channel: Simulation and Structural Studies of Polycystin-2. *Structure* **28,** 169-184.e5.

Wassenaar TA, Ingólfsson HI, Böckmann RA, Tieleman DP & Marrink SJ

(2015). Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations. *J Chem Theory Comput* **11,** 2144–2155.

Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R & Schwede T (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46,** W296–W303.

Webb B & Sali A (2016). Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Bioinformatics* **54,** 5.6.1-5.6.37.

Whorton MR & MacKinnon R (2011). Crystal structure of the mammalian GIRK2 K+ channel and gating regulation by G proteins, PIP2, and sodium. *Cell* **477,** 199–208.

Wozniak KL, Phelps WA, Tembo M, Lee MT & Carlson AE (2018). The TMEM16A channel mediates the fast polyspermy block in Xenopus laevis. *J Gen Physiol* **150,** 1249–1259.

Wu J-X, Ding D, Wang M, Kang Y, Zeng X & Chen L (2018). Ligand binding and conformational changes of SUR1 subunit in pancreatic ATP-sensitive potassium channels. *Protein Cell* **9,** 553–567.

Wu Z, Alibay I, Newstead S & Biggin PC (2019). Proton Control of Transitions in an Amino Acid Transporter. *Biophys J* **117,** 1342–1351.

Yang YD, Cho H, Koo JY, Tak MH, Cho Y, Shim W-S, Park SP, Lee J, Lee B, Kim B-M, Raouf R, Shin YK & Oh U (2008). TMEM16A confers receptor-activated calcium-dependent chloride conductance. *Nature* **455,** 1210–1215.

Ye W, Han TW, He M, Jan YN & Jan LY (2019). Dynamic change of

electrostatic field in TMEM16F permeation pathway shifts its ion selectivity ed. Swartz KJ, Aldrich R & Toombes G. *Elife* **8,** e45187.

Zerangue N, Schwappach B, Jan YN & Jan LY (1999). A New ER Trafficking Signal Regulates the Subunit Stoichiometry of Plasma Membrane K_{ATP} Channels. *Neuron* **22,** 537–548.

Zhang H, He C, Yan X, Mirshahi T & Logothetis DE (1999). Activation of inwardly rectifying K^+ channels by distinct PtdIns(4,5)P2 interactions. *Nat Cell Biol* **1,** 183–188.

Zhang M, Chen X, Shen S, Li T, Chen L, Hu M, Cao L, Cheng R, Zhao Z & Luo F (2015). Sulfonylurea in the treatment of neonatal diabetes mellitus children with heterogeneous genetic backgrounds. *J Pediatr Endocrinol Metab* **28,** 877–884.

Zhou J-J, Fatehi M & Linsdell P (2008). Identification of positive charges situated at the outer mouth of the CFTR chloride channel pore. *Pflugers Arch* **457,** 351–360.

Thank you so much!

Photo taken by: Matthew Lloyd