Multiscale Molecular Dynamics to Investigate Sodium Ion Channel Oligomerisation



William Glass

Balliol College

University of Oxford

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

Doctor of Philosophy

Michaelmas 2019

To my family,
for always looking at the Glass half full

Acknowledgements

Firstly, I would like to thank my supervisor Prof Phil Biggin who has always offered his advice and guidance on any research topic, as well as any personal matters, over the last few years. He has always engaged with me, from answering my many questions about biology and biochemistry to a whole host of matters too long to list here. His support has allowed me to challenge myself and enabled me to grow both as a scientist and as a person. I'd also like to thank Prof John McGrady for giving me the opportunity to work with Phil and pursue a research project in the biochemistry department.

Secondly, my thanks go to the many members of the Structural Bioinformatics & Computational Biology (SBCB) group. I've been lucky enough to meet a brilliant group of scientists that I'm privileged to call my friends. Unfortunately, thanking all of the people who have made working at SBCB an enjoyable experience would be far too long to fit in such a short space, so I will keep it brief. Special mentions go to Dr Irfan Alibay and Dr Aphroditi Zaki (a.k.a. the best desk neighbour) for putting up with all my questions and always being around for advice on both a professional and personal level. I'd also like to mention Dr Anna Duncan for teaching me all things coarse-grained (and our long chats about politics). Last but not least, Asst Prof Maria Musgaard for helping me when I first arrived and always offering words of wisdom (even from across the Atlantic Ocean). My thanks also extend to fellow SBCB DPhil students Joe Bluck, Chris Lindsay, Shad Velgy, Max Epstein, Marc Dämgen, Rocco Meli, Mina Moniri, and Jodie Collingridge, whose company I have always enjoyed both inside and outside the lab.

Away from the department I have built a close circle of friends who have all made this journey bearable. In particular Silvia Amabilino, Michael

Connolly, and Jon Shearer from my Centre for Doctoral Training. Although often far away, friends made during my time in Cardiff have kept me sane, namely Valerie Askin, Amy Lai, and Vicki Groves. In Oxford, I can honestly say that I would not be half the person I am today without Ricky Sengupta, Jack Evans, Melissa Bedard, Evie Foster, and Violaine André.

I would like to thank my family for their unwavering support in everything that I do. To my sisters, Naomi and Eleanor, for always looking out for their little brother. In particular to my mother Wendy, whose constant love and *joie de vivre* inspires me everyday. Lastly to my father Nigel: although he took his own life when I was a boy, I hope he would've been proud of the man I have become.

Finally, a note to the reader. Over the course of my life and academic career I have often experienced anxiety and depression. As a first generation doctoral student I've almost always doubted my capabilities and typically strived to get things done perfectly. However, being a perfectionist is simply not possible, both in life and in academia. Aiming for perfection will almost always be met with failure. To any future doctoral students reading this (or anyone else for that matter), you can rest assured that some of the most successful people I know fail often. Take care of your mental health: your perceived imperfections are often your greatest strength.

Abstract

Electrical signalling is key to a vast number of biological processes within the nervous system. For this to occur, specific types of transmembrane proteins, named voltage-gated ion channels, are utilised. In voltage-gated sodium (Nav) channels, recent advances in structural biology are beginning to solve structures in complex with so-called "auxiliary" β subunits. These transmembrane proteins are multifunctional and interact with the main α subunit of the Nav channel, in addition to having a cell adhesion function through their extracellular immunoglobulin (Ig) domain. In Nav1.5, β3 has been found to cross-link subunits within the membrane, as well as form trimers and higher order oligomers.

In this thesis, both atomistic and coarse grained (CG) molecular dynamics (MD) simulations are performed to gain an insight into the behaviour of the α subunit, the α - β subunit relationship, and β subunit dynamics in isolation.

Firstly, atomistic simulations are used to investigate a splice variant of Nav1.5. Two single point mutations that were responsible for a shift in voltage-gating characteristics were investigated and found to most likely disrupt gating through interruption of voltage-sensing arginine residues.

Secondly, a large number of simulations are performed to explore Nav β1 and β3 subunit behaviour in a simple membrane composition. Results showed large differences between β1 and β3, with the latter exhibiting increased fluctuations in Ig domain orientation. These differences are attributed to key charged residues that influence Ig domain interaction with the membrane surface. Protein - lipid interactions suggest that a conserved glutamic acid residue is important in transmembrane domain location, rather than being involved in observed oligomerisation.

Finally, the use of CG simulations examines the dependency of both α and β subunit clustering with various lipids in a complex membrane composition. Protein clustering of both subunits was primarily mediated by glycosphingolipids present in the upper leaflet. The β subunit orientations adopted three states, one of which agreed well with an atomistic trimer model of the β3 Ig domain. Large-scale simulations explored a large ($\approx$ 2 million particle) system within which multiple protein binding events occurred. Conformations present between α and β subunits agreed with experimental results from atomic force microscopy and increase our understanding of Nav subunit interaction.

Overall, the *in silico* multiscale approach adopted in this thesis aims to characterise Nav channel subunit behaviour in a variety of contexts. These simulations demonstrate the predictive power of MD in characterising channel function away from the pore, as well as informing potential future strategies against various channelopathies.

Contents

List of Abbreviations **xiii**

1 Introduction **1**

1.1 Overview . . . 2
1.1.1 Biological membranes . . . 3
1.1.2 The action potential . . . 5
1.1.3 Neuronal vs. cardiac action potential . . . 8
1.1.4 Cellular structure: Neurons vs. Cardiomyocytes . . . 10
1.2 Voltage-gated sodium channels . . . 13
1.2.1 Structure of the α subunit . . . 15
1.2.2 Structure of the β subunit . . . 18
1.3 Binding between α & β subunits . . . 20
1.4 Outline . . . 22
1.5 Work presented in this thesis . . . 23

2 Theory & methods **25**

2.1 Methods in computational chemistry . . . 26
2.2 The molecular mechanics force field . . . 29
2.2.1 Bonded interactions . . . 30
2.2.2 Non-bonded interactions . . . 32
2.3 Molecular dynamics theory . . . 34
2.3.1 Integrators . . . 35
2.3.1.1 The Verlet algorithm . . . 35
2.3.1.2 The Leapfrog algorithm . . . 35
2.3.1.3 The velocity Verlet algorithm . . . 36
2.4 System setup protocol . . . 37
2.4.1 Periodic boundary conditions . . . 37
2.4.2 Energy minimisation . . . 38

2.4.3 The thermodynamic ensemble: temperature & pressure 39
2.4.3.1 Thermostats 40
2.4.3.2 Barostats 42
2.4.4 Long-range interactions 43
2.5 Multiscale modelling 47
2.5.1 Coarse-grained molecular dynamics 48
2.5.1.1 The MARTINI force field 48
2.6 Homology model generation 52

3 Nav1.5e splice variant **55**
3.1 Introduction 56
3.1.1 Alternative splicing 56
3.1.2 Alternative splicing in Nav channels 57
3.1.3 Nav1.5e: D211K & T207S mutations 58
3.2 System setup 61
3.3 Results 63
3.3.1 Model stability & S4 movement 63
3.3.2 D211K mutant 68
3.3.2.1 Interaction with lipid 68
3.3.2.2 Residue distance analysis 69
3.3.3 T207S mutant 72
3.3.3.1 Dihedral angle analysis 72
3.3.3.2 Salt-bridge analysis 74
3.3.4 Double D211K & T207S mutant 76
3.3.4.1 D211K dynamics 76
3.3.4.2 Salt-bridge analysis 77
3.4 Discussion 79
3.4.1 D211K dynamics 79
3.4.2 T207S dynamics 81
3.4.3 Potential inhibition of S4 movement by combined D211K & T207S interactions 82
3.4.4 Conclusions 82

4 Atomistic simulations of the β3 subunit **85**
4.1 Introduction . . . 86
4.1.1 β3 Ig trimer . . . 87
4.1.2 Overview of states: a paradox . . . 90
4.1.3 A conserved glutamic acid residue . . . 91
4.2 Methods . . . 92
4.2.1 Model generation . . . 92
4.2.1.1 β1 monomer . . . 92
4.2.1.2 β3 monomer . . . 92
4.2.1.3 Glycosylated β3 monomer . . . 92
4.2.1.4 β3 ECD . . . 93
4.2.1.5 β3 trimer . . . 93
4.2.2 Molecular dynamics simulations . . . 94
4.2.3 Ig domain orientation analysis . . . 95
4.3 Results . . . 97
4.3.1 β3 Ig stability . . . 97
4.3.1.1 Monomer . . . 97
4.3.1.2 Trimer . . . 98
4.3.2 Dynamics of the full length monomer construct . . . 101
4.3.2.1 Helix analysis . . . 101
4.3.2.2 Protein - lipid interactions . . . 106
4.3.2.3 Orientation of the Ig domain . . . 110
4.3.3 An investigation of β3 glycosylation sites . . . 114
4.3.4 Dynamics of the full length trimer construct . . . 118
4.4 Discussion . . . 127
4.4.1 The varied dynamics of the Ig domain . . . 127
4.4.2 Destabilisation of C-terminal residues in the β1 & β3 helices . . . 130
4.4.3 Glycosylation: A preferred Ig orientation? . . . 133
4.5 Conclusions . . . 134

5 Coarse-grained simulations of β3 subunit oligomerisation **137**
5.1 Introduction . . . 138
5.1.1 Lipid composition: A model of an *in vivo* membrane . . . 138
5.2 Methods . . . 140
5.2.1 System setup . . . 140
5.3 Results . . . 143

5.3.1 Cluster evolution of the β3 subunit . . . 143
5.3.2 Protein - protein interactions between β3 clusters . . . 146
5.3.3 β3 protein - lipid interactions . . . 149
5.4 Discussion . . . 155
5.4.0.1 Ig dynamics in a mixed lipid membrane . . . 155
5.4.0.2 The effect of lipid clustering . . . 158
5.5 Conclusions . . . 159

6 Exploring α - β3 interactions using coarse-grained simulations 161
6.1 Introduction . . . 162
6.1.1 The significance of Nav1.5 α - β3 clustering . . . 162
6.2 Methods . . . 163
6.2.1 System setup . . . 163
6.2.1.1 The 3 $\times$ 3 α subunit grid . . . 163
6.2.1.2 The mixed α & β3 subunit grid . . . 165
6.3 Results . . . 168
6.3.1 Cluster evolution of the α subunit . . . 168
6.3.2 α - β3 subunit interactions in a large membrane system . . . 170
6.3.3 Influence of lipid interactions on α - β3 clustering . . . 177
6.3.4 Alternative binding mode of the β3 subunit . . . 179
6.4 Discussion . . . 181
6.4.1 Preferential binding of the β3 subunit . . . 181
6.4.2 Influence of the lipid membrane . . . 186
6.5 Conclusions . . . 188

7 Conclusions 189
7.1 Summary . . . 189
7.2 Future directions . . . 193
7.2.1 Protein models . . . 193
7.2.2 The influence of glycosylation . . . 194
7.2.3 The MARTINI model and increasing membrane complexity . . . 195
7.2.4 Additional auxiliary subunits . . . 196
7.3 Final remarks . . . 197

Appendices

A Mathematical background **201**
A.1 Hamiltonian mechanics 202
A.2 The Verlet algorithm 203
A.3 Tait-Bryan angle formalism 205

B Supplementary material **209**
B.1 Supplementary material for Chapter 3 210
B.2 Supplementary material for Chapter 4 211
B.3 Supplementary material for Chapter 5 212
B.4 Supplementary material for Chapter 6 213
B.5 Protein sequence alignments 215

References **216**

List of Figures

1.1 The variety of lipids found in typical eukaryotic cell membranes. . . . 4
1.2 Cycle of ion movement during an action potential. 7
1.3 Action potentials of neuronal and cardiac cells. 9
1.4 General features of a typical neuron. 11
1.5 Action potential through cardiomyocytes. 12
1.6 Structure of the Nav α subunit. 17
1.7 The structure of the β subunit. 19
1.8 Crystal structures of Nav α - β complexes 21

2.1 Schematic of a simple harmonic potential between particles i and j. . 30
2.2 Schematic of the angular potential between particles i, j, and k. . . . 31
2.3 Schematic of torsional potentials. 31
2.4 The Lennard-Jones potential. 32
2.5 The Coulomb potential. 33
2.6 A schematic of the Leapfrog algorithm in one dimension. 36
2.7 Periodic boundary conditions. 38
2.8 The Ewald sum. 45
2.9 The Verlet list. 46
2.10 Multiscale methods across various time and length scales. 47
2.11 Different particles in the MARTINI model. 49
2.12 MARTINI particle type mapping to amino acid side chains. 50

3.1 Overview of the common mechanisms of alternative splicing in gene expression. 57
3.2 The voltage-sensor (S1 - 4) of domain 1 in Nav1.5. 59
3.3 Conductance-voltage curves for the Nav1.5/mH1 and point mutations. 60
3.4 DI S3-4 multiple sequence alignment used for Nav1.5 homology model generation. 61

3.5 Mean RMSD of WT, D211K, T207S, and double D211K/T207S simulation performed. 63
3.6 Movement of arginine residues z-coordinate (Z) in the S4 helix in WT. 66
3.7 Movement of arginine residues z-coordinate (Z) in the S4 helix in the D211K mutant. 66
3.8 Movement of arginine residues z-coordinate (Z) in the S4 helix in the T207S mutant. 67
3.9 Movement of arginine residues z-coordinate (Z) in the S4 helix in the double D211K - T207S mutant. 67
3.10 Interactions of D211 / K211 with POPC headgroups. 68
3.11 Various interactions of K211 in the D211K switch system. 69
3.12 Salt-bridge interactions present in the mH1 VSDI. 71
3.13 Salt-bridge interactions present in the D211K VSDI. 71
3.14 Two poses adopted by T207. 72
3.15 Structures of threonine and serine. 73
3.16 Dihedral angles of T207 (mH1) and S207 (T207S) 74
3.17 Location of E208 in relation to T207 in mH1. 75
3.18 Histograms of E208 distances to R219 & R222 in T207S VSDI. 76
3.19 The two orientations typically adopted by K211 in the double D211K/T207S switch. 77
3.20 Histograms of E208 distances to R219 & R222 in Double D211K/T207S. 78
3.21 Dihedral angles of T207 (blue) and S207 (red), in mH1 and T207S respectively. Dihedral angles were collected over all repeats. 78
3.22 Combined E208 and K211 interaction with R219 and R222 in the S4 helix. 79

4.1 Structural features of the β3 Ig domain. 88
4.2 Schematic of the location of key trimer interface residues. 89
4.3 Overlay of the β1 model with the β3 Ig trimer. 90
4.4 Alignment of the TMD of the β1 and β3 subunits. 91
4.5 Structure of GlcNAc. 93
4.6 Schematic of the pitch angle calculation in the Ig domain of the β subunit. 96
4.7 The effects of the C48A mutation on β3 ECD monomer dynamics. . . 98
4.8 Radius of gyration of the β3 ECD trimer. 99
4.9 The ϕ and ψ angles sampled by V27 and V29 in the β3 trimer model. 100

4.10 Helix analysis of the β subunit. 103
4.11 Helix bending in the β1 and β3 subunits. 105
4.12 RMSF of the β1 and β3 subunits. 106
4.13 Protein - lipid interactions of both β1 and β3 subunits. 108
4.14 Average protein - lipid interactions visualised on the β1 and β3 subunits. 109
4.15 Distribution of pitch angles visited by the β1 and β3 subunits. 111
4.16 Binding of the β1 Ig domain to the membrane surface. 113
4.17 Distances of GlcNAc groups of glycosylated residues in the β3 Ig domain. 115
4.18 Pitch analysis of the glycosylated β3 subunit. 116
4.19 Representative snapshot of the glycosylated β3 subunit on a POPC membrane. 117
4.20 Full length β3 trimer model construction. 119
4.21 Bending of helices in the full length β3 trimer model. 121
4.22 RMSF of the full length β3 trimer model. 122
4.23 Position of E176 in the β3 trimer. 123
4.24 Protein - lipid interactions in the full length β3 trimer. 125
4.25 Distribution of T154 and Y181 Cα positions in the β3 trimer construct. 126
4.26 Alignment of the N-terminus of the β1 and β3 subunits. 129
4.27 Examination of β subunit bending. 131

5.1 CG generalised mammalian membrane composition 142
5.2 Cluster evolution of the β3 subunit in a generalised mammalian membrane of 10 μs. 144
5.3 Different states adopted by β3 clusters. 146
5.4 Protein - protein interactions in β3. 148
5.5 Radial distribution function of lipids around β3. 150
5.6 Average β3 protein - lipid interactions. 151
5.7 Regions of high interaction of the β3 subunit with GM3 153
5.8 Lipid density surrounding the β3 subunit 154
5.9 Overlay of the "up" state β3 subunit with the atomistic trimer model. 156

6.1 Construction of the large α - β3 grid system. 167
6.2 Cluster evolution of the 3 $\times$ 3 α subunit system. 168
6.3 Protein - protein interactions in the 3 $\times$ 3 α subunit system. 170
6.4 Evolution of α and β3 clusters over 17 μs. 171
6.5 Density of β3 around the α subunit. 172
6.6 Protein - protein interactions between the α and β3 subunits. 174

6.7 β3 contacts on DI & DII of the the Nav1.5 α subunit. 175
6.8 β3 contacts on DIII & DIV of the the Nav1.5 α subunit. 176
6.9 The RDF distribution of lipids surrounding the Nav1.5 α subunit in the α - β3 system. 177
6.10 Protein - lipid interactions in the Nav1.5 α subunit. 178
6.11 GM3 mediated clustering of α and β3 subunits. 180
6.12 Alignment of VSDI-IV in the Nav1.5 α subunit. 183
6.13 Alignment of the β1 and β3 TMD helical sequence. 184
6.14 Membrane distortion in the α - β3 system. 187

A.1 A rigid body with a body-fixed and space-fixed coordinate system. . . 206
A.2 Principal axes calcualted within the β subunit Ig domain. 207
A.3 The general sequence of a rotation adopting the Tait-Bryan angle formalism. 208

B.1 Salt-bridge interactions present in VSDI of the T207S system. 210
B.2 Salt-bridge interactions present in VSDI of the double D211K - T207S system. 210
B.3 β1 subunit R96 minimum distance from the membrane and RMSD. . 211
B.4 β3 clustering in a CG mixed lipid membrane. 212
B.5 Average β3 - PIP_2 interactions. 212
B.6 β3 contacts on DI & DII of the the Nav1.5 α subunit. 213
B.7 β3 contacts on DIII & DIV of the Nav1.5 α subunit. 214
B.8 Sequence alignment of Nav β1 and β3 subunits. 215

List of Tables

1.1 Localisation of Nav channels in different tissue types. 14
1.2 Localisation of β subunits. 15

2.1 Common thermodynamic ensembles. 40

4.1 Overview of atomistic β subunit simulations performed 94

5.1 Lipid composition of the generalised mammalian membrane. 142
5.2 Visual assignment of β3 states in each clustering run. 146

List of Abbreviations

AP . . . . . Action potential

ADP . . . . . Adenosine diphosphate

AFM . . . . . Atomic force microscopy

ATP . . . . . Adenosine triphosphate

BO . . . . . Born-Oppenheimer

Cav . . . . . Voltage-gated calcium channel

CAM . . . . . Cell-adhesion molecule

CG . . . . . Coarse-grained

Chol . . . . . Cholesterol

CLC . . . . . Voltage-gated chloride channel

CoM . . . . . Centre of mass

Cryo-EM . . . . . Cryogenic electron microscopy

DOPE . . . . . Discrete Optimized Protein Energy

ECD . . . . . Extracellular domain

GlcNAc . . . . . N-Acetylglucosamine

GM3 . . . . . Monosialodihexosylganglioside

GROMACS . . . . GROningen MAchine for Chemical Simulations

HF . . . . . Hartree-Fock

ICD . . . . . Intracellular domain

Ig . . . . . Immunoglobulin

Kir . . . . . Inward rectifying potassium channel

Kv . . . . . Voltage-gated potassium channel

LJ . . . . . Lennard-Jones

MD . . . . . Molecular dynamics

mH1 . . . . . Adult isoform of Nav1.5

MM . . . . . Molecular mechanics

MSA . . . . . Multiple sequence alignment

Nav . . . . . Voltage-gated sodium channel

Nav1.5e . . . . . Embryonic isoform of Nav1.5

NMR . . . . . Nuclear magnetic resonance

PA . . . . . Principal axis

PBC . . . . . Periodic boundary conditions

PC . . . . . Phosphatidylcholine

PDB . . . . . Protein Data Bank

PE . . . . . Phosphatidylethanolamine

PI . . . . . Phosphatidylinositol

PIP_2 . . . . . Phosphatidylinositol 4,5-bisphosphate

PME . . . . . Particle mesh Ewald

PM . . . . . Plasma membrane

POPC 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine

POPE 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphatidylethanolamine

POPS 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoserine

PS . . . . . Phosphatidylserine

QMEAN . . . . . Qualitative Model Energy ANalysis

QM . . . . . Quantum mechanics

RDF . . . . . Radial distribution function

RMSD . . . . . Root-mean-squared deviation

Sph . . . . . Sphingomyelin

TARP . . . . . Transmembrane AMPA receptor regulatory protein

TMD . . . . . Transmembrane domain

UC . . . . . Unit cell

VdW . . . . . Van der Waals

VGIC . . . . . Voltage-gated ion channel

V-set . . . . . Immunoglobulin-like variable domain

VSD . . . . . Voltage sensor domain

WT . . . . . Wild type

"Go placidly amid the noise and the haste,
and remember what peace there may be in silence."

— Max Ehrmann, *Desiderata*

1 Introduction

Contents

1.1 Overview 2
1.1.1 Biological membranes 3
1.1.2 The action potential 5
1.1.3 Neuronal vs. cardiac action potential 8
1.1.4 Cellular structure: Neurons vs. Cardiomyocytes 10
1.2 Voltage-gated sodium channels 13
1.2.1 Structure of the α subunit 15
1.2.2 Structure of the β subunit 18
1.3 Binding between α & β subunits 20
1.4 Outline 22
1.5 Work presented in this thesis 23

1.1 Overview

The effective initiation, propagation, and termination of electrical signals is vital for a large number of complex and finely tuned functions in the human body. Within excitable cells, these signals are generated by the coordinated movement of specific ions across the cell membrane. This movement is facilitated by a number of proteins embedded within the membrane that control ion flux in a highly regulated and distinct manner. At rest, a relatively constant concentration of ions across the membrane is achieved by integral membrane proteins that set up a resting membrane potential. However, upon excitation another class of integral membrane proteins, named voltage-gated ion channels (VGICs), can disrupt this finely tuned balance. If the local resting membrane potential undergoes depolarisation, VGICs rapidly transport ions with high specificity across the cell membrane, inducing further changes in membrane potential, that ultimately facilitate cell-to-cell communication.

Two types of VGIC, named voltage-gated potassium (Kv) and sodium (Nav) channels, are critical for the passing of potassium (K^+) and sodium (Na^+) ions across the cell membrane respectively. Structurally, Kv channels were one of the first channels to be resolved due to their relatively simple homo-tetrameric architecture [1]. In contrast, and until recently, Nav channels lagged behind in their structural determination. This delay is primarily due to the pore-forming α subunit of Nav channels being made up from one continuous peptide chain [2] that provides a real challenge for traditional crystallisation techniques. Fortunately, recent advances in cryogenic electron microscopy (cryo-EM) are now increasing the number of Nav channel structures available. Not only are these structures elucidating previously ambiguous structural details of the α subunit, but an increasing number are found in complex with additional transmembrane proteins, named β subunits [3–5]. Although not ion selective, these auxiliary subunits influence a variety of Nav channel properties such as modulation of α subunit gating, recovery from channel inactivation, and influencing the

spatio-temporal dynamics of channel clustering [6–8]. Currently, the characterisation of these β subunits remains limited and a full description of their interactions with Nav channels, other VGICs, and the various lipid species found in complex cell membranes remains an open question.

Molecular dynamics (MD) simulations offer unparalleled analysis of protein dynamics at levels of detail relevant to various experimental techniques. Many aspects of VGIC behaviour have been well characterised through this technique such as ion permeation [9] and channel clustering [10]. Not only are simulations now used to verify experimental observables, but they are increasingly able to provide new avenues for experimental research to pursue.

Within this chapter, a description of biological membrane composition and excitable cell dynamics is given. Differences between cell types and their corresponding architectures are described, before the general structures of Nav channel subunits (and protein - protein interactions therein) are introduced.

1.1.1 Biological membranes

All membranes in eukaryotic cells provide a barrier between the cells' interior and the environment in which it occupies. Additionally, membranes present within the cell serve to compartmentalise various organelles, facilitating a number of cellular processes so that it can function correctly. These thin membranes, with a non-polar region often only around 30 - 40 Å in width, are arranged as a bilayer of lipid molecules. In a plasma membrane (PM), the composition of these 'leaflets' are typically made up of glycerophospholipids, sphingolipids, and sterols whose overall distribution depends on cell and organelle type [11]. The long acyl tails of the glycerophospholipids form the hydrophobic core of the bilayer, whilst both sphingolipids and sterols contribute to its fluidity and stability [12, 13]. At the solvent exposed ends of both leaflets are the polar and charged groups of specific lipid headgroups. Between leaflets, there

is often an asymmetric distribution of lipids that leads to plasma membranes being anionic in nature [12–14]. Typically, the upper leaflet of mammalian cells is enriched with phosphatidylcholine (PC), sphingolipids such as sphingomyelin (Sph), and glycosphingolipids such as monosialodhexosylganglioside (GM3). The intracellular leaflet is enriched in phosphatidylethanolamine (PE), phosphatidylserine (PS), and phosphatidylinositols (PI) [13]. The structures of the wide variety of mammalian lipid species are shown in Figure 1.1.

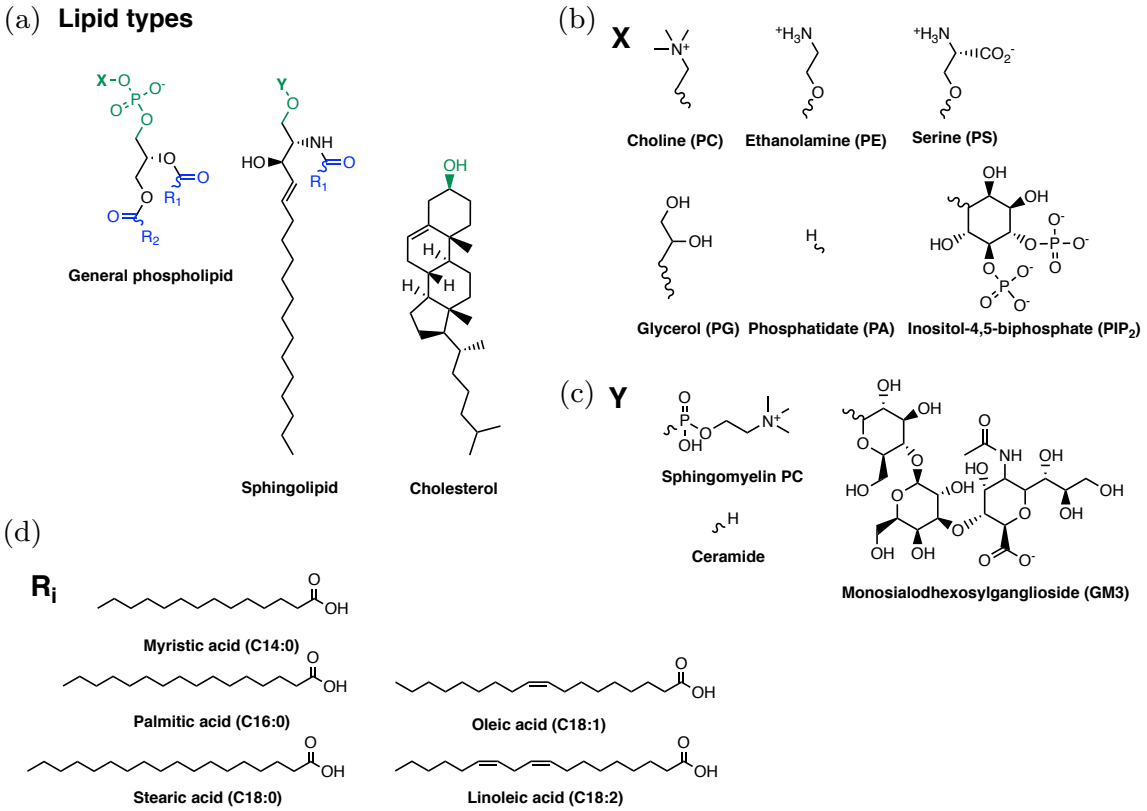


Figure 1.1: The variety of lipids found in a typical mammalian cell membrane [11, 13]. a) The general structure of phospholipid and sphingolipid lipid types found in mammalian cell membranes, as well as sterols such as cholesterol. Polar headgroups and the start of lipid tails are shown in green and blue respectively. b) & c) Various headgroup types present in phospholipids and sphingolipids. d) Common fatty acid tails, both saturated (left) and unsaturated (right), present in lipid tails where lipid numbers represent the number of carbon atoms and double bonds respectively.

1.1.2 The action potential

Excitable cells such as neuronal and cardiac cells (cardiomyocytes), found in the brain and heart respectively, utilise the movement of ions across cell membranes and cause a change in membrane potential. These changes in potential ultimately lead to the conduction of electrical signals throughout the body. Seminal work by Hodgkin and Huxley on the giant squid axon [15] modelled this movement, describing the cell membrane as a leaky capacitor capable of conducting various ions, namely Na^+, K^+, and chloride (Cl^-).

Cell membranes possess an inherent low permeability to charged ions. Typically, high concentrations of Na^+ and Cl^- ions outside (145 mM and 123 mM, respectively) and K^+ ions inside (155 mM) the cell are established [16]. This is maintained by ion pumps (transmembrane proteins that span both leaflets of the membrane) via active transport, such as sodium potassium pumps [17]. The resulting concentration gradient created is responsible for the resting membrane potential, as briefly mentioned in Section 1.1. When ions move rapidly across the membrane via ion channels, a sudden change in membrane potential is induced; the propagation of which is known as the action potential (AP). The AP works on an all-or-none principal, that is, the cell will produce the same AP regardless of stimulus strength [18]. Under no stimulus the resting membrane potential is typically between -40 and -95 mV, depending on the type of cell. Upon excitation, and if a threshold potential is reached, the cell undergoes rapid depolarisation and shifts the membrane potential to more positive values. This is followed almost instantaneously by hyperpolarisation of the cell that typically overshoots the resting potential to more negative values. Following hyperpolarisation, the cell slowly returns to its resting value during which no further APs can be produced (known as the refractory period).

The passage of Na^+ and K^+ ions across the cell membrane underpins each stage of the AP and is facilitated by Nav and Kv channels respectively, that span both

leaflets of the lipid membrane. Both Nav and Kv channels respond to changes in local membrane potential by altering their conformation to permit the flux of Na^+ and K^+ ions across the membrane. In neurons, Nav and Kv channels are closed at a resting potential of -70 mV. When the threshold membrane potential of -55 mV is reached Nav channels open and allow the high concentration of Na^+ ions outside of the cell to enter. The influx of positively charged Na^+ ions causes the membrane potential to rise rapidly. Shortly after, Nav channels undergo fast inactivation (where the pore is physically blocked by a set of conserved residues) and the passage of Na^+ ions stopped. This is followed by the opening of Kv channels, causing K^+ ions to move out of the cell to begin repolarisation. Due to the kinetics of Kv channels being slower than that of Nav, a refractory period is present that delays the return of the membrane potential to the resting potential. A visual representation of the AP process is shown in Figure 1.2.

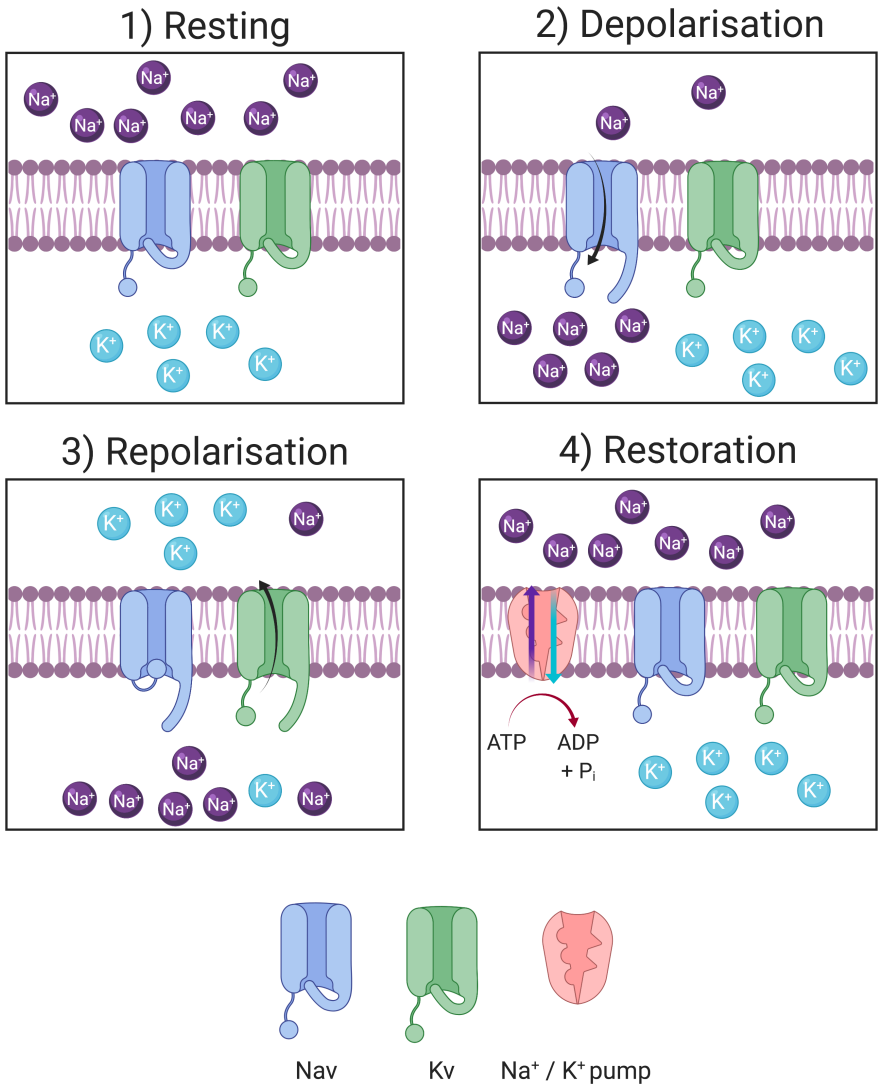
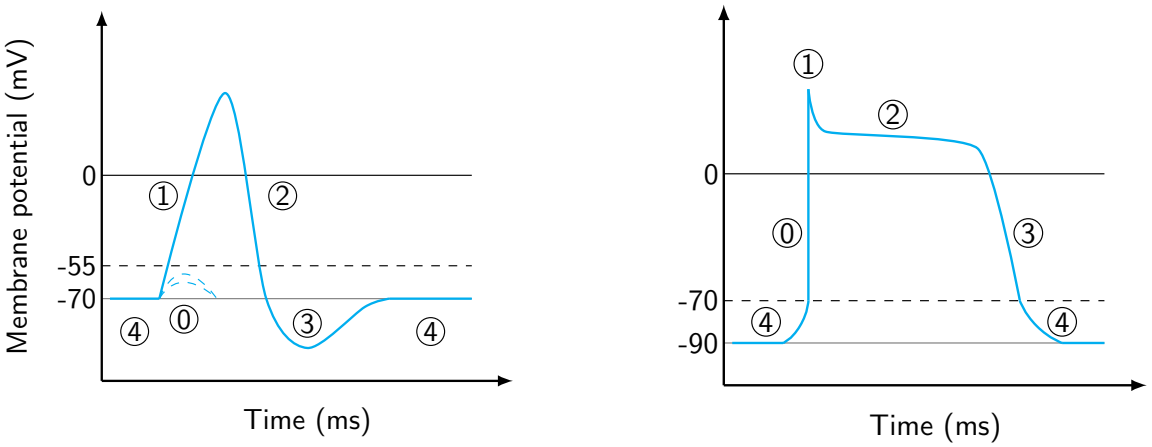


Figure 1.2: Cycle of ion movement during an action potential. (1) At the resting membrane potential both Nav (blue) and Kv (green) channels are closed. (2) Nav channels open to facilitate depolarisation of the cell, with Na^+ entering the cell. (3) Rapid repolarisation of the cell occurs with Nav channels undergoing fast inactivation (indicated via a "ball-and-chain" cartoon, where the "ball" physically blocks the pore) and Kv channels open. This causes K^+ ions to rapidly leave the cell before Kv channels undergo inactivation (not shown). (4) Eventually both Nav and Kv channels enter a closed state. After the AP is complete, the sodium-potassium pump (red) works against the concentration gradient (via the hydrolysis of ATP) to restore the resting membrane potential, this is achieved by moving three Na^+ ions out of the cell for every two K^+ ions transported into the cell. Purple and cyan arrows represent the movement of Na^+ and K+ ions against their concentration gradients respectively. ATP = adenosine triphosphate, ADP = adenosine diphosphate. Figure created with BioRender.com.

1.1.3 Neuronal vs. cardiac action potential

In cardiac cells, the AP shares similar features to the neuronal AP but also contains some important differences. These differences are clearly seen in the AP profile of atrial and ventricular myocytes[†] (Figure 1.3b). Firstly, the cardiac resting membrane potential lies at a more negative value of -90 mV. At a threshold potential of -70 mV Nav channels open, causing Na^+ ions to flow into the cell shortly followed by their rapid closure and the opening of Kv channels (analogous to the AP seen in neuronal cells). One major difference present in the cardiac AP is the subsequent plateau phase. This phase is maintained by the presence of delayed rectifier Kv channels that allow K^+ ions to leave the cell interior after a delay from the onset of membrane depolarisation. Also active during the plateau phase are long lasting voltage-gated calcium channels (L-type Cav) that cause calcium ions (Ca^{2+}) to enter the cell. The influx of Ca^{2+} ions subsequently bind to ryanodine receptors that are responsible for heart muscle contraction [19]. The balance of both the outward and inward flow of K^+ and Ca^{2+} ions keeps the membrane potential at an almost constant positive value over the course of several hundred milliseconds. Repolarisation occurs with the closing of Cav channels and the efflux of K^+ ions from the cell via both rapid and delayed Kv channels. Eventually the resting membrane potential is restored to -90 mV by Na^+ and Ca^{2+} being transported out of the cell, via ionic pumps, and K^+ restored to being inside the cell via inward rectifying potassium channels (Kir). Both the neuronal and cardiac AP profiles are shown in Figure 1.3.

[†]Other types of cardiomyocyte, called cardiac pacemaker cells, are critical in the generation of the cardiac AP that leads to coordinated contraction. However its AP is similar to that of neuronal cells (with the additional dependence on Ca^{2+} ions) and as a result will not be discussed here. Hereafter the term 'cardiomyocyte' will refer exclusively to atrial and ventricular myocytes.



(a) Neuronal

(b) Cardiac

Figure 1.3: Action potentials of neuronal and cardiac cells. (a) Typical AP profile of a neuronal cell: (4) Membrane potential at rest (-70 mV). (0) Upon an external stimulus, Nav channels open if the threshold potential (-55 mV) is reached (failed APs are shown as dashed lines). (1) Rapid depolarisation of the cell as Nav channels open, at the peak Nav channels undergo fast inactivation. (2) Almost simultaneously with Nav channel inactivation, there is rapid repolarisation of of the cell with the opening Kv channels. (3) Refractory period where a new AP cannot be initiated. (b) Typical AP profile of a atrial / ventricular myocyte: (4) Membrane potential at rest (-90 mV). (0) Nav channels open in a similar fashion to neuronal cells but at a threshold potential of -70 mV. (1) Initial repolarisation of the cell with the inactivation of Nav channels and opening of Kv channels that set the membrane potential for the plateau phase. (2) Plateau phase where Cav and Kv channels open, Ca^{2+} ions flowing into the cell cause a large intracellular release of Ca^{2+} ions through activation of the Ryanodine receptor. (3) Repolarisation of the cell to the resting potential via the closing of Cav channels and restoration via the Kv channels. **Note**: The timescales are not equal for the neuronal and cardiac AP, the neuronal AP is fast and typically occurs over a few ms whilst the cardiac AP occurs over a much longer scale of approximately 300 ms.

1.1.4 Cellular structure: Neurons vs. Cardiomyocytes

The differences in the AP profiles of neurons and cardiomyocytes is a direct consequence of their unique cellular structures. Neurons are broadly classified into three types: sensory, motor, or interneuron that play key roles in sensing, muscle contraction, and neuronal communication respectively. These simple classifications do not cover all types of neuron, of which there are a large variety [20], however the basic structure of each follows the same principles. The structure of a typical neuron is shown in Figure 1.4. Input signals to the neuron are received at dendrites that propagate the resulting change in membrane potential towards the axon hillock, a region of the central neuron body (soma) located just before the axon. Clusters of ion channels (such as Nav) at the axon hillock initiate the start of the AP if the combination of incoming potentials exceeds the threshold potential, following the all-or-none principle (as described in Section 1.1.2). The induced increase in membrane potential to more positive values causes a succession of Nav channels to open along the length of the axon and the AP to propagate towards the axon terminal. Transmission of the AP is retained via myelin-coated segments of the axon that act as insulators to prevent the leakage of current. The nodes of Ranvier occupy unmyelinated regions between neighbouring myelin sheaths and contain a high density of ion channels to ensure that the AP is conducted rapidly through "jumps" between nodes, known as saltatory conduction [21, 22]. Once the AP has reached the axon terminal, local depolarisation causes Ca^{2+} ions to be released into the presynaptic terminal, leading to the eventual release of neurotransmitters across the synaptic junction. These small molecules bind to ligand-gated ion channels in the post synaptic membrane to transmit the AP to the neighbouring neuron.

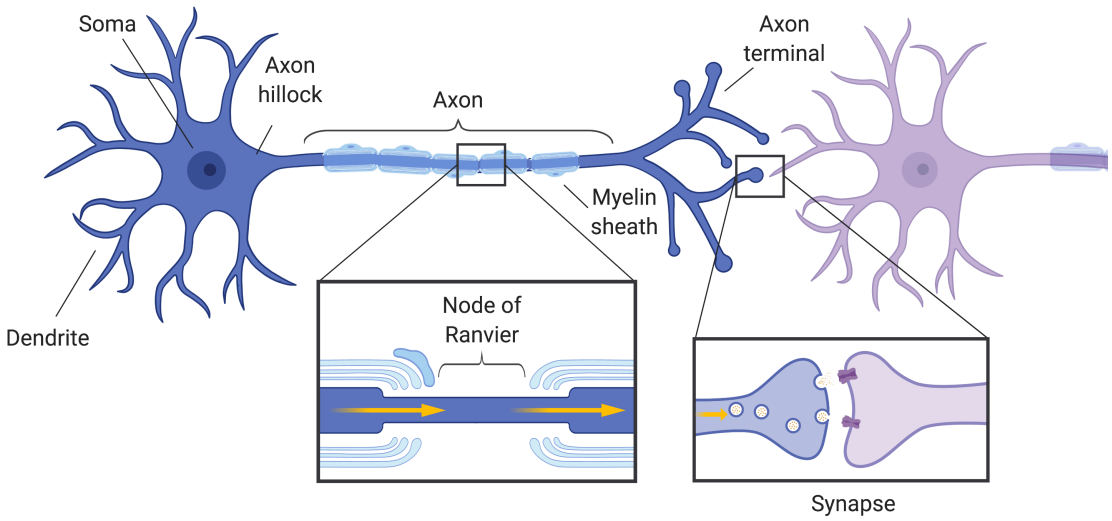


Figure 1.4: General features of a typical neuron. Yellow arrows indicate an AP travelling down the axon, with vesicles in the presynaptic terminal shown as yellow spheres. Figure created with BioRender.com.

Unlike neurons, cardiomyocytes are electrically coupled to neighbouring cells via intercalated discs. These localised regions between cells contain an assembly of proteins that form protein tunnels between neighbouring cell membranes, known as connexons [23], and form the gap junction. As a result, when a neighbouring cell is depolarised the ions are free to pass through to the next and induce further depolarisation via the activation of Nav channels exposed to the extracellular space. This facilitates the uninterrupted propagation of the AP between neighbouring cardiomyocytes and causes heart muscle to contract efficiently in a highly coordinated fashion. A diagram of features present within cardiomyocytes is shown in Figure 1.5.

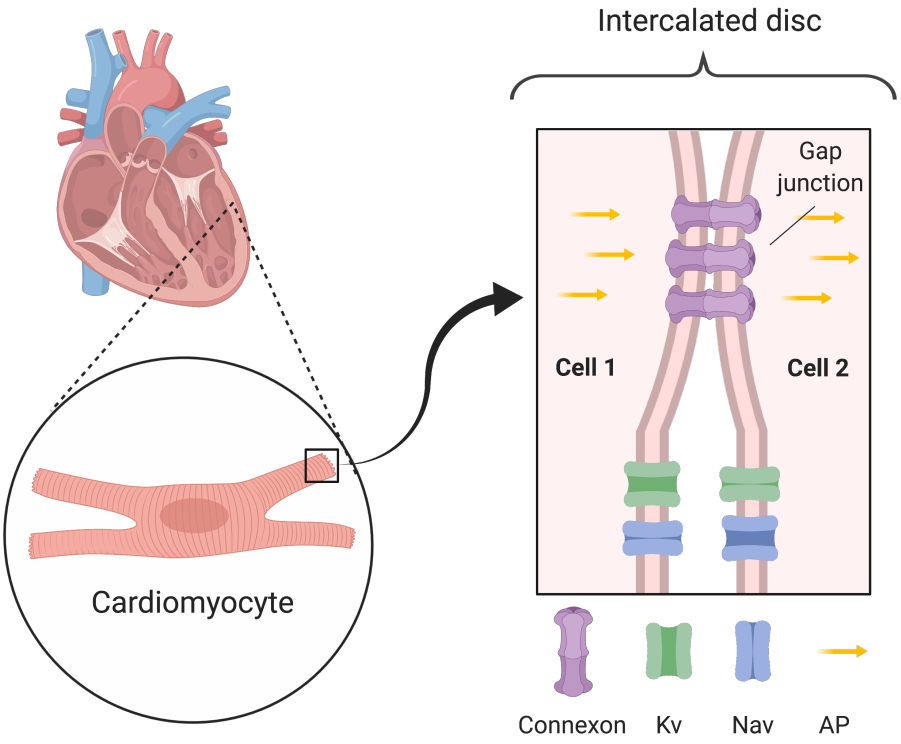


Figure 1.5: Action potential through cardiomyocytes. Yellow arrows indicate an AP travelling from one cardiomyocyte (cell 1) to another (cell 2) via a cluster of connexons (purple) at the gap junction, located within the intercalated disc. Nav (blue) and Kv (green) channels are shown in both the open and closed states between neighbouring cells. Figure created with BioRender.com.

1.2 Voltage-gated sodium channels

In both neuronal and cardiac cells, Nav channels play a critical role in depolarisation of the membrane potential. The importance of Nav channels in excitable cells is evident in their wide distribution throughout the nervous system [24]. As a result, Nav channels are also implicated in many disease and pathological states including cardiac arrhythmia [25], epilepsy [26,27], neuropsychiatric disorders [28], and chronic pain [29,30]. Historically, the family of Nav channels were one of the first VGICs to be discovered [31–34] and are regarded as the founding members of the large VGIC superfamily [35]. Each of the nine Nav channels (Nav1.X, where X = 1-9) contains an α subunit that is made up of four homologous domains [36] surrounding a central pore with a total molecular weight of 240-260 kDa [32,37].

Diversity within the Nav channel family is achieved through different isoforms that are encoded by specific genes (SCNXA where X = 1-5, 8-11), facilitating their utilisation across many different tissue types (Table 1.1). Broadly speaking, Nav channels can be separated into two groups, based on phylogenetic branching, with Nav1.1, Nav1.2, Nav1.3, and Nav1.7 being the most closely related [38] and typically found in brain tissue. The second group consists of Nav1.5, Nav1.8, and Nav1.9 that are also close phylogenetically [38] and share similar resistance to tetrodoxin [39,40]. Both Nav1.4 and Nav1.6 are present on their own phylogenetic branch, distinct from the others, as well as being distinct from themselves [38].

Nav channel	Typical localisation
Nav1.1, 1.2, 1.3	CNS [41–45]
Nav1.4, Nav1.5	Skeletal muscle [46] & Heart [47]
Nav1.6	CNS / PNS [42, 43, 48, 49]
Nav1.7, Nav1.8, Nav1.9	PNS [50–52]

Table 1.1: Localisation of Nav channels. CNS = Central nervous system, PNS = Peripheral nervous system. In-table references are adapted from Duménieu *et al.* [53], where the reader is directed for an in-depth review of Nav channel tissue expression.

Although the α subunit can function on its own, it is typically not found in isolation *in vivo*. So-called "auxiliary" β subunits are known to play multifunctional roles in Nav channel dynamics and may well offer alternative therapeutic routes for a variety of channelopathies [54]. In humans, there are four types of β subunit (β1-4) that are encoded for by the SCNXB gene (where X = 1-4). Additionally, a splice variant of SCN1B gives rise to the soluble β1B isoform [55]. As with the α subunit, various β subunits are expressed in a variety of tissue types (Table 1.2). It is known that β subunits function both independently and directly with VGICs, where they can alter their gating properties and expression. For example, β subunits have been shown to affect Nav channel inactivation kinetics [56–58] as well as channel density [58–60].

The binding of each β subunit to the pore-forming α subunit is dictated by the presence (β2 & β4) or absence (β1 & β3) of a free cysteine bond in the extracellular immunoglobulin (Ig) domain (the structure of which is described later, in Section 1.2.2), that facilitates either covalent [61, 62] or non-covalent [63–65] interactions respectively. The Ig scaffold is homologous to the variable domain (V-set) motifs present in the Ig superfamily of cell-adhesion molecules (CAMs) [60, 63, 66, 67]. The adhesive properties of the β subunit's Ig domain give rise to a number of non-conducting processes that include interactions with extracellular proteins such as tenascin-C/R [68], participation in homo- and heterophilic binding [69–71], as well as the promotion of

neurite outgrowth and cellular aggregation [6]. This is in addition to the β subunit's recruitment of the cytoskeletal protein ankyrin-G to points of cell-to-cell contact [69]. Interestingly, Nav α subunit interactions with ankyrin-G have also been shown to be required for channel clustering and efficient AP propagation at the axon initial segment [72, 73].

	β subunit				
Location	β1	β1B	β2	β3	β4
Nervous system					
Hippocampal neurons	[74]		[58, 75]	[76, 77]	[75, 78]
Cortical neurons	[79]	[80, 81]	[58, 75]	[76, 77]	[75, 78]
Purkinje neurons	[82]	[80, 81]	[58, 75]	[77]	[75, 78]
Cerebellar granule neurons	[82, 83]	-	-	-	-
Nodes of Ranvier	[71, 74]	-	[58]	-	[78]
Dorsal root ganglia	[59, 84]	[80, 81]	[59, 75, 85, 86]	[76, 77, 85, 87]	[75, 78, 85]
Heart					
Ventricular myocytes	[88]		[88]	[88]	[88]
Atrial myocytes	[89, 90]	-	[90]	[90]	[90]
Sino-atrial node	[91]	-	[91]	[91]	[91]

Table 1.2: Localisation of β subunits. Table adapted from O'Malley *et al.* [8].

1.2.1 Structure of the α subunit

The structure of the α subunit is shown in Figure 1.6. Each of the four domains contain six transmembrane helices (S1-6), where S1-4 form the voltage-sensor domain (VSD)[†] and S5-6 contribute to the pore-forming domain. The S4 helix in each VSD is lined with four positively charged residues (arginine and lysine) that are extremely sensitive to changes in local membrane potential (i.e. voltage) across the membrane and are primarily responsible for the channel's voltage-sensing. The outward movement of the S4 helix to the extracellular side of the membrane induces a conforma-

[†]For brevity, VSDs of domain X (where X = I-IV) will be abbreviated as VSDX hereafter.

tional change in the S5 and S6 helices through the S4-S5 linker, causing the pore to open. Upon pore opening, ion conduction is governed by the membrane-reentrant pore loops (P) located between the S5 and S6 helices. The P-loop of each domain contributes to the channel's selectivity filter, where aspartate (P-loop I), glutamate (P-loop II), lysine (P-loop III) and alanine (P-loop IV) form a DEKA motif that is critical in Na^+ selectivity over K^+. The region in the α subunit responsible for inactivation is the isoleucine, phenylalanine, methionine (IFM) motif, located between the S6 and S1 helices of DIII and DIV respectively [92, 93]. Although all domains are homologous, results from experiment have challenged this. In contrast to activation, where DI-III activate with similar kinetics, DIV is markedly slower and contributes most to channel inactivation [94].

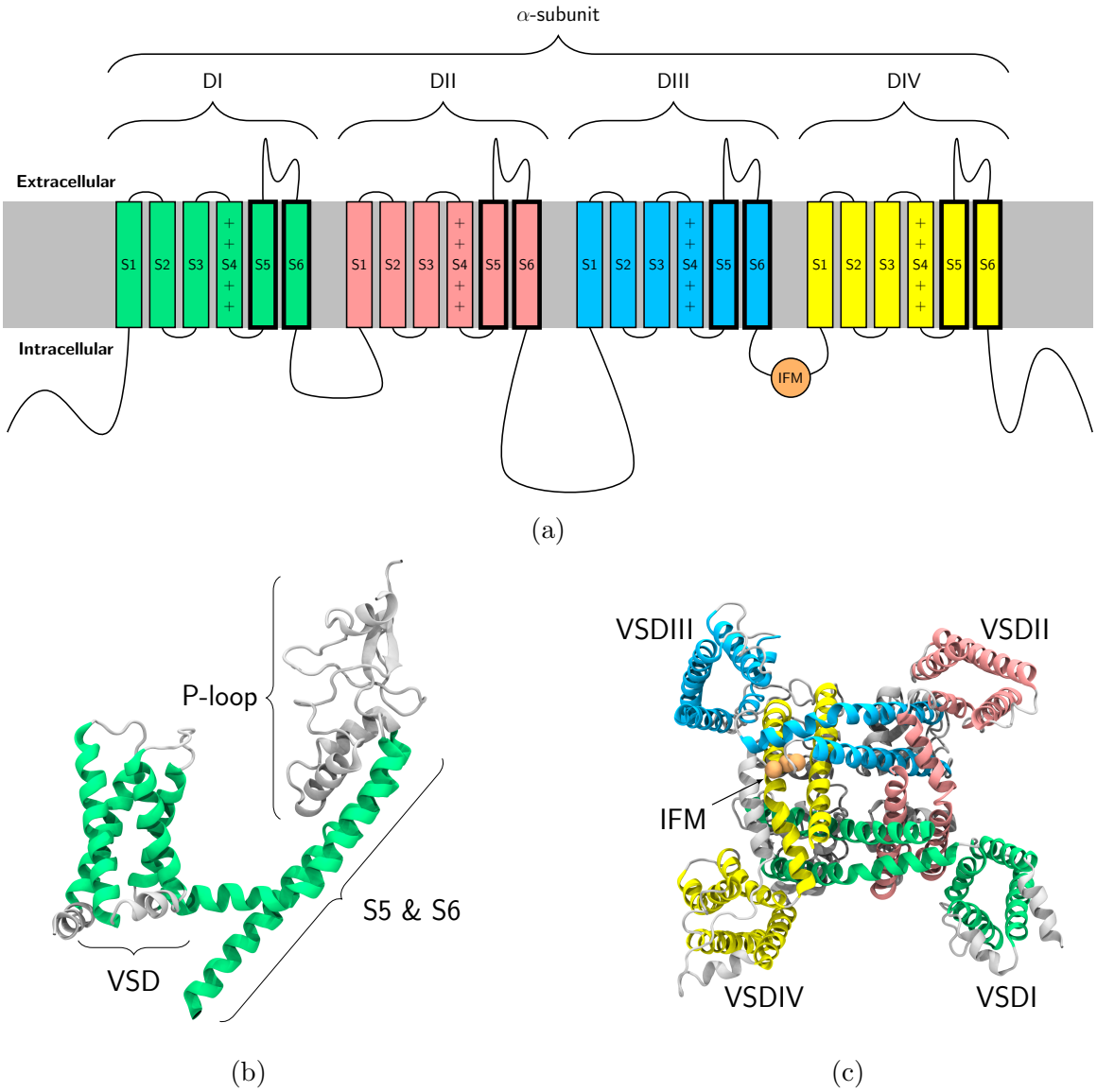
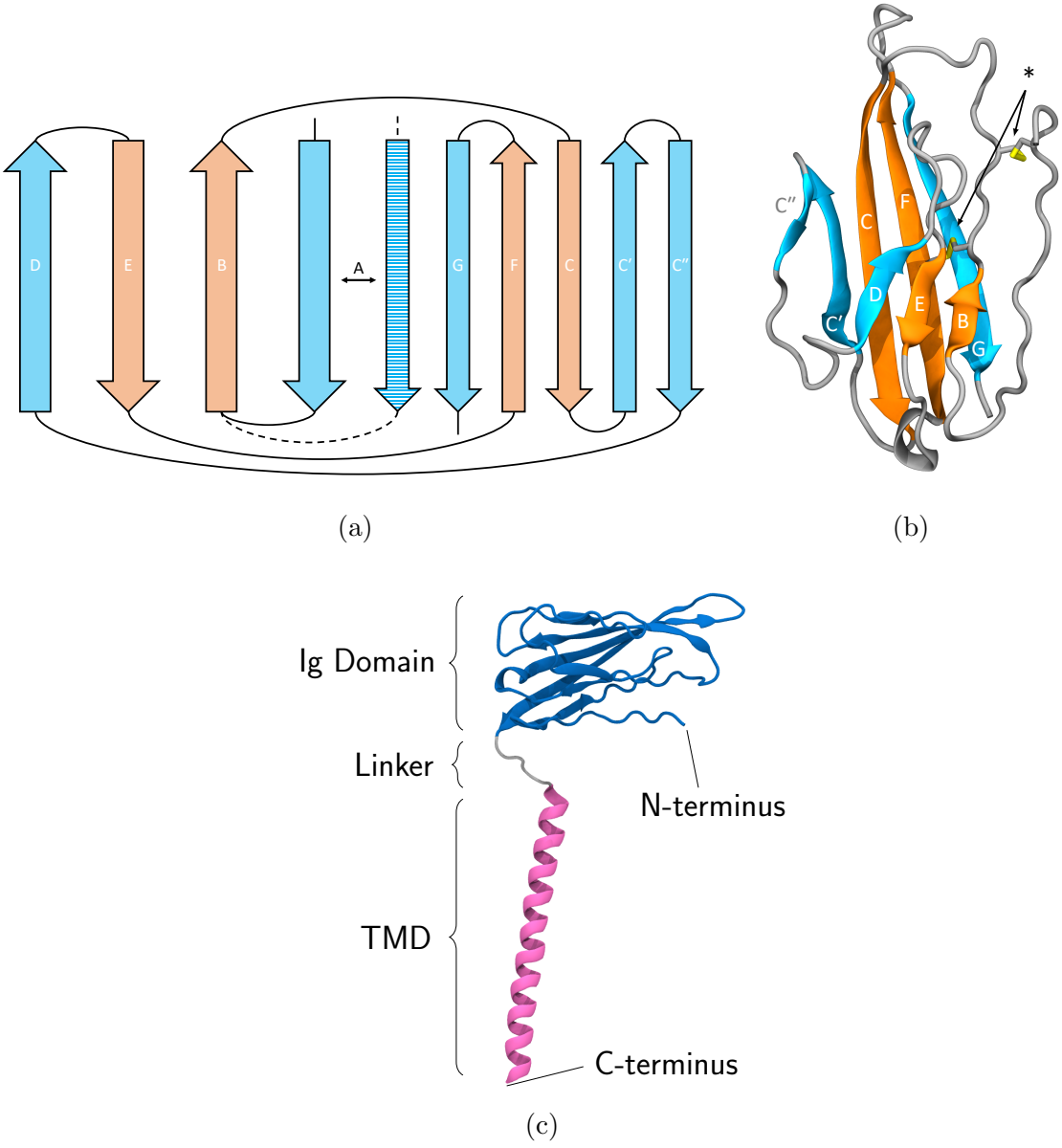


Figure 1.6: Structure of the Nav α subunit. (a) Schematic of the Nav α subunit. Domains I-IV are coloured green, pink, cyan, and yellow respectively. The positive charges present in the S4 helices in each domain are indicated with '+' symbols. The extended loop regions containing the selectivity filter and DEKA motif are shown between the pore lining S5 and S6 helices (outlined in bold) in each domain. The IFM motif (responsible for inactivation) is shown as an orange circle between the S6 and S1 helices of DIII and DIV respectively. The lipid bilayer is indicated by a grey rectangle. (b) Structure of DI with key regions indicated. (c) Intracellular view of the entire α subunit with the residues of the IFM motif indicated as orange spheres. Structures taken from Pan *et al.* [4] (β1 subunit not shown) with domains coloured as in (a), all loop regions are coloured in light grey.

1.2.2 Structure of the β subunit

With the exception of β1B, all β subunits are single-pass transmembrane proteins consisting of an Ig-type extracellular domain (ECD), with a short linker region to a single transmembrane helix in the transmembrane domain (TMD) and an intracellular domain (ICD) that is predicted to be unstructured.

Typically, the Ig-fold contains two β sheets that consist of anti-parallel β strands, these in turn form a central hydrophobic core, stabilised by a buried disulphide bond between neighbouring cysteine residues [95]. The β subunit V-set Ig domain generally follows a Greek-key topology [96,97] and due to their abundance has several labelling systems [98] that reflect the hydrogen bonding pattern of β strands. A 2D topology map of how V-set Ig domains are constructed is shown in Figures 1.7a & 1.7b. It is important to note that in β1 and β3 subunits, the A β strand is not present and instead, a conserved disulphide bridge stabilises the region. For an in-depth overview of Ig domain architecture the reader is directed to a detailed review by Bork *et al.* [98]. A short linker region, made of typically ≈ 7 residues, connects the Ig domain to the transmembrane helix. The overall structure of the Ig domain and full β subunit is shown in Figure 1.7.



(a) (b)

(c)

Figure 1.7: The structure of the β subunit. (a) Schematic of the V-set, β strand rich, Ig domain common to all Nav β subunits. Thick and narrow arrows correspond to β strands on different faces of the domain with orange and cyan colouring representing common and non-common cores across all families of Ig domain respectively. The 'A' strand is present on either face in different V-set Ig domains, as indicated by the striped and filled arrows. Image adapted from Bork *et al.* [98]. (b) The structure of the Ig domain in the β3 subunit, with conserved disulphide bridges highlighted as yellow sticks and indicated by an asterisk. Colouring and labels of the β sheets are the same as in (a) with a notable lack of the 'A' β strand. (c) Full length structure of the β1 subunit with the Ig, linker, and transmembrane domain indicated in blue, grey, and pink respectively. The ICD is predicted to be unstructured and has currently not been resolved.

1.3 Binding between α & β subunits

Until recently, the definitive structure of full length β subunits TMD was unknown, in particular the precise location of the TMD could only be predicted via the sequence similarity of related proteins. Further clarification was given by the publication of the eukaryotic Nav1.4 structure with a β1 subunit bound from Electric Eel [3], solved at a resolution of 4 Å. Here, the fully resolved β1 subunit interacts via its TMD and Ig domain with VSDIII and extracellular loops of the α subunit respectively. Shortly after, the nearly identical human structure of Nav1.4 was solved [4] with β1 again bound to VSDIII at an improved resolution of 3.2 Å. In another cryo-EM structure this time with the human Nav1.7 α subunit, not only could the position of β1 be resolved, but also the position of β2 (albeit only the Ig domain) and various toxin molecules [5]. The length of the TMD helix was revealed to be much longer than initially predicted and further clarifies potential binding sites that β subunits can occupy around the α subunit. These structures (shown in Figure 1.8) support the historic working hypothesis that VGIC structures operate as heterotrimers [34]. However, recent results from atomic force microscopy (AFM) revealed that the α subunit can in fact be decorated with several β3 subunits at once and can facilitate cross-linking events between neighbouring channels [99]. The same AFM results also highlighted the ability of α subunits to cross link with themselves (independently of the β3 subunit) and suggest an intrinsic propensity to oligomerise, a result supported by other experimental results that find α subunits can assemble and gate as dimers [100].

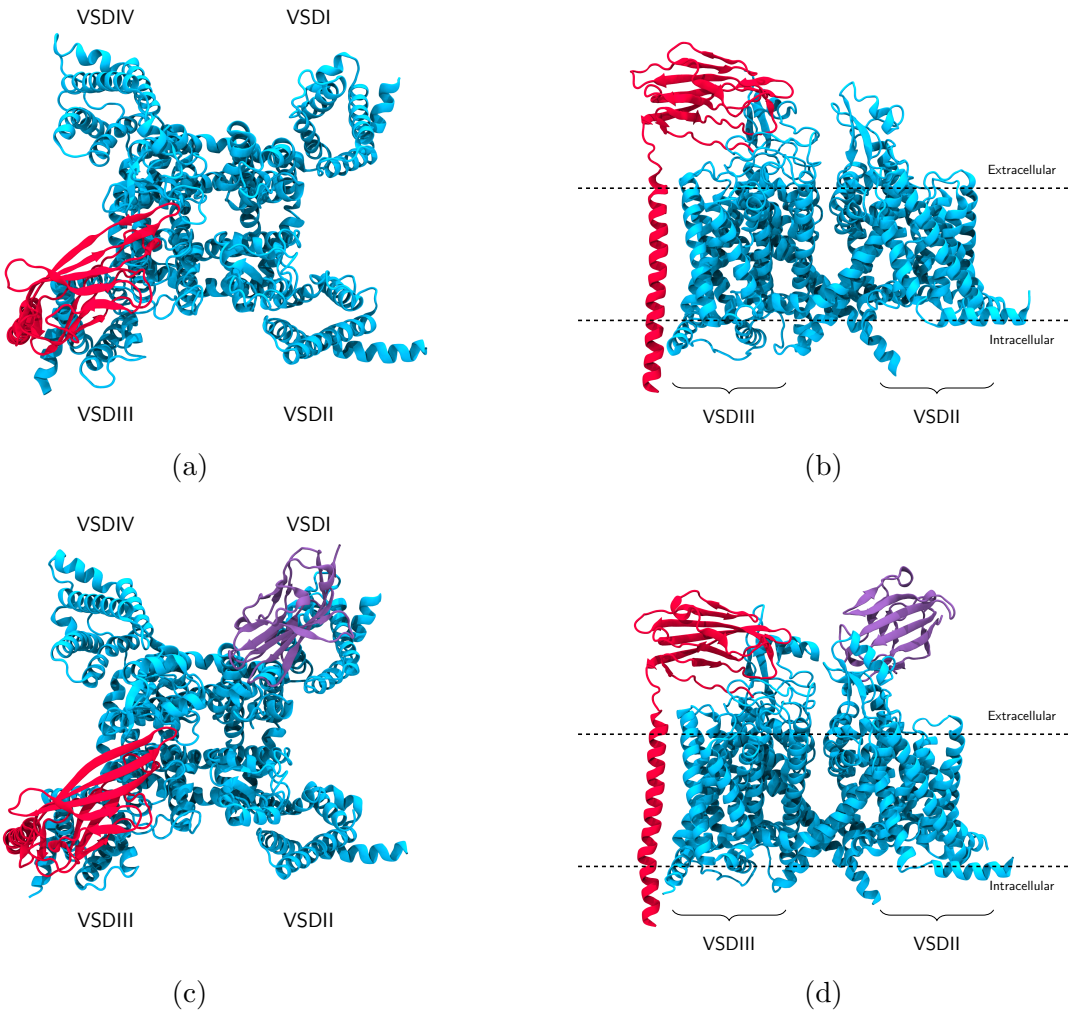


Figure 1.8: Crystal structures of Nav α - β complexes. Pan *et al.* [4]: Extracellular (a) and side view (b) of human Nav1.4 in complex with β1, bound at the VSDIII. Shen *et al.* [5]: Extracellular (c) and side view (d) of human Nav1.7 in complex with β1 and β2, bound at the VSDIII and VSDI respectively. α, β1, and β2 subunits are shown in cyan, red, and purple respectively. Dashed lines in (b) and (d) indicate the approximate location of the lipid membrane.

Understanding the interactions between the α and β subunits of Nav channels may help elucidate the structural basis for a multitude of Nav channelopathies, in particular the often hard to treat diseases of the heart and the corresponding Nav1.5 isoform. For example, people with mutations in the β3 subunit show cardiac conduction problems [6], and in mice the deletion of SCN3B leads to cardiac arrhythmias [101, 102] as well as being linked to Brugada syndrome [103].

1.4 Outline

Clearly, the interactions between the α and β subunits of Nav channel are complex, affecting both gating properties as well as the spatio-temporal dynamics of individual channels. Experimental studies have been successful in identifying the likely interaction sites between subunits. However, a detailed description of α - β interactions, and in particular β subunit dynamics in isolation of other channels, remains to be fully characterised. Investigation of such behaviour is well-suited to MD simulations that can explore the likely dynamics of individual proteins in both simple and increasingly realistic membrane compositions.

Recent evidence from Namadurai *et al.* [99] has challenged the assumption that β subunits are simply "auxiliary" and they may in fact play an important role in Nav channel localisation in various tissue types. Their work was conducted on the cardiac Nav1.5 isoform, a channel that is known to be critical in the rapid depolarisation of cardiomyocytes and initiates heart contraction. Understandably, the pore-forming α subunit of this isoform has often been treated with caution from a pharmaceutical perspective given its tissue location. An increased understanding of β subunits' individual dynamics, in addition to how they cluster both with and without α subunits, may lay the foundation for future research into alternative treatments away from the pore for a myriad of cardiac and related VGIC diseases.

The work conducted here aims to elucidate the behaviour of both the Nav1.5 α subunit and the β3 subunit in a variety of contexts. Firstly, a splice variant of Nav1.5 is examined with key single-point mutations present within the VSD that alters channel gating. Secondly, mutational studies on β3 Ig trimer are performed to address the critical importance of key residues. Thirdly, the dynamics of both the β1 and β3 subunits is investigated which, to date, have not been extensively characterised in the absence of α subunits. Finally, the higher order clustering of both α and β subunits is addressed.

1.5 Work presented in this thesis

In Chapter 2, the theoretical basis, key mathematical concepts and various modelling procedures are introduced that underpin the simulations performed within this thesis.

Chapter 3 focusses on the Nav1.5 α subunit and a particular splice variant that has been found to alter channel gating via shifting of the voltage-dependence of activation. Atomistic simulations are utilised to investigate two mutations in VSDI that were shown to be responsible for an entire observed shift in conductance-voltage recordings.

In Chapter 4, the dynamics of both the $\beta 1$ and $\beta 3$ subunits are investigated, including their likely interaction with the lipid bilayer and orientational analysis of the Ig domain. Additionally, experimental observations of the importance of a conserved disulphide bridge is confirmed along with the construction of a full length $\beta 3$ trimeric model. Protein - protein and protein - lipid interactions are characterised for all systems, as well as the potential important role glycosylation of the Ig domain plays. The stability of the TMDs of both subunits is also investigated as well as the likely contribution a conserved glutamic acid residue makes.

Chapters 5 and 6 use CG MD to investigate oligomerisation of α - α, $\beta 3$ - $\beta 3$ and α - $\beta 3$ subunit binding in a generalised mammalian membrane composition. The potential role different lipid species have on protein clustering is investigated in addition to the conformational states occupied by $\beta 3$ subunits. Binding events between the α and $\beta 3$ subunits indicate the possibility of multiple interaction sites, in agreement with experiment [99], and highlight the potential for cross-linking events between channels.

"Forget this world and all its troubles and if possible its multitudinous Charlatans–everything in short but the Enchantress of Numbers."

— Ada Lovelace

2
Theory & methods

Contents

2.1 Methods in computational chemistry 26
2.2 The molecular mechanics force field 29
2.2.1 Bonded interactions 30
2.2.2 Non-bonded interactions 32
2.3 Molecular dynamics theory 34
2.3.1 Integrators 35
2.4 System setup protocol 37
2.4.1 Periodic boundary conditions 37
2.4.2 Energy minimisation 38
2.4.3 The thermodynamic ensemble: temperature & pressure . . 39
2.4.4 Long-range interactions 43
2.5 Multiscale modelling 47
2.5.1 Coarse-grained molecular dynamics 48
2.6 Homology model generation 52

In this chapter the theoretical basis and methods used in this thesis are explained. MD at both atomistic and CG resolutions are discussed as well as the advantages and disadvantages each provides. Throughout this chapter it is assumed the reader is familiar with Hamiltonian mechanics (for a brief explanation see Appendix A.1).

2.1 Methods in computational chemistry

The structures of proteins offer a wealth of information about how amino acid sequence can translate to a three-dimensional structure [104]. However, these only offer a static view of the protein and how these structures behave in their local environment, over various time and length scales, is important to understand how they may function *in vivo*. Computational methods, often in parallel with experiment [105], can be used to probe how proteins carry out their function. As with any experiment, the choice of technique must be appropriate for the scientific question and in simulation, this comes down to the correct choice of time and length scales, as well as what is most computationally efficient.

At the smallest scale, methods based in quantum mechanics (QM) offer unparalleled detail into molecular mechanisms. At their heart, QM methods work with the electronic wavefunction, Ψ, motivated by the wave-like behaviour of the electron. The behaviour of Ψ is described by the Schrödinger equation [106]:

$$\hat{H}|\Psi\rangle = E|\Psi\rangle \tag{2.1}$$

Where $\hat{H}$ is the quantum mechanical Hamiltonian, $|\Psi\rangle$ is the state of the system, and E is the energy of the system. For a molecules with n electrons and N nuclei (with charges Z_I) the expression for $\hat{H}$ can be further expanded as:

$$\begin{aligned}\hat{H} &= \left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(\mathbf{r})\right] \\ &= -\frac{\hbar^2}{2\mu}\sum_{i=1}^{n}\nabla_i^2 - \sum_{i=1}^{n}\sum_{I=1}^{N}\frac{(Z_I e)e}{4\pi\epsilon_0 \mathbf{r}_{iI}} + \sum_{i<j}^{n}\frac{e^2}{4\pi\epsilon_0 \mathbf{r}_{ij}}\end{aligned} \tag{2.2}$$

The terms in Equation 2.2 correspond to the kinetic energy, electron - nuclear attraction, and interelectronic repulsion respectively. Where $\hbar$ is the reduced Planck's constant, μ is the reduced mass of the electron, ∇^2 is the Laplace operator, e is the charge of the electron, $\mathbf{r}_{iI}$ is the separation of the ith electron and I^{th} nucleus, ϵ_0 is the permeability of free space, and $\mathbf{r}_{ij}$ the separation of electrons i and j.

The complexity of these equations can be reduced by decoupling the molecular system into electronic and nuclear parts. The separation of electronic and nuclear motion in electronic structure calculations is known as applying the Born-Oppenheimer (BO) approximation [107] and is motivated by the large difference in mass between electrons (e_m) and nuclei (Z_m) where $Z_m : e_m \approx 2000$ [108]. Application of the BO approximation has allowed the advancement of a number of methods to calculate properties of the electronic wavefunction, such as Hartree-Fock (HF) [109, 110] or density functional theory [111, 112]. Their in-depth evaluation is beyond the scope of this thesis; for a full description of modern electronic structure theory approaches and underlying theory see a detailed perspective by Sherrill [113] and / or a comprehensive introduction by Helgaker *et al.* [114].

Biological phenomena tend to occur at time and length scales that render the use of QM methods impractical. Although highly accurate, the computational cost for such problems is high and limited to small systems typically made up of hundreds of atoms. Pioneering work by Levitt, Karplus, and Warshel aimed to "bridge the gap" using a combined QM/molecular mechanics (QM/MM) theory [115]. This culminated in the three of them winning the 2013 Nobel prize in Chemistry for "the development

of multiscale models for complex chemical systems". Briefly, properties in the region of interest are calculated using highly accurate QM methods whilst the rest of the molecule is treated with classical mechanics. The latter is described using a molecular mechanics (MM) force field (described later in Section 2.2). The QM/MM method has found particular success in computational enzymology where the active site, usually only involving a few residues, is key to biomolecular function [116].

At larger system sizes the benefits of the QM/MM approach begin to decrease as the cost of calculating a large QM region becomes computationally impractical. On the other hand if one decreases the QM region the system will need an increased MM region at the cost of decreased accuracy. At such system sizes the exact QM description of molecule may not be necessary and how its dynamics change over longer time frames as well as any large conformational changes becomes a more relevant questions. For such problems, MD simulations offer a "computational microscope" [117, 118] on processes at temporal and spatial scales that are often hard to probe experimentally. MD has become a cornerstone of modern computational research. The powerful technique has applications in many aspects of modern science from materials science [119] to protein simulation [118, 120]. One of the first simulations of a protein using such a method was in 1977 where McCammon *et al.* [121] simulated bovine pancreatic trypsin inhibitor, a small globular protein, for a total of 8.8 ps. Since then, efforts in the field coupled with the increase in computational power now allows for much larger system sizes to be routinely simulated for hundreds of nanoseconds or even microseconds [122]. This vast increase in time scales has been enhanced by the utilisation of graphical processing units that are now routinely used by modern software packages [123–127].

2.2 The molecular mechanics force field

The molecular mechanics force field is a crucial set of parameters used to define the functional form of the potential energy function, which describes the interactions between particles in the system. Beginning with a general form this can be split into the bonded and non-bonded contributions:

$$U_{Total} = \sum U(\mathbf{r}^N) \tag{2.3}$$

$$= \sum U_{bonded} + \sum U_{non-bonded} \tag{2.4}$$

Where U represents the potential energy and $\mathbf{r}^N$ the coordinates of N particles in the system of interest. The terms in Equation 2.4 can be expanded in various functional forms. Representation in accordance with the AMBER [128] potential gives:

$$\sum U_{bonded} = \sum_{bonds} k_r(r - r_0)^2 + \sum_{angles} k_\theta(\theta - \theta_0)^2 + \sum_{torsion} \sum_{n} \frac{1}{2} V_n[1 + \cos(n\tau - \delta)] \tag{2.5}$$

$$\sum U_{non-bonded} = \sum_{i} \sum_{i \neq j} 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum_{i} \sum_{i \neq j} \frac{q_i q_j}{\varepsilon_0 r_{ij}} \tag{2.6}$$

Where the terms of the bonded contribution (Equation 2.5) represent the interactions with respect to bonds, angles, and torsions. For the non-bonded contribution (Equation 2.6) each term represents the interactions through Van der Waals (VdW),

and electrostatics respectively. The following sections will now describe the theoretical basis for each of these terms and define all variables.

2.2.1 Bonded interactions

In MD simulations the covalent interactions are modelled by a simple harmonic and often based on Hooke's Law [129] due to its simplicity. Other forms such as the Morse potential [130] exist that capture the physics of bond breaking, this form is often used in the calculation of spectroscopic data where the dynamics of bond breaking is important. The potential between two particles as a result of their bond is given by:

$$U_{bond}(r_{ij}) = \frac{k^r_{ij}}{2}(r_{ij} - r^0_{ij})^2 \tag{2.7}$$

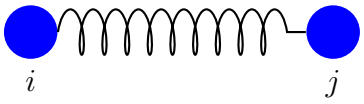


Figure 2.1: Schematic of a simple harmonic potential between particles i and j.

Where r_{ij} is the instantaneous bond length between two particles, k^r_{ij} is the covalent bond strength (i.e. spring stiffness) and r^0_{ij} is the equilibrium bond length. A similar expression can be obtained for the bond angle that now requires three particles to be defined.

$$U_{angle}(\theta_{ijk}) = \frac{k^\theta_{ijk}}{2}(\theta_{ijk} - \theta^0_{ijk})^2 \tag{2.8}$$

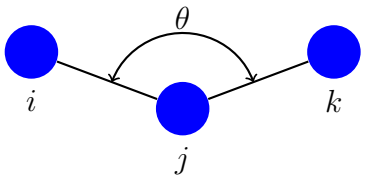


Figure 2.2: Schematic of the angular potential between particles i, j, and k.

Where θ_{ijk} is the bond angle between three particles, k^{θ}_{ijk} is the angle bending constant and θ^{0}_{ijk} is the equilibrium bond angle. Extension of the angular terms to include four particles leads to an expression for the dihedral, or torsional, angle. This is defined as the angle created between two intersecting planes, where one contains particles i, j, k whilst the other contains particles j, k, l.

$$U_{torsion} = \sum_{n} \frac{1}{2} V_n [1 + \cos(n\tau - \delta)] \tag{2.9}$$

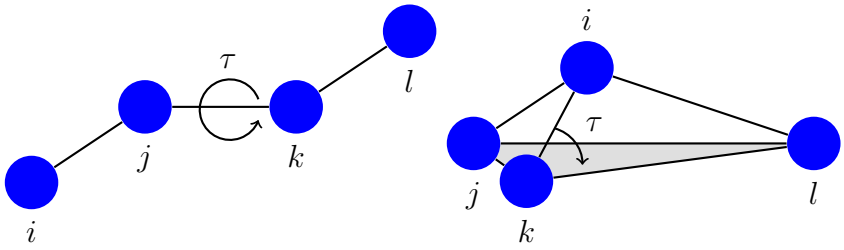


Figure 2.3: Schematic of torsional potentials with proper (left) and improper (right) dihedrals between particles i, j, k, and l.

Equation 2.9 is a Fourier-expansion of the torsional terms where V_n is the torsional barrier, δ is the phase, n is the periodicity and τ is the torsional angle. Often, to save computational effort, the torsional term is set as a truncated Fourier-series, only including a set amount of multiplicities (i.e. n values) [128].

2.2.2 Non-bonded interactions

The non-bonded interactions can be characterised by contributions made from both VdWs and electrostatic terms. Typically the VdW interactions are described with a pairwise potential between pairs of atoms as a function of their internuclear distance, this often takes the form of a Lennard-Jones (LJ) potential [131]:

$$\begin{aligned} U_{VdW}(r_{ij}) &= U_{LJ}(r_{ij}) \\ &= 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \end{aligned} \tag{2.10}$$

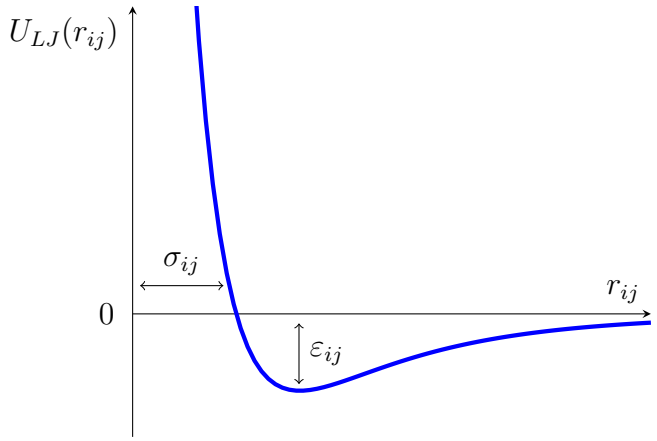


Figure 2.4: Qualitative plot of the Lennard-Jones potential.

Where r_{ij} is the interatomic distance, ε_{ij} is the well-depth and σ_{ij} is the VdW radius (the distance at which the internuclear potential is set to zero). The first and second terms in Equation 2.10 represent the short-range repulsive and long-range attractive forces respectively.

Contributions from electrostatics arise from ionic species and the uneven distribution of charge in molecular systems in the form of partial charges. The use of a non-polarisable force field fixes these values and in AMBER these are derived from restrained electrostatic-potential fitting of high-level *ab initio* QM calculations, such

as HF at the 6-31G* level of theory [132]. If needed, the partial charges can be changed during the course of a simulation as the system evolves and is implemented in a number of polarisable force fields such as AMOEBA [133,134] (for a detailed list see a topical review by Cieplak *et al.* [135]). However, with the added complexity that such models introduce this often comes with an increase in computational cost.

In the non-bonded term, electrostatic contributions can be modelled using a Coulomb potential, $U_{Coulomb}$, defined as:

$$U_{Coulomb}(r_{ij}) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \tag{2.11}$$

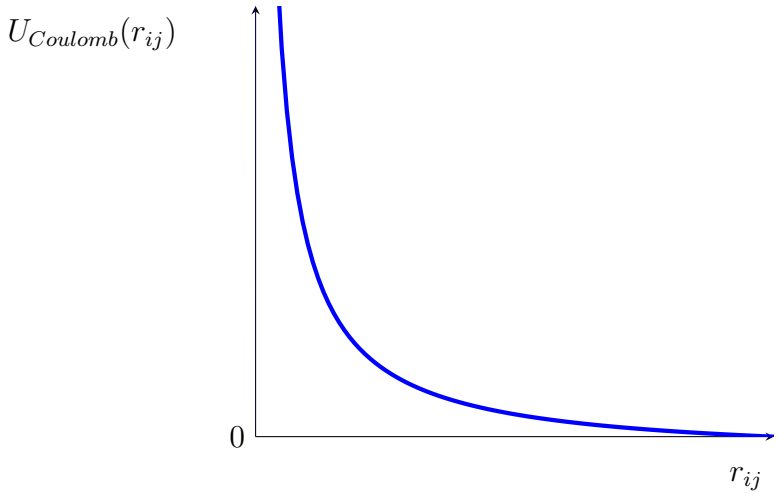


Figure 2.5: Qualitative plot of the Coulomb potential.

The $(4\pi\varepsilon_0)^{-1}$ term represents Coulomb's constant (where ε_0 is the permittivity of free space) and the q values are the charges on atoms i and j separated by distance r_{ij}.

2.3 Molecular dynamics theory

In a MD simulation, a system of N particles is evolved over time by solving Newton's equations of motion [136]. These particles interact by a predefined potential, U_{Total}, as described in Section 2.2. From this, the forces in the system may be derived (Equation 2.12) and Newton's equation of motion integrated to give the positions of all particles, $\mathbf{r}$, at some time t, with time step $\Delta t > 0$ (Equation 2.13).

$$\mathbf{F} = -\nabla U_{Total}(\mathbf{r}) \tag{2.12}$$

$$\mathbf{r}(t + \Delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{\mathbf{F}(t)}{M}\Delta t^2 \tag{2.13}$$

Where M is the mass matrix of the system. Once $\mathbf{r}(t + \Delta t)$ has been calculated, the positions are updated (i.e. the current positions become the old positions and the new positions become the current positions) and the process is repeated.

In any MD simulation the evaluation of forces is the most computationally expensive part. One way to reduce this is to evaluate the forces less frequently, if our choice of integrator allows for longer time steps the computational cost will be reduced. Another important point is that of energy conservation, which can be split into short and long-time components. Typically the Verlet (and related) algorithm(s) have moderate short-time energy drift but benefit from little long-time drift. Additionally any simulation that is repeated with the same set of initial conditions will diverge from the "true" trajectory due to errors incurred by limited machine precision. Since the aim of MD is to gather statistical data on a particular process, this need not be a huge problem as statistical averages of properties in a system will suffice. Similarly, it may be expected that integrators are time-reversible, but due to machine precision this is not the case and only in the limit of a infinite machine precision would

the trajectory be truly time-reversible. This leads to a final but important point about area preservation in phase-space. If our system is evolved over time following Hamiltonian mechanics, it would be expected that the magnitude of, for example, the volume in phase-space would remain unchanged. If this property had changed, energy conservation would be broken and lead to a drift in energy.

2.3.1 Integrators

2.3.1.1 The Verlet algorithm

Equation 2.13 outlines how the positions of particles can be found by integrating Newton's equations of motion and is more commonly known as the Verlet algorithm [137]. This formalism is readily derived from the Taylor expansion about point $\mathbf{r}$ and does not involve the velocities of each particle to be solved. One drawback to this method is the fact that velocities are only accessible once $\mathbf{r}(t + \Delta t)$ has been evaluated, this in addition to errors arising from a small $\mathcal{O}(\Delta t^2)$ term. For a full derivation of the Verlet algorithm, and how velocities may be obtained, see Appendix A, Section A.2.

2.3.1.2 The Leapfrog algorithm

An equivalent algorithm to Verlet is the Leapfrog algorithm. In this, the velocities are computed at half-integer time steps and then used to compute the new positions:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}\left(t + \frac{\Delta t}{2}\right) \tag{2.14}$$

$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}\left(t - \frac{\Delta t}{2}\right) + \Delta t \frac{\mathbf{F}(t)}{M} \tag{2.15}$$

From Equations 2.14 and 2.15 it is apparent that the positions and velocites are evaluated at different times (Figure 2.6). As a consequence the kinetic and potential energies are not directly accessible at the same time. Since the Leapfrog algorithm is derived from the Verlet algorithm it gives rise to identical trajectories. Its main benefit lies in the fact that large differences between two quantities are taken and so the chances for loss of precision are minimised [138]. All simulations in this thesis were performed using the Leapfrog algorithm.

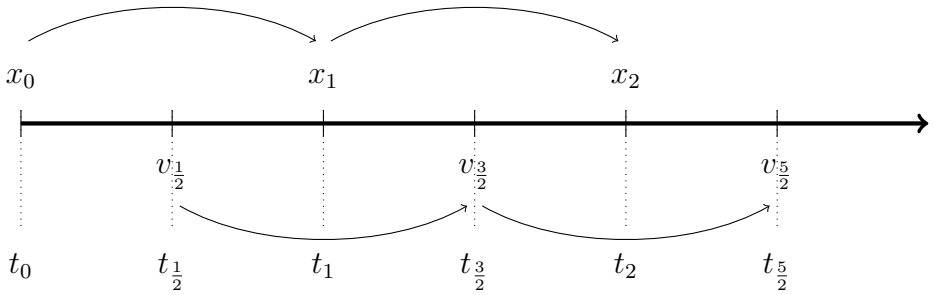


Figure 2.6: A schematic of the Leapfrog algorithm in one dimension. Positions (x) and velocities (v) are calculated at different times (t) and as such are said to "leapfrog" over one another.

2.3.1.3 The velocity Verlet algorithm

If the velocities and positions are required at the same time, it is possible to reformulate the Verlet algorithm to the velocity Verlet algorithm [139] (Equations 2.16, 2.17). In this representation the positions, velocities, and forces are all calculated at the same time. The original Verlet algorithm is recovered via removal of the velocity terms in the position calculation.

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{\mathbf{F}(t)}{2M}\Delta t^2 \tag{2.16}$$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t) + \frac{\mathbf{F}(t+\Delta t) + \mathbf{F}(t)}{2M}\Delta t \tag{2.17}$$

Due to their explicit calculation of velocities, if the user requires high precision for thermodynamic parameters (such as temperature) the velocity Verlet or Leapfrog algorithm should be used.

2.4 System setup protocol

In this section the typical steps and considerations taken to prepare a system for simulation will be outlined. This is a generalisation and, where necessary, specific details of each simulation setup performed in this work will be outlined in the relevant chapters.

2.4.1 Periodic boundary conditions

All systems in MD simulations will contain boundaries that, if not accounted for, can lead to unphysical behaviour [140]. One way to mitigate this problem is to use periodic boundary conditions (PBC) and defining the system as the unit cell that contains replicas surrounding it in all dimensions to create an infinite lattice (Figure 2.7). This is easily applied in three dimensions where periodic images are repeated in the x, y, and z dimensions. PBC conditions are usually employed with the minimum image convention that ensures that only interactions between a particle and the closest periodic image of its neighbours are considered [141].

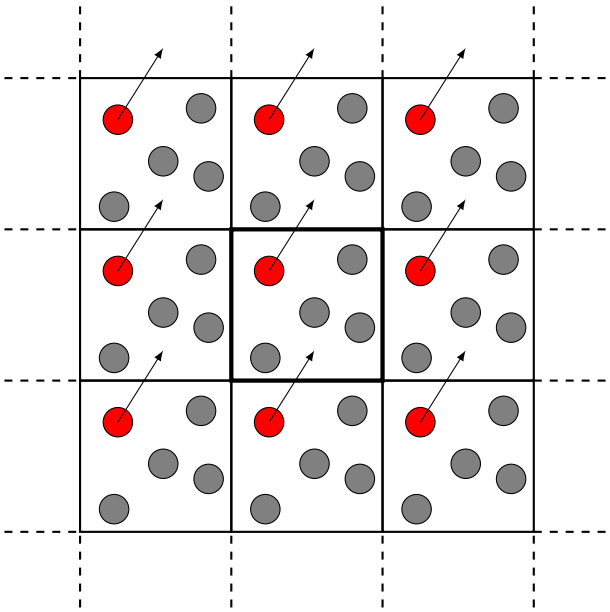


Figure 2.7: A 2D illustration of periodic boundary conditions. The system under investigation (centre, bold boundary) is surrounded by periodic images of itself. If a particle (red) moves out of the simulation box its periodic image enters at the other side (arrow).

2.4.2 Energy minimisation

Once a system is setup an energy minimisation step is needed to remove any high-energy interactions between particles as defined by the force field. Typically, a steepest decent algorithm is followed and the system coordinates adjusted to reduce the overall net forces acting upon them. This method is repeated until some user-defined energy cut-off or maximum step number is reached. Once minimised, the system is ready to undergo a series of equilibration steps.

2.4.3 The thermodynamic ensemble: temperature & pressure

Any thermodynamic system can be described by the number of particles in the system (N), the temperature (T), and the pressure (P). These parameters describe macroscopic observables of the system that are determined by underlying microscopic details, or microstates. A large number of unique microstates, each with their own positions ($\mathbf{q}$) and momenta ($\mathbf{p}$), occupy a multidimensional space known as phase space ($\mathbf{\Gamma}$). Different points in phase space can reproduce the same macroscopic observable (A_{obs}) and so as the system evolves over time it should pass through all points contained within $\mathbf{\Gamma}$, as described in Equation 2.18. The collection of all points in phase space is known as an ensemble. If averaged over infinite time this should represent the single system and is often referred to as the ergodic hypothesis. A set of different macroscopic observables will give rise to different ensembles, which are detailed in Table 2.1. Each ensemble will have an associated density of phase space points ($\rho(\mathbf{\Gamma})$) that represents the probability of finding the system in a state at points $\mathbf{p}, \mathbf{q}$. In a simulation, time is discretised and extension to infinite time is unfeasible. To circumnavigate this problem, the ergodic hypothesis is employed and the average over ensembles (Equation 2.19) can be used instead of averaging over time. As a result, the amount of sampling in phase space is often a major bottleneck in MD simulations.

$$
\begin{aligned}
A_{\text{obs}} = \langle A \rangle_{\text{time}} &= \left\langle A\Big(\mathbf{\Gamma}(t)\Big) \right\rangle_{\text{time}} \\
&= \lim_{t_{\text{obs}} \to \text{inf}} \frac{1}{t_{\text{obs}}} \int_0^{t_{\text{obs}}} A\big(\mathbf{\Gamma}(t)\big) \mathrm{d}t \qquad (2.18) \\
&= \langle A \rangle_{\text{ensemble}} \qquad (2.19)
\end{aligned}
$$

Ensemble	Abbreviation	Fixed parameters
Microcanonical	NVE	Number of particles, volume, energy
Canonical	NVT	Number of particles, volume, temperature
Isothermal-Isobaric	NPT	Number of particles, pressure, temperature
Grand canonical	μVT	Chemical potential, volume, temperature

Table 2.1: Common thermodynamic ensembles.

2.4.3.1 Thermostats

Control of temperature can be achieved through the coupling of the system to a heat bath. One of the most simple ways to achieve this is through direct rescaling of the particle velocities to a reference temperature as the simulation progresses:

$$\frac{\mathrm{d}T(t)}{\mathrm{d}t} = \tau_t^{-1}(T_{\mathrm{ref}} - T(t)) \tag{2.20}$$

Where $T(t)$ is the instantaneous temperature at a particular time step, T_{ref} is the reference temperature and τ_t is the coupling of the system to the bath. This method is motivated by the direct relationship of the kinetic energy of the system to its instantaneous temperature (Equation 2.21) and is the algorithm used by the Berendsen thermostat [142, 143].

$$\begin{aligned} E_k &= \frac{1}{2}\sum_{i=1}^{N} m_i v_i^2 \\ &= \frac{3}{2} N k_B T \end{aligned} \tag{2.21}$$

Where E_k is the kinetic energy, m_i and v_i are the mass and velocity of particle i respectively, N is the number of particles in the system, k_B is the Boltzmann constant, and T is the temperature.

Although the Berendsen thermostat is useful for getting systems to a target temperature, it has been shown that the correct canonical ensemble is not repro-

duced [144, 145], in addition to the introduction of artefacts. For these reasons it is typically only used for the initial equilibration stages of a simulation.

The velocity rescale thermostat [146] (Equation 2.22) is an extension of the Berendsen thermostat that does reproduce the canonical ensemble by introduction of a stochastic term. Instead of using a strict reference, the temperature is periodically drawn from the canonical equilibrium distribution from the kinetic energy of the system.

$$\mathrm{d}K = \tau_t^{-1}(K_{ref} - K)\mathrm{d}t + 2\sqrt{\frac{KK_{ref}}{N_f}}\frac{\mathrm{d}W}{\sqrt{\tau_t}} \tag{2.22}$$

Where K is the kinetic energy of the system, N_f is the number of degrees of freedom, and W a Wiener process (the stochastic term).

Once a system has been equilibrated, an extended system thermostat (where the system is in contact with an external heat bath) such as the Nosé-Hoover thermostat [147, 148] can be utilised. This introduces additional degrees of freedom to the Hamiltonian of the system. The central idea is to incorporate the coupling of the heat bath as an integral part of the system by the introduction of a virtual variable s, with an associated mass Q and velocity $\dot{s}$. The corresponding Hamiltonian now takes the form:

$$H(\mathbf{q}, \mathbf{p}, s, p_s) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{q}) + \frac{p_s^2}{2Q} + gkT\ln s \tag{2.23}$$

Where the first and second terms describe the kinetic and potential energy of the system, respectively. The third term represents the kinetic energy of the virtual system s and the final term is the potential of s (where g represents the degrees of freedom in the system). The Nosé-Hoover thermostat does reproduce the canonical ensemble but suffers from being nonergodic for some small systems [149].

Sometimes it may be appropriate to use a stochastic approach, such as the Langevin

thermostat. In comparison to Nosé-Hoover, stochastic methods introduce a random force into the equations of motion as well as a frictional term that is proportional to the particles velocities [150]. At each time step the thermostat adds this frictional term to the equations of motion, as well as a random force drawn from an appropriate probability distribution.

2.4.3.2 Barostats

In a similar fashion to the Berendsen thermostat [142], the pressure can also be rescaled in accordance to a reference pressure bath:

$$\frac{\mathrm{d}\mathbf{P}(t)}{\mathrm{d}t} = \tau_p^{-1}(\mathbf{P}_{ref} - \mathbf{P}(t)) \tag{2.24}$$

This is the Berendsen algorithm, where $\mathbf{P}(t)$ is the pressure at a particular time step, $\mathbf{P}_{ref}$ is the reference pressure, and τ_p is the coupling of the system to the bath. The algorithm scales the system coordinates ($\mathbf{r}$) and volume (V) every n_{PC} step according to the matrix μ [151] (Equations 2.25, 2.26 & 2.27).

$$\mathbf{r}' = \mu\mathbf{r} \tag{2.25}$$

$$V' = \mu^3 V \tag{2.26}$$

where the prime notation represents the scaled variable, and μ is defined as:

$$\mu_{ij} = \delta_{ij} - \frac{n_{PC}\Delta t}{3\tau_p}\beta_{ij}\{P_{ref_{ij}} - P_{ij}(t)\} \tag{2.27}$$

where β_{ij} corresponds to the isothermal compressibility (for water at 1 atm at 300 K this is 4.6×10^{-10} Pa^{-1}).

For simulations in water, the pressure coupling will be isotropic (i.e. the same value for all box vectors), however if the system is completely anisotropic μ takes the form:

$$\mu = \begin{pmatrix} \mu_{xx} & \mu_{xy} + \mu_{yx} & \mu_{xz} + \mu_{zx} \\ 0 & \mu_{yy} & \mu_{yz} + \mu_{zy} \\ 0 & 0 & \mu_{zz} \end{pmatrix}$$

In membrane simulations the x and y box vectors are often coupled and scaled by the same amount whilst the z vector (defined as normal to the membrane) will be scaled independently. This is known as semiisotropic coupling.

The Berendsen algorithm for pressure coupling is useful in reproducing the average pressure of the system but fails to sample from the correct NPT ensemble. To sample from such an ensemble, the Parrinello-Rahman algorithm [152, 153] can be used in a similar fashion to that of the Nosé-Hoover thermostat.

2.4.4 Long-range interactions

On average, the most computationally demanding part of any MD simulation is calculation of the non-bonded interactions. The number of particles (N) in any MD simulation can be large. Since each particle, by definition, must interact with every other, the resulting calculation scales as N^2 and thus becomes unfeasible to run on current computer architectures. In order to tackle this problem several schemes have been developed that reduce computational demand whilst retaining the accuracy of the simulation.

As seen in Section 2.2.2, the non-bonded interactions can be split into contributions made from both VdW and electrostatics. The LJ potential, that describes the

VdW forces, contains a r^{-6} term that captures the behaviour of the long-range attractive forces. This term notably decays rapidly towards zero as the particle-particle distance increases. Intuitively, several cut-off schemes have been introduced to only consider particles within a particular distance for LJ calculation. The simplest of these is to use a "hard" cut-off where the potential function ($v_c(r_{ij})$) is truncated at a defined distance (r_c), although this may introduce slight errors it dramatically reduces the number of calculations needing to be performed at each step (Equation 2.28).

$$v_c(r_{ij}) = \begin{cases} v(r_{ij}) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \tag{2.28}$$

Since particles at the boundary of such a scheme may experience unphysical forces during the simulation, a shifted-force potential (v_{SF}) can be used (Equation 2.29). Here the potential is shifted by v_c and a linear term added so that the derivative at r_c is zero [154].

$$v_{SF}(r_{ij}) = \begin{cases} v(r_{ij}) - v_c - (r_{ij} - r_c)\left(\frac{\mathrm{d}v(r_{ij})}{\mathrm{d}r_{ij}}\right) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \tag{2.29}$$

The benefit of the above approach is that now the discontinuity (as in the "hard" cut-off case) is now only present in the gradient of the force. This conserves energy and removes any instabilities within the equations of motion and as a result is commonly used in the CHARMM [155–158] and MARTINI [159–166] force fields (described later in Section 2.5.1.1).

Since the electrostatic potential falls as r^{-1} with increasing distance, the use of simple cut-off schemes may lead to artefacts in the force calculation [167]. To tackle this, it is common to use Particle Mesh Ewald (PME) summation [168]. By use of a simple identity, the r^{-1} term in the Ewald summation [169] can be expressed as:

$$r^{-1} = \frac{f(r)}{r} + \frac{1 - f(r)}{r} \tag{2.30}$$

Equation 2.30 represents the two terms that need to be calculated. The first is in direct space and corresponds to the central box, represented by the point charge distribution screened by Gaussian distributions ($\rho^q(r)$). The second is in reciprocal space and is constructed as a charge distribution of the same shape and sign as $\rho^q(r)$, as shown in Figure 2.8 and can be solved quickly by Fourier transforms.

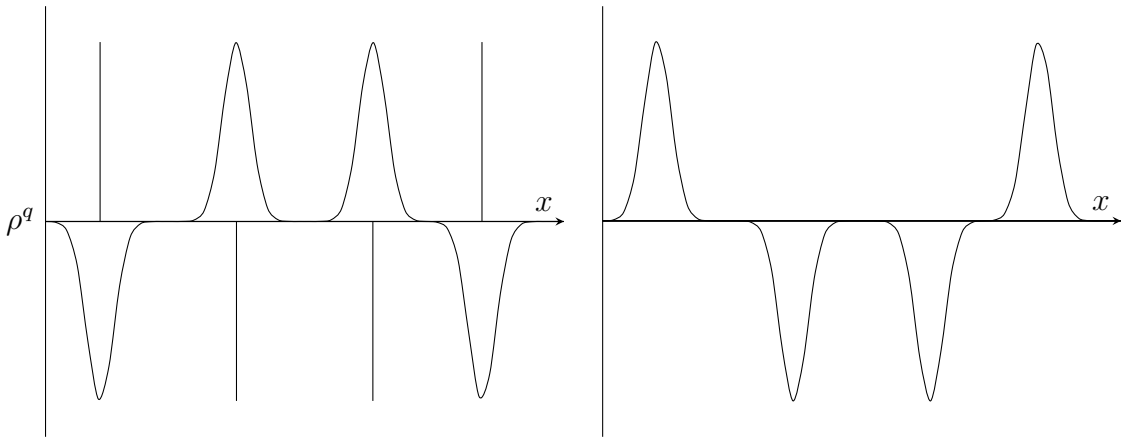


Figure 2.8: One-dimensional illustration of the Ewald sum. *Left*: Point charges with screening distribution. *Right*: Cancelling distribution. Image adapted from [170].

Unfortunately, Ewald summation scales as $\mathcal{O}(N^{\frac{3}{2}})$ and so is computationally inefficient. Instead, PME uses a grid based approach that replaces the continuous point-charges by grid-based charges. This is to facilitate the use of Fast Fourier Transforms that increase the speed of this calculation and improves scaling to $\mathcal{O}(N\log(N))$.

An alternative to the use of PME is that of the reaction-field approach. This method treats the Coulombic contributions inside a spherical region of the system explicitly, but once past a certain cut-off there is an assumed continuum of uniform dielectric constant [171, 172]. Although this approach saves computational expense there are issues when a particle crosses the cut-off boundary. As discussed by Allen & Tildesley [173], when such an event occurs there is a discontinuous jump in energy

between the cavity and in the reaction-field contribution. These do not cancel exactly and can lead to poor energy conservation.

Typically, PME is the method of choice for electrostatic contributions and on modern computer architectures its evaluation (and associated Fast Fourier Transforms) perform very efficiently. This now leads to a new bottleneck: calculation of the contribution of electrostatics from every atom with every other in the system.

An additional tool to increase efficiency is the use of neighbour lists, the most common of which is the Verlet list, shown schematically in Figure 2.9. The approach stores the neighbours in a list (the Verlet list) within a certain radius (r_l) of particle i. In the next calculation step if the maximum displacement of the particles is less than $r_l - r_c$, where r_c is a radius surrounding particle i that is $< r_l$, then only particles contained within the Verlet list are considered. As soon as one of the particles moves a distance greater than $r_l - r_c$ then the list is updated. Although using such a scheme scales as N^2, the Verlet list contains considerably fewer particles than the entire system and its use on systems with a large number of particles (i.e. > 100) is justified.

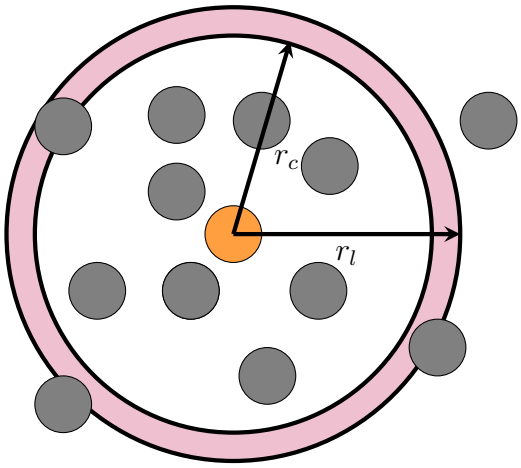


Figure 2.9: The Verlet list: A particle i (orange) only interacts with particles j (grey) within r_c. The Verlet list contains particles that are within a sphere of $r_l > r_c$ (pink). Image adapted from Frenkel & Smit [174].

2.5 Multiscale modelling

Biological processes take place over various spatio-temporal scales and in order to tackle this a number of computational techniques have been developed [175–177] (Figure 2.10). At the smallest scale the natural choice are QM calculations where the behaviour of electrons is key. On a larger scale, all atom models are more suitable and lend themselves well to the study of protein dynamics. As system size increases CG models may be utilised that incorporate several atoms into one representative particle reducing the number of degrees of freedom that need to be calculated. Similarly, mesoscale models decrease the number of particles again, whilst at the furthest end of the spectrum there are continuum models where the collective motion of particles is more important than individual interactions. The spirit of multiscale modelling is to use the most suitable level of theory for the scale that a given problem pertains to.

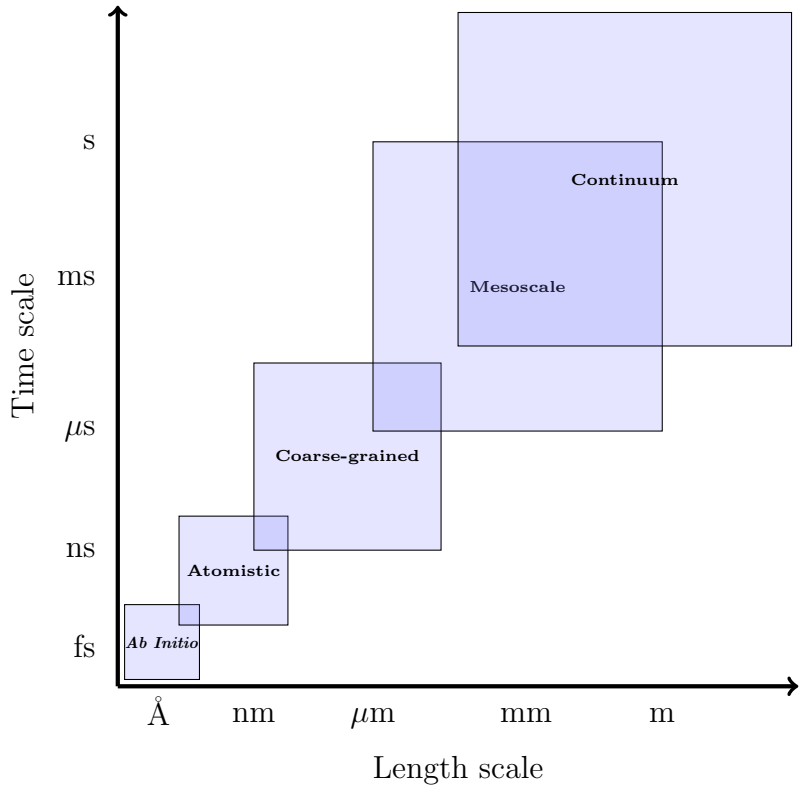


Figure 2.10: Multiscale methods across various time and length scales.

The benefit of a multiscale approach increases computational efficiency when choosing the right level of theory to use for a particular problem. For biological questions concerning, for example, an enzyme's active site it would be appropriate to use QM (or QM/MM) methods [178–180]. Measuring protein folding lends itself well to an all-atom approach, but tackling problems relating to protein oligomerisation or membrane undulation often requires a CG description [10, 181, 182]. Finally, large scale collective motion would be captured well using a mesoscale model [183].

2.5.1 Coarse-grained molecular dynamics

Typically atomistic simulations are the primary tool to study protein dynamics, however as system size grows the computational cost increases and as a result a number of CG force fields have been developed. The underlying principle is a simple combination of several atoms to form a single particle. The effect of this reduces the number of degrees of freedom and particle number contained within the system of interest. Although this will clearly induce a loss of accuracy there is an increase in computational efficiency and the level of coarse-graining ultimately depends on the biological problem of interest.

2.5.1.1 The MARTINI force field

The MARTINI force field is one of the most popular CG force fields used for biomolecular simulation. Initially it was developed for lipid only simulations by the groups of Marrink and Tieleman [162]. Subsequent releases increased the models use for work with proteins [160], DNA/RNA [163, 164], carbohydrates [165], polymers [166], and nanoparticles [184, 185].

At the centre of the MARTINI model are four particle types: polar (P), nonpolar (N), apolar (C), and charged (Q). For more complex chemical moieties, such as structures containing rings or small molecules, an S-type particle is used. Additionally,

each of these particle types has its own subtype that allows for greater flexibility and more accurate chemical descriptions (Figure 2.11). The basic design principle of the MARTINI model is a simple 4-to-1 mapping of heavy particles (as well as any bound hydrogens) to one coarse-grained bead. In the case of S-type particles, the mapping is altered to an increased resolution of up to 2-to-1. The various combinations of particle types and subtypes gives total of 18 fundamental particles to construct representations of molecules from and are crucial in accurately modelling amino acid side chains (Figure 2.12). How these particles interact can be found in the interaction matrix outlined by Marrink *et al.* [161].

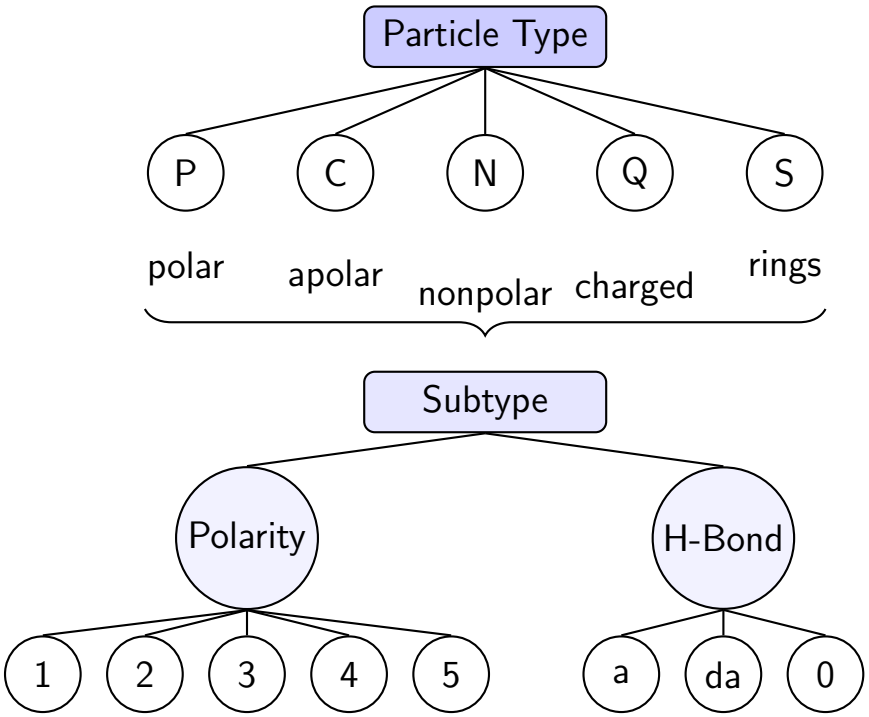


Figure 2.11: Different particles in the MARTINI model. Various interactions between particles can be made by alternative combinations of particle type and subtypes. Polarity varies from very weakly polar (1) to very strongly polar (5). Hydrogen bond ability is defined as either acceptor (a), donor/acceptor (da) or neither (0).

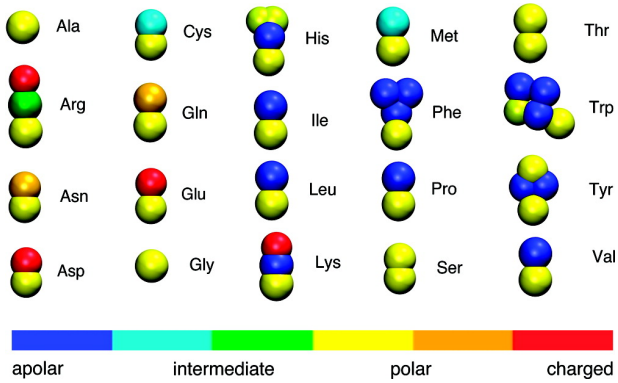


Figure 2.12: MARTINI particle type mapping to amino acid side chains. Image adapted from [160].

Bonded & non-bonded interactions

Bonded interactions in the MARTINI force field are typically designed using structural data derived from atomistic simulation using a standard set of potential functions as outlined in Section 2.2.1. When simulating proteins the atomistic structure from the Protein Data Bank (PDB) [186, 187] can be used as a reference. This is usually used in combination with an elastic network approach, such as ElNeDyn [188], that constrains the protein's secondary structure to a particular conformation. The ElNeDyn methodology applies a set of harmonic potentials (i.e. springs) between the backbone particles of the protein within a certain distance cut-off. In this approach the backbone particles are centred on the Cα atoms, whereas in the MARTINI force field the backbone particles are centred on the centre of mass of the backbone atoms (N, Cα, C, O). To solve this the ElNeDyn model adapts the MARTINI bonded parameters (as well as mapping for aromatic residues). For a given ElNeDyn model the cut-off, R_C, is set to 0.8 - 1.0 nm with a force constant, K_{spring}, of 500-1000 kJ mol^{-1} nm^{-2}. All CG simulations in this thesis used $R_c = 0.9$ nm and $K_{spring} = 500$ kJ mol^{-1}.

Non-bonded interactions are described using a 12-6 LJ potential with a well depth of $\varepsilon_{ij} = 2.0$-5.6 kJ mol^{-1} for interactions between polar and apolar groups or strongly polar groups respectively, thus mimicking the hydrophobic effect. Typically, all par-

ticles in the model have the same effective size of $\sigma = 0.47$ nm, however ring-like molecules use a sightly reduced value of $\sigma = 0.43$ nm and ε_{ij} is scaled to 75% of its original value. Charged interactions within the non-bonded parameters are captured using charged particle types, Q, that interact via a Coulombic energy function. These interactions are implicitly screened with a dielectric constant of $\varepsilon = 15$. In order to mimic the effect of distance-dependent screening, the MARTINI model uses the non-bonded potential energy functions in a shifted form, with a cut-off interaction distance $r_{cut} = 1.2$ nm, the LJ potential shifted from 0.9 nm to r_{cut}, and the electrostatic potential shifted from 0.0 nm to r_{cut}.

Water model

In MARTINI, four atomistic water molecules are represented by one coarse-grained particle. Clearly, reducing the complexity of solvent interactions will alter the effects of polarization and screening in the system when compared to an all-atom system. To address these issues, a number of alternative water models have been developed, namely the polarizable and continuum models. In the polarizable MARTINI model [189, 190], additional charged particles (WP and WM, carrying charges $+q$ and $-q$ respectively) are bound via a harmonic potential to the central W particle. Their rotation is controlled by a harmonic angle potential that allows the dipole moment to vary. Although useful, the cost of simulating extra interactions that arise from such a model is often high in terms of computational power. Since the philosophy of any coarse-grained technique is to simplify the interactions, it is often unnecessary to use such a model. If solute-solvent interactions are of critical importance to the system, an atomistic representation will yield far more accurate results. For very large systems it may be appropriate to use a continuum based model such as Dry MARTINI [191]. In this approach, all water molecules are modelled implicitly and so the computational cost is reduced. For these systems, macroscopic effects are much

more important and so the fine details of solute-solvent interactions are not necessary in order to capture the correct physics of the process of interest. For all MARTINI simulations in this work the standard water model was used.

2.6 Homology model generation

Over recent years there has been a large increase in the number published of high resolution protein structures deposited in the PDB, with $\approx 40,000$ between 2015 - 2018 [192]. Several methods are utilised by experimental groups such as X-ray crystallography, nuclear magnetic resonance (NMR), and cryo-EM. Often, a combination of several methods, sequence knowledge, and related domain structures are needed to produce refined atomic structures of proteins. Although powerful, sometimes it is not possible to experimentally resolve the entire structure of a particular protein and often certain regions are missing. For example, if the protein under study is large and has several domains, only a few of these may be resolved and it is up to the user to reassemble these. Additionally if there are regions of the protein that are highly flexible, such as loops, these may not be resolved (in the case of NMR this is not a problem as an ensemble of loop structures are captured).

Fortunately, a number of software packages exist to model in missing residues of a protein based on a solved structure (the template) and its sequence (the target). Usually a protein will be part of a much wider family of related proteins with conserved residues that are important to their biological function. These patterns in conservation can be captured using a multiple sequence alignment (MSA) of the template, target, and a number of closely related proteins in the same family. The MSA can be used to include the newly aligned sequence of the target into the modelling software of choice and a series of homology models created. In this thesis the MUSCLE [193, 194] web server was used to create all MSAs and homology models

were constructed using Modeller (v9.16) [195].

Once created, the quality of each model must be assessed and given a score. Scoring can be based on a number of statistical potentials, one of which is Discrete Optimised Protein Energy (DOPE) function [196] that is implemented within the Modeller software package. DOPE is an atomic distance-dependent statistical potential that measures the quality of a generated model based on a native reference state as a sphere of uniform density and finite size. Additional assessment can be made by composite scoring functions such as Qualitative Model Energy ANalysis (QMEAN) [197, 198]. This function works by defining five geometrical descriptors used to assess a particular model that include torsional angles, a secondary structure specific distance-dependent pairwise residue potential, a solvation potential, and two terms to assess agreement between predicted vs. calculated secondary structure and solvent accessibility. All protein models constructed in this thesis were assessed using both DOPE and QMEAN scoring functions.

"Everything has an end,
and you get to it if you only keep all on."
— Peter Waterbury, in *The Railway Children*, by E. Nesbit

3
Nav1.5e splice variant

Contents

3.1 Introduction 56
3.1.1 Alternative splicing 56
3.1.2 Alternative splicing in Nav channels 57
3.1.3 Nav1.5e: D211K & T207S mutations 58
3.2 System setup 61
3.3 Results 63
3.3.1 Model stability & S4 movement 63
3.3.2 D211K mutant 68
3.3.3 T207S mutant 72
3.3.4 Double D211K & T207S mutant 76
3.4 Discussion 79
3.4.1 D211K dynamics 79
3.4.2 T207S dynamics 81
3.4.3 Potential inhibition of S4 movement by combined D211K & T207S interactions 82
3.4.4 Conclusions 82

This chapter is based on the manuscript: *Alternatively-spliced isoforms of the cardiac Nav1.5 sodium channel modify channel activation by distinct structural mechanisms*, Adamo S. Mancino, William G. Glass, Yuhao Yan, Philip C. Biggin and Derek Bowie. ***in prep***.

3.1 Introduction

In this section, the concept of alternative splicing is introduced and how this applies to Nav channels. More specifically, focus is drawn upon a specific splice variant and relevant single-point mutations highlighted by Mancino *et al.* [199] that were then simulated using MD.

3.1.1 Alternative splicing

In eukaryotes, a particular gene is constructed from coding and non-coding regions. During transcription, all regions are included to form pre-mRNA, after which a number of post-transcriptional modifications take place, with one such process being RNA splicing. This removes non-coding regions from pre-mRNA whilst retaining and ligating coding regions, these regions are named introns and exons respectively. The splicing of pre-mRNA is catalysed by five small nuclear ribonucleoproteins and a number of associated proteins [200]. These make up a huge and dynamic ribonucleoprotein complex [201], named the spliceosome. The simple removal of introns from pre-mRNA is named consecutive splicing. However, more commonplace is alternative splicing that is regulated by the spliceosome.

Alternative splicing plays a crucial role in dramatically increasing the number of mRNA isoforms formed from a single gene. This allows for increased diversity in the proteome of a specific organism. In humans, this allows for the approximately 20,000-30,000 protein-coding genes [202] to increase the protein isoforms they can produce

more than tenfold [203]. Typically this process is tightly regulated and depends on factors such as cell-type and cell-cycle [204]. Figure 3.1 details the variety of mechanisms used in alternative splicing. In humans, exon skipping is most prevalent, however a variety of other mechanisms exist such as mutually exclusive exons, intron retention, and alternative 5′ and 3′ splice sites.

Once mature mRNA is formed after (consecutive or alternative) splicing, it can then undergo translation to eventually form a functional protein. Although similar, a protein that has been formed via alternative splicing will have a different amino acid sequence to that of its canonical form.

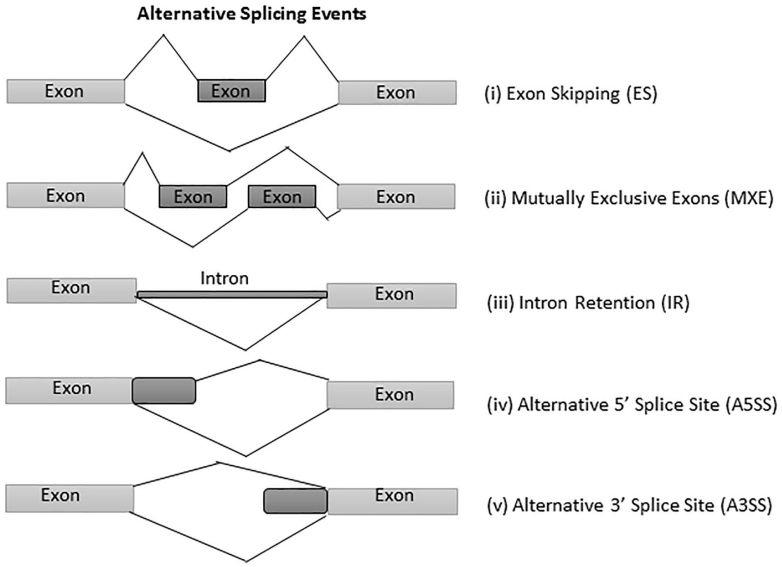


Figure 3.1: Overview of the common mechanisms of alternative splicing in gene expression. In humans, exon skipping is the most prevalent form of alternative splicing. Image adapted from Chaudhary *et al.* [205].

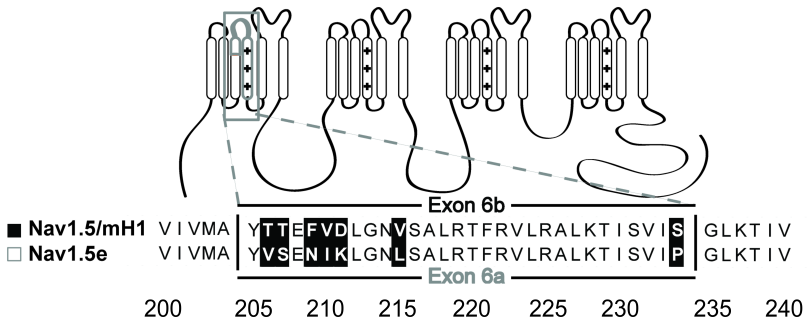
3.1.2 Alternative splicing in Nav channels

In addition to specific Nav channel isoforms (described in Section 1.2), increased diversity in channel behaviour can be achieved through alternative splicing. Each of these isoforms possesses their own unique alternatively spliced variant, but there are common splicing features shared across the family of channels. In six of the Nav

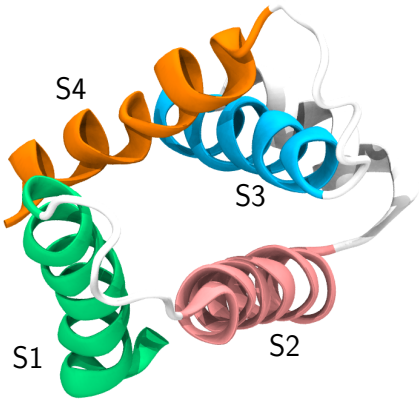
channel isoforms there are two forms of the gene that encode for the S3-S4 helices in VSDI and in Nav1.5 these are exon 6b and 6a (or the analogous exon 5a and 5b in Nav1.1, Nav1.2, Nav1.3, Nav1.6, and Nav1.7). Inclusion of exon 6a encodes for the "adult" form, whereas inclusion of exon 6b gives rise to the "embryonic" form. Alternative splicing in the S3-S4 region of the VSDI has a profound affect on channel gating properties. In Nav1.1, Nav1.2, and Nav1.7, inactivation of the channel is altered [206–208] whilst in Nav1.5 the voltage-dependence of activation is shifted [209].

3.1.3 Nav1.5e: D211K & T207S mutations

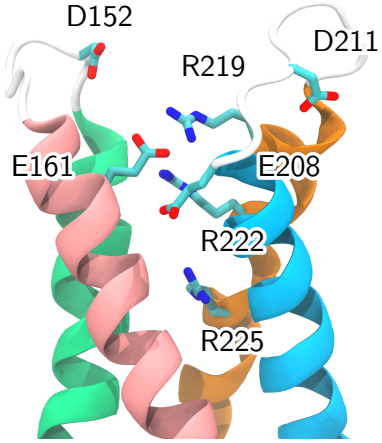
The aspartic acid to lysine switch at position 211 (D211K) has been shown by Onkal *et al.* [209] to be responsible for the entire shift in Nav1.5e. However, the molecular mechanism that causes the observed depolarised activation profile remains elusive. Interestingly, there are around seven amino acid differences between the human (mH1) and embryonic (Nav1.5e) isoforms of Nav1.5 (Figure 3.2). In order to investigate the impact that any of these other amino acids had on channel activation, in addition to D211K, Mancino *et al.* [199] ran single point mutations of each of the seven residues in both the forward (mH1 $\rightarrow$ Nav1.5e) and reverse (Nav1.5e $\rightarrow$ mH1) direction.



(a) Exon 6



(b) Top down

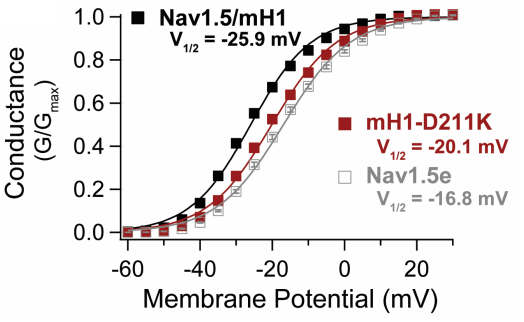


(c) Side on

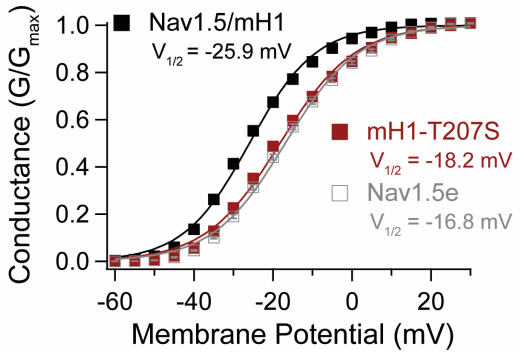
Figure 3.2: The voltage-sensor (S1 - 4) of domain 1 in Nav1.5. (a) Diagram of exon 6 in both mH1 and the Nav1.5e splice variant with a total of seven amino acids differ between the two, reproduced with permission from Mancino *et al.* [199]. (b) Top down view of the VSDI. S1 (green), S2 (pink), S3 (cyan), and S4 (orange) are shown as helices, with linker regions shown in white. (c) Side on view of the VSDI with the same colouring as in (b), key residues highlighted in a stick representation (T207 is not shown for clarity). Both (b) and (c) are taken from the Nav1.5 homology model constructed from PDB code: 5X0M (see Section 3.2).

Surprisingly, and contradicting results from Onkal *et al.* [209], Mancino *et al.* [199] found that the D211K mutation in both the forward and reverse direction caused only a partial shift in activation (Figure 3.3a). A full shift was recorded for the forward Nav1.5e-T207S mutation (Figure 3.3b), whereas only a partial shift was seen for the reverse mH1-S207T mutation, suggesting that position 207 of exon 6 is important in the differences seen between the mH1 and Nav1.5e gating profiles. Additionally,

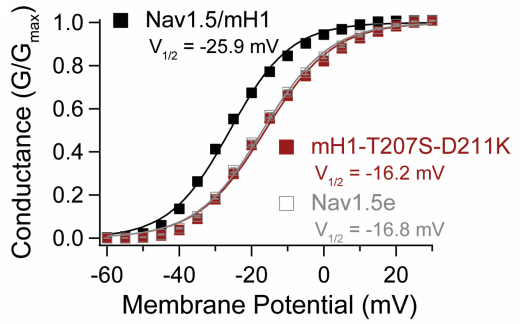
applying the forward and reverse double mutations at positions 211 (ASP → LYS) and 207 (THR → SER) gave an activation profile that overlapped with the one formed by Nav1.5e (Figure 3.3c), i.e. a complete exon swap. Conductance-voltage curves for each forward mutation are shown in Figure 3.3. This suggests that both positions 207 and 211 in exon 6 are sufficient and necessary for the differences seen between the two splice variants. In order to investigate the contribution that the T207S and D211K mutations make to the structural stability of VSDI, MD simulations were utilised.



(a) mH1-D211K



(b) mH1-T207S



(c) mH1-T207S-D211K

Figure 3.3: Conductance-voltage curves for the Nav1.5/mH1 and forward point mutations. (a) and (b) show the single-point D211K and T207S mutations respectively. A full shift is achieved when both mutations are present, as shown in (c). All data reproduced with permission from A. Mancino [199].

3.2 System setup

The recent crystal structure of NavPaS (PDB Code: 5X0M) [210] was used to construct a model of Nav1.5/mH1. A total of 200 models were created using the multiple sequence alignment of full length human Nav1.5 with Nav1.X (where X=1-9) and the 5X0M crystal structure sequence (Figure 3.4). The sequence identity between the NavPaS structure and Nav1.5 was 45.1%. All sequences were obtained from UniProt [211].

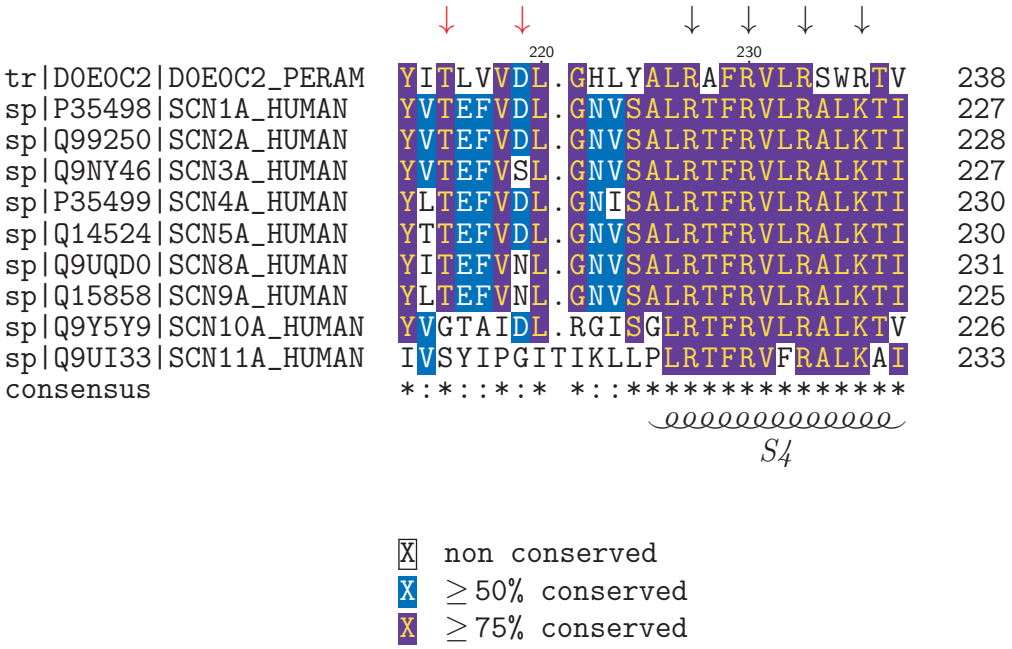


Figure 3.4: DI S3-4 multiple sequence alignment used for Nav1.5 homology model generation. Both T207 and D211 in Nav1.5 are indicated by red arrows whilst conserved positive arginine and lysine residues are indicated with black arrows. The approximate location of the start of the DI S4 helix is also shown.

Each model was created and scored using the DOPE function [196] in the Modeller (v9.16) software package [195]. The 10 best models were ranked using the QMEAN scoring function [197] in SWISS-MODEL [212] and the final model chosen with the highest QMEAN score. All mutations in the final model were performed using PyMol (v1.4) [213].

All simulations were performed using GROMACS 2018 [214] with the following force fields: AMBER ff99SB-ILDN [215] for protein and ions, SLipids [216–218] for lipids, and TIP3P [219] for water. Each system was prepared by overlaying the protein in a 1-palmitoyl-2-oleoyl-sn-glycero-phosphocholine (POPC) bilayer and using the InflateGRO methodology [220] for embedding into the membrane. In-house scripts were used for final adjustment of the embedded channel. The system was solvated with water and neutralised with Na^+ and Cl^- ions at a concentration of 150 mM. Energy minimisation was performed using the steepest decent algorithm. Equilibration steps consisted of a short (1 ns) and long (5 ns) equilibration whilst position restraining the Cα atoms with a force constant of 1000 kJ mol^{-1} nm^{-2}. All simulations were carried out in the NPT ensemble. The temperature and pressure were set to 323 K and 1 bar with semi-isotropic pressure coupling and isothermal compressibility set to 4×10^{-5} bar^{-1}. Temperature and pressure were controlled using the Berendsen thermostat and barostat [142] in the short equilibration with coupling constants set to 0.5 and 5.0 ps respectively. In the long equilibrations and production runs the Nosé-Hoover thermostat [147, 148] and Parrinello-Rahman barostat [152] with coupling constants of 0.8 and 5.0 ps respectively were used. The VdW cut-off was set to 12 Å with dispersion correction for energy and pressure. Electrostatic interactions were treated using the smooth PME methodology with a real space cut-off and grid spacing of 12 Å and a grid spacing of 1.6 Å used to obtain the reciprocal energy term in k-space with cubic interpolation. Bonds containing hydrogen were constrained using the LINCS algorithm [221]. The Verlet cut-off scheme was used with a buffering tolerance of 0.005 kJ mol^{-1} ns^{-1} per particle to construct the pair list. A total of four independent simulations were run for 400 ns each and coordinates written out every 2.5 ps. Analysis was performed using MDAnalysis [222, 223] and in-house scripts. All visualisations were performed in VMD [224]. All graphs were produced using the Seaborn [225] and Matplotlib [226] Python libraries.

3.3 Results

3.3.1 Model stability & S4 movement

The D211K, T207S, and double D211K/T207S systems were constructed from the homology model of mH1. To measure the overall stability of the subunit the root mean squared deviation (RMSD) was calculated from the backbone atoms C, N, and Cα for the entire protein (Figure 3.5). All systems remain stable over 400 ns with an RMSD between 6-7 Å.

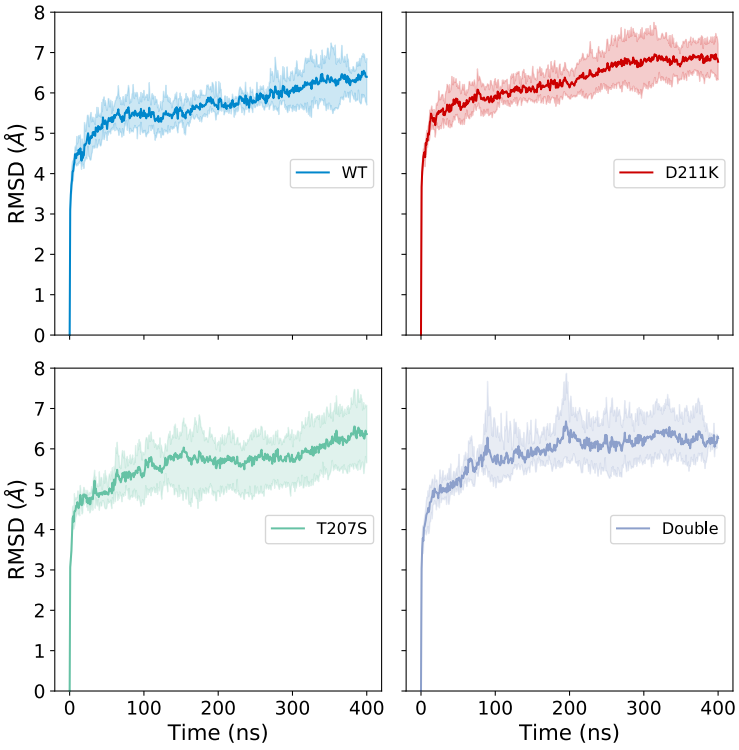


Figure 3.5: Mean RMSD of WT (blue), D211K (orange), T207S (green), and double D211K/T207S (purple) system over 400 ns. Standard deviations are shown as shaded areas.

As discussed in Chapter 1, the S4 helix is lined with positively charged residues, with the movement of this helix being key to pore opening and channel function. Both the D211K and T207S mutations are present at the top of VSDI and are close to the arginine residues lining S4 (R219, R222, and R225). To investigate the movement

of these residues in each system, the z-coordinate from the centre of mass of the terminal nitrogen atoms (NH1 and NH2) was taken and compared to the first frame in each simulation (Figures 3.6, 3.7, 3.8, & 3.9). Although hard to determine, the crystal structure was resolved at zero voltage and as a result the S4 helices would be expected to be close to the extracellular side of the membrane (i.e. close to the final stages of S4 movement towards the extracellular side of the membrane). Conversely in the resting state (i.e. at a resting membrane potential) the S4 helix would be expected to be "down", that is to say, closer to the intracellular leaflet. Due to the difficulty in assigning definitive stares, the focus taken in these studies is made on the S4 residues propensity to move towards or away from the extracellular side of the membrane plane. The movement of these residues in mH1 is highly dynamic. Over the relatively short simulation time of 400 ns there was not a clear pattern of R219, R222, and R225 movement, with residues tending to fluctuate by only a few ångströms from their starting value.

Comparing to the D211K, T207S, and double D211K/T207S mutants to mH1, it can be seen that the likelihood across repeats for these residues to move towards the ECD is somewhat reduced. The reduction in movement was most pronounced for the D211K system where in runs 2-4 (Figures 3.7b, c, & d), where there is either little net movement or the displacement of residues towards the ICD when compared to mH1. In contrast, run 1 (Figure 3.6a) exhibits a very slight movement of all residues towards the extracellular region of around 3 Å.

For the T207S system, this trend of restricted movement continues. All repeats display little change from the first frame positions, with only run 3 (Figure 3.8c) exhibiting pronounced movement downwards of R225 to around 5 Å.

In the double D211K/T207S runs (Figure 3.9) the trend is somewhat unclear. In run 3 (Figure 3.9c), there is little movement of all three arginine residues throughout the course of the simulation. Here, there is a combined steric blocking from D211K

and increased interaction with E208, as will be examined further in Section 3.3.4. In run 2 (Figure 3.9b) there is minimal movement of R219 and R225, however R222 is forced downwards and towards the intracellular side. Interestingly, run 1 (Figure 3.9a) shows a general movement of all residues to the extracellular side of the VSD, whilst run 4 (Figure 3.9d) shows the opposite. In run 4 there are a number of conformations visited, the two extremes of which are shown in Figure 3.19 (Section 3.3.4), where increased interaction of D211K with lipid and / or the increased interaction of E208 with S4 arginine residues may restrict S4 movement. An interaction was counted if ≤ 5 Å from a POPC phosphorus atom.

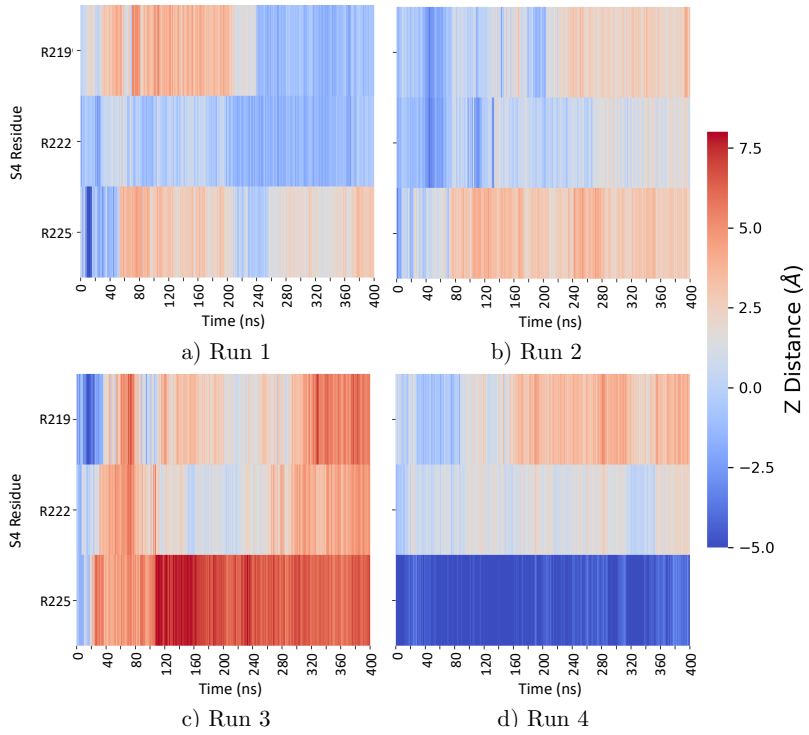


Figure 3.6: Movement of arginine residues z-coordinate (Z) in the S4 helix in mH1.

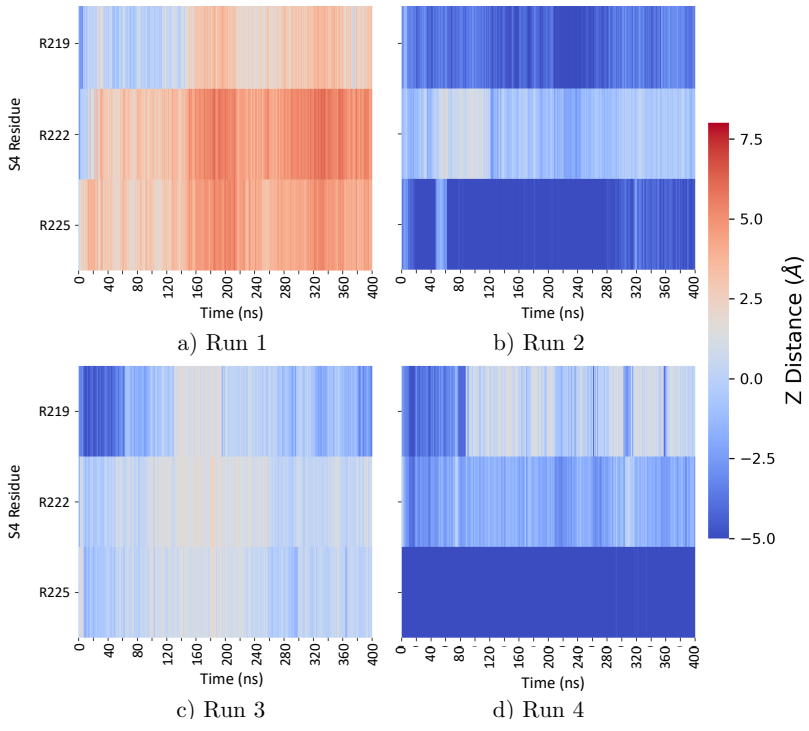
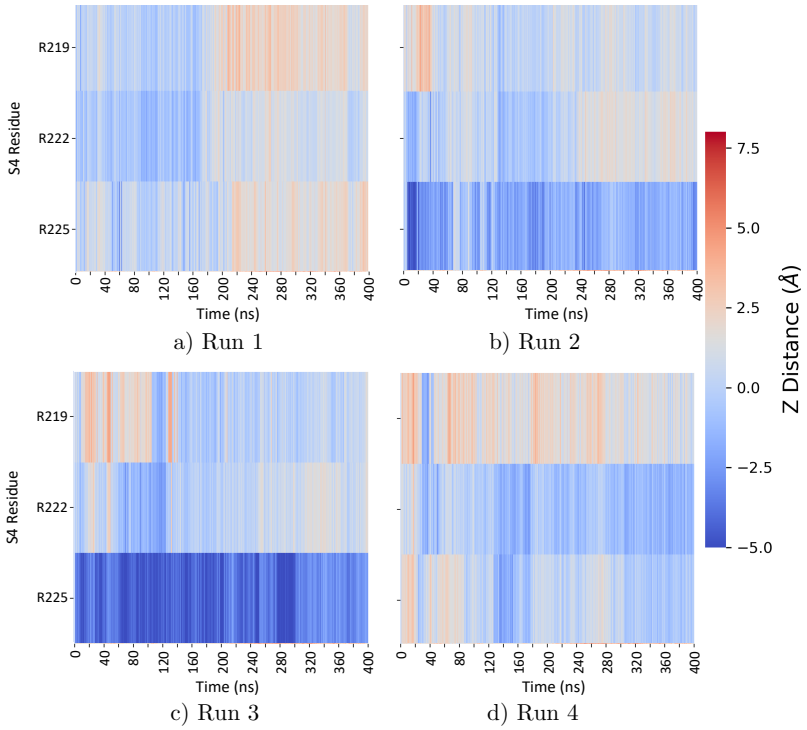


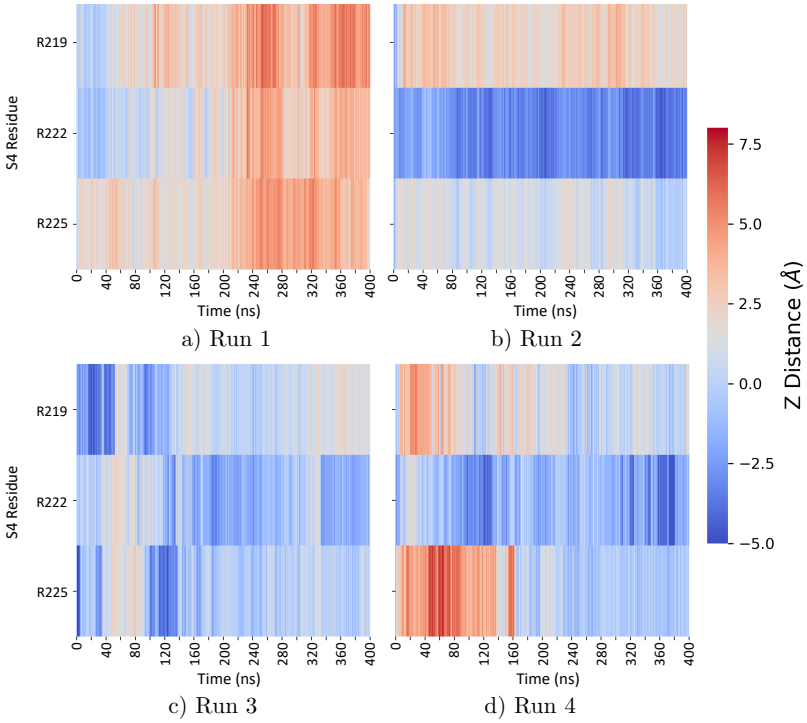
Figure 3.7: Movement of arginine residues z-coordinate (Z) in the S4 helix in the D211K mutant.



a) Run 1 b) Run 2

c) Run 3 d) Run 4

Figure 3.8: Movement of arginine residues z-coordinate (Z) in the S4 helix in the T207S mutant.



a) Run 1 b) Run 2

c) Run 3 d) Run 4

Figure 3.9: Movement of arginine residues z-coordinate (Z) in the S4 helix in the double D211K - T207S mutant.

3.3.2 D211K mutant

3.3.2.1 Interaction with lipid

The D211K mutation is located in the linker region between the S3-S4 helices in VSDI (see Figure 3.2c). Over the four repeats there was significant interaction of D211K with the POPC lipid membrane (Figure 3.10). All distances were taken from the terminal oxygen(s) / nitrogen of D211 / K211 and the phosphate group of POPC. In D211 of mH1, this was taken from the centre of mass of the terminal oxygens.

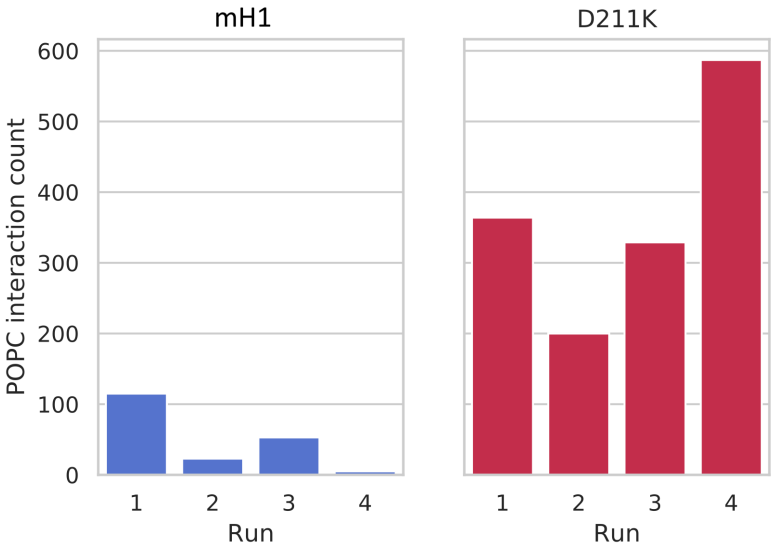
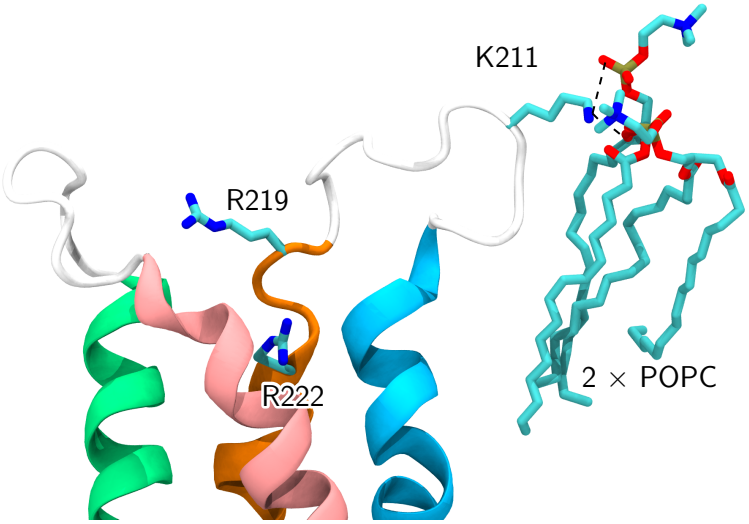
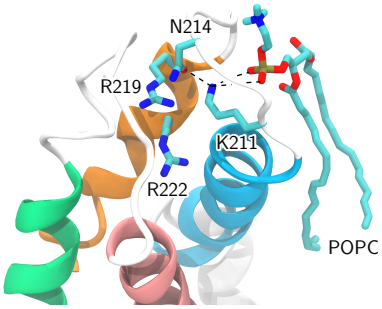


Figure 3.10: Interactions of D211 / K211 with POPC headgroups. *Left*: mH1, *Right*: D211K.

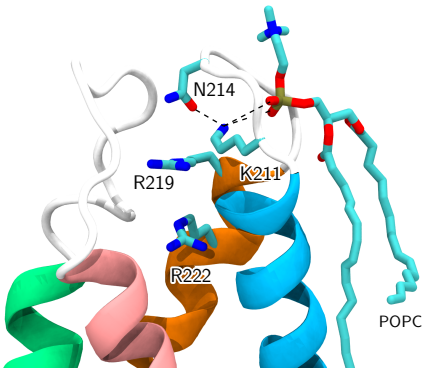
Interaction with lipid headgroups is most pronounced in run 4 where D211K dramatically opens the top of the VSD (Figure 3.11a). In run 2 the converse is true (Figures 3.11b & 3.11c), where there is decreased interaction with POPC. However, in this run D211K positions itself "inside" the VSD, stabilised by the nearby N173 residue and a single POPC lipid. Positioned here, D211K is directly above the topmost arginine residue (R219) in the S4 helix.



(a) Interaction of D211K with the headgroup of two POPC lipids.



(b) Top down



(c) Side on

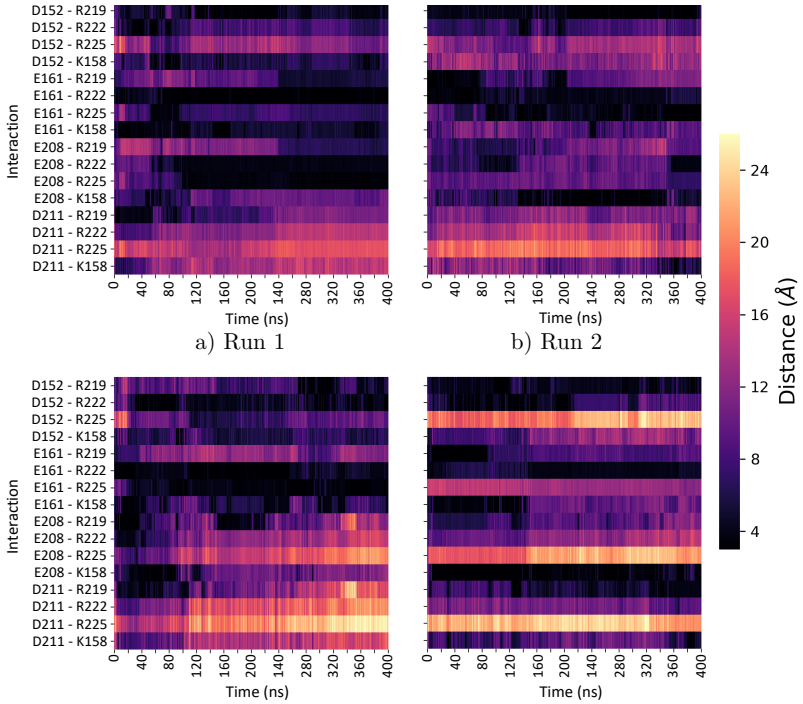
Figure 3.11: Various interactions of K211 (dotted lines) in the D211K switch system. (a) Interaction of D211K with the headgroup of two POPC lipids in run 4. (b) & (c) Interaction of D211K at the top of the VSD in run 2. D211K positions itself at the top of the VSD and is stabilised by an oxygen in the nearby N214 residue as well as a POPC phosphate headgroup.

3.3.2.2 Residue distance analysis

In order to assess the change in electrostatic contacts within the VSD between mH1 and D211K the distances over the course of each simulation run were measured. The centre of mass of the terminal atoms of each of the negative and positively charged residues was taken, and the distance matrix between all other charged residues cal-

culated. Interactions present in mH1 and D211K are shown in Figures 3.12 & 3.13 respectively. Of the three arginine residues in the S4 helix, R219 is the closest to the extracellular region and in close proximity to residue 211 (located in the S3-S4 linker).

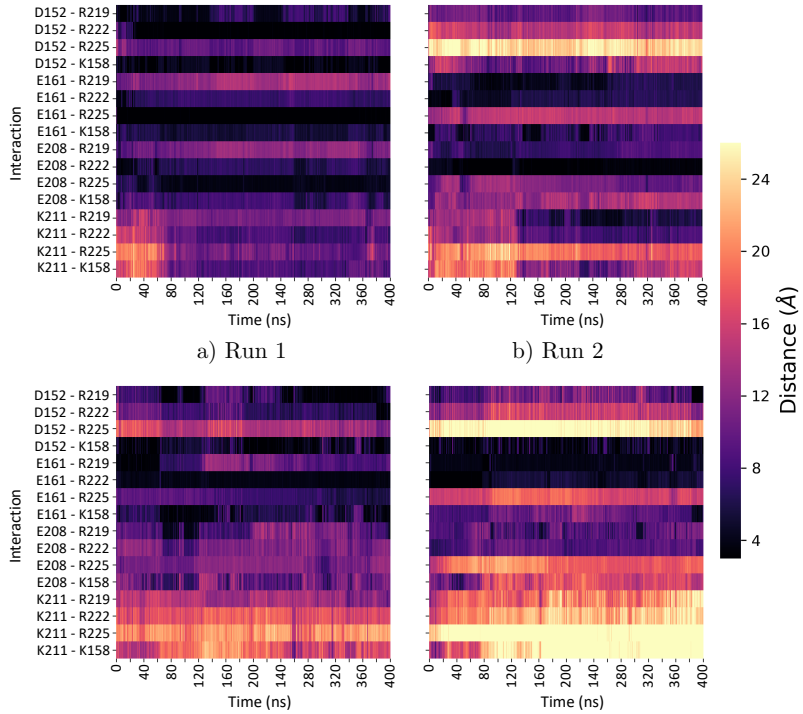
In mH1, R219 interacts strongly with D152 (located in the S1-S2 linker) with a mean distance of below 5 Å in three of the four repeats. This trend is followed with E161 (located at the top of the S3 helix), albeit at a slightly increased and variable distance across repeats. The interaction between E208 (located at the top of the S3 helix) follows a more convincing trend across repeats, with a fairly consistent interaction of below 10 Å. With the D211K mutation, a positive lysine residue is introduced in the S3-S4 linker and so in close proximity to R219. In D211K the average distance between D152 and R219 slightly increases in runs 1 and 3, with a more dramatic change seen in runs 2 and 4 where the distances increase to around 10 Å. This difference can be attributed to the steric blocking of R219 by K211 in run 2 and the "pulling apart" of the top of the VSD in run 4, making interaction with D152 less probable. Similarly the interactions between D152 and R225 also follow the same pattern. With E161 and E208 distances there is little change when compared to mH1.



a) Run 1 b) Run 2

c) Run 3 d) Run 4

Figure 3.12: Salt-bridge interactions present in the mH1 VSDI.



a) Run 1 b) Run 2

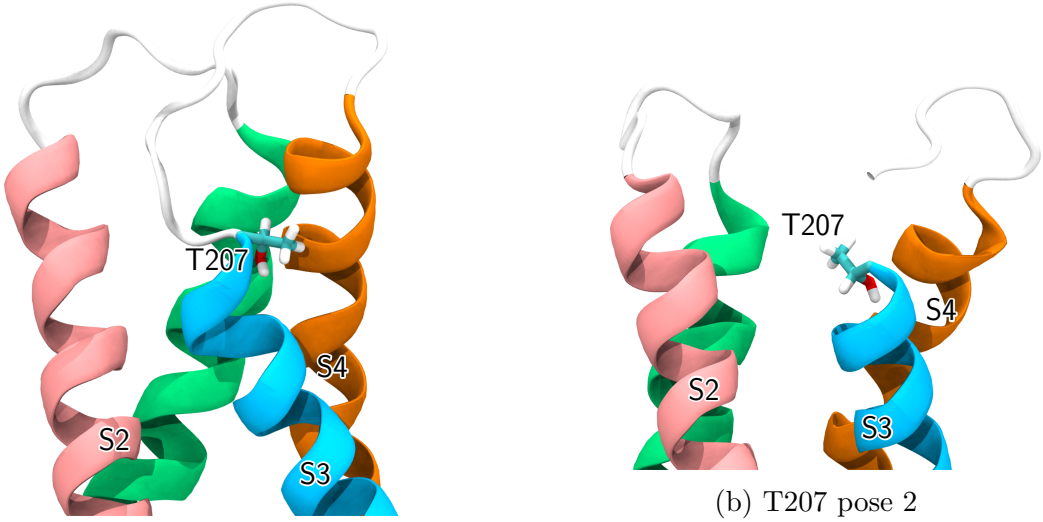
c) Run 3 d) Run 4

Figure 3.13: Salt-bridge interactions present in the D211K VSDI.

3.3.3 T207S mutant

3.3.3.1 Dihedral angle analysis

In mH1, the T207 residue contains a hydrophobic methyl group that preferentially orientates itself to face outwards and away from the centre of the VSD helices, as shown in Figure 3.14.



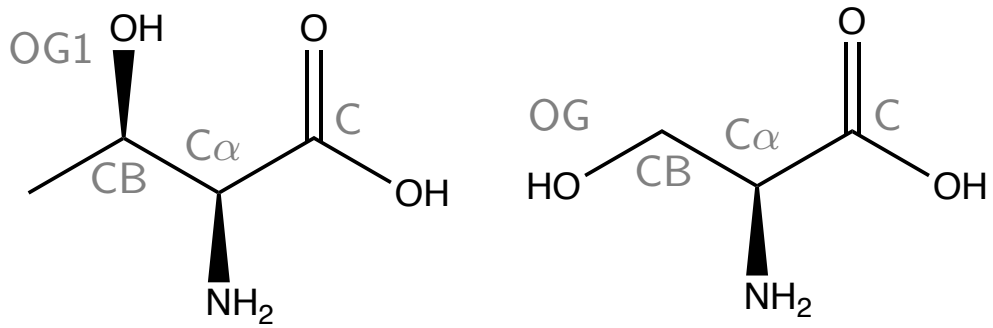
(a) T207 pose 1

(b) T207 pose 2

Figure 3.14: Two poses adopted by T207 in mH1 simulations. (a) The methyl group is pointing away from the VSD and towards the lipid membrane. (b) The methyl group is pointing towards the centre of the VSD, part of the S3-S4 linker removed for clarity. S1-4 helices are shown in green, pink, cyan, and orange respectively, with linker regions shown in white.

This outward facing mode is preferential due to the hydrophobic effect present between the methyl group and POPC lipid tails. Additionally, the hydroxyl group (present in both threonine and serine) will preferentially take part in hydrophilic interactions. Since the inner lining of the VSD is exposed to water and ions from the extracellular and intracellular regions, the hydroxyl group would be expected to prefer orientating itself towards this region. Loss of the methyl group (as in T207S) would not only disrupt this balance of interactions, but also decreased branching of the side chain would be expected to lead to an increase in conformations sampled

by S207. Indeed, it became clear that the loss of the methyl group caused S207 to become much more flexible over the course of each repeat simulation performed. To quantify the preferred orientation T207 adopts in mH1 (and similarly S207 in T207S) the dihedral angle (χ_1) was measured. χ_1 was defined between the atoms C, Cα, CB, and OG1 in T207 and atoms C, Cα, CB, and OG in S207 (Figure 3.15).



(a) Threonine (b) Serine

Figure 3.15: Structures of threonine and serine with atoms used for dihedral angle calculations shown.

A comparison of dihedral angles (χ_1) adopted in all runs of mH1 and T207S is shown in Figure 3.16. In mH1, the dihedral angle of T207 occupies a relatively narrow region of values with the highest density present at $\chi_1 \approx 38°$ with a much lower density of spread between $-20° \leq 0° \leq 20°$. In T207S, χ_1 samples a much wider range with a bimodal distribution. Although density remains in the $\chi_1 \approx 38°$ region it is much less compared to mH1 with a much wider distribution of angles sampled between $-40° \leq 0° \leq 20°$.

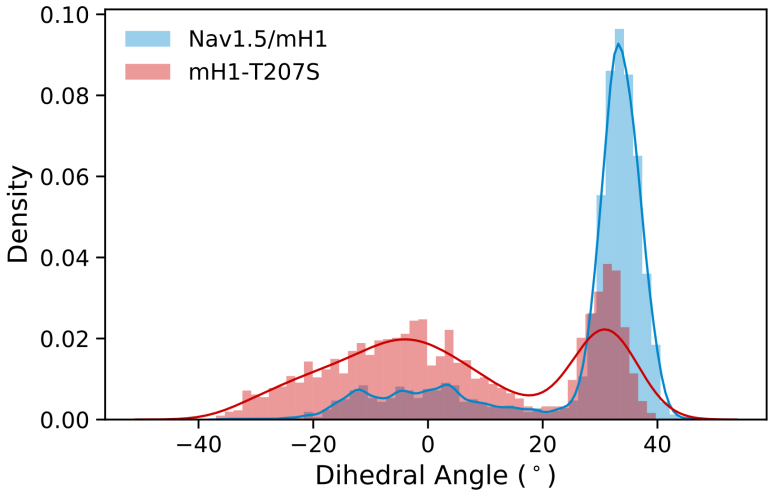


Figure 3.16: Dihedral angles of T207 in mH1 (blue) and S207 in T207S (red). Dihedral angles were collected over all repeats.

3.3.3.2 Salt-bridge analysis

The increase in conformations sampled by S207 in T207S may destabilise the top of the S3 helix and potentially disrupt any salt-bridges formed with the VSD. One residue that is in close proximity to position 207 is E208, located in the S3-S4 linker region (Figure 3.17). The destabilisation caused by the increased flexibility of S207 in T207S may cause a change in the interactions E208 is involved in. To investigate this, salt-bridge measurements were analysed within the VSD for each run (see Appendix B, Figure B.1) with a particular focus made on the distance between E208 and the positively charged residues in the S4 helix, notably R219 and R222.

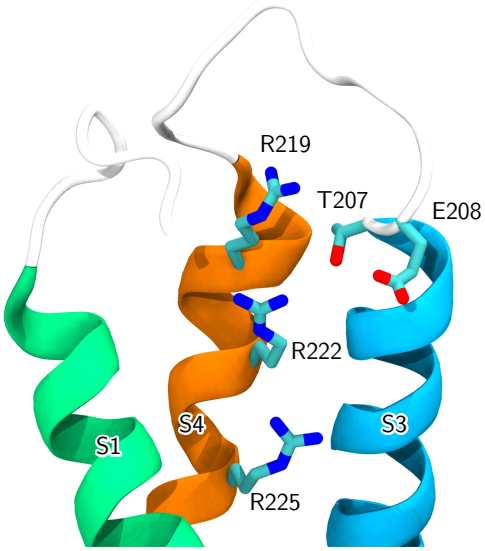
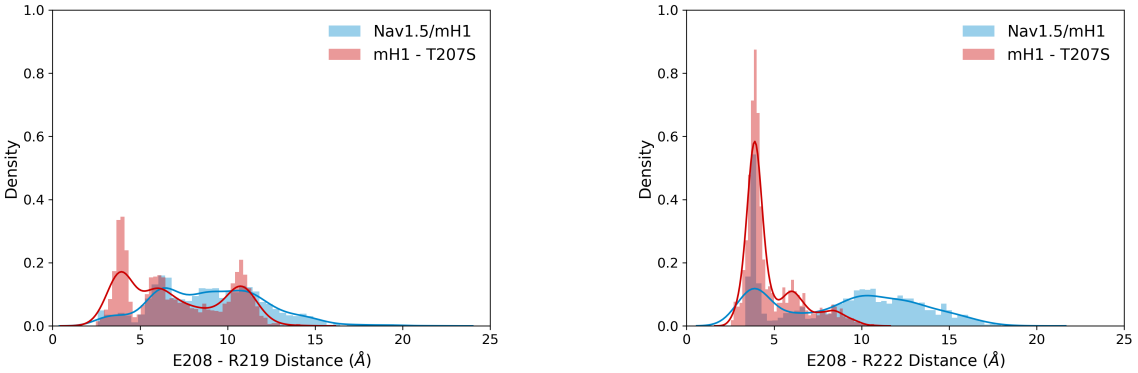


Figure 3.17: Location of E208 in relation to T207 in mH1. The S1 helix is shown in green, S2 is removed for clarity, S3 and S4 helices are shown in cyan and orange, linker regions are shown in white. Relevant residues are shown as sticks with hydrogens removed.

Both R219 and R222 are located at the top of the S4 helix and are close to E208. In mH1, E208 seems to take part in transient interactions with the positively charged residues in the S4 helix (Appendix B, Figure B.1). It may be expected that the interactions between E208 and R219 would be more frequent and closer in length than with R222 due to their spacial proximity. However, between repeats of the mH1 system it seems that the interactions are non-specific, with only a slight preference shown for R222. In contrast, there is a subtle change in salt-bridge interaction with the T207S switch. In these runs the interaction between E208 and R222 is persistent, with an average distance consistently below 6 Å.

The transient interaction between E208 and R219 can be seen in Figure 3.18a, where a histogram of distances collected over all repeats shows very little difference between mH1 and T207S, possibly due to R219 being further away than R222 in the VSD as well as exhibiting increased dynamics due to being closer to the extracellular region. The same analysis was performed for the E208 - R222 distance, and conversely there is a noticeable difference between the E208 - R222 distance in mH1 and T207S

(Figure 3.18b). In T207S the distance E208 - R222 distance is typically around a value of 4 Å, exhibiting a much more restricted set of distances when compared to mH1.



(a) E208 - R219

(b) E208 - R222

Figure 3.18: Histograms of E208 distances to R219 & R222 in T207S VSDI.

3.3.4 Double D211K & T207S mutant

3.3.4.1 D211K dynamics

As highlighted in Section 3.1.3, both T207S and D211K in exon 6 are necessary for the differences seen in activation between mH1 and Nav1.5e. Although subtle, there is a disruption to the top of the VSD where now the combined effects of the D211K and T207S switch are realised. The introduction of lysine at position 211 again causes a general increase in distance between K211 and the positively charged arginine residues that line the top and middle of the S4 helix, as seen in the single D211K switch in Section 3.3.2. This is most notable in run 4, where the top of the VSD opens up with a persistent interaction of K211 with the phosphate headgroups of surrounding POPC lipid (Figure 3.19a). However, K211 is dynamic and in the same run adopts a "snorkelling" orientation, protruding towards the ECD. This allows E208 to interact with the arginine residues at the top of the S4 helix and is accompanied by increased distortion at the top of the S3 helix (Figure 3.19b).

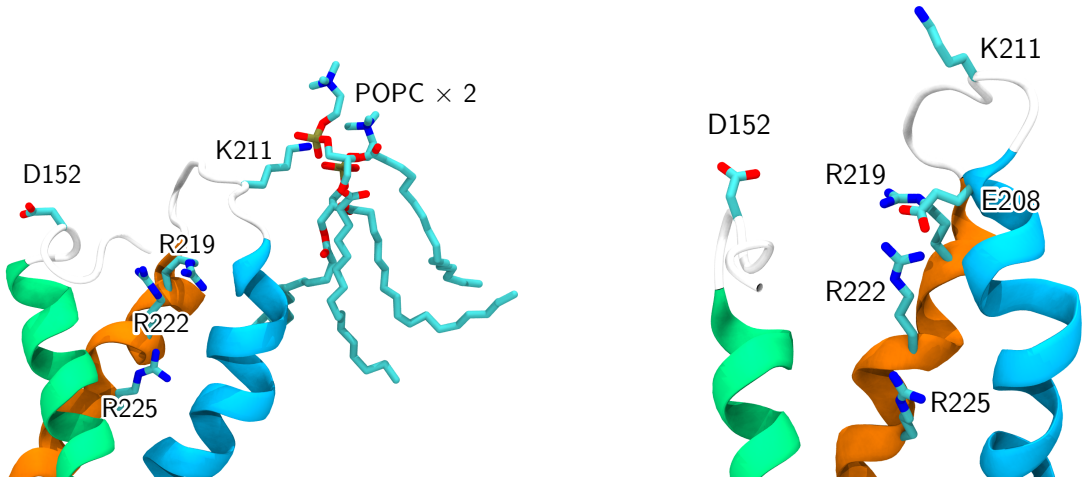
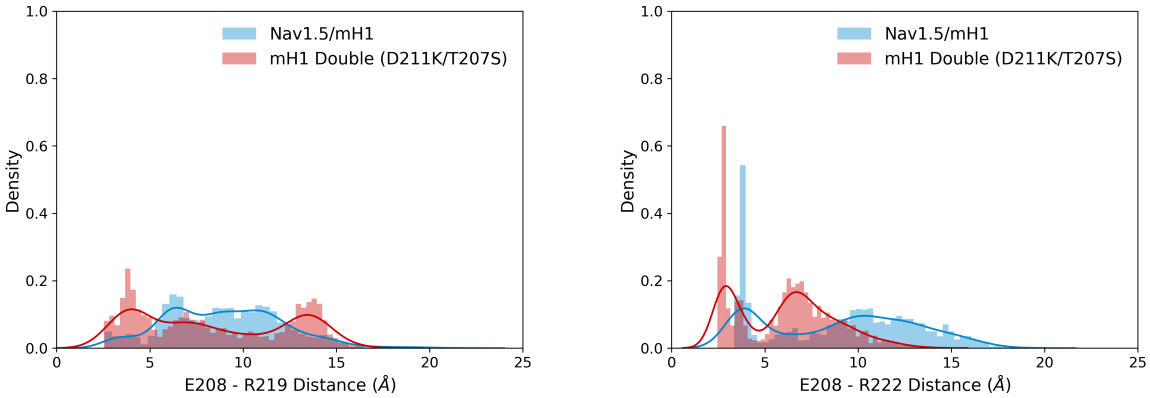


Figure 3.19: The two orientations typically adopted by K211 in the double D211K/T207S switch. (a) K211 interacts strongly with the phosphate group of the surrounding POPC lipid. (b) K211 protrudes upwards, towards the extracellular domain, with increased interaction of residue E208 and the arginine residues at the top of the S4 helix.

3.3.4.2 Salt-bridge analysis

In the single T207S switch, there is increased interaction of residue E208 with the arginine residues in S4. In the double D211K/T207S switch this pattern is present again. All repeats exhibit a general decrease in distances between E208 and the S4 arginine residues, notably at the top of the S4 helix (see Appendix B, Figure B.2 for the full set of salt-bridge distances in each repeat). This change is subtle and is most easily seen through a histogram of distances between E208 and R219 / R222, as shown in Figure 3.20. Here, the change in distances between E208 and R219 is not so obvious, however the change in distances between E208 and R222 is much clearer. Both of these trends follow the same pattern as seen in the single T207S switch (see Figure 3.18). Upon the double D211K/T207S switch, the distance between E208 and R222 tends to decrease across repeats to a value of approximately 2.5 Å, in addition to a second distance just over 5 Å. This is in contrast to mH1, where the distance is

highly variable, sampling distances over the range of ≈ 4-17 Å.



(a) E208 - R219

(b) E208 - R222

Figure 3.20: Histograms of E208 distances to R219 & R222 in Double D211K/T207S.

By analysing the dihedral angles visited by T207S in the double D211K/T207S the same general trend of a wider range of χ_1 sampled (as seen in the single T207S switch) is observed. The double D211K/T207S switch gives rise to a bimodal distribution of χ_1 values (Figure 3.21), centred around 38° and -18° that correspond to pose 1 and pose 2 (as detailed earlier in Figure 3.14). The range of each χ_1 distribution is roughly the same as in the single T207S switch (Figure 3.16).

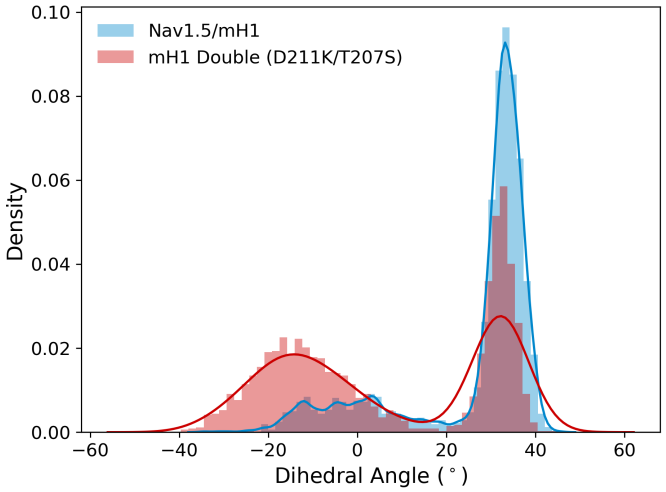


Figure 3.21: Dihedral angles of T207 (blue) and S207 (red), in mH1 and T207S respectively. Dihedral angles were collected over all repeats.

A possible combined effect of the double D211K/T207S switch can be seen in Figure 3.22. In this orientation, K211 is orientated towards and inside the VSD, potentially blocking S4 movement via unfavourable electrostatic repulsion between K211 and R219 / R222. Favourable electrostatic interaction is achieved through the attraction between E208 and R219 / R222.

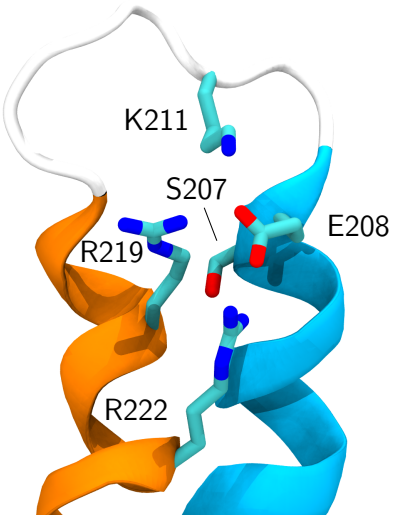


Figure 3.22: Combined E208 and K211 interaction with R219 and R222 in the S4 helix. Snapshot taken from run 3 of the double D211K/T207S simulation. S3 and S4 helices shown in blue and orange respectively, with the linker region shown in white.

3.4 Discussion

3.4.1 D211K dynamics

Position 211 is located in the S3-S4 linker region of VSDI. This region sits on the extracellular side of lipid membrane and adopts a loop structure. In this study, MD simulations revealed that the interactions of D211 in mH1 with the positively charged residues of the S4 helix were generally specific to R219 and R222 that line the top of the S4 helix. However, these were typically unstable and over the course of each run D211 tended to partake in several different interactions within the VSD. Analysis of R219, R222, and R225 reveals a general movement towards the extracellular region

that may be promoted, in part, by the presence of D211 in the S3-S4 loop. This was supported by lipid analysis of D211, where there is generally little interaction with the surrounding membrane, firstly due to the afore mentioned favourable charge interaction and secondly due to the short carbonyl chain in aspartic acid that restricts its reach out towards the surrounding membrane.

As highlighted in Section 3.1.3, the D211K in both the forward and reverse direction only caused a partial shift in activation, in disagreement with Onkal *et al.* [209], who proposed that the entire shift was due to the D211K mutation alone. In the simulations performed here, there is indeed a difference when compared to mH1. For instance, lipid contact analysis reveals a much higher frequency of contact of K211 compared to D211 in mH1. The longer carbonyl chain of K211 allows the residue to extend further out into the surrounding membrane and so increases the likelihood for the positively charged -NH_3^+ headgroup to interact with the negatively charged phosphate groups in the surrounding POPC bilayer. Contact of K211 with other charged residues in the VSD is altered too, there is a general increase in distance between the positively charged arginine residues at the top of the S4 helix.

Initially, it is sensible to assume this is due to the positive - positive charge repulsion between the arginines of S4 and K211 in the S3-S4 linker. However, simulations reveal that there may be multiple mechanisms through which S4 motion is restricted. For example, in run 4 there is large distortion of the top of the VSD due to K211 interacting strongly with the POPC lipid but in run 2 it appears that K211 enters the VSD and is stabilised by N214 and a POPC molecule, allowing it to position itself above R219 (Figure 3.11). In both of these runs there is restricted movement of the S4-lining arginines, with some residues moving downwards and towards the intracellular region.

These findings, although subtle, are in agreement with the results found from Mancino *et al.* [199]. In this study, simulations of the D211K switch suggest that K211

is much more flexible than D211 in mH1. The alternative conformations adopted by K211 in the S3-S4 linker are an example of this region's flexibility within the VSD. When K211 is orientated away from the VSD helices there is pronounced interaction with the surrounding lipid bilayer, however when orientated towards the VSD the positively charged -NH_3^+ group is able to sit above the S4 helix and stabilised by lipid species and / or surrounding residues.

3.4.2 T207S dynamics

The full shift in activation seen in Nav1.5e experimentally [199] was only achieved through the additional mutation T207S. The structural basis for how this switch could recover the Nav1.5e shift seems obscure, however simulations show that the single T207S switch causes the top of the S3 helix to become more unstable. This instability can be attributed to the loss of the methyl group in the threonine side chain, allowing the residue to explore an increased conformational space when compared to mH1. In T207, the hydroxyl group is positioned in such a way as to promote helix stability through hydrogen bonds with the helix backbone atoms, whilst the bulkier methyl group points out towards the membrane. In S207, this bulky side chain is lost and dihedral angle analysis reveals that roughly two rotomeric states are occupied. Although the loss of a methyl group is a relatively minor change, the increased instability at the top of the S3 helix allows the neighbouring E208 residue to become more flexible. Distance analysis between E208 and the positively charged arginine residues reveals no significant change in interaction with R219, however the overall interaction with R222 is increased with a bond distance of around 4 Å being most favourable. This increased interaction between E208 and R222 may inhibit the overall movement of the S4 helix by pinning R222 within the VSD. Analysis of the translocation the S4 arginine residues supports this, with neither promotion or inhibition of R222 movement over the course of each simulation.

3.4.3 Potential inhibition of S4 movement by combined D211K & T207S interactions

As stated previously, the combined D211K and T207S switches are required to produce the full depolarised shift in activation. Simulations of the double D211K/T207S switch show the presence of multiple conformations in VSDI that may inhibit the movement of the S4 helix and subsequent opening of the ion conducting pore in Nav1.5e.

As seen in the single D211K, there is substantial interaction of D211K with the surrounding lipid bilayer in double D211K/T207S. When not interacting with POPC, K211 exhibits highly dynamic behaviour above the VSD that can be attributed to its longer carbonyl tail. Interestingly, in one run, K211 faces towards the VSD, interacting with E208 that stabilises R219, R222, and K211 at the top of the S4 and S3 helices where it remains for almost a third of the trajectory. This is somewhat analogous to the single D211K case, where K211 enters the VSD and is stabilised in a position above the S4 helix by N214. Focusing on the T207S mutation in the double D211K/T207S, there is again a bimodal distribution of dihedral angles visited as seen in the single T207S switch. This is also accompanied by a bimodal distribution in the E208-R222 distances, with a distance of around 4-5 Å preferred, as well as a bond distance between 5-10 Å.

3.4.4 Conclusions

The D211K switch removes the presence of a negative charge in the S3-S4 linker and replaces it with a highly mobile, positively charged lysine. At first, it may be sensible to assume that the electrostatic repulsion between K211 and R219 at the top of the S4 helix would be responsible for the inhibition of S4 movement and eventual pore opening. In this work, it has been shown that the situation may be more complex.

Firstly, the D211K induces steric strain to the top of the VSD through increased interaction with the surrounding lipid bilayer, granted by the low-branched nature of the lysine side chain. Secondly, when facing towards the VSD, K211 is able to extend further over the S4 helix. This conformation is not necessarily stable and in simulations is typically stabilised by surrounding residues or lipid molecules. Upon the introduction of the T207S mutation, the E208 residue is allowed to interact more frequently with residues in the S4 helix, as well as stabilising K211 inside the VSD.

Overall, the introduction of both the D211K and T207S switches disrupts the carefully tuned electrostatics within the VSD. The movement of the S4 helix is crucial in the ion channel's ability to sense a change in voltage and its movement relies upon interactions with negatively charged residues, predominantly located within the S2 helix [227, 228]. The swapping of charge in the S3-S4 linker, as well as the rotomeric instability caused by T207S in the S3 helix, alters the overall structure of the VSD and impinges S4 movement through a variety of conformational changes.

More generally, the specific activation order and importance of each VSD in Nav channels remains an active area of research. Recently DIV has been shown to be important in channel inactivation [94], whilst VSDs in domains I-III are responsible for channel activation in the order DIII $\rightarrow$ DI $\rightarrow$ DII, as verified by voltage-clamp fluorometry [229, 230]. This demonstrates DI's central role in the channel's ability to function correctly and may explain why there is such a high level of splicing in the DI S3-4 region across the family of Nav channels [206–209].

For future studies, it is worth highlighting that the model used in this work was based on a relatively low quality crystal structure isolated from cockroach at 3.8 Å resolution [210]. More recent human structures such as those found by Pan [4] and Shen *et al.* [5] (described earlier in Chapter 1) may be more suitable in model construction, as well as having a higher sequence identity with human Nav1.5.

"You've got to think, eat, live, and breath it."
— Wendy (mother)

4

Atomistic simulations of the β3 subunit

Contents

4.1 Introduction 86
4.1.1 β3 Ig trimer 87
4.1.2 Overview of states: a paradox 90
4.1.3 A conserved glutamic acid residue 91
4.2 Methods 92
4.2.1 Model generation 92
4.2.2 Molecular dynamics simulations 94
4.2.3 Ig domain orientation analysis 95
4.3 Results 97
4.3.1 β3 Ig stability 97
4.3.2 Dynamics of the full length monomer construct 101
4.3.3 An investigation of β3 glycosylation sites 114
4.3.4 Dynamics of the full length trimer construct 118
4.4 Discussion 127
4.4.1 The varied dynamics of the Ig domain 127
4.4.2 Destabilisation of C-terminal residues in the β1 & β3 helices 130
4.4.3 Glycosylation: A preferred Ig orientation? 133
4.5 Conclusions 134

Note: Parts of this chapter are based on the open access publication [231]: W. G. Glass, A. L. Duncan, and P. C. Biggin, Computational Investigation of Voltage-Gated Sodium Channel β3 Subunit Dynamics, *Frontiers in Molecular Biosciences*, vol. 7, p. 40, 2020. **DOI**: https://doi.org/10.3389/fmolb.2020.00040

4.1 Introduction

The α subunit of Nav channels has been heavily studied due to its critical role in the initiation of the action potential. As discussed previously, the movement of the S4 helices in each of the four VSDs are ultimately responsible for the central pore opening and the subsequent influx of sodium ions into the cell. In Chapter 3, it was shown that electrostatic disruption in the VSD can cause a change in the finely tuned conformational dynamics.

On a structural level, the exact location of where β subunits bind to the α subunit, and facilitate any change in gating dynamics, is uncertain. Additionally, the exact ratio of α:β subunits is also not very well characterised and may vary depending on tissue type and the cellular environment [232]. Evidence from fluorescence studies suggests that both β3 and β1 can bind to the α subunit and may alter the rate of fast inactivation through interaction with VSDIII and VSDIV respectively [233]. The advent of cryo-EM is beginning to improve our understanding of how the α and β subunits may interact. Recently, structures from Yan and Pan [3, 4] (PDB codes: 5XSY and 6AGF respectively) exhibit the β1 subunit bound to Nav1.4 in both Electric Eel and Human variants. In both of these structures the β1 subunit is bound to VSDIII (between DIII and DIV) and close to the IFM motif that, in combination with DIV, is critical to fast inactivation [4, 92, 94, 234].

Increasingly, there is evidence to suggest that Nav channels may in fact operate in higher order complexes, such as dimers [100], and can also form complexes with many

other proteins involved in a variety of signalling pathways [235]. As outlined earlier (Chapter 1, Section 1.2), β subunits also function independently as CAMs [60,75,236] and may play a role in Nav channel clustering at the nodes of Ranvier [71] to promote the propagation of the AP. Additionally, evidence from Namadurai *et al.* [99] suggests that the β3 subunit can oligomerise to form higher order structures and even cluster multiple α subunits on the cell membrane.

The exact dynamics of β subunits remain to be fully characterised in isolation. In this study, the isolated β3 trimer and more recent full-length β1 subunit are used as a starting point for atomistic MD simulations. These offer a highly detailed description of how these related proteins may behave *in vivo* and help to explain the structural basis of observed higher order organisation.

4.1.1 β3 Ig trimer

The ability of β subunits to form higher order oligomers gained substantial experimental grounding with the crystallisation of the β3 Ig domain. Namadurai *et al.* [99] characterised, by atomic force microscopy (AFM) and ultracentrifugation, the Ig domain in a variety of oligomeric states. These included monomers, dimers, and trimers (Figure 4.1a). Interestingly, the α subunit forms dimers and trimers independently, with such structures more prevalent when in the presence of β3 subunits. Higher order β3 oligomers also tend to cross-link α subunits (Figures 4.1b, 4.1c, 4.1d) and increase clustering of protein species on the plasma membrane.

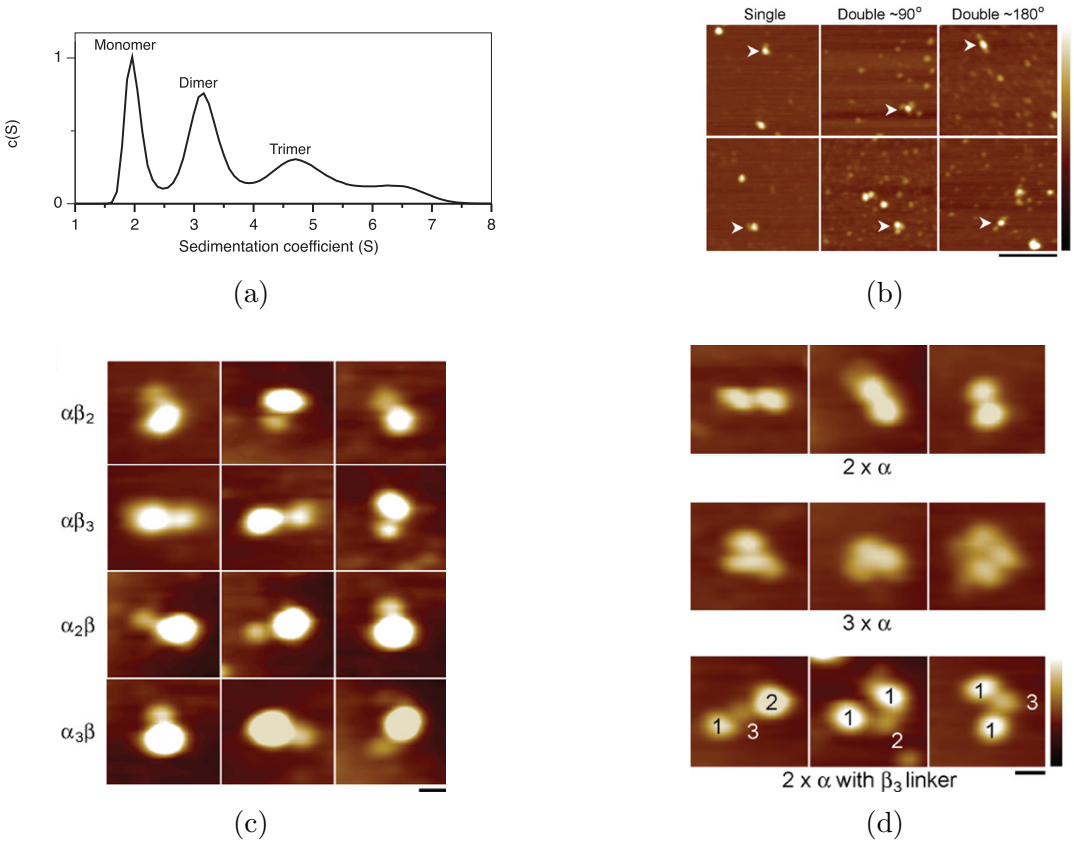


Figure 4.1: Structural features of the β3 Ig domain. (a) Component sedimentation coefficient distribution with monomeric, dimeric, trimeric (and larger) species. (b) Low magnification AFM images of a sample containing coisolated Nav α/β3 subunits. Arrowheads indicate α subunits singly decorated by β3 subunits (left panels) or doubly decorated at either $\approx 90°$ (centre panels) or $\approx 180°$ (right panels). Scale bar = 200 nm; height scale = 0-5 nm. (c) Examples of decoration events involving multiples of either α subunits or β3 subunits, as indicated. (d) Examples of double (upper panels) and triple (centre panels) α subunits and of α subunits connected by β3-subunit oligomers (lower panels). The numbers in the lower panels indicate the stoichiometries of the various particles. Scale bar = 20 nm; height scale = 0-3nm. All images adapted from Namadurai *et al.* [99].

At the centre of the crystallised β3 trimer, three hydrophobic residues (V27, V29, and P30) are responsible for the favourable interaction between each chain. Typically, the first 10 amino acids of Ig domains are stabilised by an antiparallel β sheet, but in β3 this is not the case. The region is stabilised, and constrained, by an intramolecular disulphide bond (C26-48). Increased stability is also coupled with the E28-R139 intramolecular salt-bridge that forces the key hydrophobic residues to point out and

away from the main body of the Ig domain, causing them to be surface-exposed which may promote the observed oligomerisation. Other contributions to stability include the formation of hydrogen bonds within the Ig domain between R124 and T58. These interactions are highlighted in Figure 4.2, for an in-depth overview of all interactions the reader is directed to Table 3 in the original paper [99].

Namarudai *et al.* highlight the importance of the C26-48 bond using a C48A mutant. Critically, the increased flexibility granted by the C48A mutation causes the hydrophobic patch responsible for the favourable interaction at the trimer interface to become unstable. The increased flexibility results in a reduced number of observed β3 oligomers, with those that are present in a "splayed" state.

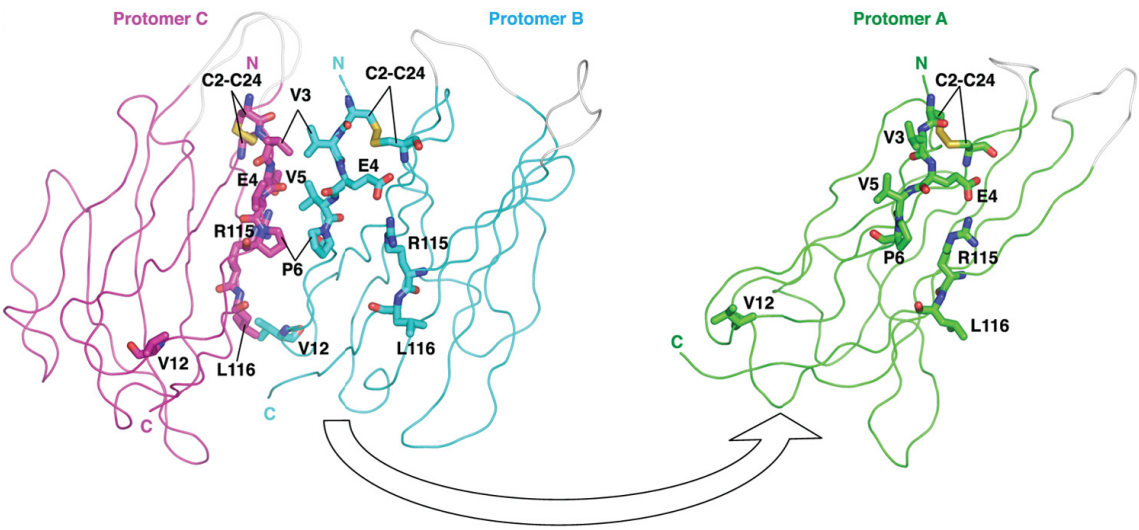


Figure 4.2: Schematic of the location of key trimer interface residues. For clarity, protomer A (green) is removed from the rest of the trimers by application of 180° rotation around its centre of mass and rightward translation. Image adapted from Namadurai *et al.* [99] with residue numbering -24 of the canonical sequence.

4.1.2 Overview of states: a paradox

The recent structures of the Nav α / β subunit complexes revealed that the β1 Ig domain adopts a conformation such that the long axis of the Ig β sheets sits parallel to the membrane surface. When bound to the pore-forming α subunit, this orientation allows residues 20-27, that are analogous to the hydrophobic Ig interface responsible for observed β3 olgomerisation, to interact with the top of the VSDIII. When extending this bound conformation to β3, it becomes apparent that if the β3 subunit adopted the same trimeric structure as observed for the Ig domain only, its interaction with the membrane would most likely require some substantial conformational re-arrangement. Figure 4.3 indicates the unrealistic conformation adopted by the β1 subunit if it were to form a similar trimer to that found in β3.

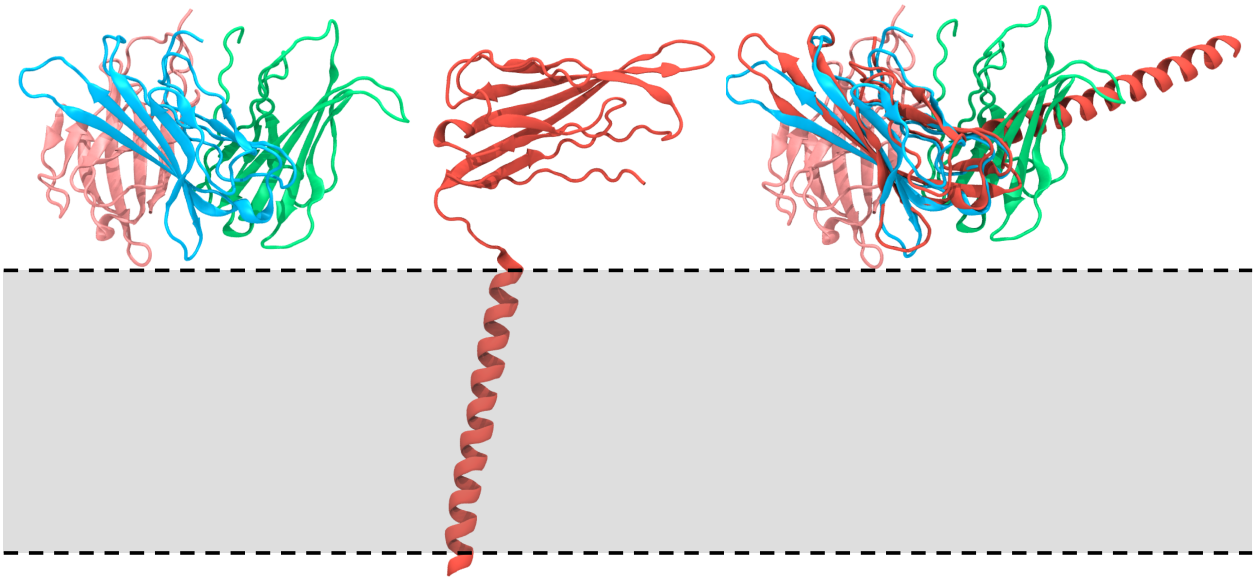


Figure 4.3: Overlay of the β1 model with the β3 Ig trimer. *Left*: β3 Ig trimer model. Separate chains are shown in green, cyan, and pink. *Centre*: β1 model. *Right*: Overlay of the Ig domains of the β1 subunit and the trimeric structure of the β3 Ig domain. The approximate position of the membrane is shown as a grey rectangle bound by dashed lines.

This paradox highlights the need for, at least, the $\beta 3$ subunit to change its orientation whilst embedded in the membrane. Critically, this relies upon the flexibility and dynamics of the short linker region between the top of the TMD and the start of the Ig domain. This is primarily due to the Ig domain essentially acting as a rigid body, since it contains a large number of highly stable β sheets.

4.1.3 A conserved glutamic acid residue

The trimeric crystal structure of the Ig domains does not include the single-pass TMD and the extent to which the TMD is involved in binding of multiple β subunits (as well as α subunit cross-linking) is unclear. Furthermore, across the β subunit family a conserved glutamic residue (E177/E176 in $\beta 1$/$\beta 3$ respectively) is present in the middle of the predicted transmembrane helix. In $\beta 3$, this is predicted to be located between residues 160 and 180, positioning E176 close to the intracellular side of the helix. The Pan *et al.* [4] structure reveals that the analogous E177 in $\beta 1$ is in fact present in a much longer helix between residues 155 and 190, around 15 residues longer than in $\beta 3$. Interestingly, E177 is positioned facing away from the α subunit, suggesting it may not be directly involved in α-β binding. The alignment of the $\beta 1$ helical sequence from Pan *et al.* [4] and the predicted helical sequence of $\beta 3$ is shown in Figure 4.4.

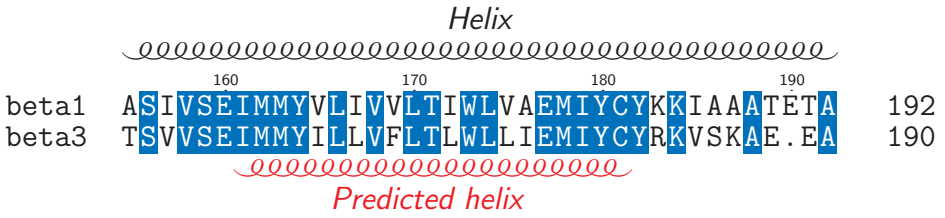


Figure 4.4: Alignment of the TMD of the $\beta 1$ and $\beta 3$ subunits. Location of resolved helix by Pan *et al.* [4] is shown in black and the predicted $\beta 3$ helix location is shown in red. Numbering of $\beta 1$ residues is shown above both sequences.

4.2 Methods

4.2.1 Model generation

4.2.1.1 β1 monomer

The model used for β1 simulations was constructed directly from the Nav1.4-β1 structure [4] (PDB code: 6AGF) since all residues had been resolved.

4.2.1.2 β3 monomer

The recent crystal structure of the β1 subunit in the human Nav1.4-β1 complex [4] was used to construct a model of the human β3 subunit. A total of 200 models were created using the sequence alignment of full length human β3 with the crystal structure sequence (sequence identity = 48.8%). All sequences were obtained from UniProt [211]. Each model was created and scored using the DOPE function [196] in the Modeller (v9.16) software package [195]. The 10 best models were ranked using the QMEAN scoring function [197] in SWISS-MODEL [212] and the final model chosen with the highest QMEAN score.

4.2.1.3 Glycosylated β3 monomer

The Ig domain of the β3 subunit contains four predicted N-linked glycosylation sites [211] at N95, N109, N113, and N121. Using the generated β3 monomer (see Section 4.2.1.2), a single N-Acetylglucosamine (GlcNAc) species (structure shown in Figure 4.5) was added to each of the four asparagine residues with the recently developed *doGlycans* package [237], utilising the GLYCAM06h force field [238]. Each GlcNAc moiety added was in the β anomeric form and in the energetically favourable $^4C_1$ conformation. This conformation was based on the two resolved GlcNAc residues in the Pan *et al.* [4] structure at N113 and N121. The final glycosylated model was visually inspected using VMD [224] for any steric clashes, of which there were none.

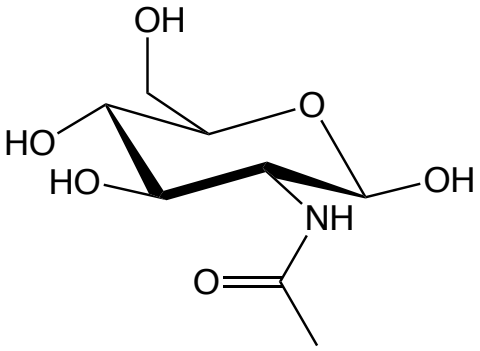


Figure 4.5: Structure of GlcNAc.

4.2.1.4 β3 ECD

The crystal structure of the β3 subunit [99] (PDB code: 4L1D) was used as a template to construct a model of the trimeric extracellular region. A total of 200 models were created and each model was scored using the DOPE function [196] in the Modeller (v9.16) software package [195]. The 10 best models were ranked using QMEAN [197] in SWISS-MODEL [212] and the final model chosen with the highest QMEAN score. The monomeric β3 ECD was created by extraction of one chain from the trimer model. All mutations were performed in Pymol (v1.4) [213].

4.2.1.5 β3 trimer

Over the course of the simulations of the β3 monomer model (see Section 4.2.1.2), a large variety of conformations were visited. To construct the β3 trimer model a frame from the first run was taken at a pitch angle of 44.7°, with the Ig domain approximately perpendicular to the plane of the membrane. This was aligned with each chain of the trimeric β3 ECD crystal structure (see Section 4.2.1.4). After the model was constructed, it was checked for steric clashes, of which there were none. All alignments were performed in PyMol (v1.4) [213].

4.2.2 Molecular dynamics simulations

Simulations were performed using GROMACS 2018 with the AMBER ff99SB-ILDN force field. All full length protein models were embedded in a POPC membrane, prepared using the InflateGRO methodology [220] and in-house scripts for final adjustments. Each system was solvated with the TIP3P water model [219] and neutralised with Na^+ and Cl^- ions at a concentration of 150 mM. Energy minimisation was performed using the steepest decent algorithm. Equilibration steps consisted of a short (1 ns) and long (5 ns) equilibration whilst position restraining the Cα atoms with a force constant of 1000 kJ mol^{-1} nm^{-2}. All simulations were carried out in the NPT ensemble. The temperature and pressure were set to 300 K and 1 bar using the Nosé-Hoover thermostat and Parrinello-Rahman barostat with coupling constants of 0.8 and 5.0 ps respectively.

A summary of all simulations performed is outlined in Table 4.1. In the following results sections the reader is referred to this table for details of each simulation unless where otherwise stated.

System	Domains	Box type	Box dimensions (nm)	Duration (ns)
β1 M	Full length	Rectangular	9 × 9	25 × 400
β3 M	Extracellular	Triclinic	-	10 × 200
β3 T	Extracellular	Triclinic	9.3 × 9.3	25 × 200
β3 T	Full length	Rectangular	13 × 13	4 × 400
β3 M	Full length	Rectangular	9 × 9	25 × 400
β3 (GlcNAc) M	Full length	Rectangular	9 × 9	6 × 400

Table 4.1: Overview of atomistic β subunit simulations performed. M = monomer and T = trimer. Box dimensions correspond to the x and y dimensions only.

4.2.3 Ig domain orientation analysis

Assessment of the Ig domains favoured orientation in the constructed full length β subunit models was achieved by calculating the principal axes (PAs) at each frame. These were then used to find and measure the Tait-Bryan angles (see Appendix A, Section A.3) using the basis $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ as a reference (where $\mathbf{e}_x = [1, 0, 0]$, $\mathbf{e}_y = [0, 1, 0]$ and $\mathbf{e}_z = [0, 0, 1]$). In order to calculate the PAs, the centre of mass was calculated for any residue involved in secondary structure within the Ig domain (i.e. the β-sheets and 3-10 helices). This was chosen so as to minimise any noise associated with flexible loop movement over the course of each simulation. The PAs $\mathbf{p}_1$, $\mathbf{p}_2$, and $\mathbf{p}_3$ were obtained via the diagonalization of the moment of inertia tensor, I, derived as shown in Equations 4.1 & 4.2.

$$\Lambda = U^T I U \tag{4.1}$$

$$I = \sum_{i=1}^{N} m_i \left[(\mathbf{r}_i \cdot \mathbf{r}_i) \sum_{\alpha=1}^{3} \mathbf{e}_\alpha \otimes \mathbf{e}_\alpha - \mathbf{r}_i \otimes \mathbf{r}_i \right] \tag{4.2}$$

Where m_i and r_i are the mass and coordinates of the ith particle respectively. $\mathbf{e}_\alpha$ (with $\alpha = 1, 2, 3$) represents the orthogonal Cartesian basis vectors, and $\otimes$ represents the tensor product. Finally, $U = (\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3)$ and Λ is a diagonal matrix of eigenvalues that correspond to the principal moments of inertia.

At every frame, the PAs were used to define a rotation matrix and from this the Tait-Bryan angles computed. Using an intrinsic rotation formalism of ZYX, the yaw, pitch, and roll angles were defined. An example schematic of the pitch angle calculation between a frame and the starting structure is shown in Figure 4.6. All angle analysis was produced from in-house python scripts using the MDAnalysis 0.19.2 [222, 223] module (see github.com/glass-w/protein_orientation). All graphs were produced using the Seaborn [225] and Matplotlib [226] Python libraries.

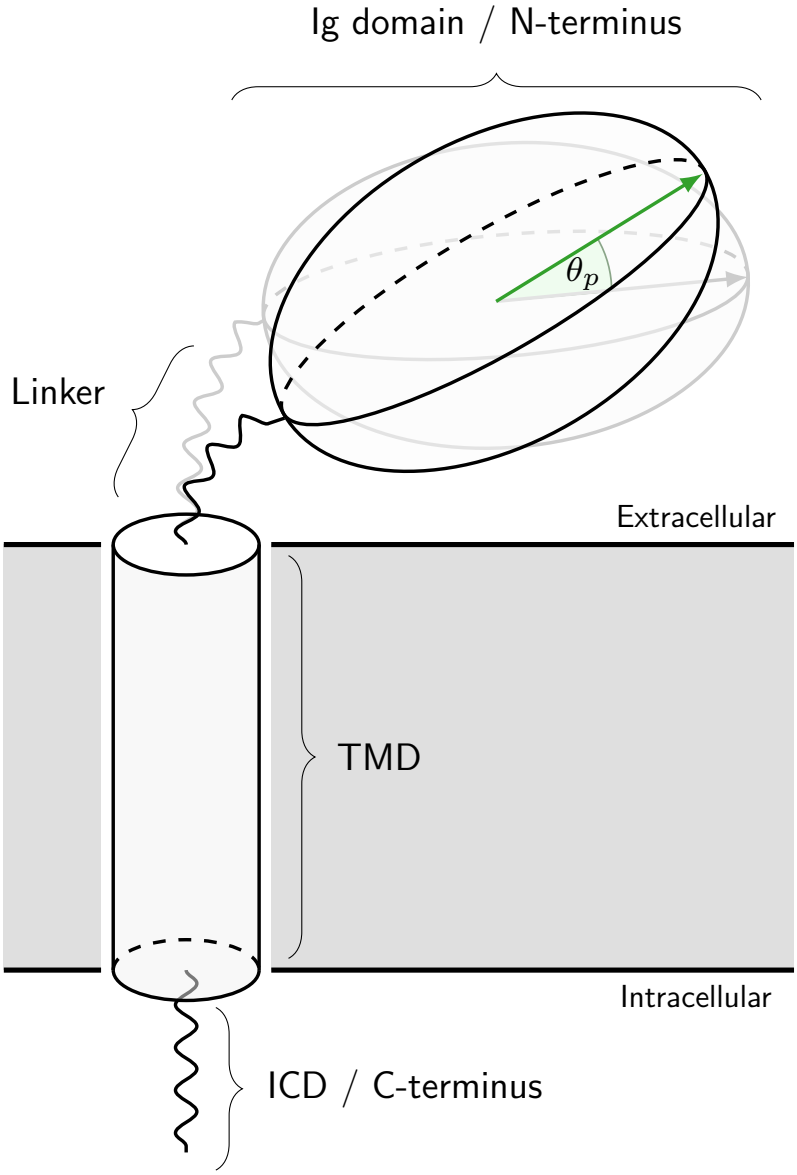


Figure 4.6: Schematic of the pitch angle calculation the Ig domain of the β subunit. The crystal structure conformation is indicated in light grey and the current frame as solid black lines. The Ig domain is represented as an ellipsoid, the short linker and unstructured ICD domain are shown as wavy lines, and the TMD is represented as a cylinder. The approximate position of the membrane is shown in thick black lines and grey rectangles. The pitch angle, θ_p, is defined as the angle between the principal axis (thick green arrow) of the current frame and a reference (the basis vectors e_x, e_y, and e_z, not shown) following the Tait-Bryan angle formalism.

4.3 Results

4.3.1 β3 Ig stability

The crystallised Ig trimer is stabilised by hydrophobic contacts between each of the three domains. As described in Section 4.1.1, critical to this interface are the hydrophobic V27, V29, and P30 residues. In order for these residues to remain in their established conformation they rely on the C26-48 disulphide bond. As found by Namarudai *et al.* [99], when this is disrupted through a C48A mutation the number of observed oligomers decreases on the plasma membrane, with those that do form in a "splayed" state. To address why this occurs, MD simulations were performed on both the monomeric and trimeric Ig domains of β3. Models were constructed as outlined in Section 4.2.1.4 to produce both the WT and the C48A mutant.

4.3.1.1 Monomer

The WT monomer systems remained stable over the course of each 200 ns run, with RMSD values centred around 15.4 Å. However, in C48A the destabilising effect caused by the loss of the C26-48 disulphide bridge causes a large conformational change in several repeats, most notably in run 9 (Figure 4.7b, 4.7d) where residues 24 to $\approx$ 34 move away from the Ig domain and explore a variety of conformations within solvent. When assessing the radius of gyration (R_G) across all repeats, there is a general trend for the C48A mutation to destabilise the Ig domain. Calculation of R_G was performed using all Cα atoms positions. Figure 4.7a indicates a rightward shift in R_G values with a significant amount of density still overlapping with WT values. Additionally, the root mean squared fluctuation (RMSF) of each residue in the Ig domain (Figure 4.7c) shows a general increase for residues $\approx$ 24 to 30. Other residues show little change in RMSF between WT and C48A.

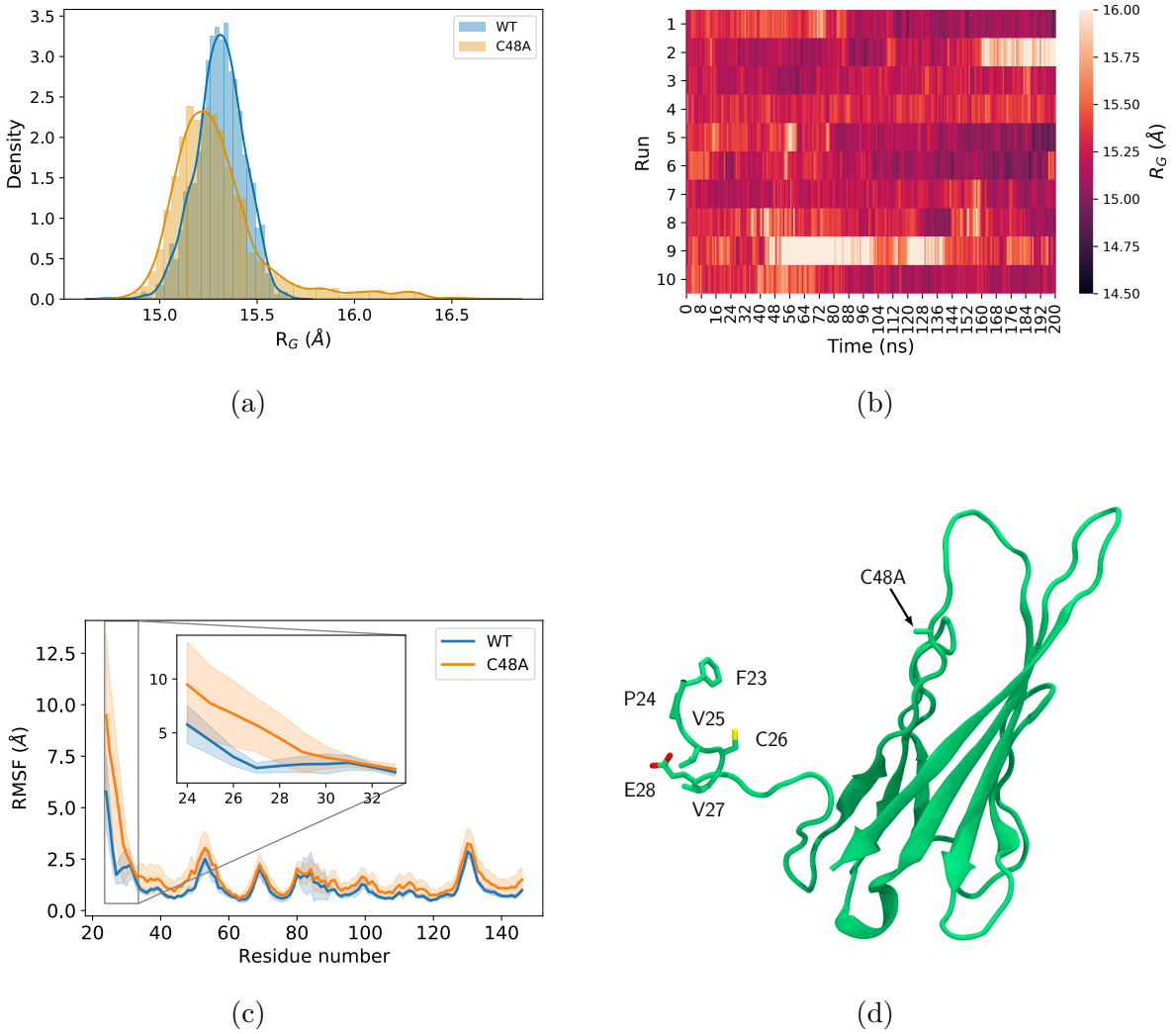
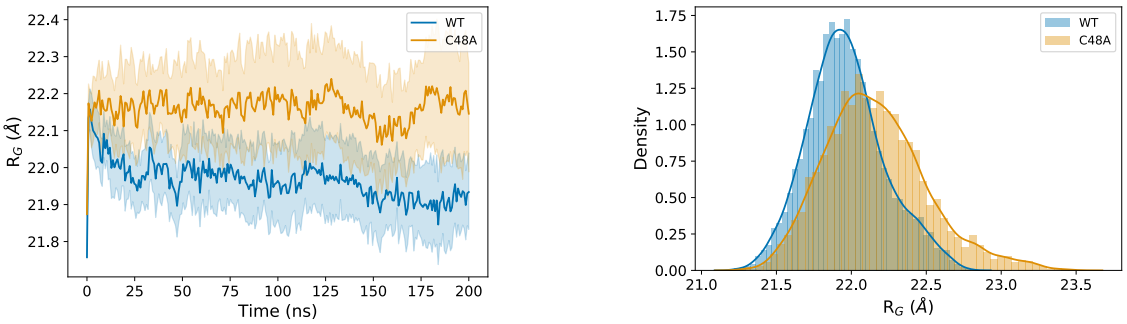


Figure 4.7: The effects of the C48A mutation on β3 ECD monomer dynamics. (a) Distribution of R_G values for both WT and C48A. (b) Heatmap of R_G values in C48A, large values correspond to the region defined between residues 24 to $\approx$ 34 breaking away from the Ig domain. (c) RMSF of residues in the β3 Ig domain. Inset focused on the first few residues where destabilisation occurs. (d) Snapshot of β3 monomer with residues 24 - 30, and C48A, shown as sticks.

4.3.1.2 Trimer

Similar analysis was performed on the β3 trimer. Both WT and C48A β3 subunits were reported to form oligomers on the plasma membrane where the C48A trimers exhibited a "splayed" state. In complex with other Ig domains, the effect of the C48A mutation may be examined both intra- and intermolecularly. As shown previ-

ously, the C48A mutation causes destabilisation of the residues responsible for this oligomerisation but how these residues interact assuming the same trimeric scaffold remains unknown. To quantify the splaying observed by Namarudai *et al.* [99], the R_G values for both the β3 WT and C48A trimer were measured, shown in Figure 4.8.

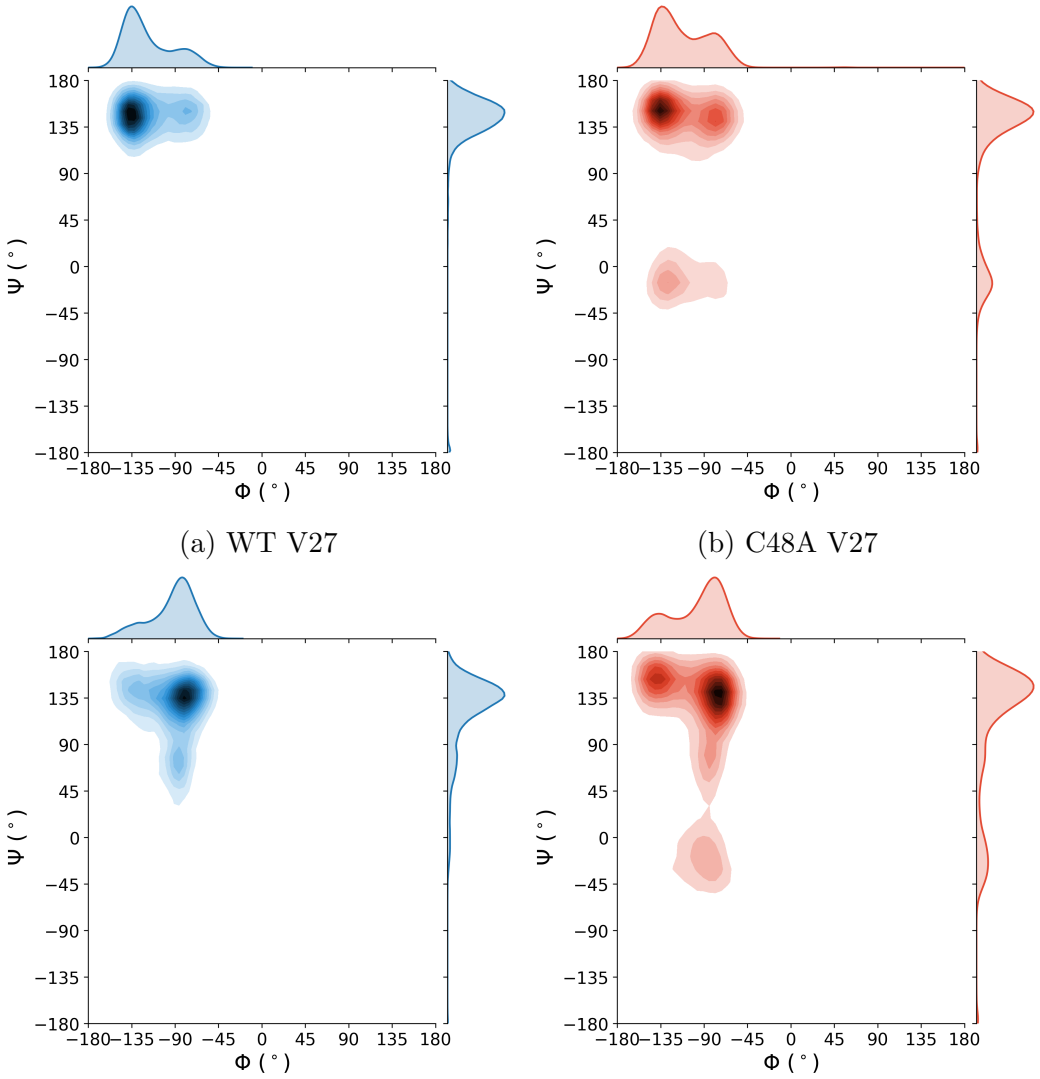


(a) (b)

Figure 4.8: Radius of gyration of the β3 ECD trimer. (a) Variation of R_G with time across 25 repeats for both WT and C48A. (b) Distribution of R_G values between WT and C48A.

The R_G of WT tends to decrease over the course of the 200 ns simulation while C48A remains consistently higher. It is important to note that WT β3 oligomers were found to be present in a number of oligomeric states on the plasma membrane, with the trimeric state being less favourable than the dimer or monomer. The overlap between between both WT and C48A R_G values supports this, indicating that even in WT there is some inherent instability in some of the runs.

The β3 trimer is conformationally restricted, especially at the trimer interface where residues V27, V29, and P30 in each chain are in close contact through favourable hydrophobic interactions. Due to the steric hindrance induced by the trimeric construct close to C26, the effects of the C48A mutation were also assessed using the torsional angles, ψ and ϕ, in the form of a Ramachandran plot (Figure 4.9). The ψ and ϕ angles for P30 were not calculated due to the residue's incorporation of its side chain with the protein backbone.



(a) WT V27

(b) C48A V27

(c) WT V29

(d) C48A V29

Figure 4.9: The ϕ and ψ angles sampled by V27 and V29 in the $\beta 3$ trimer model.

In WT, the torsional angles for V27 are $90 \leq \psi \leq 180$ and $-180 \leq \phi \leq -45$, for V29 there is an increase in spread with $0 \leq \psi \leq 180$ and $-180 \leq \phi \leq -45$. Although subtle, the changes in torsional angles sampled by both valine residues highlight the region's increased flexibility as a result of losing the C26 - C48 disulphide bond.

4.3.2 Dynamics of the full length monomer construct

Structures of the β1 subunit have only been isolated whilst in contact with the central α subunit [3–5]. Additional experimental techniques such as AFM, as discussed in Section 4.1.1, have also been successful in identifying the multi-oligomeric states that the β3 subunit may adopt *in vitro*. The trimeric crystal structure only contains the extracellular domain and the question of how the TMD may interact within the membrane, both as a monomer or as a higher order complex, remains to be fully characterised.

Models of both subunits were constructed as described in Section 4.2.1. In order to probe the likely conformation of the TMD and ECD in both β1 and β3 a total of 25 unbiased, atomistic MD simulations were performed for a total of 400 ns each. The large number of repeats enables a higher degree of sampling from the conformational landscapes of both subunits. Additionally, a direct comparison can be made to see if there are any conserved structural changes that may promote or hinder oligomerisation.

4.3.2.1 Helix analysis

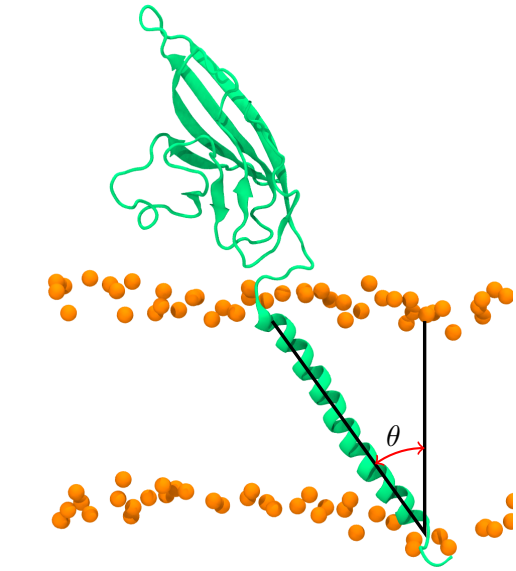
In the Nav1.4 - β1 crystal structure, the TMD of β1 is almost perpendicular to the membrane plane, with residues in the helix directly interacting with the S2 helix of the VSDIII. To address the dynamics of the TMD in both β1 and β3, the tilting and bending were analysed. For analysis, the HELANAL library [239] was used from MDAnalysis.

Briefly, four contiguous Cα atoms are used as a sliding window across the helix, this defines local helix origins and calculates properties such as local bend and helix shape. The helix tilt is defined through a vector of best fit through the calculated helix origins. The angle between this vector and the vector approximately perpendicular to the membrane plane (i.e. [0, 0, 1]) was used to define the helix tilt (Figure 4.10a).

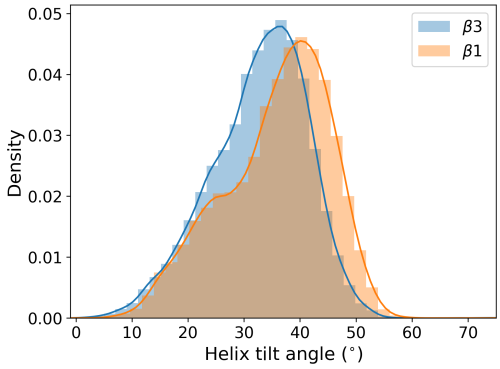
Analysis was run over all repeats for both the $\beta 1$ and $\beta 3$ subunits.

The tilt of both the $\beta 1$ and $\beta 3$ helices is not static and in fact occupies a somewhat large distribution of angles as shown in Figure 4.10b. A tilt angle of 0° corresponds to the helix aligned perpendicular to the bilayer, whilst an angle of 90° corresponds to the helix aligned parallel to the bilayer. The $\beta 3$ subunit tilt exhibits a peak centred around $\approx 36-38°$ whilst the $\beta 1$ subunit shows a slight increase in tilt value of around $\approx 40-42°$ with slight shoulder at $\approx 25°$.

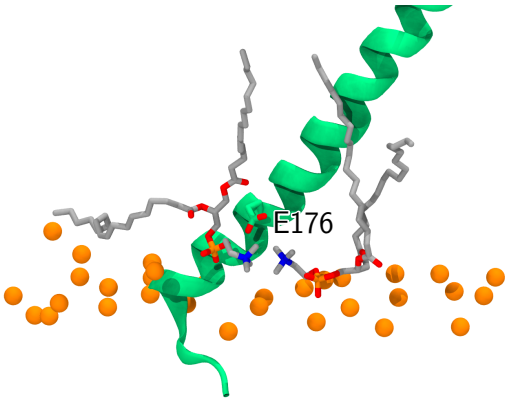
The hydrophobic effect is a well known driving force for the stability and positioning of TMDs embedded in a lipid bilayer [240]. Additionally, charged residues present at either end of the TMD, in this case the α helix, contribute to the position of the domain in the bilayer through electrostatic interactions with the headgroups of the POPC bilayer. As described in Section 4.1.3, prior to the release of the cryo-EM structure from Pan *et al.* [4], sequence analysis highlights a somewhat out of place conserved glutamic acid residue (E177/E176 in $\beta 1/\beta 3$ respectively) in the middle of the TMD where its functional role, if any, remains uncertain. In the cryo-EM structure this residue is placed at the bottom of the S2 helix, close to the headgroups of the inner leaflet. To assess the interaction of this residue with the membrane, the minimum distance between the centre of mass of the terminal oxygen atoms in E177/E176 and all POPC nitrogen atoms was measured. The favourable interaction between the carboxylic group of the glutamic acid and the ammonium group of POPC positions the residue close to the intracellular domain (Figure 4.10c). The distributions of minimum distances between E177/E176 and POPC are shown in Figure 4.10d. Both the $\beta 1$ and $\beta 3$ subunits exhibit essentially identical distribution profiles with a large peak centred at ≈ 4.5 Å and a smaller shoulder at ≈ 6.5 Å.



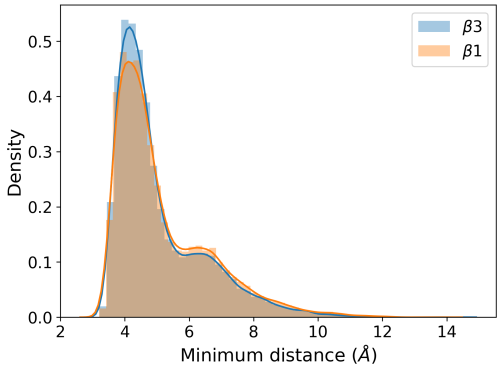
(a)



(b)



(c)



(d)

Figure 4.10: Helix analysis of the β subunit. (a) Schematic of the angle used to measure the tilt angle in the TMD of the β subunit. Phosphorus atoms of the POPC bilayer are shown as orange spheres. (b) Distribution of helix tilt angles between $\beta 1$ and $\beta 3$. (c) Snapshot of E176 - POPC interaction in $\beta 3$. (d) Distribution of the minimum distance between the conserved glutamic acid residue (E177/176) $\beta 1/\beta 3$ and POPC headgroups.

The position of E177/E176 during simulations seems to fluctuate, however the most persistent interaction seems to be close to the lipid headgroups of the intracellular leaflet. The slight shoulder in Figure 4.10d is indicative of the E177/E176 being positioned just inside the bilayer. Coupled to this favoured orientation is significant helix tilting, presumably to accommodate itself within the membrane and increase hydrophobic matching between TMD residues and the POPC lipid tails.

Throughout each simulation of both subunits the TMD helices did not retain their secondary structure. Notably, there was increased distortion and bending in the helix at both the extracellular and intracellular ends. To quantify this, the local bending angle of each residue in the TMD (residues 154-191 in β1 and 156-185 in β3) was calculated using the same HELANAL script described previously. Although not exact to each residue, the local bending angle gives information about the degree of distortion about a residue and in this case is used to highlight loss or depletion of secondary structure. Figures 4.11a & 4.11c show examples from one run of the β1 and β3 subunit respectively. Over the course of 400 ns, both subunits show bending at the ends of the helix, with β1 containing localised bending at the intracellular end, namely residues K183 - A188. Similarly, there is a large increase in local bend angle for residues K183 to S185 in β3 as well as distortion around residues V156, V157, and S158. Averaging across all repeats, these trends can be seen in Figures 4.11b & 4.11d.

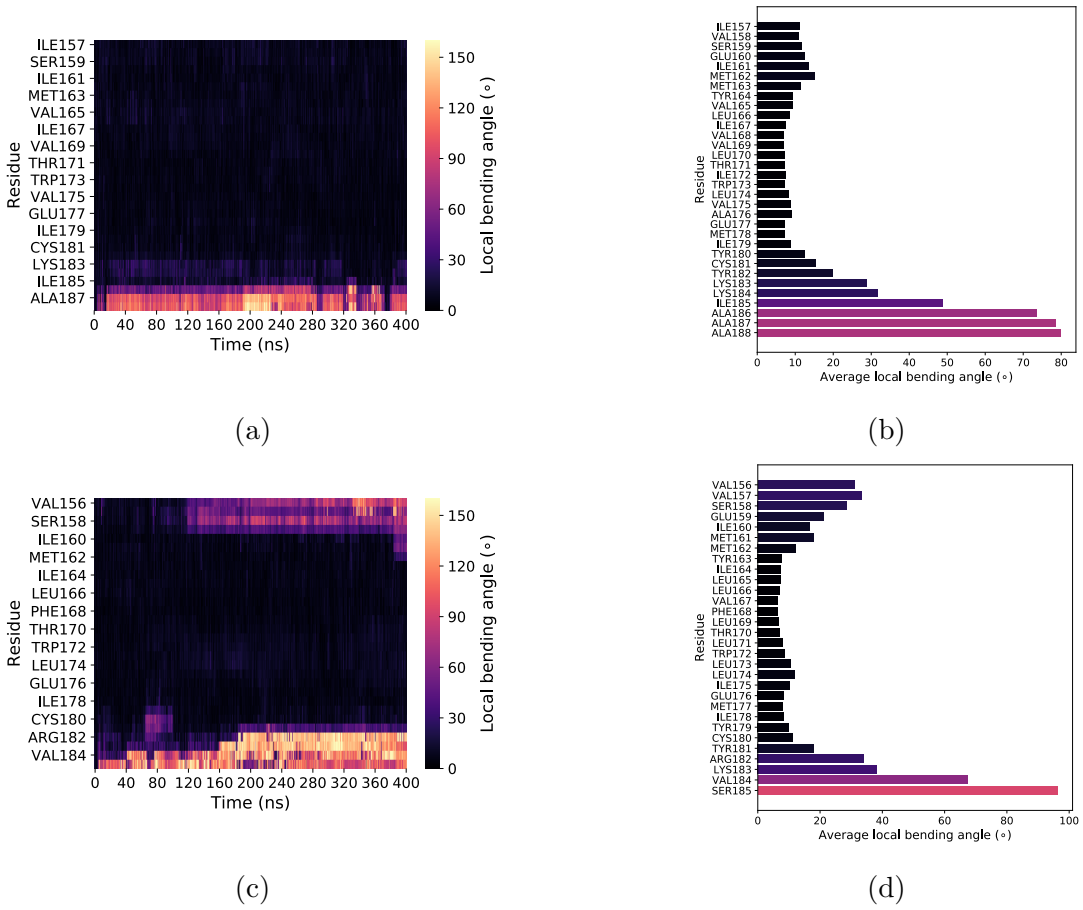
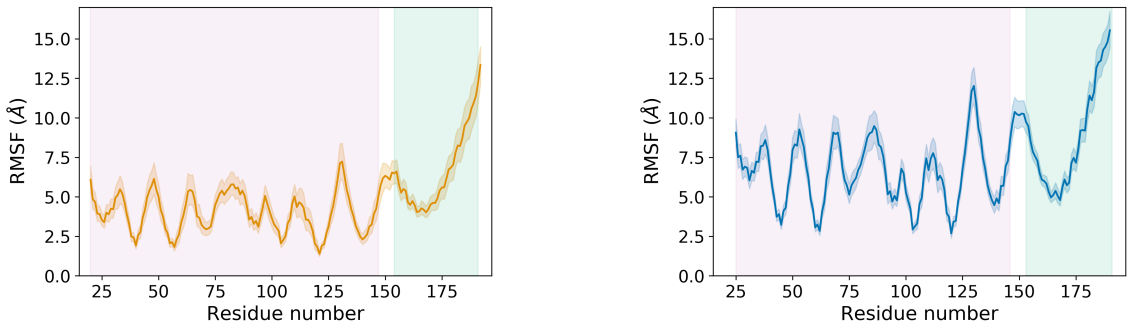


Figure 4.11: Helix bending in the $\beta 1$ and $\beta 3$ subunits. Example of local bending angles in one run of (a) the $\beta 1$ and (c) $\beta 3$ subunit. Average local bending angles in (b) the $\beta 1$ and (d) $\beta 3$ TMD helix across all repeats.

Coupled with the increased local bending angle in both subunits, the RMSF also increases towards the end of the TMD helix in both $\beta 1$ and $\beta 3$ (Figure 4.12). The increase in RMSF close to the C-terminus is expected due to the region's close proximity to the charged lipid headgroups and may reflect the beginning of the predicted disordered intracellular domain. Overall, the $\beta 1$ subunit has a lower RMSF across all residues when compared to $\beta 3$. In contrast, all residues in $\beta 3$ exhibit an increased RMSF when compared to $\beta 1$ in both the Ig domain and TMD.



(a) β1 (b) β3

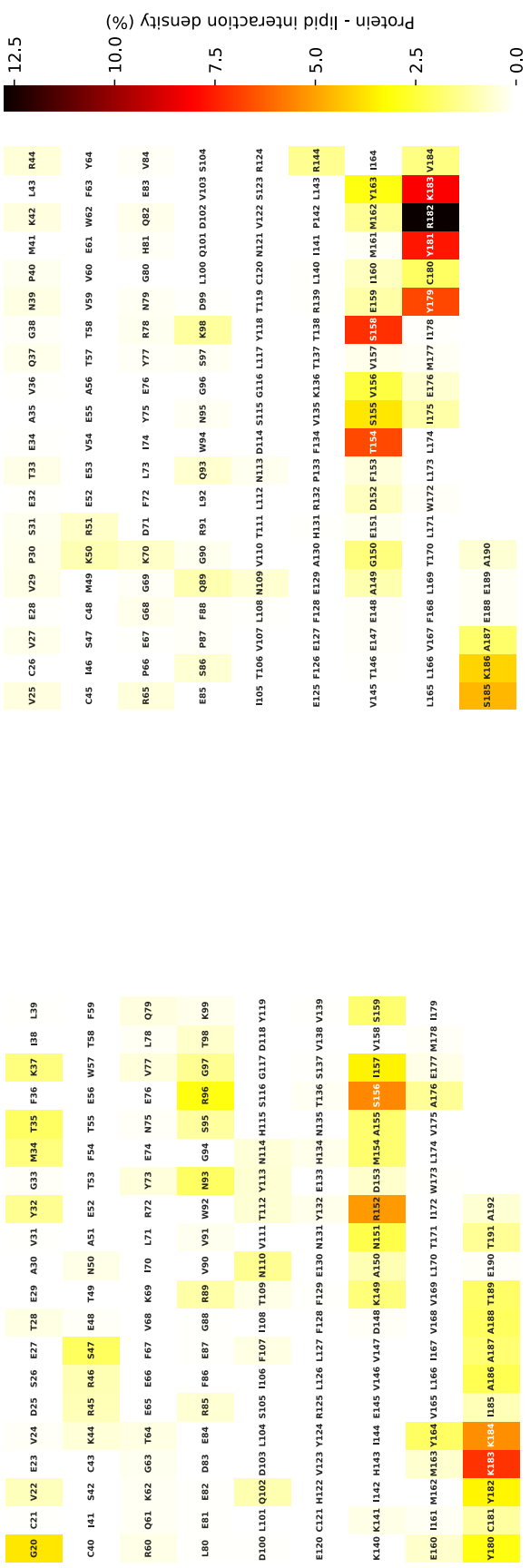
Figure 4.12: RMSF of both (a) the β1 and (b) β3 subunits. The residues contained within the Ig and TMD are indicated by light pink and light green boxes respectively.

4.3.2.2 Protein - lipid interactions

The Ig domains of the both the β1 and β3 subunit interact with the membrane in which they are embedded. To determine which residues within these domains are important in membrane interactions, analysis of contacts made between the Ig domains and the POPC headgroups was performed. The distances between the centre of mass of side chain atoms in the β subunit and the phosphorus atoms of the POPC membrane were calculated over all frames. A residue was deemed to be in contact with the membrane if the minimum distance was ≤ 5 Å. The resulting protein - lipid contact information was then averaged across all 25 repeats for both subunits.

By mapping the interactions to the sequence of each subunit model (Figure 4.13), the preference for β1 to bind to the membrane becomes clear. There are favoured hotspots of interaction between the charged residues K37, K44, K45, R46, R85, R89, N93, and R96 in the Ig domain with little spread over residues when compared to β3. In the β3 subunit, interactions within the Ig domain are much more sparse, reflecting the dynamic nature of membrane binding. Regions of relatively high interaction consisted of charged residues K50, R51, R65, K70, and K98. As expected, additional protein - lipid interactions are present at both ends of the TMD region of both subunits. In β1, these consisted of charged residues K149, N151, R152, D153, E160

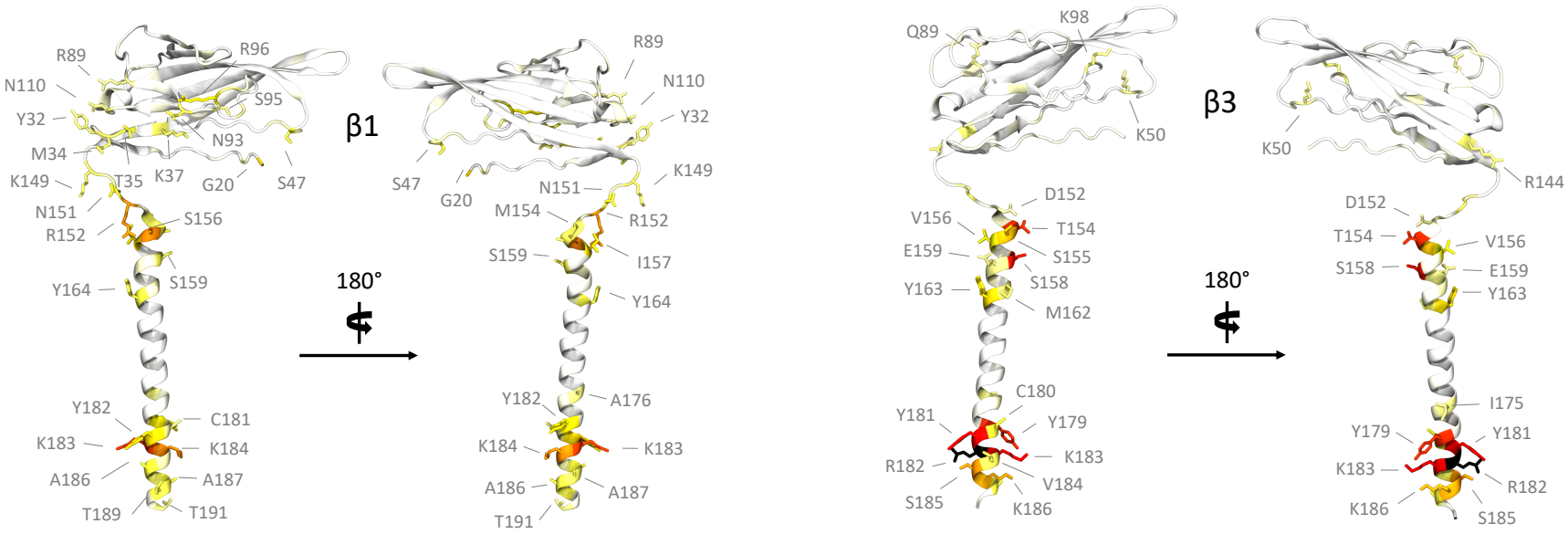
on the extracellular side whilst E177, K183, and K184 account for interactions on the intracellular side. A similar picture is seen in β3, with charged residues D152 and E159 at the extracellular end and E176, R182, K183, and K186 towards the intracellular region. Interactions with the intracellular side of the TMD seem to be somewhat more frequent in β3 and coupled with a decrease in interaction at the extracellular side when compared to β1. This may reflect the decreased stability at this end of the TMD due to the highly flexible linker region. In β1, there are more residues interacting with the membrane, possibly due to the restricted movement of the linker due to stronger binding of the Ig domain to the membrane surface.



(a) $\beta 1$

(b) $\beta 3$

Figure 4.13: Protein - lipid interactions of both $\beta 1$ and $\beta 3$ subunits. (a) Interactions of $\beta 1$ with POPC headgroups. Hotspots are located at charged residues K37, K44, R46, R85, R89, N93, and R96 in the Ig domain. (b) Interactions of $\beta 3$ with POPC headgroups. In contrast to $\beta 1$, more residues in the Ig domain come into contact with the membrane but not as frequently. Hotspots are located at charged residues K50, R51, R65, K70, and K98 in the Ig domain.



(a) $\beta 1$ (b) $\beta 3$

Figure 4.14: Average protein - lipid interactions visualised on (a) the $\beta 1$ and (b) $\beta 3$ subunits. Side chains of residues with high interaction are shown as sticks. Colour scale is from least interacting (white) to most interacting (black), as indicated in the heatplots in Figure 4.13

4.3.2.3 Orientation of the Ig domain

The dynamics of the Ig domain was assessed in terms of pitch angles, as described in Section 4.2.3. Analysis was performed using 25 repeats of 400 ns and each trajectory was post processed to centre the protein, with the membrane plane kept in the approximate starting xy plane throughout. A pitch angle of 0° corresponds to the Ig domain approximately parallel to the membrane plane. Conversely, a pitch angle of 90° corresponds to the pitch vector approximately perpendicular to the membrane plane with the long axis of the Ig domain pointing up, and away from the membrane. Finally, a negative pitch angle corresponds to the long axis of the Ig domain pointing down towards the membrane surface. Comparison between both subunits (Figure 4.15) shows that the β1 subunit exhibits a tight clustering of pitch angles close to 0°, indicating a favoured orientation of the Ig domain close to the membrane. The sampling of pitch angles in the β3 subunit is much more varied, with angles spread across a wider range centred roughly around 30°. Structurally, the restricted sampling of angles in β1 reflects the orientation of Ig domain closer to that of the bound β1 crystal structure whereas in β3 the number of confirmations visited is much larger with no single conformation favoured over another.

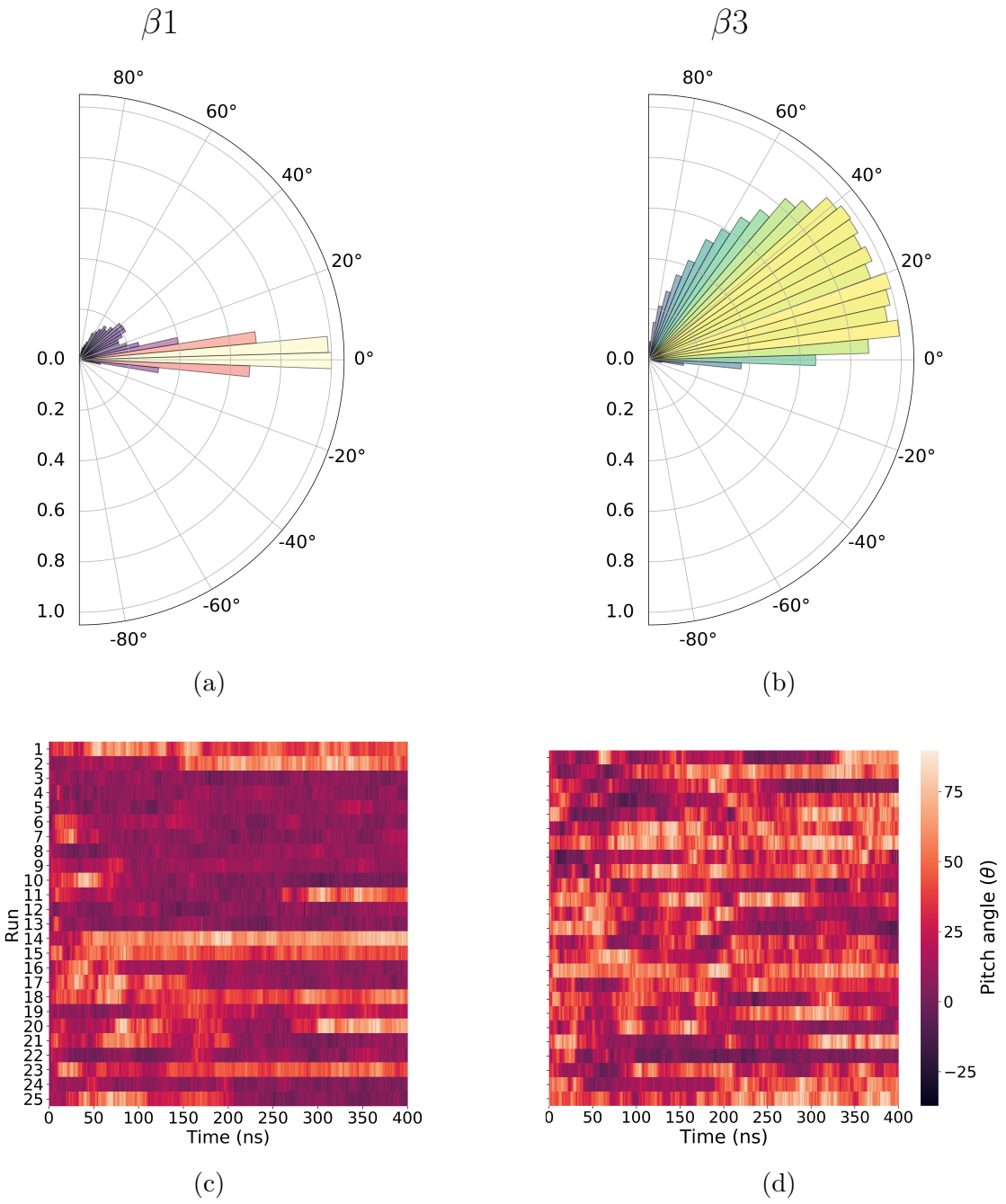
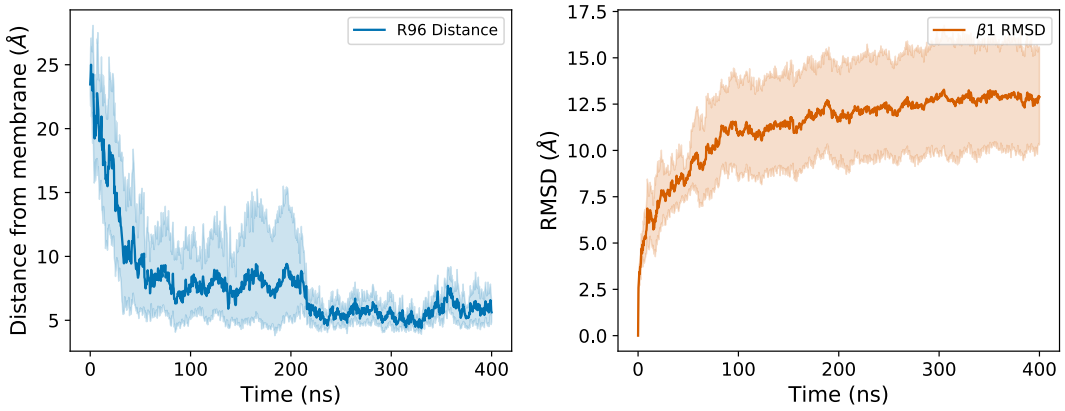


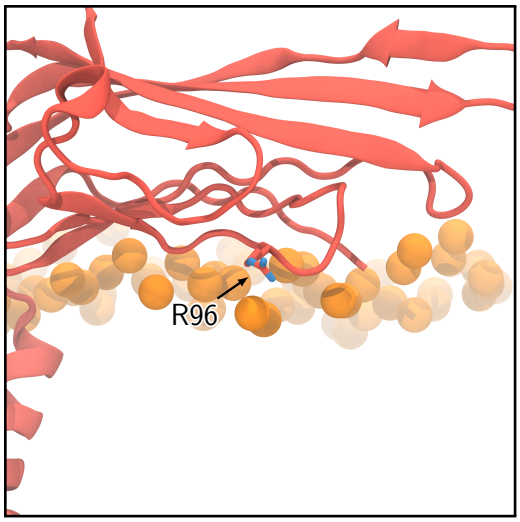
Figure 4.15: Distribution of pitch angles visited by the $\beta 1$ and $\beta 3$ subunits. Polar histogram of pitch angles in the (a) $\beta 1$ and (b) $\beta 3$. Heatplots of pitch angle in each repeat for the (c) $\beta 1$ and (d) $\beta 3$ subunits.

The restricted movement of the β1 Ig domain can be understood in terms of the increased membrane interaction as described in Section 4.3.2.2. Within the Ig domain of β1 there is a region of high interaction[†] at R96, located between the D and E β strands. A closer look using the MSA between the β1 and β3 subunits (Appendix B.5, Figure B.8) reveals that this residue is not conserved in β3. In the simulations performed of the β1 subunit, 60% of the runs (R_{bound}) feature the Ig domain binding to the membrane surface, where it remains for the rest of the simulation. Visual inspection of R96 across these runs reveals that it essentially "pins" one face of the Ig domain to the membrane surface. Analysis of these R_{bound} runs were assessed in terms of both the R96 residue's distance to the membrane and the RMSD of the β1 subunit. The centre of mass of the terminal nitrogen atoms in R96 was calculated at each frame and the minimum distance between that and the phosphorus atoms of the POPC lipids recorded. At approximately 200 ns, there is a significant reduction in R96-POPC distance in all R_{bound} simulations, with a slight levelling off of the RMSD at $\approx$ 12.5 Å, as shown in Figure 4.16a. A snapshot of R96 bound to the membrane is shown in Figures 4.16b & 4.16c. The R96 distance and RMSD of all β1 runs is shown in Appendix B, Figure B.3.

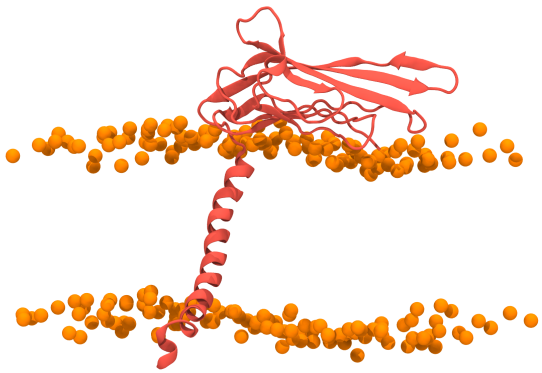
[†]After the peer review process for the publication by Glass *et al.* [231], results presented in this section have been revised. The reader is referred to the publication for the updated results.



(a)



(b)



(c)

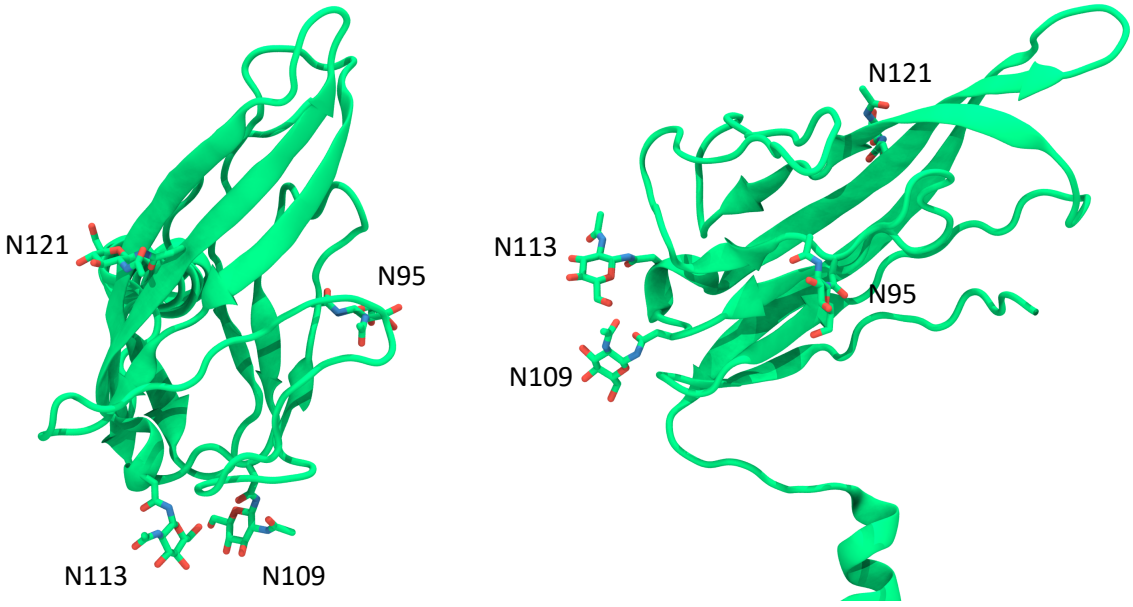
Figure 4.16: Binding of the β1 Ig domain to the membrane surface†. (a) The distance of R96 from the membrane over the course of a simulation (*left*) with an associated plateau of the whole β1 subunit RMSD (*right*). (b) Final frame of one β1 run in which the Ig domain is bound to the membrane. (c) Location of R96 in the bound Ig domain with the residue highlighted as sticks and phosphorus atoms of POPC represented as orange spheres.

†After the peer review process for the publication by Glass *et al.* [231], results presented in this section have been revised. The reader is referred to the publication for the updated results.

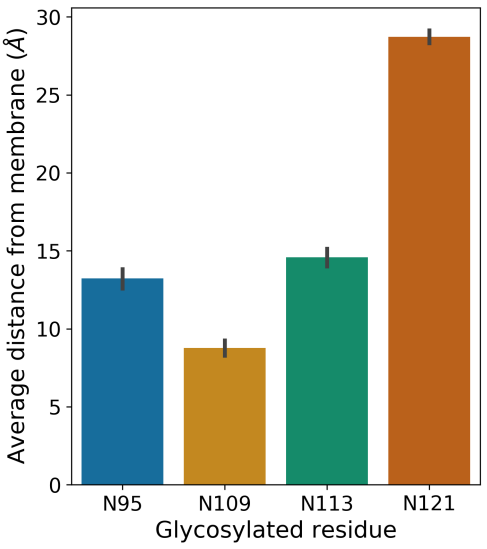
4.3.3 An investigation of β3 glycosylation sites

The addition of glycosylation sites to the Ig domain of β3 was investigated using the model as described in Section 4.2.1.3. Briefly, a single GlcNAc species was added to each of the four predicted N - linked glycosylation sites: N95, N109, N113, and N121 as a β anomer based on the two GlcNAc residues at N13 and N121 conformations in the Pan *et al.* [4] structure. Of the four sites, two are positioned on each face of the Ig domain with one located on β strand F (N121), whilst the other is located on the D β strand (N95). The remaining two sites are located at the N-terminal end of the E β strand (N109) and the other located in a short 3-10 helix between the E and F β strands (N113). Spatially, N109 and N113 are towards one end of the Ig domain and near to the short linker region, causing them to be close to the membrane surface at the start of the simulation. For definitions of the β strand labelling system, the reader is referred to Chapter 1, Figure 1.7.

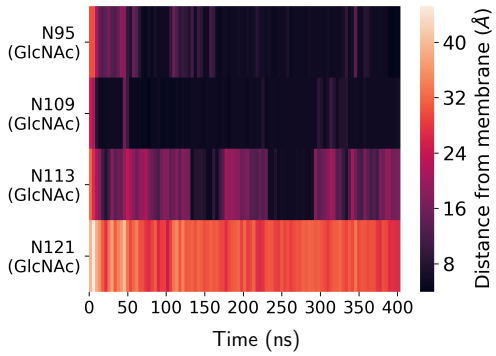
In order to assess what impact, if any, the glycosylated residues had on the orientation of the Ig domain their proximity to the membrane surface was assessed over the course of each 400 ns simulation (Figure 4.17). The centre of mass of each GlcNAc species (excluding hydrogen atoms) was calculated and used to find the minimum distance made between all phosphorus atoms of the POPC membrane. Across all repeats, glycosylated N121 exhibits an increased distance from the membrane when compared to the other glycosylated residues with an average distance of 28.7 Å. Residues N95, N109, and N113 were all closer to the membrane with average distances of 13.2, 8.8, and 14.6 Å respectively.



(a)



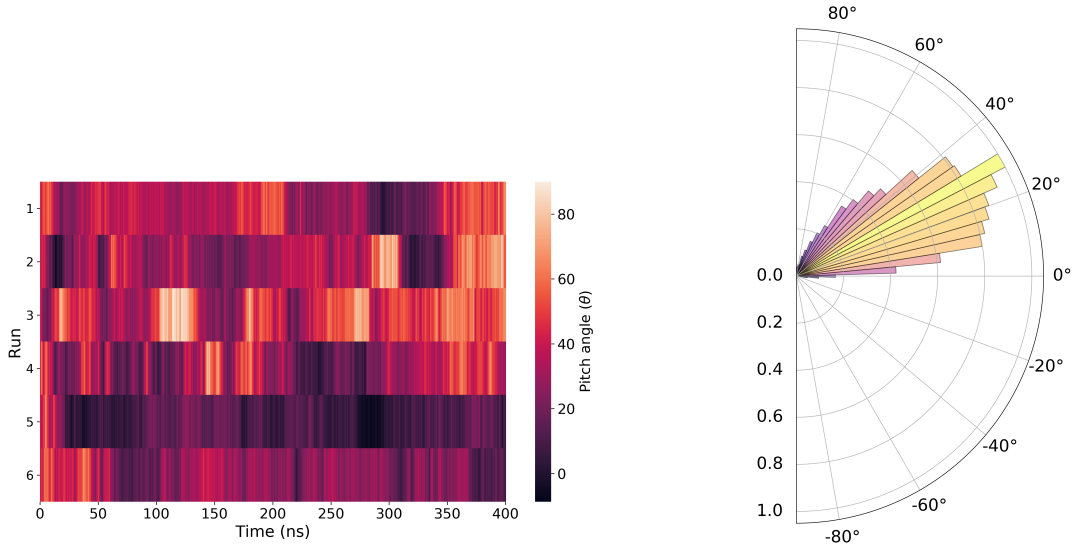
(b)



(c)

Figure 4.17: Distances of GlcNAc groups of glycosylated residues in the β3 Ig domain. (a) Top down (*left*) and side (*right*) view of the Ig domain with glycosylated asparagine residues indicated in a stick representation. Hydrogen atoms are omitted for clarity. (b) Average distances of GlcNAc groups from the membrane surface. (c) Typical variation of GlcNAc - membrane distances over 400 ns.

In order to gain a better understanding of how the Ig domain behaved on the membrane, the pitch angles of the Ig domain were calculated using the same pitch analysis used in Section 4.3.2.3. When calculating the centre of mass the GluNAc, rings were not included. Exclusion of the GluNAc rings was chosen in order to make a direct comparison between both the glycosylated and non-glycosylated species without any difference between the calculation of the principal axes. Figure 4.18a shows the change in pitch angle of the Ig domain over each run of the glycosylated β3 subunit. Over each repeat there is clear variability, with some repeats sampling a variety of angles over the course of 400 ns whilst others (e.g. runs 5 & 6) tend to sample less. Analysis of pitch angles sampled over all repeats (Figure 4.18b) reveals that although variable, the bulk of pitch angles favoured by the Ig domain are between $20 - 40°$.



(a) (b)

Figure 4.18: Pitch analysis of the glycosylated β3 subunit. (a) Pitch angles sampled over 400 ns. (b) Histogram of pitch angles across all repeats.

Combining the results from the glycosylated residue proximity to the membrane and the most sampled pitch angles, there is a tendency for one region of the Ig domain to be closer to the membrane. On average, the GlcNAc residues present at N109 and

N113 remain close to the membrane surface. The decrease in distance between N95, located on one side of the Ig domain, and the POPC headgroups reveals a slight tendency for the Ig domain to roll to one side allowing the face of the Ig domain defined by β strands D, E, B, and A to also be in close proximity to the membrane surface. When orientated away from the membrane, the glycosylated β3 subunit adopts a conformation as shown in Figure 4.19. This orientates the Ig domain to be almost perpendicular with the membrane plane and exposes the N-terminal residues that are critical to the formation and stability of the observed β3 trimer.

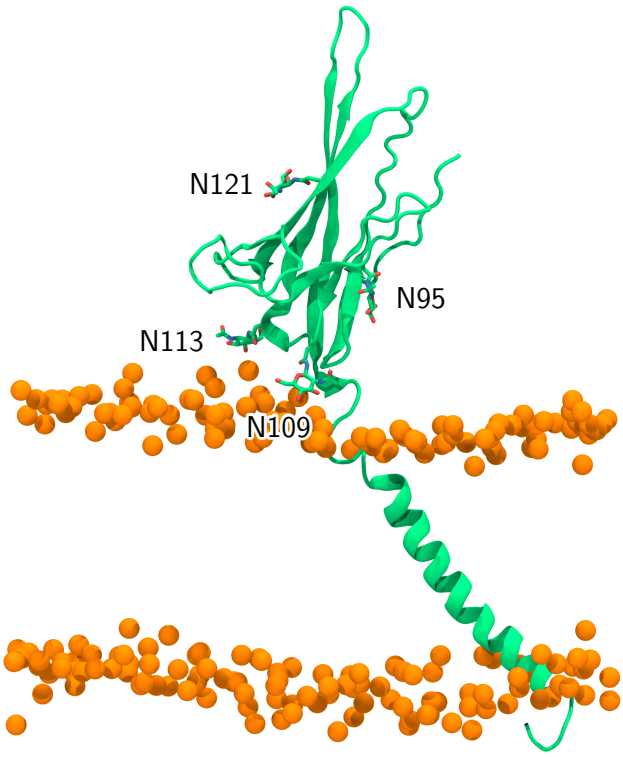


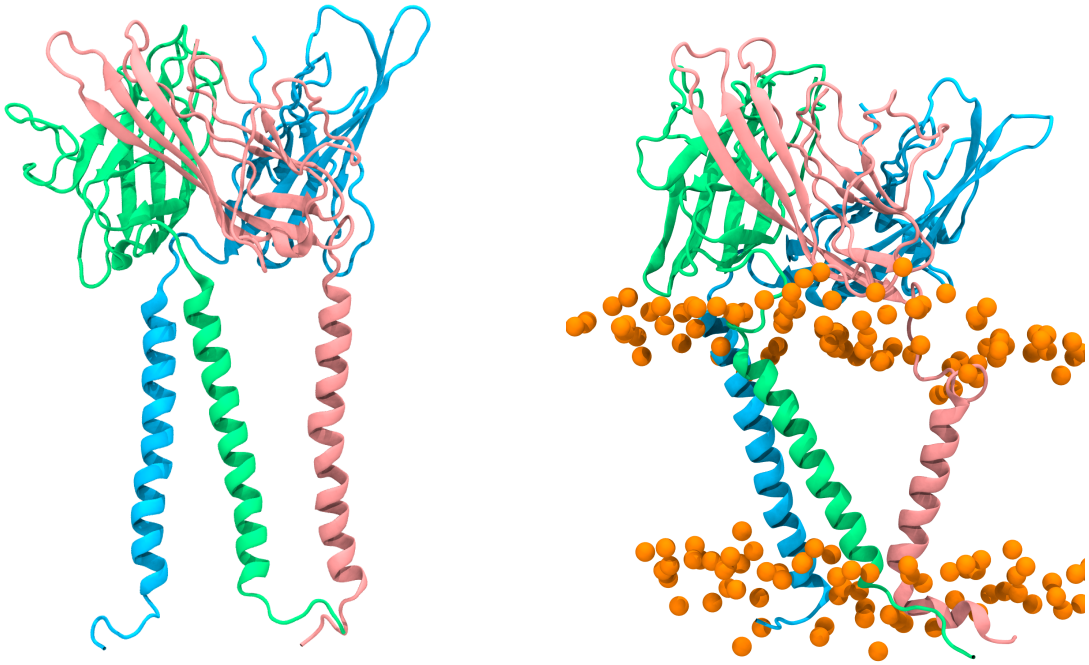
Figure 4.19: Representative snapshot of the glycosylated β3 subunit on a POPC membrane. Asparagine (and GlcNAc) residues are shown as sticks with hydrogen atoms removed for clarity. The POPC phosphorus atoms are shown as orange spheres.

4.3.4 Dynamics of the full length trimer construct

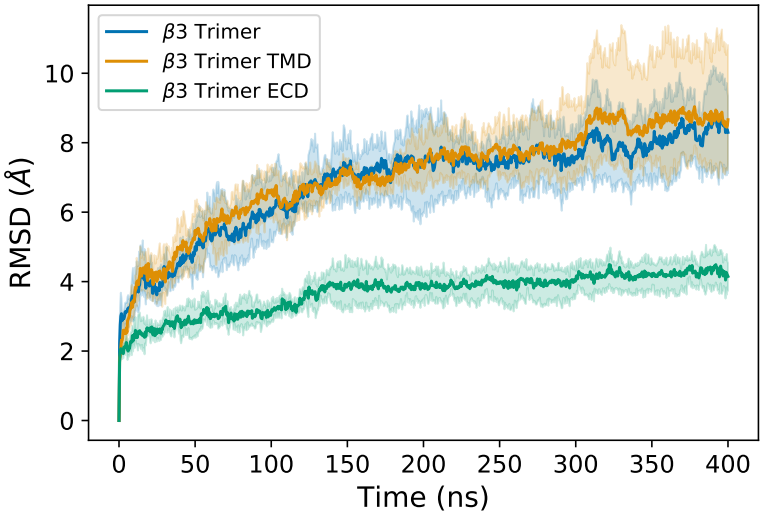
As discussed in Section 4.1.2, when the β3 ECD trimer and β1 crystal structures are overlaid there is an unfeasible overlap between the Ig domain that results in the α helix of the TMD to orientate itself parallel to where the membrane would be located. In order for the full length β3 model to adopt an orientation that would accommodate the trimeric scaffold, a significant change in Ig domain orientation would be required. As a consequence of the large number of repeats performed for β1 and β3 in Section 4.3.2, a number of conformations are adopted by both subunits on the POPC membrane. To assess what contribution, if any, the TMD of the β3 subunit makes to trimer stability, a full length trimeric model was constructed from a snapshot in one repeat of the monomeric β3 subunit (see Section 4.2.1.5 for details of model construction).

At this point it is important to note that this was only carried out for the β3 subunit. Although both the β1 and β3 subunit share a high sequence similarity there remains to be substantial evidence that β1 subunits can form higher order structures on the cell membrane *in vivo*.

The choice of monomer model was made visually, with reference to the pitch angles sampled by each repeat. The final monomer conformation was chosen with an Ig domain pitch angle of 44.7° and the hydrophobic V25, V27, and P30 patch exposed. In order for the TMD helices of each chain in the β3 trimer to accommodate the restrained Ig domain conformation, the helix of each chain is located directly below the Ig domain of its neighbouring chain. The final trimeric model is shown in Figure 4.20a. A total of 4 $\times$ 400 ns simulations were performed, in all repeats the trimeric construct remained intact (Figure 4.20b). Calculation of the RMSD of all chains indicated that the ECD remained stable over the entire simulation, whilst the dynamics of the TMD helices was much more variable and contributed to the large shift in total RMSD, as shown in Figure 4.20c.



(a) (b)



(c)

Figure 4.20: Full length β3 trimer model construction. (a) Side on view, with different chains indicated by the colours green, pink, and cyan. (b) Example of a final conformation of the β3 trimer model after 400ns. Protein chains are coloured as in (a) with phosphorus atoms of the POPC bilayer shown as orange spheres. (c) The RMSD over all repeats of the β3 trimer, with different regions of the protein indicated.

When constructing the model, the conformation of the helices had to be oriented perpendicular to the membrane plane so that the base of the Ig trimer could effectively "sit" on the surface of the membrane. In terms of RMSD, over the course of the simulation the Ig trimer seems to not be influenced heavily by the presence of the POPC membrane. The membrane restricts tumbling that the Ig trimer can undergo and throughout each simulation the complex remains orientated in approximately the same starting orientation. The stability of the full length β3 trimer supports the experimental observation of trimers present on the cell membrane [99], however the nature of the TMD dynamics remains to be verified experimentally. As seen previously in Section 4.3.2.1, the helix of monomeric β1/β3 subunits adopts a large tilt of around 38°. In the constructed trimer model, the helices cannot adopt this somewhat artificial conformation and in fact adopt a tilt angle close to 0°. Given this, it would be expected that there may need to be substantial helix reorientation when forming higher order oligomeric structures *in vivo*. Measuring the RMSD of all TMD helices reveals that these domains contribute the most to the overall increase in RMSD of the whole protein complex, indicative of a large conformational change. Across each repeat there is indeed a change in conformation of each helix, with substantial tilting from the starting orientation. Although each helix is identical there seems to be no definitive preferred tilt angle adopted across all three domains. Analysis of the tilt angle in each helix over all repeats is shown in Figure 4.21.

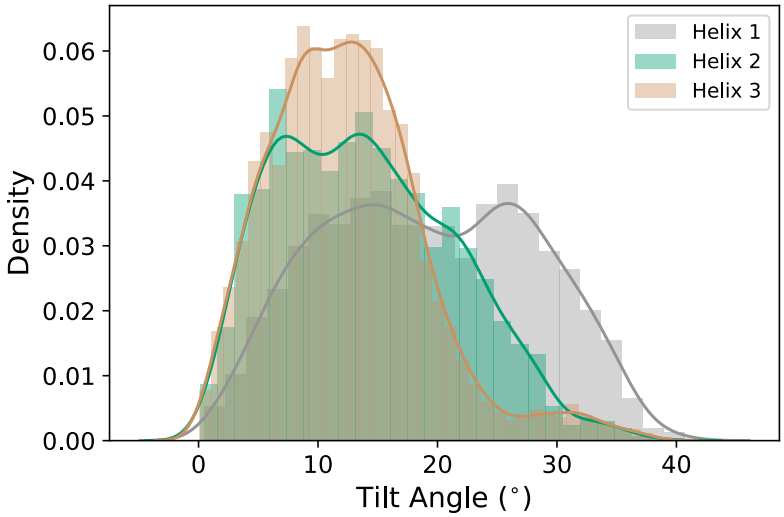


Figure 4.21: Bending of helices in full length β3 trimer model.

One helix exhibits a roughly bimodal distribution, with centres at $\approx 12°$ and 27° whilst the remaining helices have a slightly more restricted distribution centred at $\approx$ 12°. This is in stark contrast to the tilt angles sampled in the monomeric constructs, where a single helix sampled a variety of angles with a peak centred between 36-38° (Section 4.3.2.1, Figure 4.10b). In addition to helix tilting, there is also an associated instability at either end of the helix, as observed in the β3 (and β1) monomers. Although all helices in the trimeric construct are identical, each helix undergoes bending to varying extents, both within the same run and between repeats.

For each chain, the RMSF is particularity stable within the Ig domains, however there is a sharp increase at the N-terminal end of the TMD as seen in the monomeric β3 case (Figure 4.22). Additionally, all chains within the trimer exhibit smaller RMSF values when compared to those in the monomer simulations. This trend continues, with both the Ig domain pitch angle and TMD movement being more restricted in the full length β3 trimer model when compared to the full length β3 monomer.

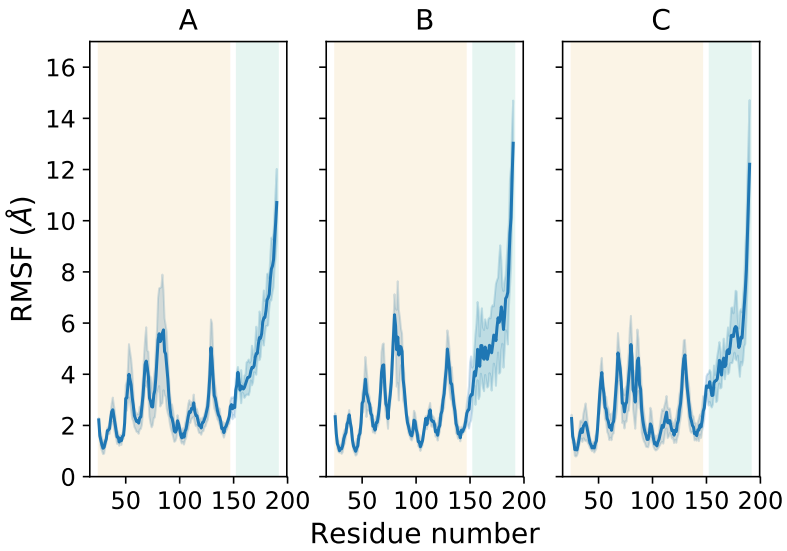


Figure 4.22: RMSF of the full length β3 trimer model. Different chains are indicated as A, B, and C. The Ig and TMDs are indicated by light orange and light green boxes respectively.

To assess the influence that the conserved E176 residue had on trimer stability, its position in the membrane was measured. As seen in the monomers of both β1 and β3, the glutamic acid residue interacts strongly with the POPC lipids of intracellular leaflet (Figure 4.10d). The density of the three E176 residues and the POPC phosphorus atoms along the z-coordinate was measured over the course of the 400 ns simulation and is shown in Figure 4.23c. Although each helix in the trimer undergoes substantial movement, when they are in close proximity the E176 residues are orientated out and away from each other (Figures 4.23a & 4.23b). Coupled with an increased interaction with the membrane, the behaviour of the E176 residue suggests a role primarily involved with membrane insertion and location, rather than interactions between β3 subunits.

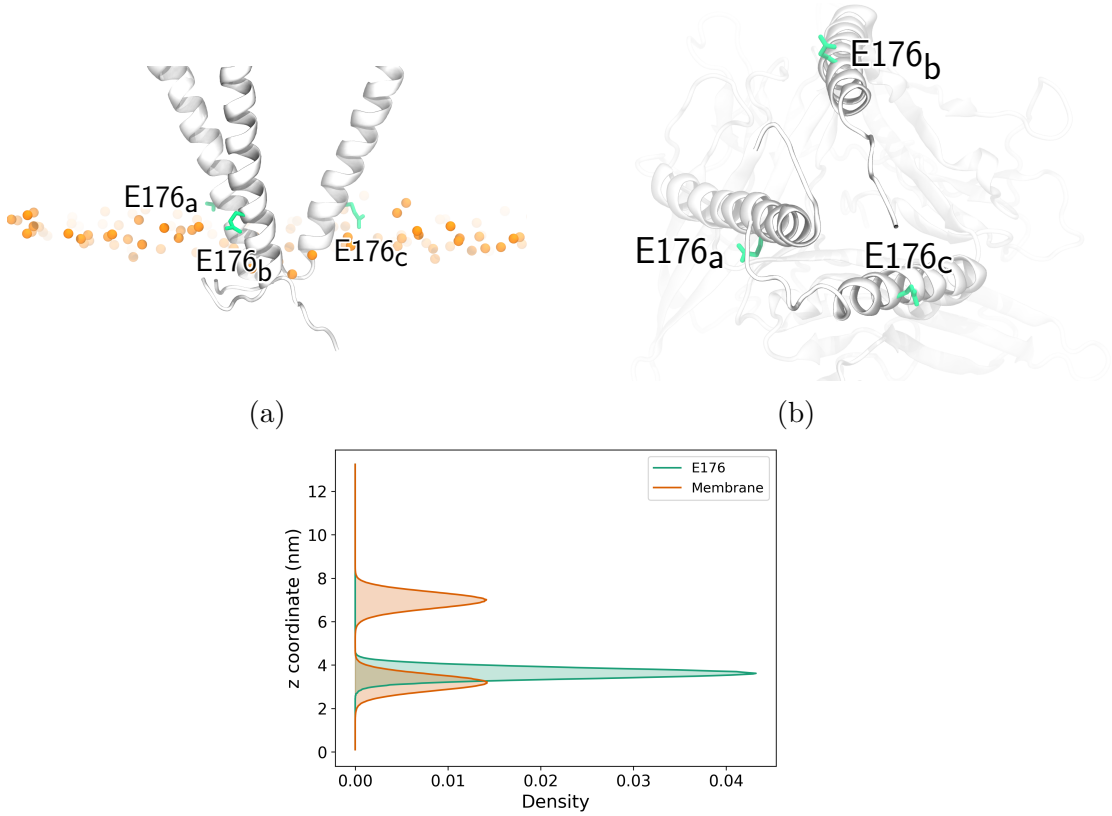


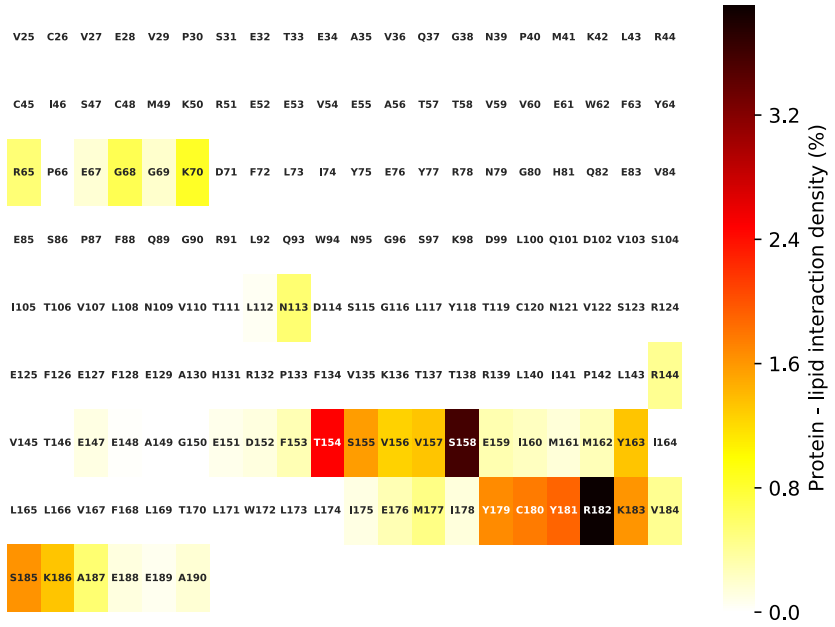
Figure 4.23: Position of E176 in the β3 trimer. (a) Side on and (b) intracellular view of the C-terminus of the β3 trimer. E176 residues are shown in green and orientated towards the phosphorus atoms of the POPC lipids (orange spheres). TMD helices are shown in white. (c) Density of E176 and POPC along the z-axis of the the simulation box. Protein chains are indicated in lower case letters.

Protein - lipid interactions

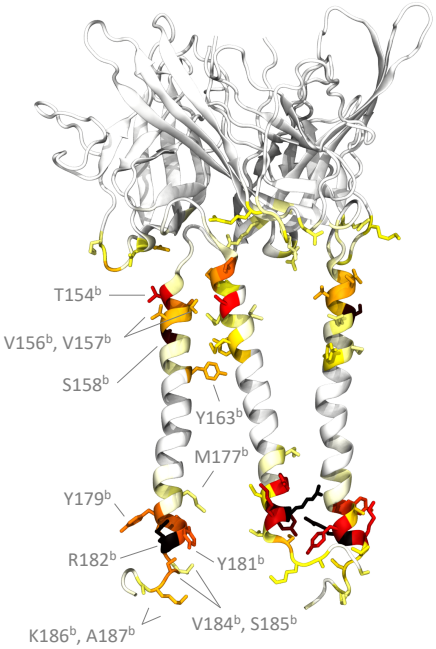
How the full length trimer model interacts with the membrane was assessed in the same way for the monomeric species in Section 4.3.2.2. Protein - lipid contacts were analysed over each of the 400 ns trajectories for each of the three chains. A residue was deemed to be in contact with another residue if the minimum distance between the centre of geometry of its terminal side chain atoms and the phosphorus atom of a POPC molecule was ≤ 5 Å. A heatplot of lipid contacts and the protein sequence is shown in Figure 4.24a. Residues in high contact with POPC include R65, E67, G68, K70, and N113 in the ECD as well as T154, V156, V157, S158, Y163, Y179, C180,

Y181, R182, K183, S185, and K186 in the TMD (Figure 4.24a).

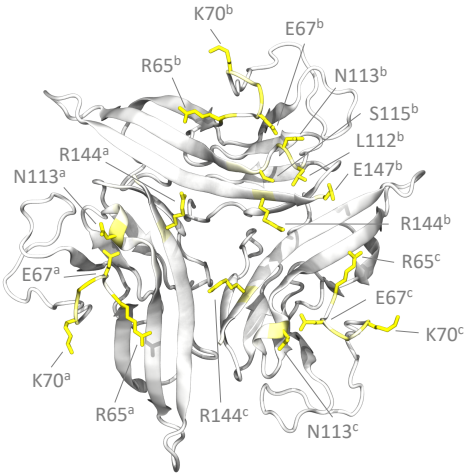
Visualising these interactions reveals that the majority of the lipid interactions are clustered, as expected, at either end of the TMD helices (Figure 4.24b). At the intracellular end of each helix, there is a loss of secondary structure with the resulting residues exhibiting partial interaction with the inner leaflet of the POPC membrane. Towards the N-terminus, and extracellular side of the protein, the base of the Ig trimer interacts with the membrane surface through a small number of residues (Figure 4.24c). Within the Ig trimer and across all chains the positively charged residues R65 and K70 exhibit marked interaction with the membrane surface. Additionally R144, located towards the C-terminus of the Ig domain and positioned close to the centre of the trimer also interacts with POPC lipids located close to the short linker region. In the small 3-10 α helix, located between the E and F β strands, the polar N113 residue shows similar high interaction with the membrane.



(a)



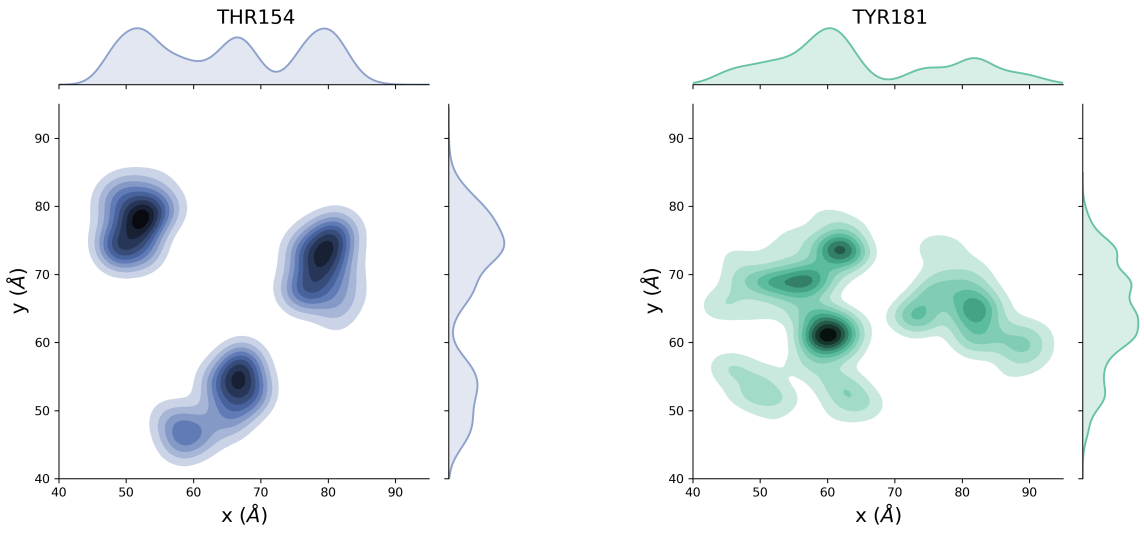
(b)



(c)

Figure 4.24: Protein - lipid interactions in the full length β3 trimer. (a) Heatmap of one β3 chain (chain B) in the trimeric construct. (b) Protein - lipid contacts mapped on to the β3 trimer model. Key interactions on one chain (Chain B) are highlighted. (c) View from the underside of the Ig domains (towards the ECD), regions of high contact in each Ig domain are indicated. Protein chains are indicated in superscript. Residues with high lipid interaction are shown as sticks. The colourscale in (b) and (c) is the same as shown in (a).

Over the course of the 400 ns simulation the Ig domains undergo little conformational change, reflecting the same stability seen in the ECD only simulations in Section 4.3.1.1. As discussed previously, the TMD helices tilt is restricted in the trimeric construct. To assess the movement of the helices at both the extracellular and intracellular ends the positions of Y181 and T154 in each helix were analysed. Both Y181 and T154 were chosen from analysis of the final frame of each repeat since they did not undergo significant helical distortion over the course of each simulation. The positions sampled of each residue's Cα atom are shown in Figure 4.25. At the extracellular side of the TMD and defined by the positions of T154 in Figure 4.25a, the residues move in a somewhat restricted fashion, primarily due to their close proximity to the Ig domain. The converse is true at the intracellular side of the TMD (Figure 4.25b), where the movement of Y181 suggests that the residues sample a large number of conformations in an unspecific manner, reflecting this region's increased instability as shown previously from RMSF calculations (Figure 4.22).



(a) (b)

Figure 4.25: Distribution of T154 and Y181 in the β3 trimer construct. (a) 2D histogram of the x and y coordinates of the T154 Cα atoms located towards the extracellular end of the β3 helices. (b) 2D histogram of the x and y coordinates of the Y181 Cα atoms located towards the intracellular end of the β3 helices. Data was collected over each repeat.

4.4 Discussion

4.4.1 The varied dynamics of the Ig domain

The isolated β3 trimer crystal structure from Namadurai *et al* [99] revealed the key role that the hydrophobic V27, V29, and P30 Ig domain residues play in providing an interface upon which other Ig domains can bind. As highlighted briefly in Section 4.1, the Ig domain in β3 differs slightly from typical features present in most Ig domains [98], with the first $\approx$ 10 N-terminal residues not being part of an antiparallel β sheet conformation. Instead, the C26-48 disulphide bond replaces the expected β sheet motif. The C48A mutation performed by Namadurai *et al.* [99] highlights the bond's importance in facilitating oligomer formation. Similar studies on Basigin (a dimer consisting of two Ig domains) describes similar destruction of N-terminal disulphide bridges which leads to a reduction in ability to form higher order oligomers, albeit in the central core of the Ig scaffold [241].

In Section 4.3.1.1, the importance of this bond was verified through the C48A mutation in the β3 monomer, where results indicate an increase in R_G values when compared to WT. Although in some cases subtle, the overall effect of the C48A mutation causes destabilisation in the critical region containing the hydrophobic residues responsible for observed oligomerisation. Interestingly, not all C48A repeats undergo such dramatic conformational changes, with a significant amount corresponding typically to WT. This should not come as a complete surprise, since experimentally the higher order structures of β3 still formed even with the C48A mutation in lower cluster numbers and in "splayed" conformation [99]. Taking the results of monomeric simulations into account, it can be hypothesised that on the plasma membrane some Ig domains simply cannot form oligomers due to critical residues (V27, V29, and P30) becoming destabilised to such an extent as to be distant in space from the Ig body, causing the Ig domain to become partially disordered. The partially disordered Ig

domains may simply take longer to orientate themselves in the correct position on the plasma membrane, with an increased entropic cost for oligomer formation. The C48A repeats that exhibit WT-like properties (i.e. little disorder at the N-terminus and subsequent low R_G values) may in fact become partially disordered at longer simulation time scales.

Extension of the C48A mutation to the trimeric construct in Section 4.3.1.2 revealed a subtle change in splaying of the complex when compared to monomeric simulations. Critically, it is important to note that the C48A mutation was applied to the already constructed β3 trimer model and so it would be expected that there is some inherent stability associated with the starting complex. This was verified with only a subtle, and much less dramatic, increase in R_G values, with no large conformational change in the N-terminus of the Ig domains when compared to the monomeric system. Torsional angle analysis verified that there is increased flexibility of V27 and V29 with C48A, suggesting that the interface of the trimer is more flexible and may be prone to dissociation, given longer sampling time.

In summary, the simulations performed of the monomeric and trimeric β3 Ig domain in this Chapter reinforce the argument that the C26-48 disulphide bond is critical to the overall stability of the Ig scaffold.

As described in Section 4.1, the β subunits of Nav channels are single pass type-1 transmembrane proteins. The construction and simulation of the full length β1 and β3 subunits revealed that although similar in sequence and assumed α subunit binding characteristics, their behaviour on the membrane is markedly different. Simulations performed in this chapter suggest that the Ig domain of the β3 subunit is much more dynamic than β1 through pitch angle analysis. Across each simulation of the β3 subunit a number of pitch states are sampled, with only a few corresponding to the Ig domain's long axis parallel to the membrane plane. Conversely, the β1 subunit simulations provide evidence for more restricted Ig motion, with the long axis of

the Ig domain parallel to the membrane plane in 60% of the simulations performed and the non-conserved R96 seemingly responsible for the favoured Ig - membrane interaction[†]. The extra positive charge in this region may encourage the observed increase in propensity for the β1 Ig domain to interact with the POPC headgroups. Since R96 is not present in β3, the Ig domain is less tightly bound to the headgroups and so may contribute to the observed increase in sampled pitch angles. Additionally, the binding of R96 to the membrane surface not only orientates the long axis of the Ig domain parallel to the plane of the membrane, it also restricts the movement of V21 and V22 whose equivalent positions in β3 are responsible for trimer stability. Analysis of the MSA in the N-terminus of the Ig domain (Figure 4.26) also highlights a key aspartic acid residue (D25) in β1 that in β3, at the equivalent position, is a proline (P30). If the β1 Ig domain became unbound to the membrane and adopted an orientation similar to that seen in the β3 trimer structure it may be hypothesised that it could form similar oligomeric states. However, these would be predicted to be much less stable due to the presence of the negatively charged D25 close to the interface.

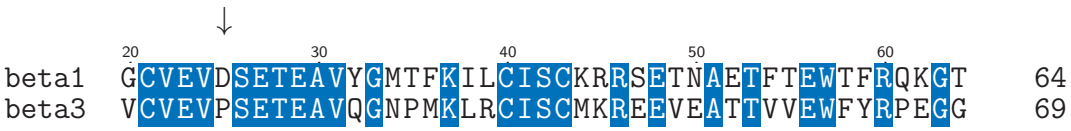
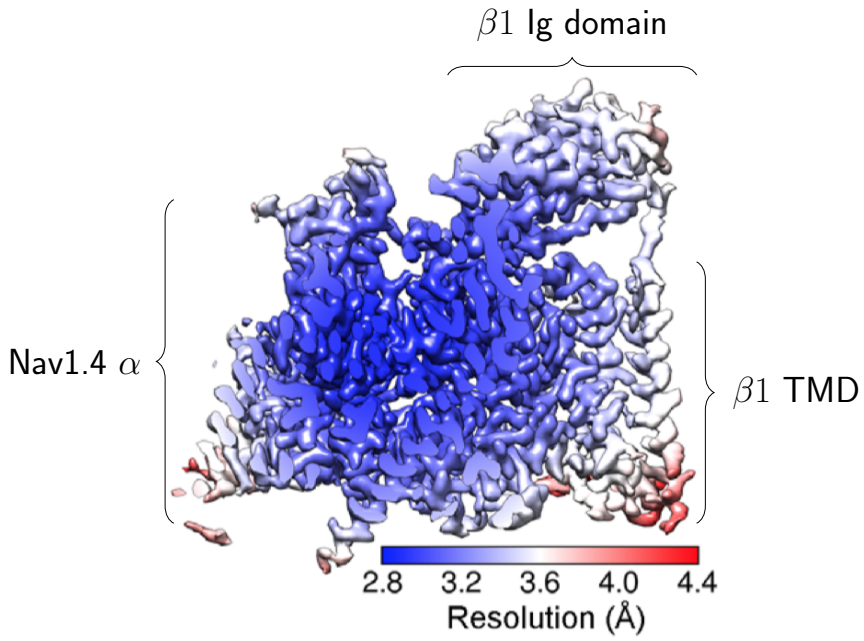


Figure 4.26: Alignment of the N-terminus of the β1 and β3 subunits. An arrow indicates the D $\rightarrow$ P switch from β1 $\rightarrow$ β3. Numbering for β1 residues is shown above alignment.

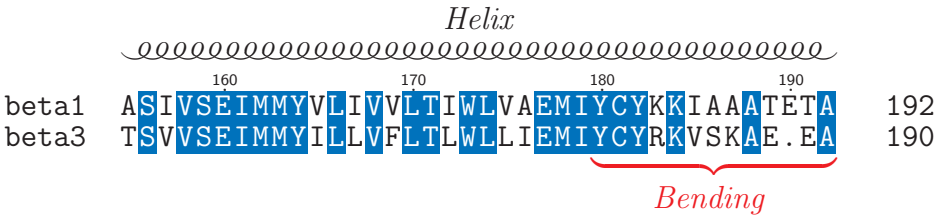
[†]After the peer review process for the publication by Glass *et al.* [231], the discussion presented in this section has been revised. The reader is referred to the publication for an updated discussion.

4.4.2 Destabilisation of C-terminal residues in the β1 & β3 helices

Across all simulations of the β1 and β3 subunit there is substantial destabilisation of the C-terminal end of the TMD. Through measurement of the local bending angle in Section 4.3.2.1, both the β1 and β3 undergo extreme bending with average values between 40-80°. This is coupled with large RMSF values in the TMD when compared to the rest of the β subunit domains. A key question to address is why there is such a large degree of distortion in this region in both subunits? It is important to note that both models of the β1 and β3 subunit were based on the structure determined by Pan *et al* [4]. A closer look at the electron density map (Figure 4.27) reveals that the resolution from residues Y180-A192 in β1 (corresponding to Y179-A190 in β3) is $\geq$ 4 Å.



(a)



(b)

Figure 4.27: Examination of β subunit bending. (a) The electron density in the Pan *et al.* structure, image adapted from [4]. (b) Alignment of the β1 and β3 TMD. Regions that experience a large degree of bending are indicated in red braces. Numbering for β1 residues is shown above alignment.

Furthermore, these residues also account for high protein - lipid counts as seen in both monomeric and trimeric models. A somewhat large tilt [242] in both β subunits accompanies the increased C-terminal residues interaction with the lipid bilayer and most likely reflects the TMD reducing hydrophobic mismatch. Thus, it can be hypothesised that the region defined as the end of the TMD helix may in fact be shorter than what the structure by Pan *et al.* [4] defines.

When β3 is present as a full length trimer, the degree of TMD tilting is reduced when compared to the full length monomeric model studied in Section 4.3.2. The

restricted conformational freedom in the full length trimer can be attributed to the pre-formed Ig trimer restricting the short linker motion that ultimately influences the positions of the TMD helices. Analysis of the protein - lipid interactions in the full length trimer reveals that although tilting is reduced, the C-terminal residues that exhibit high contact with lipid as a monomer also experience increased interaction when part of the trimeric construct. A major difference between monomeric and trimeric $\beta 3$ lipid interactions is in the Ig domain. As part of a trimer, the Ig domain is no longer able to sample large pitch states due to the favourable hydrophobic interactions at the trimer interface (located at the N-terminus of each chain). As such, the lipid interaction between the Ig trimer is markedly reduced when compared to the monomer, with only a few charged residues (R65, E67, and K70) interacting with the membrane surface. Interestingly, if the $\beta 3$ subunit were to interact with the VSDs of the pore-forming α subunit in a similar fashion to $\beta 1$ [4], there would need to be substantial rearrangement of the Ig domains. This leads to the question of what, if any, the role that the TMD helices could play in $\alpha - \beta$ and / or $\beta - \beta$ interactions? Lipid contact analysis in the TMD reveals that there is little difference between the trimeric and monomeric constructs. Close to the Ig domain the restriction imposed via the stable Ig trimer reduces translational motion, whereas at the intracellular end the translational motion is much more dynamic with no clear preference for residue - residue interactions between chains. These results suggest that the TMD of the $\beta 3$ subunit does not have an overall stabilising effect on the $\beta 3$ trimer and in fact may only be required for correct positioning within the membrane. Additionally when the helices of each $\beta 3$ chain are in close proximity, the conserved E176 residue appears to be orientated away from the trimer centre and preferentially interact with the POPC membrane.

4.4.3 Glycosylation: A preferred Ig orientation?

Carbohydrates facilitate and participate in a plethora of biological functions [243, 244]. Critically in proteins, and Ig domains in particular, glycosylation sites serve as the basis for large and complex carbohydrate structures that can influence protein function [245]. In both the β1 and β3 subunit there are a total of four glycosylation sites all located in the Ig domains as illustrated in Section 4.3.3. Simulations of the glycosylated β3 subunit performed in this chapter made use of the recently developed *DoGlycans* [237] software package. Of the many types of glycosylation that can occur [246], the *DoGlycans* package was used to model the N-linked glycosylation present in β3 with a single GlcNAc moiety to each glycosylation site. In the non-glycosylated form (i.e. WT) of β3, it was shown that the Ig domain visits a number of conformations on the membrane surface (see Section 4.3.2.3). Surprisingly, when modelling just one GlcNAc at each glycosylation site, the orientation of the Ig domain is somewhat restricted when compared to WT. Although the number of simulations of the glycosylated β3 subunit is smaller than those performed in WT, there is a general trend for the Ig domain to be orientated away from the membrane. Contact analysis between the GlcNAc rings and the lipid bilayer reveals that there is a preference for one face of the Ig domain to interact with the membrane surface (glycosylated N95 in the D β strand). Additionally, glycosylated N109 and N113 interact persistently with the membrane. Combining the afore mentioned asparagine interactions there is an increased propensity for the glycosylated Ig domain to expose N-terminal residues. The residues located in the N-terminus are either exposed laterally on the membrane (through increased interaction between the membrane and N95) or, as revealed by pitch analysis, are exposed to a greater degree with the long axis of the Ig domain perpendicular to the membrane plane (with increased membrane interaction between both N109 and N113).

These observations are intriguing and suggest that glycosylation may play an important role in Ig domain orientation. However, it is important to highlight some limitations when modelling glycosylation sites. Firstly, in this study only a single GlcNAc was added to each glycosylation site in the Ig domain. The addition of a single GlcNAc was motivated from the experimental structure from Pan *et al.* [4] where GlcNAc molecules are resolved at positions N109 and N113. Due to the refinement process of membrane proteins, the bulk of sugar moieties are often lost, whereas *in vivo* the length and branching of glycans is often complex and can have a measurable effect on protein dynamics [247]. Luckily, with increasing experimental resolution we are now beginning to elucidate glycosylation sites more readily [248] and perhaps one day reveal more complex glycan architectures. Secondly, the choice of force field used to model carbohydrate dynamics is often non trivial [249, 250]. Furthermore, carbohydrates contain a large number of rotatable bonds that increase the accessible conformational landscape, often requiring long simulation times to gain sufficient sampling [251]. Although the modelling of carbohydrates is not simple, the results presented here suggest that even simple models of GlcNAc may represent a change in the dynamics of the β3 conformation on the cell membrane.

4.5 Conclusions

In summary, the dynamics of Nav channel β subunits are highly variable. Simulations performed in this chapter have outlined that although similar in binding to the α subunit, there is an inherent difference in behaviour between the β1 and β3 subunits when in isolation on the cell membrane. Critically, the non-conserved R96 residue on one side of the Ig domain in β1 seems to be responsible for an increased propensity for the Ig domain to be bound to the membrane surface. In contrast, the Ig domain of the β3 subunit is much more flexible. When present as a monomer, β3

can expose hydrophobic residues contained within the N-terminus of the Ig domain that are responsible for observed oligomerisation, in a conformation that can also be stabilised by incorporation of glycosylation sites. As a full-length trimer, the β3 subunit is stable with the TMD helices seemingly not involved in additional stabilisation. The intriguing E177/E176 in β1/β3 located in the middle of the TMD seems to in fact be closer to the intracellular membrane surface than initially predicted from sequence analysis in all models tested. More generally, the results presented here offer new insights into β subunit dynamics and act as the starting point to further our understanding of Nav channel modulation and clustering.

"You're stronger than you think, you've got this Broski."

— Eleanor (sister)

5

Coarse-grained simulations of β3 subunit oligomerisation

Contents

5.1 Introduction 138
5.1.1 Lipid composition: A model of an *in vivo* membrane 138
5.2 Methods 140
5.2.1 System setup 140
5.3 Results 143
5.3.1 Cluster evolution of the β3 subunit 143
5.3.2 Protein - protein interactions between β3 clusters 146
5.3.3 β3 protein - lipid interactions 149
5.4 Discussion 155
5.5 Conclusions 159

5.1 Introduction

As described in Chapter 1, membrane proteins are not present in isolation *in vivo*. In eukaryotic cells, the cell membrane is a crowded and dynamic space with a large variety of lipid and protein species, as well as higher order and complex biological scaffolds. VGICs are no exception and typically Cav, Kv, and Nav channels are often found in combination with a variety of proteins [252] within the membrane and cytoskeleton [100, 253–257]. As well as influencing channel gating, these proteins can often function independently [258, 259].

In Chapter 4, the dynamics of the "auxiliary" β3 subunit was examined. Atomistic simulations revealed that the Ig domain undergoes a variety of pitch states with populations exposing the hydrophobic patch responsible for the stability of the crystallised ECD trimer [99]. In this chapter, the use of CG molecular dynamics is utilised to study the higher order oligomerisation observed in Nav β3 subunits. A CG approach (described in Chapter 2, Section 2.5.1.1) is adopted to access longer time and length scales than a typical atomistic simulation. Additionally the use of a CG model, and in particular the MARTINI force field, allows for the construction of an increasingly complex membrane consisting of a variety of relevant lipid species.

5.1.1 Lipid composition: A model of an *in vivo* membrane

Lipids constitute the main building blocks of cellular membranes in which membrane proteins carry out various functions [260]. Compositional diversity in the membrane varies greatly between organism, cell type, and cell leaflet [261]. This is in addition to diversity at a chemical level that is achieved through the fine tuning of specific lipid groups [11] which results in thousands of possible lipid types via alteration of head groups and / or aliphatic chains [262]. In mammalian cells, differences in lipid composition occur between the plasma membrane (PM), endoplasmic reticulum, mi-

tochondria, and the nucleus [12, 13]. Within the PM of mammalian cells, the typical lipid species are glycerolipids (65%), sphingolipids (10%), and sterols (25%) [13]. The extracellular and intracellular PM leaflets are enriched with different lipids to varying degrees, causing an inherent asymmetry within the bilayer [263]. Typically, the extracellular leaflet is enriched in PC, sphingolipids such as Sph, and glycosphingolipids such as GM3. Conversely, the intracellular leaflet is enriched with PE, PS, and PI. The exact distribution of cholesterol in the PM is somewhat elusive, with evidence suggesting its presence in both the extracellular [12] and intracellular [264, 265] leaflets. Simulations of a PM model suggest that the exact distribution is almost symmetrical between leaflets, with a slight preference for the extracellular side, demonstrating cholesterol's fast "flip-flop" rate between leaflets [266].

Due to the complex nature of lipid composition and chemical diversity in eukaryotic membranes, it remains a challenge to accurately model realistic *in vivo* membranes. To date, various CG MD studies have utilised so-called "generalised" PM models to characterise various properties such as general membrane biophysics, lipid - lipid, and lipid - protein clustering [266–269] as well as more complex membrane topologies [270]. The lipid composition used in the CG simulations presented in this chapter is taken from the work of Koldsø *et al.* [267] and includes POPC, 1-palmitoyl-2-oleoyl-glycero-3-phosphatidylethanolamine (POPE), Sph, GM3, Chol, 1-palmitoyl-2-oleoyl-glycero-3-phosphatidylserine (POPS), and 1-palmitoyl-2-oleoyl-glycero-3-pho-sphatidylinositol-4,5-bisphosphate (PIP_2) lipids. For an overview of the general chemical structures of each lipid the reader is guided to Chapter 1 (Figure 1.1).

5.2 Methods

As seen in Chapter 4, atomistic simulations of membrane proteins offer highly detailed insight into biomolecular dynamics. In the case of the β subunit, the orientation of the Ig domain was assessed in atomistic detail, with precise measurements of protein - protein and protein - lipid contacts that ultimately relate to biological function *in vivo*. However, sampling of relevant states remains a bottleneck in atomistic MD and although a number of methods exist to increase sampling [271] it is often necessary to exchange molecular accuracy for feasible simulation length. The use of CG MD is one such way to achieve this, being particularly well suited for studying biological processes at increasing time and length scales (see Chapter 2, Section 2.5.1). To study the effect of β3 clustering, the MARTINI CG force field was used.

5.2.1 System setup

The full length atomistic model of the β3 subunit (generated in Chapter 4, Section 4.2.1.2) was converted to a CG representation using the martinize.py script (v2.4) [159] to generate topology and structure files with the MARTINI (v2.2) force field [159–166] and an ElNeDyn elastic network [188]. The resulting CG β3 model was then embedded in a generalised mammalian plasma membrane composition, performed using the *insane* methodology [272] with a specified ratio of lipids, detailed in Table 5.1. A 6×6 grid of β3 subunits was constructed by tiling the single β3 subunit embedded in a mixed membrane in both the x and y direction using the *genconf* GROMACS tool. The resulting system contained 36 copies of protein with a total of 10,080 lipid species, with x and y box dimensions equal to 52 nm. Energy minimisation was performed with the steepest decent algorithm. The resulting system was solvated using the non-polarisable MARTINI water model with Na^+ and Cl^- ions at a concentration of 150 mM. Equilibration steps consisted of a short (20 ns) and

long (100 ns) equilibration whilst position-restraining backbone particles with a force constant of 1000 kJ mol^{-1} nm^{-2}. The orientation of each protein was randomised in the membrane by position-restraining one backbone particle located in the short linker region (G150) with a force constant of 1000 kJ mol^{-1} nm^{-2} and another long (100 ns) equilibration performed with no restraints on any other particles. All simulations were carried out in the NPT ensemble at a temperature of 323 K and pressure of 1 bar. Covalent bonds were constrained to equilibrium values using the LINCS algorithm [221] and the leap-frog algorithm used with a 20 femtosecond timestep. The V-rescale [146] temperature and Berendsen [142] pressure couplings were used for short equilibrations, with coupling constants of 1.0 and 8.0 ps, respectively. The V-rescale temperature and Parrinello-Rahman [152] pressure couplings were used for long equilibrations, with coupling constants set to 4.0 ps and 8.0 ps, respectively. A total of 3×10 μs independent simulations were run from each of the randomised protein orientation starting systems. All simulations were performed using GROMACS 2019 [214, 273–278].

Analysis of cluster sizes, protein - protein, and protein - lipid interactions was performed using in-house scripts utilising the MDAnalysis [222, 223], NetworkX [279] and NumPy [280] Python libraries. Radial distribution functions were calculated using the GROMACS *gmx rdf* tool. Visualisation of each system was performed using VMD 1.9.2 [224]. All graphs were produced using the Seaborn [225] and Matplotlib [226] Python libraries.

	Lipid ratio						
Membrane leaflet	POPC	POPE	Sph	GM3	Chol	POPS	PIP_2
Upper	40	10	15	10	25	-	-
Lower	10	40	-	-	25	15	10

Table 5.1: Lipid composition of the generalised mammalian membrane from Koldsø *et al.* [267] used in CG simulations of the β3 subunit. Where POPC = 1-palmitoyl-2-oleoyl-glycero-3-phosphatidylcholine, POPE = 1-palmitoyl-2-oleoyl-glycero-3-phosphatidylethanolamine, Sph = sphingomyelin, GM3 = monosialodihexosylganglioside, Chol = cholesterol, POPS = 1-palmitoyl-2-oleoyl-glycero-3-phosphatidylserine, and PIP_2 = 1-palmitoyl-2-oleoyl-glycero-3–phosphatidylinositol-4,5-bisphosphate.

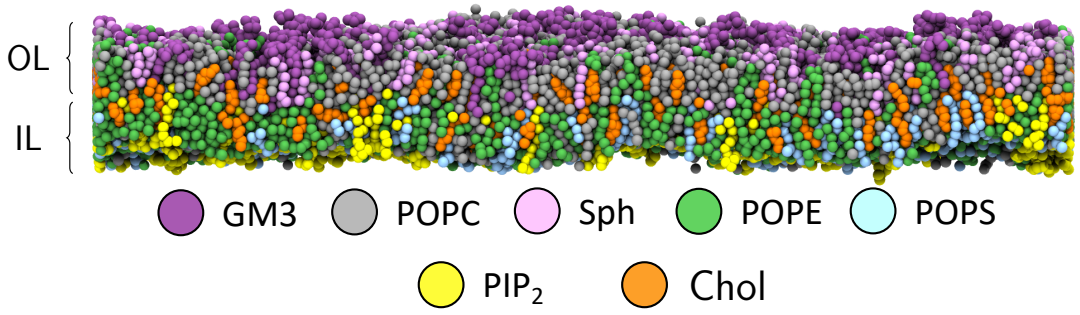
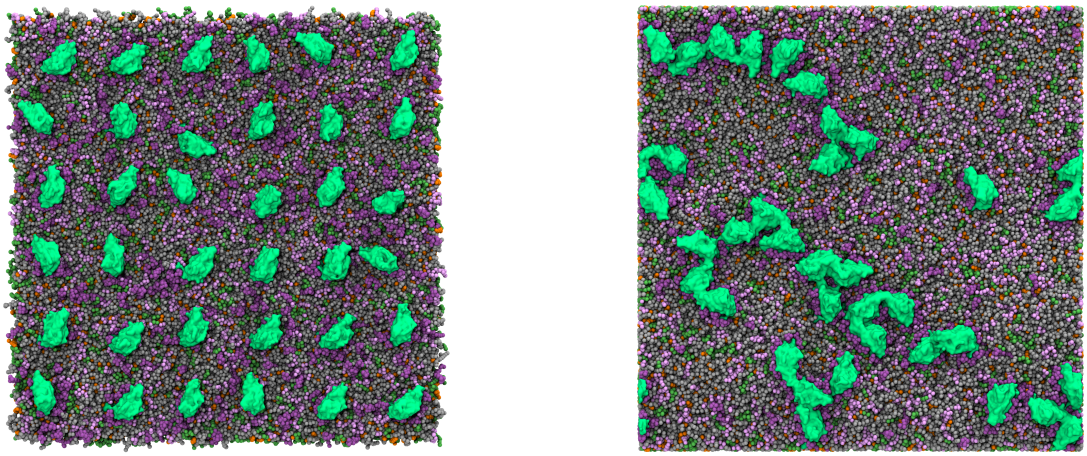


Figure 5.1: CG generalised mammalian membrane composition. Lipid species are coloured separately and shown as spheres in the membrane cross-section, OL = outer leaflet, IL = inner leaflet. Lipid labels are the same as in Table 5.1.

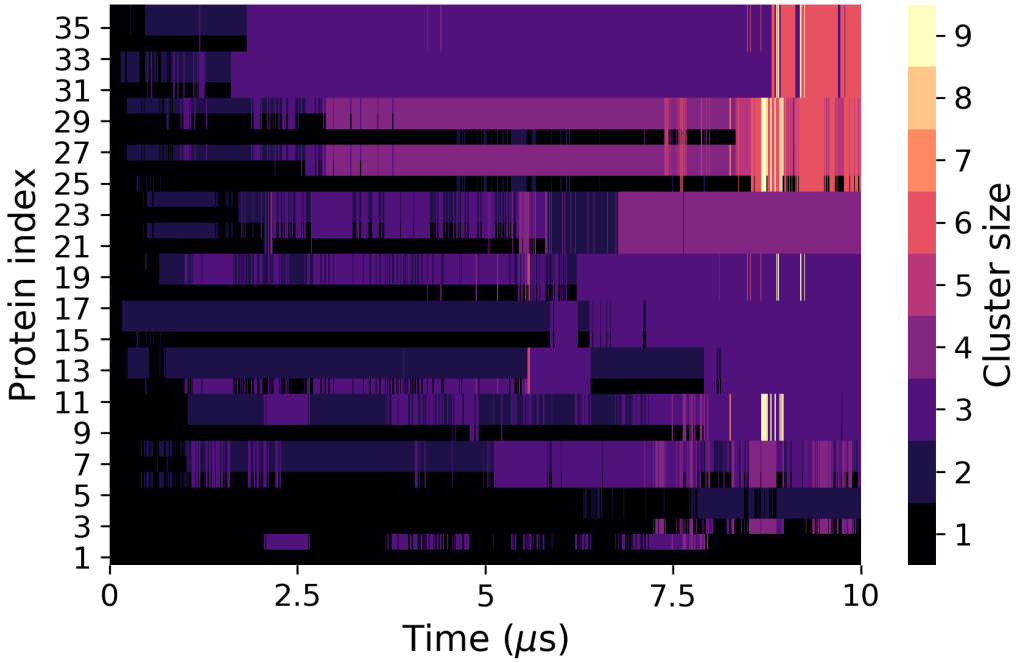
5.3 Results

5.3.1 Cluster evolution of the β3 subunit

Each starting 6×6 β3 system contained a unique distribution of Ig domain orientations after randomisation as well as different lipid distributions. Over the course of each 10 μs simulation of the large β3 system, higher order oligomers were formed (Figure 5.2b). As seen in Chapter 4, the β subunit is particular flexible at the linker region. In order to mitigate any unrealistic clustering analysis, residues 25-150 were used to define the protein centre. These residues were chosen as they make up the bulk of the Ig domain that undergoes little conformational change and are structurally rigid due to the presence of several β strands. The centre of geometry defined by these residues in the Ig domain was used to calculate the distance between each protein copy at every frame. Proteins were considered to be part of a cluster if the centre of geometry of the Ig domain was within a radius of 45 nm of another β3 Ig domain. Additionally, each trajectory was post-processed to ensure that each protein copy was made whole across periodic boundaries. Over the course of each 10 μs, cluster sizes generally increased. The most prominent protein cluster size was a monomer or dimer, with significant populations involved in higher order clusters, as shown in Figure 5.2c.



(a) (b)



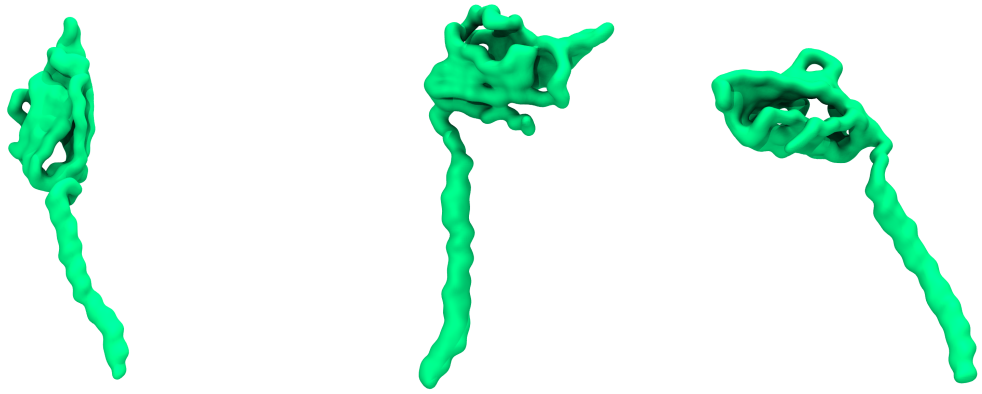
(c)

Figure 5.2: Cluster evolution of the β subunit in a generalised mammalian membrane over 10 μs. (a) Extracellular view of the starting arrangement of β3 subunits (shown in green) at $t = 0$ μs. Each β3 subunit's orientation has been randomised prior to the 10 μs run. (b) Extracellular view of the final arrangement of β3 subunits as part of higher order oligomers. (c) Evolution of cluster size that each β3 monomer is a part of over 10 μs. Colour of lipid species is the same as shown in Figure 5.1.

In all CG β3 runs, similar clustering patterns were formed (see Appendix B, Figure B.4 for snapshots of other repeats). In the 25 $\times$ 400 ns atomistic simulations of the β3 monomer (see Chapter 4), a variety of pitch states are visited by the Ig domain. The protein - POPC interactions in these atomistic simulations was non-specific in the Ig domain, reflecting the Ig domain's dynamic motion on the membrane surface. Although adequate to study short time-scale motion, a simple POPC bilayer does not reflect the true lipid complexity present in eukaryotic cell membranes. In order to assess whether the introduction of a more complex CG membrane influences the pitch and overall conformation of the β3 subunit, the conformation of each protein copy in the final frame of each simulation was extracted. Across all CG simulations performed there was a distribution of approximately three Ig domain orientations, coupled with more variability in the TMD. The majority of Ig domains were present in a "down" state, resembling that of the β3 model constructed from the α - β1 crystal structure, with the N-terminal 'A' β strand orientated above VSDIII of the α subunit (as described in Chapter 1.3). Conversely, another population exhibits an "up" state of the Ig domain, with the long axis of the Ig domain approximately perpendicular to the plane of the membrane. Also present were Ig domains in a "rolled" state with one face of the Ig domain bound to the membrane and exposing the hydrophobic patch responsible for the observed trimerisation in the Namarudai *et al.* [99] crystal structure. All protein copies were inspected visually and assigned a state based on the afore mentioned descriptions. A summary of states in the final frame for each run and representative structures are shown in Table 5.2 and Figure 5.3.

	Number of states		
Run number	Up	Down	Rolled
1	6	12	18
2	2	7	27
3	2	9	25

Table 5.2: Visual assignment of $\beta 3$ states in each clustering run.



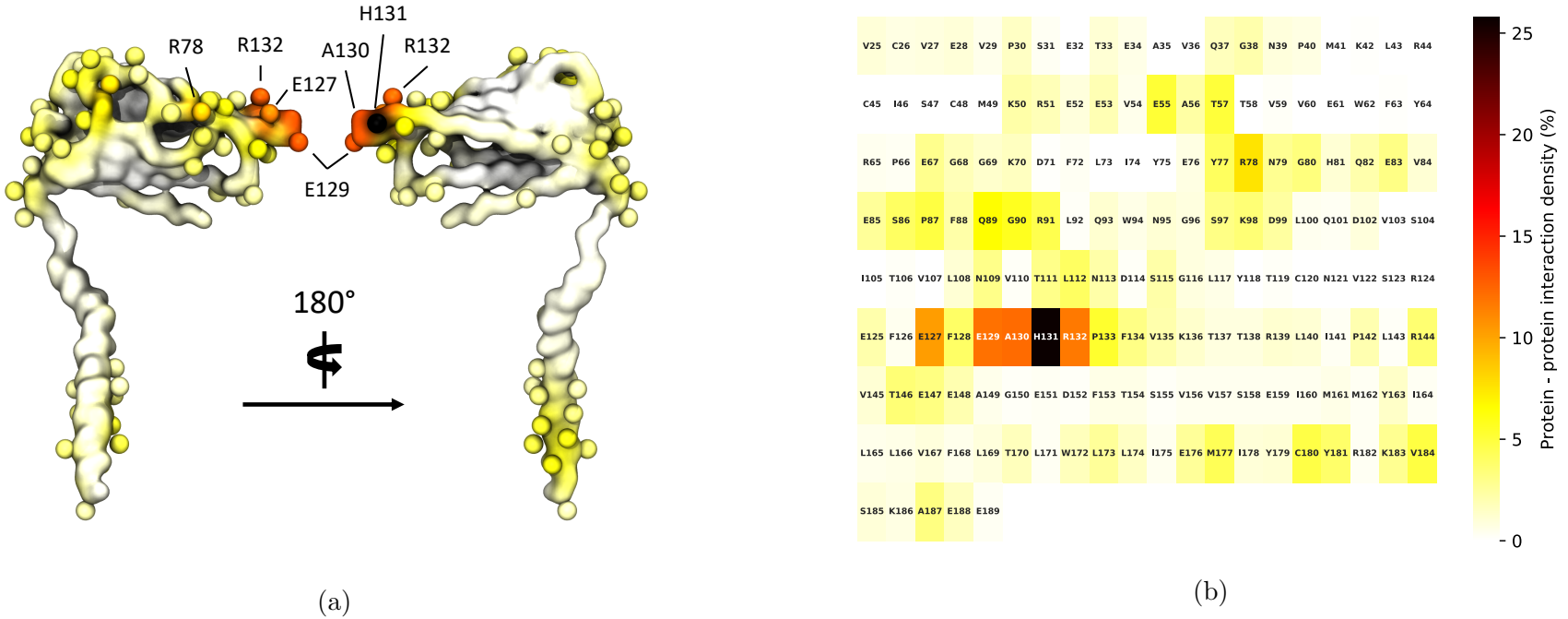
(a) Up (b) Down (c) Rolled

Figure 5.3: Different states adopted by $\beta 3$ clusters.

5.3.2 Protein - protein interactions between $\beta 3$ clusters

To assess the likely protein - protein interfaces between $\beta 3$ subunits, protein contacts were calculated across all repeats using in-house scripts described in Section 5.2.1. At each frame, protein residues were considered to be in contact if their distance was $\leq$ 8 Å. Protein - protein contact analysis was repeated for each run and the average interaction for each residue recorded. The contact results for each run were combined to give an overall average of protein - protein contacts and each residue assigned a percentage of the total protein - protein interaction. The interactions were then visualised on both the structure and sequence of the $\beta 3$ subunit, as shown in Figure 5.4. Overall, there are high regions of contact on the Ig domain and C-terminal end of the TMD. In the Ig domain, one face seems to be favoured with most contact

appearing at the "tip" (i.e. at the end of the long axis, between the F and G β strands) of the Ig domain. At the opposing end of the Ig domain there are high regions of contact, albeit slightly more sparse. Regions of contact in the TMD are located close to the intracellular end of the helix with high areas of contact partially favouring one side consisting of the conserved E176 residue as well as residues M177, C180, Y181, K183, and V184. Taking a closer look at contact regions on the β3 sequence (Figure 5.4b) reveals other residues that are in close contact with other protein copies at a lower interaction density. At the N-terminus, the residues responsible for the observed trimerisation in the β3 Ig domain (V25, C26, V27, E28, V29, and P30) exhibit a low interaction density. Other regions are located on the surface of the body of Ig domain with low but numerous contact sites between residues Q82-D99 that make up the end of the Ig domain closest (spatially) to the start of the short linker region.



(a) (b)

Figure 5.4: Protein - protein interactions in $\beta 3$. (a) Average $\beta 3$ protein - protein interactions. The first side chain particle of residues that account for more than 1.5% of the total protein - protein interactions are shown as spheres. (b) Average $\beta 3$ protein - protein interactions heatmap overlaid with amino acid sequence.

At first glance, the protein - protein contacts made between CG copies of the β3 subunit seem somewhat counterintuitive when considering the trimeric crystal structure isolated by Namadurai *et al.* [99]. The key hydrophobic patch that enables the favourable interaction between the Ig domains seems to have low interaction density, with most of the protein - protein contacts between subunits resembling an "end-on" orientation at the extrema of the long axis of the Ig domain. However, as highlighted in Figure 5.3 the β3 Ig domain is present in a number of states over the course of a trajectory, analogous to those seen in atomistic simulations. Clearly, alternative Ig domain orientations will lead to alternate binding patterns between subunits. As seen in Chapter 4, atomistic simulations revealed that the β3 Ig domain occupies a variety of pitch states but critically these simulations are performed in a relatively simple POPC membrane. Conversely, in CG simulations the membrane is highly heterogeneous and contains a multitude of different lipid species that ultimately have an effect of how the Ig domain orientates itself on the membrane.

5.3.3 β3 protein - lipid interactions

Over the course of each 10 μs simulation, the β3 subunit interacts with the variety of different lipid species present in the CG membrane. To quantify any preferential binding the protein has to a particular lipid, the radial distribution function (RDF) of lipid species around the protein copies was computed. Briefly, the RDF gives an estimate for the likelihood of finding a particle at distance (r) from a reference point, in this case the reference is the β3 subunit. Out of the seven different lipid types constituting the generalised mammalian membrane, only GM3, PIP_2, and cholesterol interact strongly with the β3 subunit. Of species in the outer leaflet, GM3 shows the strongest preference with a slow decay towards bulk (i.e. $g(r) = 1$), as shown in Figure 5.5.

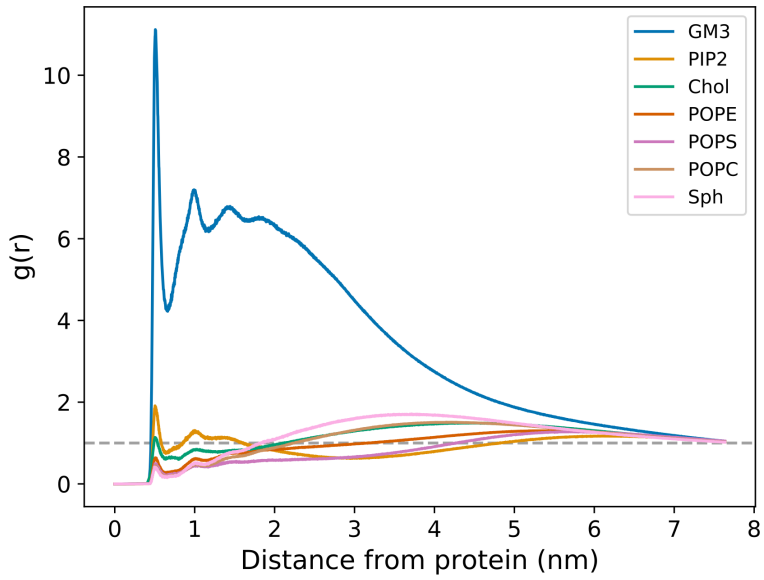


Figure 5.5: Radial distribution function of lipids around β3.

In the lower leaflet, PIP_2 has the highest $g(r)$, whilst Chol (known to flip-flop between membrane leaflets) shows a slight increase in probability close to β3. The resulting POPE, POPS, POPC, and Sph species all have a measured $g(r) < 1$ close to the subunit $g(r)$ increasing gradually to bulk at larger r values. Mapping the protein - lipid interactions on to the starting model of the β3 subunit (Figure 5.6) reveals that GM3 effectively coats one side of the Ig domain, as well as persistent interaction with the short linker region at the top of the TMD. With PIP_2, located exclusively in the lower leaflet, the density is concentrated at the C-terminus. Due to the negatively charged headgroup of PIP_2, interactions are concentrated at the positively charged R182, K183, and K186 residues (see Appendix B, Figure B.5). Cholesterol binds non-specifically across most residues present in the TMD helix. POPE exhibits density at both ends of the TMD and a small amount on the Ig domain, with a slight preference for the same face as seen in GM3. As with PIP_2, there is concentrated density of POPS at the C-terminus whilst both POPC and Sph preferentially interact with residues at the top of the TMD.

Figure 5.6: Average β3 protein - lipid interactions. (a) Face 1 of the β3 subunit. (b) Face 2 of the β3 subunit. The first side chain particle of residues that account for more than 2.5% of the total protein - lipid interactions are shown as spheres. Colour scale is from least (white) to most interacting (black). All lipid abbreviations are the same as described in Table 5.1

Visual inspection of all lipid interaction sites revealed that one face of the Ig domain preferentially interacts with the membrane, as shown in Figure 5.6a. As explained earlier, there are a number of states that the Ig domain occupies on the membrane, with the "rolled" state most abundant (Table 5.2 & Figure 5.3c). The interaction sites across all lipids in the Ig domain of β3 show that the face of the Ig domain that points downwards and towards the membrane is the same face that accumulates most lipid interaction from GM3, POPC, and Sph. Mapping the β3-GM3 interactions onto the protein sequence reveals several regions of the subunit that interact most frequently (Figure 5.7). Some of the highest regions of interaction with GM3 include residues at the N-terminus of the Ig domain from C26 to V54. In the β1 - Nav1.4 crystal structure, this region is in close proximity to the top of VSDIII and positioned just above the membrane plane. Additional high interaction sites occur at R78, N79, G80, and H81 that make up the loop region between β strands[†] C$''$ $\rightarrow$ C$'$. Similarly in the loop between the E $\rightarrow$ D β strands, residues 96, 97, 98, and 99 also interact with GM3 and are close to the predicted N-linked glycosylation N95 site. Within the short linker region between the Ig domain and the TMD (residues E147 $\rightarrow$ D152) and at the start of the TMD helix there are regions of high interaction at E148 $\rightarrow$ F153.

[†]The reader is guided to Chapter 1, Figure 1.7 for a reminder of Ig β strand labelling convention.

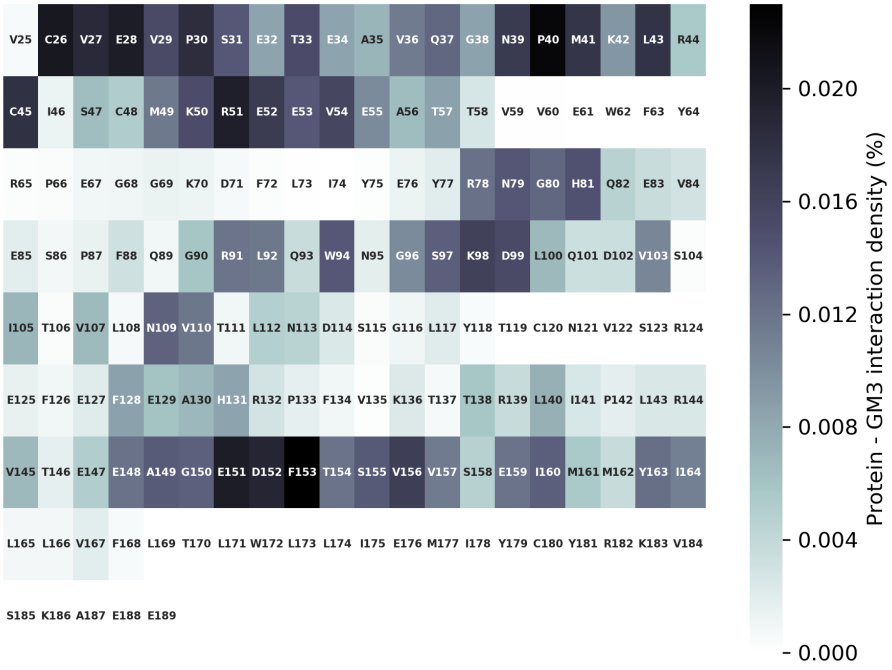


Figure 5.7: Regions of high interaction of the β3 with GM3.

The density of each lipid type surrounding each β3 protein was calculated by aligning the protein coordinates to C26 (located at the N-terminus at the start of the 'A' β strand) at every frame and for each protein copy. The density of each lipid species surrounding the protein within a 12 nm radius was recorded from the positions of the CG headgroup particle. The distribution of lipids surrounding the β3 subunit at various angles was also recorded to give a rough indication as to whether any side of the protein was favoured for lipid binding. With reference to the RDF (Figure 5.5), the two most interacting lipid species (GM3 and PIP_2) are shown in Figures 5.8a & 5.8b. The clustering of GM3 typically accumulates below and to one side of the Ig domain, as well as surrounding the extracellular end of the TMD helix. Analysis of GM3 density at various angles (with C26 taken as 0°) reveals a slight preference for one side of the β3 subunit and reflects the accumulation of GM3 on one "face" of the Ig domain, as seen in Figure 5.6. Conversely, PIP_2 interactions are localised to the C-terminal end of the TMD (Figure 5.8c). Unsurprisingly, highly interacting residues

include the positively charged K183 and K186 residues that favourably interact with the anionic PIP_2 headgroup. Additionally, residues V184, A187, and E188 also show high interaction density that may be a reflection of their close proximity to K183 and K186. There seems to be little angular preference for where PIP_2 binds to the TMD with all angles equally sampled over the course of the trajectory (Figure 5.8d). The full set of PIP_2 interactions are shown on the β3 sequence in Appendix B (Figure B.5).

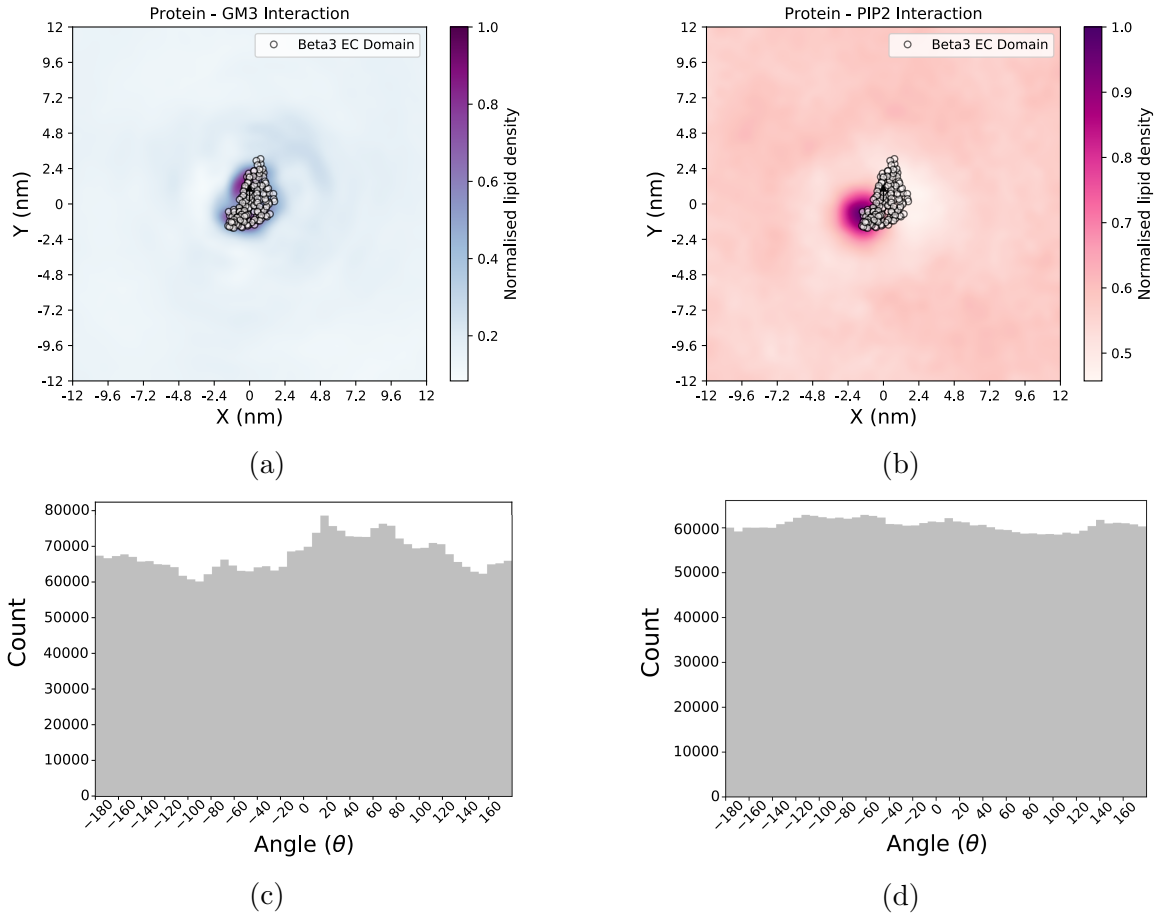


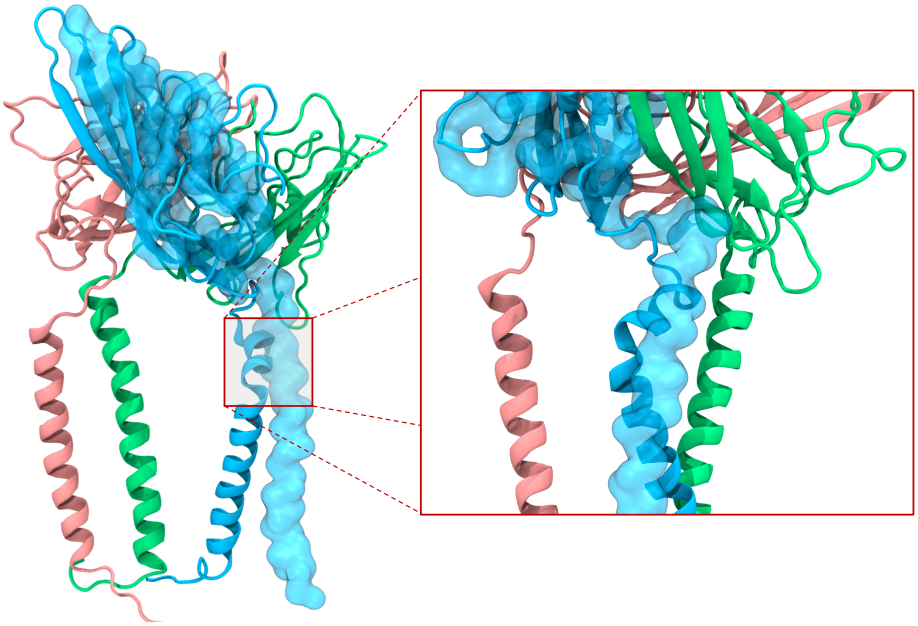
Figure 5.8: Lipid density surrounding the β3 subunit. (a) GM3 density. (b) PIP_2 density. (c) and (d) show the lipid count at various angles around the protein. The solid black arrow in (a) and (b) represents the vector that each β3 subunit was aligned to at every frame and points from the centre of mass of the Ig domain to residue C26, located at the N-terminus.

5.4 Discussion

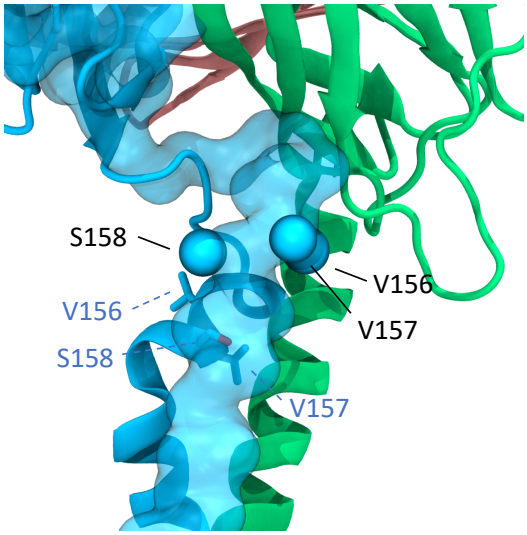
5.4.0.1 Ig dynamics in a mixed lipid membrane

The clustering of the β3 subunit on a generalised mammalian membrane gives rise to varying degrees of cluster size within which a series of different protein conformations were visited. Most prevalent of these was the Ig domain positioned in a "rolled" state with the face containing β strands B, D, and E preferentially bound to the membrane, leading to slight exposure of the hydrophobic patch responsible for observed trimerisation. Interestingly, the population of "up" states of the Ig domain is somewhat reduced when compared to atomistic simulations performed in Chapter 4, Section 4.3.2. However, one must consider the lipid environment in which these simulations were performed. As described in Section 5.2.1, the CG membrane used contains a variety of different lipid species other than the simple POPC used in all atomistic simulations. The accumulation of GM3 on one side of the Ig domain, as well as on the short linker region, seems to restrict the motion it can undergo. It is most likely due to GM3's ability to extend beyond the membrane plane (attributed to its large sugar headgroup) that allows it to interfere with the dynamics of the Ig domain. This may go some way to explain why other lipids tend to only interact sparsely with the Ig domain and are restricted mainly to TMD interactions.

Although few in number, a population of β3 orientation states were observed over the course of each simulation in an "up" state that resembled an Ig orientation compatible with the potential to form trimer structures as seen in the Namadurai *et al.* [99] structure. The alignment of this conformation with the final frame of the full length β3 trimer model (see Chapter 4, Figure 4.20b) reveals that this orientation does not form any major steric clashes between the Ig domain and TMD (Figure 5.9a).



(a)



(b)

Figure 5.9: Overlay of "up" state $\beta 3$ subunit with the atomistic trimer model. (a) Side on view of the full length, atomistic $\beta 3$ trimer model with the CG I$\beta 3$ "up" conformation superimposed on one chain (transparent and cyan). Inset shows local bending in the TMD of one $\beta 3$ monomer. Separate chains are coloured cyan, green, and pink. (b) Residues close to the observed bending in the TMD in both the atomistic trimer (shown as sticks with blue labels) and the"up" $\beta 3$ conformation (shown as cyan spheres with black labels).

Overlaying the "up" conformation with the full length trimer model appears to satisfy the steric restraints imposed by the Ig domain present as a trimer. However, visual inspection of the TMD (shown in Figure 5.9a) indicates that the start of the TMD helix is close to the neighbouring Ig domain. Importantly, and as found in atomistic simulations shown in Chapter 4 (Section 4.3.4), after 400 ns of simulation the top of the TMD helix in β3 distorts significantly close to residues V156, V157, and S158. In the CG β3 model, the same amount of distortion is not seen and instead the helical structure is retained (Figure 5.9b), most likely due to the secondary structure restraint imposed by the use of the ElNeDyn network. The imposed restriction on secondary structure movement may also explain why the N- and C-termini of the TMD remain relatively stable throughout each 10 μs simulation, as opposed to the unravelling seen in AT simulations (Chapter 4, Section 4.3.2). It is also worth noting that the sedimentation coefficient distribution analysis performed by Namadurai *et al.* [99] (Chapter 4, Figure 4.1a) indicates that the trimeric form of β3 is not as abundant as the dimeric and monomeric forms. Additionally, the use of CG MD in this study has allowed for the analysis of a cumulative 30 μs of simulation time which, although long by modern standards, may not be sufficient for the Ig domain to encounter other protein copies in the correct orientation to facilitate trimer formation in the crystal pose.

Other contributions not studied here include the effect of any glycosylation sites on the Ig domain. As studied in Chapter 4, the Ig domain is heavily glycosylated with four sites (N95, N109, N113, and N121). The position of glycosylation sites is not expected to influence the ability of the β3 Ig domain to form higher order oligomers since their positions are located away from the hydrophobic interface, as demonstrated by the trimeric crystals structure isolated by Namadurai *et al* [99]. Interestingly, one of the glycosylation sites at N95, is on the same face as the increased protein - GM3 lipid interaction site. Additional glycosylation sites N109 and N113 are

located at the E $\rightarrow$ G β strands and close in space to the short linker. The favourable interaction between GM3 and any CG glycosylated asparagine residues may help to further stabilise certain orientations of the Ig domain on the membrane and can be hypothesised to correspond to the "up" and "rolled" states defined earlier. However, one must exercise caution when modelling sugars, as highlighted in Chapter 4, Section 4.2.1.3. This is especially apparent in the MARTINI model, where the inclusion of large polysaccharide chains results in excessive aggregation [281].

5.4.0.2 The effect of lipid clustering

As described by Koldsø *et al.* [267], the CG lipid bilayer is highly heterogeneous and contains a variety of lipid - lipid interactions, the most prevalent of these being the observed clustering of GM3. In all CG simulations, GM3 appears to coat each β3 subunit copy, forming localised lipid clusters and promoting the association between Ig domains. The ability for GM3 to form lipid rafts is well known [282] and although not at the same scale as an *in vivo* membrane, the CG simulations performed here hint at GM3 clustering as a driving force for protein association. In the closely related β1 subunit, results have shown that it is preferentially found in lipid rafts where it acts as a substrate for sequential cleavage by β-(BACE) and γ-secretases, with links to the pathology of Alzheimer's disease [283]. This may well explain the relatively large cluster sizes seen over the course of the 10 μs simulations. GM3 tends to act as a bridge between β3 subunit monomers and although close to each other (i.e. at or below the 45 nm cut-off), the proteins may not necessarily be in direct contact. This is not to say that protein contacts do not form. As described earlier, the protein - protein interaction is concentrated at the "tip" of the Ig domain and may reflect the limited regions available for contact making whilst sterically blocked by the large GM3 sugar groups. Interestingly, and from a functional viewpoint, the auxiliary β1 subunit of the Ca^{2+} and voltage-gated potassium (BK) channel has been found to retain its

fine-tuning of channel gating properties only in the presence of negatively charged lipids in the membrane [284]. Since at physiological pH the sialic acid component of the GM3 headgroup will be negatively charged, it is tempting to speculate that the GM3 clustering seen in the simulations performed in this chapter of the Nav β3 subunit may influence the α subunit in a similar way.

5.5 Conclusions

In this chapter CG MD has been utilised to probe the observed β3 clustering in a generalised mammalian membrane. A total of 36 β3 subunits were found to cluster in the membrane, forming large oligomeric structures. Analysis of protein - protein contacts revealed that the interaction sites are localised in the Ig domain, with a small contribution made from the key hydrophobic residues responsible for stabilising the trimeric crystal structure. A closer look at the surrounding lipids revealed that GM3 effectively coats one face of the Ig domain, restricting its movement on the membrane whilst simultaneously promoting GM3 (and so protein) aggregation. This was in contrast to atomistic simulations where the Ig domain of the β3 subunit was much more dynamic and highlights the importance of membrane composition in protein behaviour.

The clustering of the β3 subunit is not present in isolation and how this clustering may influence the spatial density of the pore-forming α subunit is the subject of Chapter 6.

"Done is better than perfect."

— Naomi (sister)

6

Exploring α - β3 interactions using coarse-grained simulations

Contents

6.1 Introduction 162
6.1.1 The significance of Nav1.5 α - β3 clustering 162
6.2 Methods 163
6.2.1 System setup 163
6.3 Results 168
6.3.1 Cluster evolution of the α subunit 168
6.3.2 α - β3 subunit interactions in a large membrane system . . 170
6.3.3 Influence of lipid interactions on α - β3 clustering 177
6.3.4 Alternative binding mode of the β3 subunit 179
6.4 Discussion 181
6.4.1 Preferential binding of the β3 subunit 181
6.4.2 Influence of the lipid membrane 186
6.5 Conclusions 188

6.1 Introduction

In Chapter 4, the dynamics of the β3 subunit in isolation was found to be highly variable with a large range of pitch angles sampled. However, in Chapter 5 the introduction of a more complex lipid membrane altered the conformations visited by β3 and suggested a much more restricted set of Ig pitch angles that are present in broadly three states (labelled "up", "down", and "rolled"). As described in Chapter 1, the extent to which the β3 subunit influences the spatio-temporal dynamics of the Na^+ conducting α subunit remains to be fully characterised. Although there are an increasing number of Nav channel structures with β subunits bound [4, 5], the full extent of interactions between cardiac Nav1.5 channels and auxiliary subunits via cross-linking is not clear. In this chapter, large scale simulations are performed on a $\cong$ 2 million CG particle system to investigate the clustering of α and β3 subunits using the same lipid composition as described in Chapter 5.

6.1.1 The significance of Nav1.5 α - β3 clustering

It is known that for the efficient initiation of the action potential, VGICs preferentially cluster at the axon initial segment and the Nodes of Ranvier [285]. As explained in Chapter 1, the nodes of Ranvier are located between myelinated regions of the axon body where they help to propagate the AP. As discussed in a recent review by Freeman *et al.* [286], studies have typically shown that Nav1.1, Nav1.2, Nav1.6, Nav1.7, Nav1.8, and Nav1.9 preferentially cluster at these nodes [287, 288]. The absence of Nav1.5 is predominantly due to its presence in cardiac myocytes, whose overall architecture differs to that of skeletal muscle and nerve cells (see Chapter 1, Section 1.1.4). Instead, cardiac myocytes are connected through gap junctions that facilitate rapid depolarisation from adjacent cells, initiating the opening of neighbouring Nav1.5 channels. Increasing evidence suggests that these cardiac Nav channels do not func-

tion in isolation and can interact with copies of other Nav1.5 α subunits [100, 257] as well as a variety of other protein species, including β subunits [289]. If Nav1.5 channels are similar to those present at the nodes of Ranvier, their clustering with auxiliary subunits may enhance AP propagation and impact overall cardiac function *in vivo*.

MD simulations performed in this chapter aim at elucidating the possible binding site(s) of the β3 subunit, address possible cross-linking of α subunits, and measure the effect of protein - lipid interactions in a complex lipid membrane.

6.2 Methods

6.2.1 System setup

6.2.1.1 The 3 $\times$ 3 α subunit grid

The Nav1.5 sequence was aligned with the known sequences of several Nav channels (SCNXA, where X = 1,2,3,5,7,8,9,10,11) using the EMBL-EBI MUSCLE web server [193, 194]. The aligned Nav1.5 sequence was extracted and the recent Nav1.7 cryo-EM structure (PDB: 6J8G) solved by Shen *et al.* [5] was used as a template to construct a total of 200 homology models of the Nav1.5 α subunit (sequence similarity = 76.2 %). Each model was created and scored using the DOPE function [196] in the Modeller (v9.16) software package [195]. The 10 best models were ranked using the QMEAN [197] scoring function in SWISS-MODEL [212] and the final model chosen with the highest QMEAN score.

Once created, the atomistic Nav1.5 α subunit model was embedded in a POPC membrane, prepared using the InflateGRO methodology [220] and in-house scripts for final adjustments. All atomistic simulations used the AMBER ff99SB-ILDN force field [215]. The system was solvated using the TIP3P water model [219] and neutralised

with Na^+ and Cl^- ions at a concentration of 150 mM. Energy minimisation was performed using the steepest decent algorithm. Equilibration steps of each system consisted of a short (1 ns) and long (5 ns) equilibration whilst position restraining the Cα atoms with a force constant of 1000 kJ mol^{-1} nm^{-2}. All simulations were carried out in the NPT ensemble. The temperature and pressure were set to 310 K and 1 bar using the Nosé-Hoover thermostat [147, 148] and Parrinello-Rahman barostat [152] with coupling constants of 0.8 and 5.0 ps respectively. In order to relax the model in the membrane, a production run of 200 ns was undertaken without restraints using the Nosé-Hoover thermostat and Parrinello-Rahman barostat.

The relaxed α subunit structure was converted to a CG representation using the martinize.py script (v2.4) [159] to generate a topology and structure files with the MARTINI (v2.2) force field [159–166] and an ElNeDyn elastic network [188]. The resulting CG α model was then embedded in the same generalised mammalian membrane composition used in Chapter 5, obtained using the *insane* methodology [272]. A 3 $\times$ 3 grid of α subunits was constructed by tiling the single α subunit embedded in a mixed membrane in both the x and y direction using the *gmx genconf* GROMACS tool. The resulting system contained 9 copies of protein with a total of 7,137 lipid species and box dimensions of 47 nm in both x and y. Energy minimisation was performed with the steepest decent algorithm. The system was then solvated using the using the non-polarisable MARTINI water model and neutralised with Na^+ and Cl^- ions at a concentration of 150 mM. Another round of energy minimisation was performed, again using the same steepest decent algorithm. Equilibration steps consisted of a short (20 ns) and long (100 ns) equilibration whilst position restraining backbone particles with a force constant of 1000 kJ mol^{-1} nm^{-2}. The orientation of each protein was randomised in the membrane by position restraining one backbone particle with a force constant of 1000 kJ mol^{-1} nm^{-2} and another long (100 ns) equilibration performed. This particle (S1744) was chosen as its position in the S6 helix

of DIV was close to the centre of the channel as well as being roughly central with respect to the membrane z-axis. All simulations were carried out in the NPT ensemble at a temperature of 323 K and pressure of 1 bar. Covalent bonds were constrained to equilibrium values using the LINCS algorithm [221] and the leap frog algorithm used with a 20 femtosecond timestep. The V-rescale [146] temperature and Berendsen [142] pressure couplings were used for short equilibrations, with coupling constants of 1.0 and 8.0 ps respectively. The V-rescale temperature and Parrinello-Rahman [152] pressure couplings were used for long equilibrations, with coupling constants set to 4.0 ps and 8.0 ps respectively. A total of 3 $\times$ 10 μs independent production run simulations were performed from each of the randomised protein orientation starting system. All simulations were performed using GROMACS 2019 [214, 273–278]. MDAnalysis [222, 223] was used for analysis and all graphs were produced using the Seaborn [225] and Matplotlib [226] Python libraries.

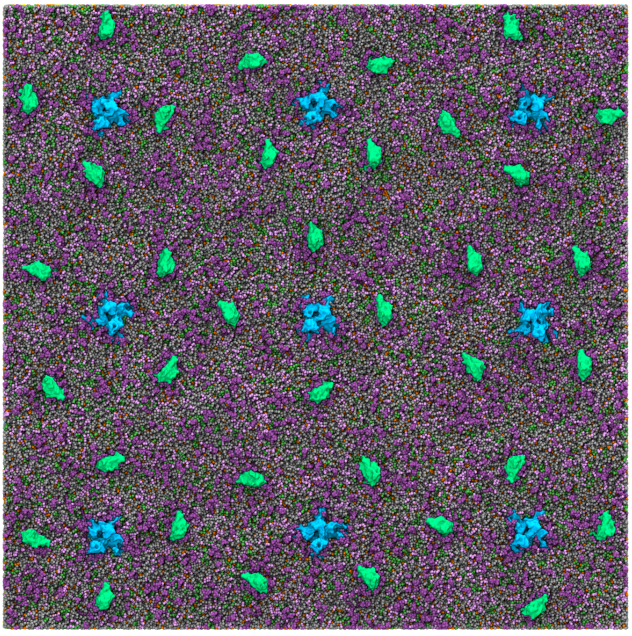
6.2.1.2 The mixed α & $\beta 3$ subunit grid

A large CG grid containing a mixture of α and $\beta 3$ subunits was constructed using the monomeric CG models developed in Section 6.2.1.1 and Chapter 4, Section 5.2.1, respectively. A system containing one α subunit and four $\beta 3$ subunits was constructed using successive iterations of the *gmx editconf* GROMACS tool and labelled the unit cell (UC). The same system was constructed at a 45° angle to the reference ($UC_{45°}$), which was created in order to minimise the distance between protein copies. The final large scale grid of α and $\beta 3$ subunits was constructed by combining both the UC and $UC_{45°}$ in three rows (Figure 6.1a). The resulting grid of α and $\beta 3$ subunits was embedded in the same generalised mammalian membrane composition as used in Chapter 5, again using the *insane* methodology [272]. The embedded system was solvated with the non-polarisable MARTINI water model, neutralised with Na^+ and Cl^- ions at a concentration of 150 mM and energy minimised using the steepest de-

cent algorithm. This was followed by a short (20 ns) equilibration whilst position restraining backbone particles. The system contained 9 copies of the α subunit and 36 copies of the $\beta 3$ subunit, with a total of 43,666 lipid species and box dimensions of 102 nm in both x and y (Figure 6.1b). A long equilibration (100 ns) was performed with a reduced time step of 10 femtoseconds whilst position restraining the x and y coordinates of all protein backbone particles, allowing each protein to relax in the z dimension. Protein orientations were randomised by a further long equilibration step for 1 μs whilst position restraining the backbone particles of residues S1744 and G150 (located in the short linker region, as described in Chapter 5, Section 5.2.1) in the α and $\beta 3$ subunits respectively. All position restrains were performed with a force constant of 1000 kJ mol^{-1} nm^{-2}. All simulations were carried out in the NPT ensemble at a temperature of 323 K and pressure of 1 bar. Covalent bonds were constrained to equilibrium values using the LINCS algorithm [221] and the leap frog algorithm used with a 20 femtosecond time step. The V-rescale [146] temperature and Berendsen [142] pressure couplings were used for short equilibrations, with coupling constants of 1.0 and 4.0 ps respectively. The V-rescale temperature and Parrinello-Rahman [152] pressure couplings were used for long equilibrations, with coupling constants set to 1.0 ps and 12.0 ps respectively. A total of 3 $\times$ 17 μs independent production run simulations were performed from the starting system shown in Figure 6.1. All simulations were performed using GROMACS 2019 [214, 273–278]. Visualisation of each system was performed using VMD 1.9.2 [224]. All graphs were produced using the Seaborn [225] and Matplotlib [226] Python libraries.

$UC_{45°}$	UC	$UC_{45°}$
UC	$UC_{45°}$	UC
$UC_{45°}$	UC	$UC_{45°}$

(a)



(b)

Figure 6.1: Construction of the large α - β3 grid. (a) A schematic illustrating the large CG system setup that creates a chequerboard pattern with each unit cell at a 45° orientation to its neighbour. (b) An extracellular view of the final starting structure after protein randomisation containing 9 α subunits (cyan) and 36 β3 subunits (green) embedded in a generalised mammalian membrane (lipid colouring is the same as in Chapter 5, Figure 5.1), solvent removed for clarity.

6.3 Results

6.3.1 Cluster evolution of the α subunit

The evolution of α subunit clusters over each 10 μs simulation in the 3 $\times$ 3 system was analysed using the same methodology as outlined in Chapter 5, Section 5.3.1. Briefly, the positions of residues 117-236 , 699-822, 1188-1317, and 1516-1640 were used to define the centre of geometry of the α subunit. These residues were chosen on the basis that they encapsulate VSDs I-IV of the α subunit, additionally they contain few residues above or below each leaflet of the membrane and so did not distort the centre of geometry to be out of the xy plane. Neighbouring α subunits were considered to be part of a cluster if the distance, defined by the calculated centre of geometry, was $\leq$ 11 nm. Over the course of 10 μs, clusters of α subunits formed with a maximum cluster size of 4.

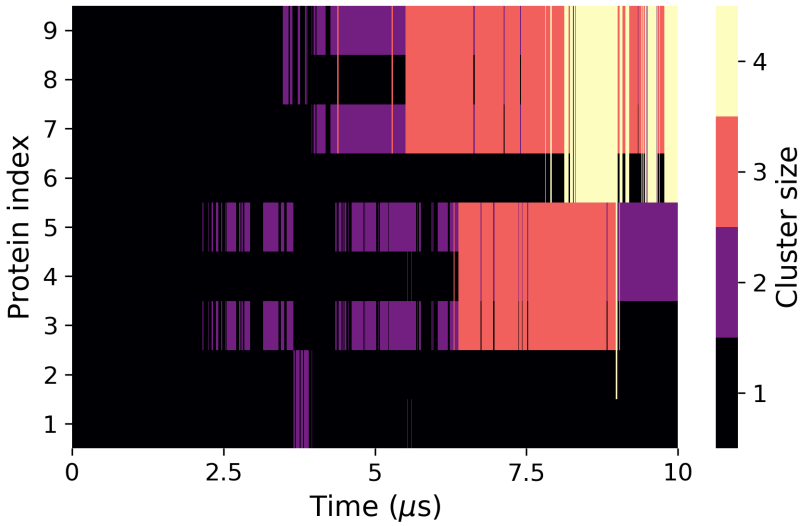
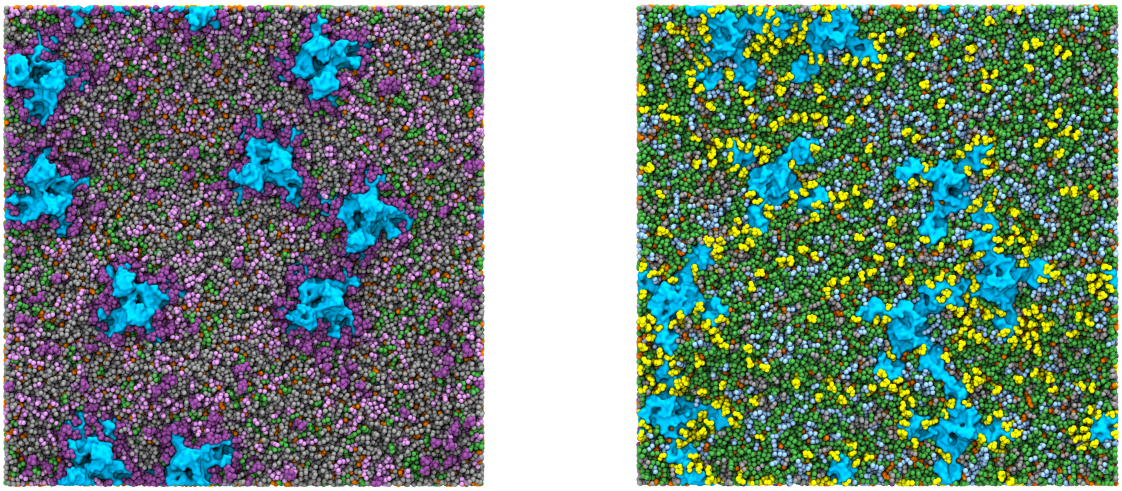


Figure 6.2: Cluster evolution of the 3 $\times$ 3 α subunit system.

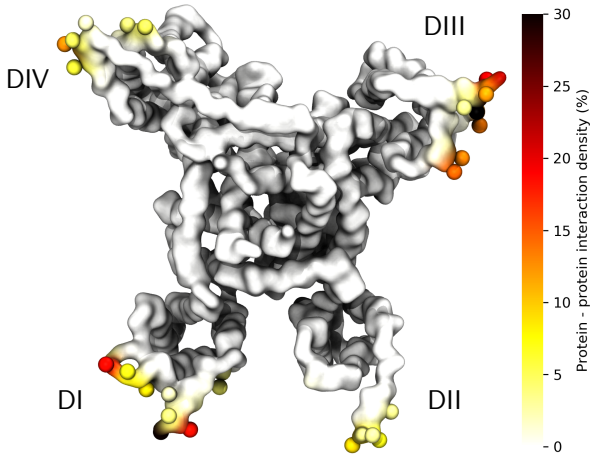
As might be expected, the interaction sites between clusters were exclusively located on the four VSDs. In each VSD, the interactions are located primarily on residues in the cytoplasmic leaflet of the membrane with the exception of T1290,

L1291, and G1292 on VSDIII and the lone Y149 on VSDI. Generally, the residues with the highest levels of interaction are those which occupy protruding regions away from the central α subunit body and so interact first with neighbouring subunits. Residues of VSDII and VSDIV seem to preferentially be in contact with one another (Figure 6.3). The slight preference for VSDII and VSDIV interactions in the cytoplasmic leaflet leads to "chains" of α subunits (Figure 6.3b). When viewed from the extracellular domain, the same α subunits are not in contact and instead are predominately mediated by GM3 (Figure 6.3a), similar to results shown in Chapter 5, Section 5.3.2.

The lipid clustering around the α subunit in the 3 $\times$ 3 system was identical to that of the α subunit in the mixed α - β3 system (described later in Section 6.3.3). Briefly, the α subunit exhibited a similar protein - lipid RDF as was seen for the CG β3 clustering in Chapter 5, Section 5.3.3. GM3 preferentially clusters around the top of each VSD, as well as some loop regions in the ECD, and seems to inhibit residues in the extracellular leaflet making direct contact. The anionic PIP_2 lipid binds exclusively to residues at the base of each VSD in the cytoplasmic leaflet, binding preferentially at VSDs II and IV.



(a) (b)



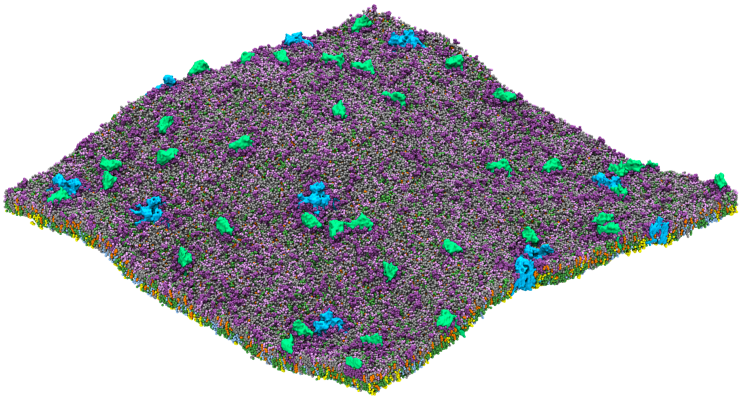
(c)

Figure 6.3: Protein - protein interactions in the 3 × 3 α subunit system. (a) Extracellular and (b) intracellular view of the 3 × 3 α subunit systems at 10 μs. Colours of lipids are the same as in Chapter 5, Figure 5.1. (c) Highly interacting residues between α subunits, averaged over all runs (viewed from the intracellular side). Residues that account for more than 1.5% of the total protein - protein interactions are shown as spheres.

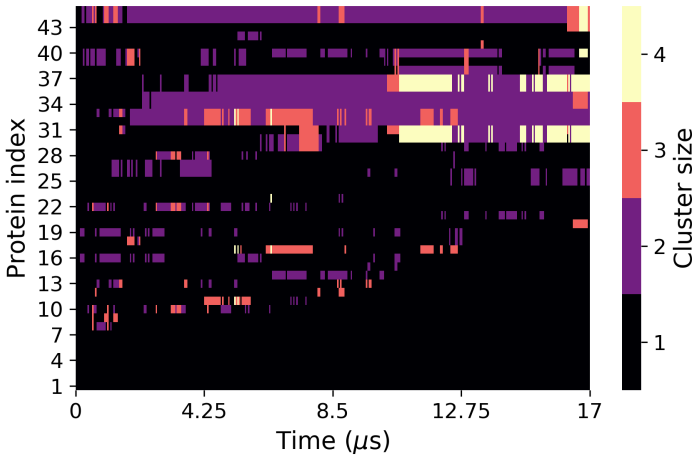
6.3.2 α - β3 subunit interactions in a large membrane system

The clustering of both the α and β3 subunits with their own protein copies in the large CG system was assessed using the cut-offs between subunits as described previously

(i.e. 11 nm and 4.5 nm for the α and β3 subunits respectively). A cluster of α and β3 was defined if the distance between both species' centres of geometry was $\leq$ 6 nm. After 17 μs of simulation, higher order clusters formed between the α and β3 subunits with the majority of clusters present as dimers with a small fraction of higher order oligomers. In addition to α - β3 clusters, several copies of the β3 subunit formed small clusters over the course of each simulation, similar to those seen in Chapter 5. A snapshot of the final frame of one 17 μs run is shown in Figure 6.4 together with the typical evolution of protein clusters.



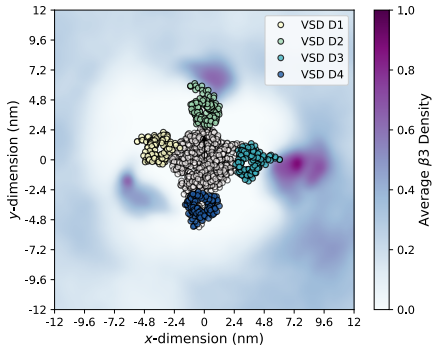
(a)



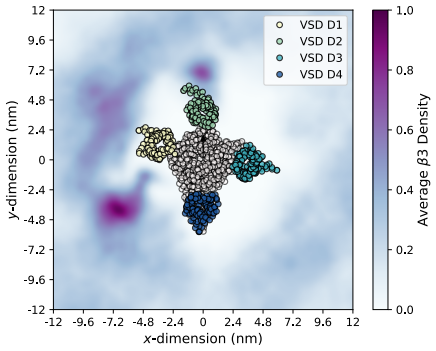
(b)

Figure 6.4: Evolution of α and β3 clusters over 17 μs. (a) Snapshot after 17 μs of simulation where clusters of α (cyan) and β3 subunits (green) are present in the mixed lipid membrane. (b) Evolution of cluster sizes over 17 μs.

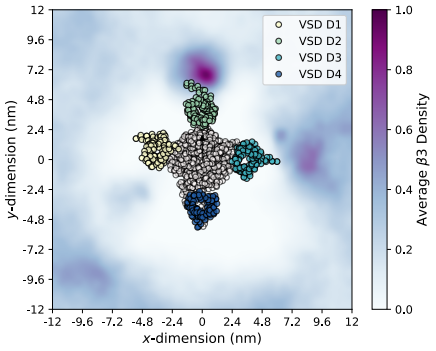
The density of β3 subunits surrounding the α subunit was analysed for every frame. For each of the nine α subunit copies in the system, the β3 density was recorded within a 12 nm radius where each α subunit copy was aligned to I717 (located in the S1 helix of VSDII). Generally there was a depletion of β3 density surrounding the S5-S6 pore-forming helices and instead all density was concentrated at the VSDs. Between the three runs performed there were variable sites of β3 density spread across each of the four VSDs. Typically, the density was found to be highest on VSDIII with lower densities at VSDI and VSDII. There was little to no density of VSDIV, known to be important in fast inactivation of the pore after depolarisation. Densities of β3 surrounding the α subunits in each run are shown in Figure 6.5.



(a) Run 1



(b) Run 2



(c) Run 3

Figure 6.5: Density of β3 around the α subunit. β3 density is averaged across all frames of a 17 μs run, each VSD is colour coded and indicated in the legend. All α subunit copies are aligned from the centre of mass of the collective VSDs to I717 located in S1 of VSDII (black arrow).

When comparing the protein binding sites on the α subunit to the density plots shown in Figure 6.5, there is some discrepancy in VSD preference. However, and as seen in Chapter 4, the β3 subunit is not a rigid protein and the large Ig domain undergoes significant conformational changes on the membrane. The size ratio between the Ig domain and the TMD (see Chapter 1, Figure 1.7c) causes a slight bias in the calculation of β3 density, with the Ig domain seemingly making up most of the observed interactions around the α subunit when viewed from the extracellular side. In fact, when analysing residues that exhibit high contact on both the α and β3 subunits, it becomes clear that most are concentrated on the TMD of both proteins, with contacts typically being hydrophobic in nature (Figure 6.6).

Inspection of the binding sites between the α and β3 subunits reveals a somewhat conflicting preference for VSDI and VSDII when compared to the binding site of β1 to VSDIII in the Shen *et al.* structure [5]. In the simulations performed, binding between subunits occurs predominantly on the S2 helix, the S2/S3 linker, and at the short linker region before the beginning of S1 helix, in agreement with the Shen *et al.* structure [5] structure, albeit not localised to VSDIII. The β3 subunit binds with both VSDI and VSDII preferentially, with fewer interactions present in VSDIII and VSDIV (Figure 6.6a, 6.7, & 6.8). Most β3 contacts are present on all domains, with both DI and DII showing the highest levels of contact at hydrophobic residues H184, V200, and T207 in DI, and C429, C430, P431, L432, and W433 in DII (Figures B.6). Conversely in DIII and DIV β3 interacts primarily between the linker regions of the S2-S3 linker in VSDIII and the S6-S1 linker between domains DIII and DIV. In β3, protein - protein contacts are concentrated in the TMD, again with the hydrophobic residues accounting for most of the α subunit interaction (Figure 6.6c). The conserved E176 residue located at the C-terminal end of the TMD accounts for approximately 14% interaction density with the α subunit, whilst other residues such as M177, C180, and Y181 exhibit lower protein - protein contact values. The Ig domain of β3 makes

contact with the extracellular loops of the α subunit through residues at the "tip" located between β strands B $\rightarrow$ C (E53) and F $\rightarrow$ G (E129, A130, and H131).

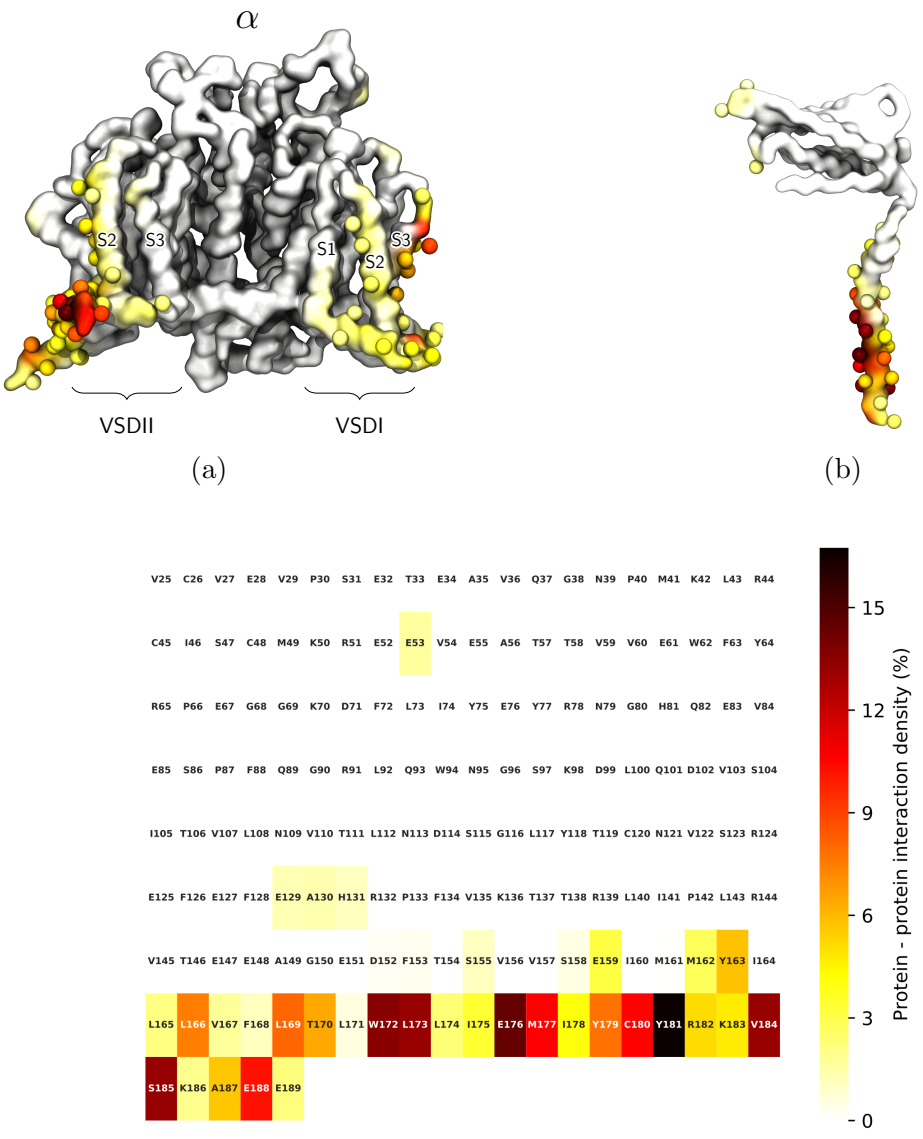
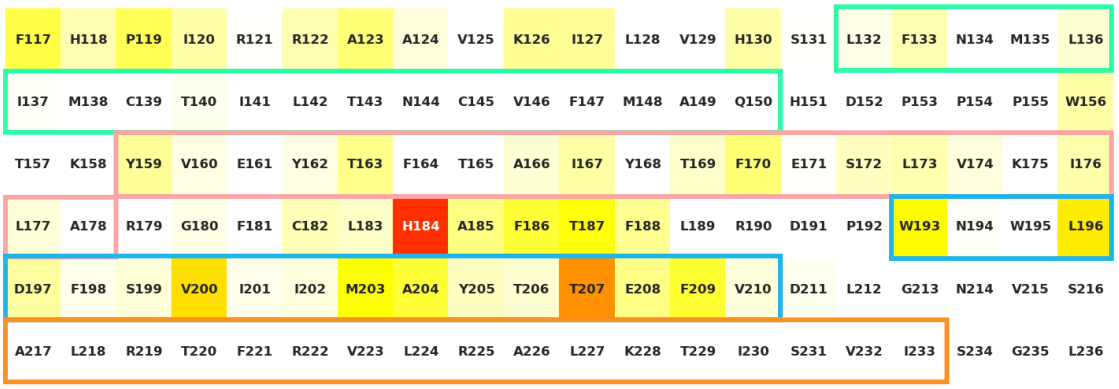
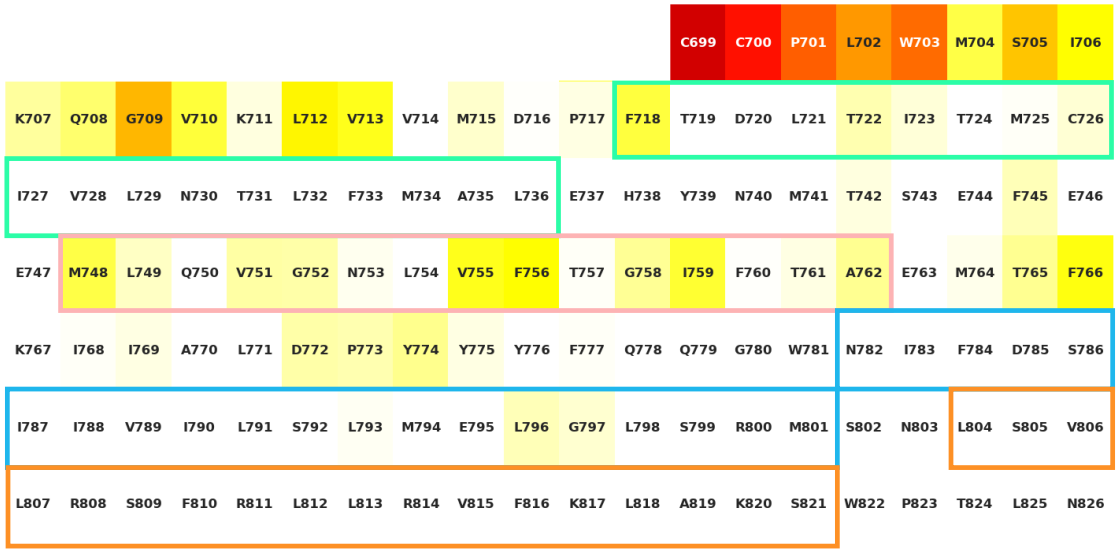


Figure 6.6: Protein - protein interactions between the α and β3 subunits. Regions of high contact of the β3 subunit on (a) the α subunit and (b) the β3 subunit. High levels of contact are located on the helices of VSDI and VSDII in the α subunit (in particular S2 as well as the pre-S1 and S2/S3 linker regions, in agreement with the Shen *et al.* structure [5] at VSDIII) and the TMD of β3. The first side chain particle of residues that account for more than 1.5% of the total protein - protein interactions are shown as spheres. (c) Heatmap of high contact regions on the β3 subunit. Regions of β3 contact on each domain of the α subunit are shown in Figures 6.7 & 6.8.



(a) DI



(b) DII

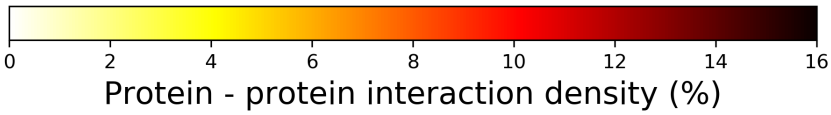


Figure 6.7: β3 contacts on DI & DII of the the Nav1.5 α subunit, the S1-4 helices in each domain are represented by boxes coloured green, pink, cyan, and orange respectively. Full length heatmap sequences of each domain can be found in Appendix B, Figure B.6.



(a) DIII



(b) DIV

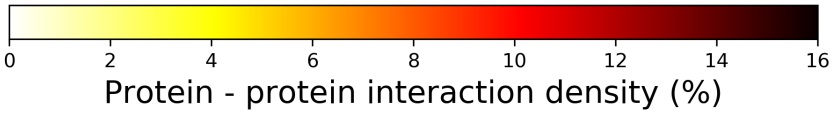


Figure 6.8: β3 contacts on DIII & DIV of the the Nav1.5 α subunit, the S1-4 helices in each domain are represented by boxes coloured green, pink, cyan, and orange respectively. Full length heatmap sequences of each domain can be found in Appendix B, Figure B.7.

6.3.3 Influence of lipid interactions on α - β3 clustering

In Chapter 5, the clustering of multiple β3 subunits was found to be dependent primarily upon GM3 and PIP_2, with each favouring binding to the Ig domain and base of the TMD ,respectively. The RDF distribution of lipids surrounding the α subunit (Figure 6.9) is similar to that of β3 with respect to the proximity of GM3, however there is a much more pronounced interaction with PIP_2 and cholesterol. As shown in Figure 6.10, there is significant GM3 clustering at the top of all VSDs in the α subunit with additional interactions within the loop regions close to the selectivity filters. In contrast to β3 lipid interactions (examined in Chapter 5, Section 5.3.3) the α subunit exhibits a slight increase in cholesterol interactions that tend to coat the S1-S6 helices in all domains.

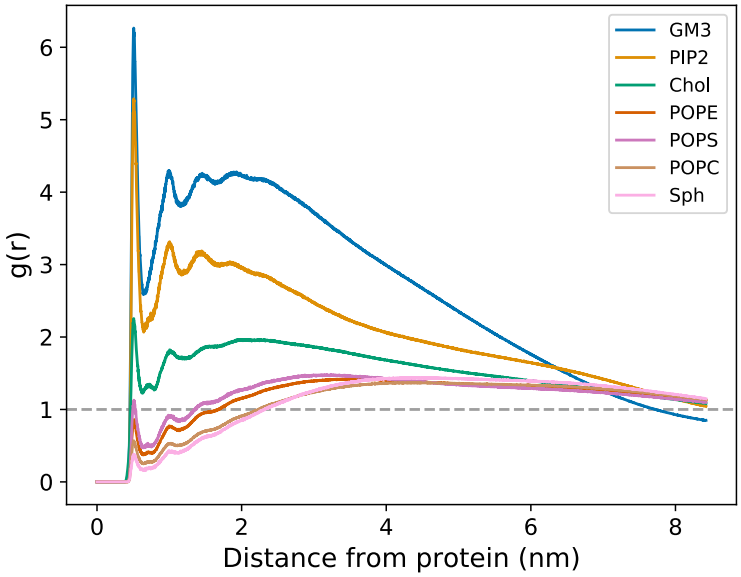


Figure 6.9: The RDF distribution of lipids surrounding the Nav1.5 α subunit in the α - β3 system. When compared to β3 (Chapter 5, Figure 5.5), there is a notable increase in PIP_2 and cholesterol proximity.

Figure 6.10: Protein - lipid interactions in the Nav1.5 α subunit. (a) Extracellular view of GM3 and (b) cholesterol interactions and (c) the intracellular view of PIP_2 interactions. (d), (e) & (f) show interactions of GM3, cholesterol, and PIP_2 respectively on the transmembrane domain of the α subunit. The first side chain particle of residues that account for more than 2% of the total protein - protein interactions are shown as spheres.

6.3.4 Alternative binding mode of the β3 subunit

Structures of the closely related β1 subunit have shown it to be bound preferentially at VSDIII, with the Ig domain interacting with extracellular loops [3–5]. Results from Chapters 4 & 5 indicate that the short linker region, and by extension the Ig domain, of β3 explores a variety of conformational states. To investigate whether poses similar to that of β1 in crystal structures were captured in the CG simulations, each cluster of α and β3 subunits was investigated. As seen in the previous section (and Chapter 5, Section 5.3.3), the clustering of both α and β3 subunits is heavily mediated by GM3 and to a lesser extent PIP_2 and cholesterol. This trend continues between clustered α and β3 subunits, with GM3 mediating contacts between both subunits.

Visual inspection was performed on the clusters of proteins in the final frame of each 17 μs repeat. Figure 6.11b shows examples of various binding modes of β3 to the α subunit. Typically, the Ig domain is orientated away from the body of the α subunit with the bulky GM3 headgroup present between both protein species. The Ig domain remains bound to the surface of the membrane with the same side (consisting of β sheets D, E, and B) that partially exposes N-terminus of the Ig domain away from the membrane plane (Figure 6.11b i). Binding modes in which more than one β3 subunit was associated with the VSDs of the α subunit were also present and display a similar orientation of the Ig domain positioned away from the α subunit body (Figure 6.11b ii). Additional higher order structures of β3 were also found, this time with an Ig orientation similar to that seen in the β1 bound crystal structure (Figure 6.11b iii).

(a) (b)

Figure 6.11: GM3 mediated clustering of α and β3 subunits. (a) Extracellular view of clustered α and β3 subunits in one run on the cell membrane (shown as a white surface), with GM3 headgroups (purple spheres) in close proximity to both protein species. (b) Snapshots of a binding events between both α and β3 subunits. The α and β3 subunits are shown in cyan and green respectively.

6.4 Discussion

6.4.1 Preferential binding of the β3 subunit

When in isolation, the α subunits do not exhibit high levels of oligomerisation. In the simulations performed here there was only slight interaction between protein copies in the intracellular leaflet, predominately at the base of each VSD. The interaction between VSDs is unsurprising since these regions encompass roughly the vertices of a square defined by the overall architecture of the channel when viewed from the extracellular (or intracellular) side. It is important to note that in the homology model created not all flexible loop regions are included, since these regions were not resolved in the cryo-EM structure. In particular the long intracellular loop between DI-DII is not present and, in addition to being heavily phosphorylated [290], has been shown to play an important role in $\alpha - \alpha$ subunit binding through interactions with cytoplasmic proteins such as 14-3-3 [100].

In contrast to tetrameric Kv channels, it is important to note that although homologous, all of the domains in Nav1.5 are not separate monomers. Instead, the α subunits of Nav channels are better described as pseudo symmetric. Although similar in amino acid sequence, the VSDs of DI-DIV are not identical which may lead to differences in protein - protein binding patterns. Simulations performed in this chapter seem to suggest that the β3 subunit binds preferentially to VSDI and VSDII, with a decreased interaction with the VSDIII and VSDIV. Analysis of the specific helices that β3 interacts with reveals a slight preference for S2 and S3 as well as the S2-S3 cytoplasmic loops in DI. In VSDII, β3 also interacts preferentially with the S2 helix with additional contacts at the short helical region before the S1 helix defined by residues C699 to V713.

By aligning the sequence of each VSD in Nav1.5 (Figure 6.12) it is clear that there is some level of conservation. However, focussing on the S2 helix in VSDI some of

the high contact residues are not well conserved across domains, such as Y159 and F170. Interestingly, when looking at areas in VSDI that make up very high regions of contact, such as H184 and T187, there is also little conservation. In VSDII, regions of high interaction seem to be concentrated at the short helical segment located just before the S1 helix and include residues C699, C700, P701, L702, W703, S705, and G709. Additionally, many hydrophobic contacts are made, similar to those seen in VSDI, within the S3 helix. Interaction with the short helical segment before S1 is also evident in VSDIII. Since VSDIII is the same binding site as seen in the Shen *et al.* [5] and Pan *et al.* [4] structures of bound β1 (located between SI and S2), it is somewhat reassuring that β3 binds in same place. Residues in VSDIII such as G1118, K1189, V1190, W1191, and W1192 make up high regions of contact with the β3 subunit. This pattern is again seen in VSDIV, albeit to a lesser extent, with residues P1513, L1514, N1515, K1516, and Y1517.

The binding events shown in Figure 6.11 are reminiscent of those seen between the β1 subunit and VSDIII in both the Shen *et al.* [5] and Pan *et al.* [4] structures. The simulations performed here capture binding events at other VSDs (namely VSDI and VSDII) that adopt similar poses in the TMD regions of both subunits and as seen in the afore mentioned structures. It may be that current experimentally resolved structures are capturing the most stable conformation of α - β subunit interaction at VSDIII.

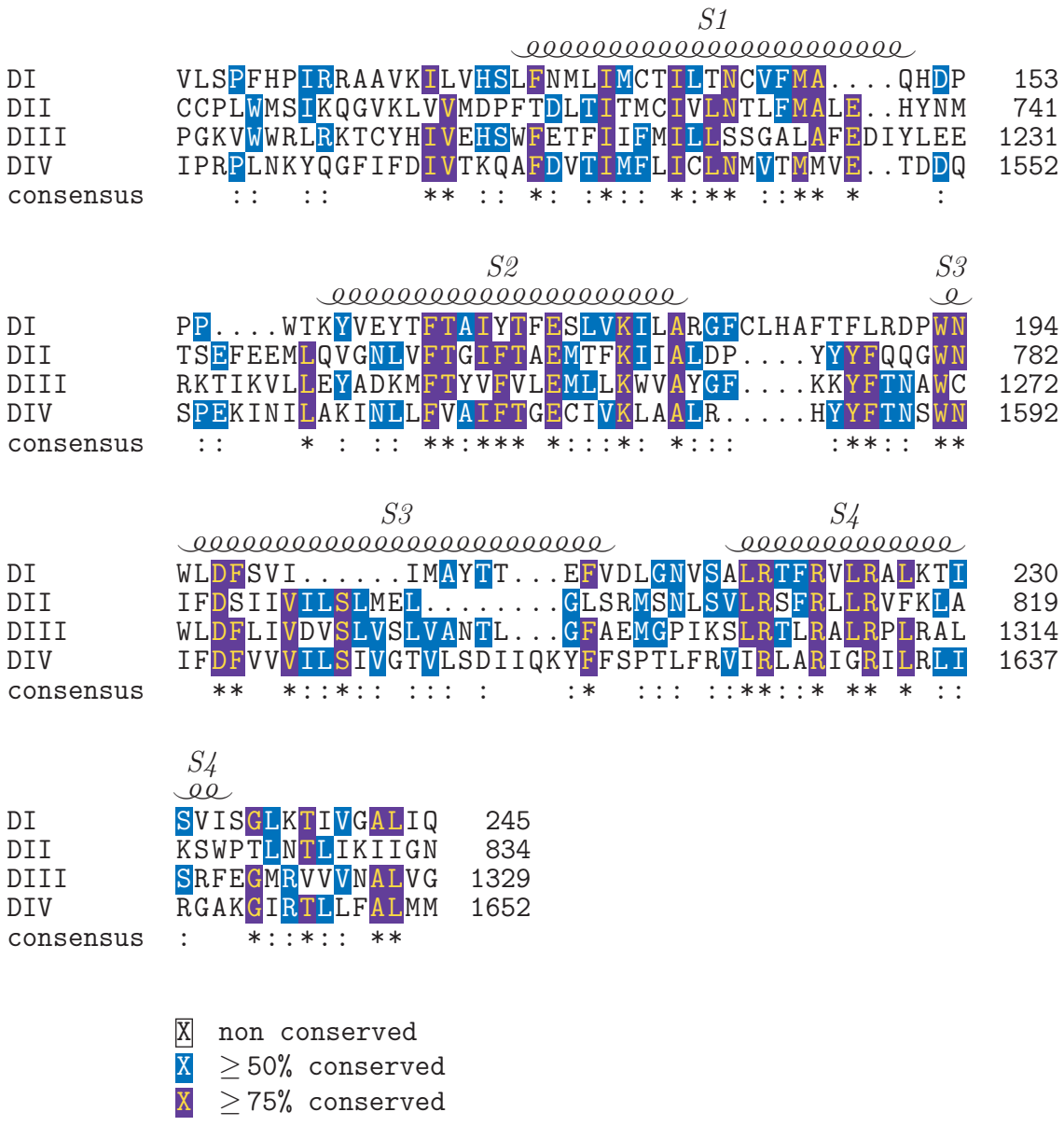


Figure 6.12: Alignment of VSDI-IV in the Nav1.5 α subunit. Approximate helical regions are indicated by spirals and labelled accordingly.

As described in Chapter 1, Nav1.5 contains four homologous domains that have been shown to bind multiple β3 subunits at once [99]. In order for this binding pattern to occur the β3 subunit must bind to multiple VSDs. This is in contrast to structures of β1, where the subunit tends to be positioned at VSDIII only. Somewhat conversely to experimentally resolved structures, results presented in this chapter

seem to suggest that the β3 subunit binds preferentially to VSDI and VSDII with low (but not zero) interaction on VSDIII and VSDIV. At this point an important question must be highlighted: should β1 and β3 be expected to interact with the α subunit in the same way or at the same location? Both subunits are homologous (with 51% sequence identity) and may be expected to bind to similar VSDs. The binding face of the β3 subunit seems to match well with the equivalent residues in the Shen *et al.* structure [5] (Figure 6.13), with most interactions being hydrophobic in nature. Interestingly, the conserved E176 residue exhibits high contact with the α subunit. However, it is important to note that E176 is aligned close to the intracellular side of the VSD and in fact interacts mostly with the short helical segment present before the S1 helix (analogous to the equivalent E177 in β1).

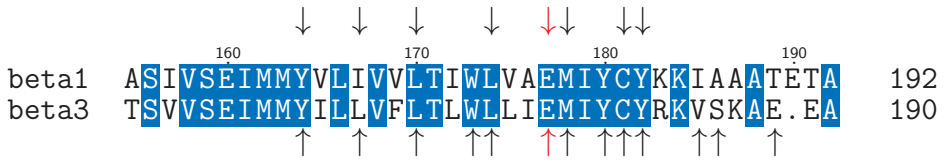


Figure 6.13: Alignment of the β1 and β3 TMD helical sequence. Residues present at the binding face of β1 to the α subunit in the Shen crystal structure [5] are indicated by down arrows. Residues in β3 exhibiting regions of high interaction ($\geq$ 6% interaction density) with the Nav1.5 α subunit in simulations are indicated by up arrows. E177 / E176 residues in β1 / β3 are highlighted in red. Residue numbering is shown for β1 and conserved residues are shown in blue.

Both β1 and β3 have been shown to be able to bind to Nav1.1 at the C-terminus of the α subunit [64], in close proximity to VSDIV. The location of this binding site is further supported by studies using chimeric α subunits that suggest β1 interaction with the extracellular loops of S5-S6 in DIV (as well as DI) [291, 292]. Results from fluorescence studies suggest that both subunits can be bound at once, with the β1 subunit bound between the cleft formed by VSDIII and VSDIV to modulate inactivation at VSDIV and β3 bound at the cleft between DII and DIII to influence channel

kinetics [233].

Altogether, the binding preference of both the β1 and β3 subunit is unclear. Recent crystal structures have revealed that VSDIII is the domain of choice for β1, however this is at odds with other experimental findings and hints at a more dynamic role of α and β subunit interaction. Although there does seem to be a slight preference for β3 binding to VSDI and VSDII (and the S2 helix therein), results presented in this chapter suggest that the binding of the β3 subunit is generally not restricted to any one VSD and in fact adopts similar binding poses at various sites. There also seems to be a preference for the Ig domain of β3 to be orientated in such a way as for it to extend away from the α subunit. In this outward facing position, it may be possible for the Ig domain to overcome the favourable interaction with the GM3 headgroup and orientate the long axis of its Ig domain to be almost perpendicular to the membrane plane (as seen in Chapter 5). In this position, it could either interact with the extracellular loops of the α subunit or outwardly, with another β3 subunit bound in a similar fashion to a neighbouring α subunit. Binding as the latter could be akin to interactions seen in the trimeric β3 crystal structure resolved by Namadurai *et al.* [99].

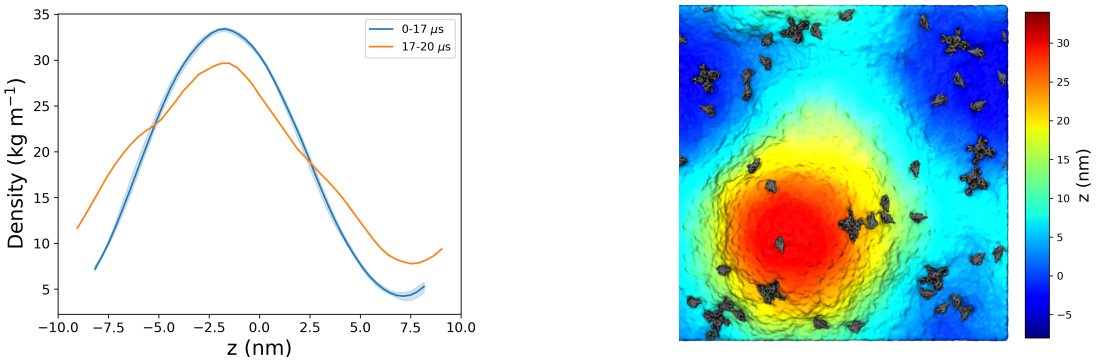
In all simulations performed here, such cross-linking events are rare and a closer look at the simulation setup is required. Although the constructed system studied is large and made up of $\cong$ 2 million CG particles, there is still an issue of sufficient sampling. Each run performed resulted in a cumulative simulation time of $\geq$ 30 μs, with a low number of clusters formed. This is primarily due to the main barrier of cluster formation being a combined result of initial protein - protein separation and sufficiently long simulation time, both of which are non-trivial to address. It is also important to note that interactions not included in the systems studied here include those with the cytoskeleton, through proteins such as ankyrin-G which are known to play an important role in Nav localisation [293].

6.4.2 Influence of the lipid membrane

Results in this chapter, as well as those seen in Chapter 5, show that the surrounding membrane model used in this study influences protein - protein interaction between both the α and β3 subunits. Overall, there is a tendency for GM3 to mediate clustering with smaller contributions present from PIP_2. The clustering of GM3 around the extracellular side of each VSD is intriguing since, as speculated in Chapter 5, the head group of GM3 is negatively charged at physiological pH and could potentially influence the local electric field in each VSD. Indeed, several experimental [294–297] and computational [298–300] studies have examined the influence that the surrounding lipid bilayer may have on VSD dynamics. Compared to β3, the α subunit exhibits a greater degree of interaction with PIP_2 that is located on the cytoplasmic side of each VSD. Since Nav channels are insensitive to PIP_2 binding, these results suggest that it is most likely due to the protein's relative size rather than any intrinsic functional role, as found in other ion channels [301]. In contrast to the β3 subunit, there is increased interaction between the α subunit and cholesterol, with preferential binding to the transmembrane domains as well as penetration to the pore-forming S5 and S6 helices in DI-IV. Although lipid interactions within the plasma membrane are complex [302], it seems tentative to suggest that the combined interaction of both gangliosides and cholesterol with Nav1.5 may indicate in the formation of nanodomains, both of which have been found *in vivo* [303] as well as in model membranes [304]. Interestingly, cholesterol is well known to make up large constituents of caveolae [305] (a type of lipid nanodomain that forms invaginations in the cell membrane) that cluster ion channels important for cardiomyocyte function [306].

Despite the membrane in the mixed $\alpha - \beta$3 subunit system is large, the protein density is sparse in comparison to Nav channel clusters *in vivo*, such as at the nodes of Ranvier where protein particle density can be $\approxeq$ 1000-2000 μm^{-2} [16, 307]. Recent simulations of large CG systems have highlighted the importance of protein density

in the evolution of cluster size, protein diffusion, membrane stiffness, and membrane undulation [10, 181, 183, 267, 308]. Such studies reveal that membranes with low protein populations exhibit large undulations and highlight that, although complex lipid compositions have increased our understanding of lipid dynamics [266, 309], the effect of protein crowding and membrane biophysics remains to be fully characterised and continues to be an active area of research [267, 310, 311]. In light of this, simulations performed of the $\approxeq$ 2 million particle CG system in this chapter reveal that over the course of each 17 μs trajectory the plasma membrane remains relatively flat with only slight undulation (Figure 6.14a). However, extension of one run to 20 μs reveals more dynamic motion with areas of high and low distortion relative to the starting membrane plane (Figure 6.14b).



(a) (b)

Figure 6.14: Membrane distortion in the α - β3 system. (a) Membrane density of phosphate atoms with respect to the z-axis, averages taken over 0-17 μs for all runs. Additionally, one run was extended to a total of 20 μs where the membrane exhibits increased undulation. (b) Extracellular view of the membrane at 20 μs with regions of high and low distortion indicated.

The increased undulation seen in the 20 μs simulation suggests the system constructed may not be crowded enough, coupled with the complete lack of cytoskeleton interactions as mentioned previously. However, without sufficient evidence from experiment it is difficult to conclude whether the undulations seen in such large simula-

tions truly reflect what is present *in vivo*. Future advances in imaging techniques may help to elucidate the distortion of a variety of membranes, for example techniques such as 3D electron tomography that have recently characterised complex Golgi structures in neurons [312]. Continuing research into the complex relationship between proteins, lipids, and membrane curvature [313] may help to improve *in silico* studies in the future at longer time and larger length scales.

6.5 Conclusions

In this chapter the clustering of the Nav1.5 α subunits has revealed that protein - protein contacts are centralised to the extremities of the cytoplasmic side of each VSD. There is a clear lipid interaction, with GM3 and PIP_2 binding most favourably in the extracellular and intracellular leaflets respectively, in addition to unspecific binding of cholesterol across the TMD. This trend continues in large-scale systems consisting of around 2 million CG particles, where clusters of both α and β3 subunits exhibit multiple binding modes, the most prevalent of which include the Ig domain orientated away from the α subunit body. GM3 tends to coat both subunits in the extracellular leaflet and seemingly stabilises a "rolled" conformation of the Ig domain on the membrane surface (as found in Chapter 5). With the limitation of sufficient sampling, the exact nature of β3-mediated binding between α subunits remains to be fully characterised. However, the flexibility of the β3 Ig domain, as found in Chapter 4, does indicate that transition to an "up" state from a "rolled" state is not insurmountable and highlights the complex interplay between protein behaviour and membrane composition. Finally, the results presented in this chapter serve as a starting point for investigations into α and β3 subunit interactions in realistic membrane compositions.

"We can only see a short distance ahead, but we can see plenty there that needs to be done."

— Alan Turing

7 Conclusions

7.1 Summary

Prior to the work conducted in this thesis, MD simulations of Nav β subunits were limited to those in complex with the pore-forming α subunit [314]. For the first time, an in-depth study of non-covalently bound β subunits has been conducted at both atomistic and CG levels with key differences between the organisation on, and interactions with, the membrane highlighted. The use of large-scale CG MD simulations has proposed likely binding patterns between Nav subunits that are in agreement with experimental results. Incorporation of both atomistic and CG descriptions of

Nav channel systems has facilitated the comprehensive analysis of protein dynamics and the effects of particular lipid species on observed protein clustering.

In Chapter 3, the importance of the S4 helix in channel activation was examined in the context of a Nav1.5e splice variant. Atomistic simulations characterised the dynamics of two single point mutations located in the VSDI S3-S4 linker. Increased steric strain, electrostatic interactions with the surrounding lipid headgroups, and increased serine χ_1 angle sampling in the VSD were shown to be the most likely causes for the shift in channel activation seen experimentally. However, it must be noted that there were significant differences in observed dynamics of both mutants across individual runs. One explanation is that, although long by modern standards, the performed simulations have insufficiently sampled the phase space of the system. This is compounded by the high flexibility of the S3-S4 linker that will more than likely adopt a large variety of conformations. There is also the need for higher quality VGIC structures, such as Nav1.5, from which to construct homology models and increase the confidence in results obtained from MD simulations.

In Chapter 4, atomistic simulations were performed again to analyse the orientational preference of the $\beta 1$ and $\beta 3$ subunits' Ig domain and determine the overall dynamics of each subunit in isolation of the α subunit. To achieve this, an algorithm based on the Tait-Bryan angle formalism was developed. Results showed a variety of pitch angles sampled by the $\beta 3$ Ig domain and examined the tentative role glycosylation plays in stabilisation of an Ig conformation perpendicular to the membrane plane. Analogous measurement in the related $\beta 1$ subunit revealed a large preference for the Ig domain to stay bound and parallel to the membrane, with the non-conserved R96 residue identified as important for binding with phosphate lipid headgroups[†]. Construction of a trimeric $\beta 3$ model suggests that a conserved glutamic acid residue, previously thought to stabilise the observed trimer crystal structure,

[†]After the peer review process for the publication by Glass *et al.* [231], results presented in this section have been revised. The reader is referred to the publication for the updated results.

may be more likely to assist in TMD orientation within the lipid membrane. Additionally, the instability of the intracellular domain of both subunits calls into question the secondary structure assignment in the corresponding cryo-EM map from which the models were built.

In both Chapters 5 and 6, CG systems of the α and $\beta 3$ subunits were constructed, as well as a large scale system comprising of a mixture of both. The longer time and length scales granted by this CG approach allowed analysis of higher order clusters between subunits to be obtained at the residue level. In all CG runs, the lipid composition had a profound effect on Ig orientation, with three typical conformations adopted, in addition to influencing the clustering of subunits. Clusters of $\beta 3$ subunits were found to be mediated strongly by GM3 lipids present in the generalised mammalian membrane, causing the Ig domain of $\beta 3$ to orientate itself and present one "face" bound to the membrane surface. A variety of $\beta 3$ conformations were captured with a few orientated in such a way as to promote oligomerisation between the N-terminus of neighbouring subunits. In large-scale simulations with both α and $\beta 3$ subunits present, clusters formed in a variety of α:$\beta 3$ ratios. Ig domains were seen to bind in a number of different orientations, both in agreement with α subunit bound $\beta 1$ crystal structure pose, as well as orientated in such a way as to potentially promote cross-channel binding.

Results presented in this thesis highlight the importance of considering ion channel function as a composite of both the pore-forming and various auxiliary subunits. Historically, experiment and MD simulations have typically focussed on the passing of charged ions through the central pore of VGICs. However, and as demonstrated here, the lipid composition and presence of different protein species can have profound effects on protein conformation and, potentially, channel behaviour. The influence of subunits on gating and spatio-temporal dynamics are now drawing increased atten-

tion, with the structure of many ion channels in complex with auxiliary subunits now being solved (discussed further in Section 7.2). The dynamics of individual subunits in the absence of their canonical pore-forming subunit remains an open question in the field. It would be interesting to see whether a general pattern of subunits containing an Ig-like CAM domain, such as glialCAM [315, 316] (see Section 7.2.4), behave in a similar fashion as Nav β subunits presented here. It is tentative to suggest that the CAM domain in such subunits may be necessary and sufficient to induce channel clustering due their inherent adhesive properties.

Work here focussed on the cardiac Nav1.5 isoform. In contrast to the well known clustering of Nav channels in neurons [286], clustering of Nav1.5 in cardiomyocytes is somewhat poorly understood. If clusters of Nav1.5 are present in cardiomyocytes, the likelihood of a successful AP being generated will be increased and ultimately influence the behaviour of heart contraction. Additionally, and as outlined in Chapter 1, different Nav channel isoforms are present in a number of other excitable cells. It would not seem unreasonable to suggest that due to the β subunits' almost ubiquitous expression with all Nav channels, similar binding patterns may be present in a variety of different cell types.

Another interesting question remains as to how different VGICs interact and cluster together, such as Nav1.5 and K_{ir} that are implicated in cardiac excitability and arrhythmogenesis [317, 318]. Overall, studies into VGIC subunit behaviour and the binding patterns of higher order VGIC clusters may help to inform future drug developments away from the pore.

7.2 Future directions

The use of a multiscale MD approach to investigate protein oligomerisation, as presented here, has shown that such methodologies can be a used to probe the behaviour of protein - protein interactions in realistic model membrane compositions. However, looking to the future there are a number of avenues that are worth pursuing. The following sections cover topics pertaining to both results presented in this thesis as well as to the field more generally.

7.2.1 Protein models

A protein model is only as good as the determined structure from which it is constructed. In order to accurately describe protein motion via MD simulations, high resolution structures are required. For example, in Chapter 4 the C-terminus of both the $\beta 1$ and $\beta 3$ subunits underwent significant bending and when compared back to the cryo-EM map, these regions are of lower overall quality when compared to the rest of the protein complex. Traditional X-ray crystallography requires the protein of interest to be crystallised, however this is often difficult to achieve with integral membrane proteins [319]. Thankfully, recent advancements in cryo-EM have circumnavigated this problem and allow the determination of near-atomic resolution protein structures close to their native conformation [320]. The field is moving rapidly with higher resolution structures being deposited each year, with almost 4000 structures available in the PDB as of 2019 [192]. This is also coupled with simulation software now beginning to incorporate model refinement via cryo-EM maps [321]. Cryo-EM is beginning to elucidate protein complexes in their native environment and seems to be only hampered by advancements in sample preparation and data acquisition [322]. To better ascertain the likely binding mode between Nav subunits, the use of cryo-EM in determining more $\alpha - \beta$ complexes at ever higher resolution will be indispensable in

validating results from simulation and advancing our understanding of higher order Nav oligomerisation.

7.2.2 The influence of glycosylation

In addition to high resolution structures, there is an increasing need to derive the correct glycosylation patterns on glycoproteins. In Nav β subunits, the Ig domain contains four predicted glycosylation sites and in Chapter 4 the effects of glycosylation on the β3 Ig domain were studied. Simulation results showed that there was a pronounced effect on Ig domain orientation, with only a single GlcNAc moiety modelled at each site. In MD simulations the glycosylation sites are often ignored due to their large degree of flexibility and subsequent cleavage to facilitate the crystallisation process. This is in addition to the variable chain lengths, complex ring-puckering energy landscape [251], and often intricate branching patterns found in a number of glycoproteins. The lack of structural glycoprotein data makes accurate modelling of glycosylation sites difficult to validate and as a result the development of accurate carbohydrate force fields has progressed slowly in comparison to other biomolecules. Nevertheless, progress has been made in force fields such as GLYCAM and GROMOS [238, 323, 324] that offer a large variety of carbohydrates, coupled with user-friendly software such as the DoGlycans package [237] and CHARMM GUI [325] for glycoprotein construction. In this work, investigation of the glycosylation sites in the β1 subunit were unable to be performed. In future work it would be interesting to compare results to that of the β3 subunit. Additionally for both subunits, determination of a higher order glycosylation pattern, if any, would assist our understanding of how the Ig domain is orientated on the PM.

Such glycosylation patterns can often become increasingly large in scale and as such CG models may be more appropriate. Fortunately, the MARTINI force field has also developed force field parameters to extend the model to carbohydrates [165].

However, care must be taken when interpreting results from such models. Although the loss of resolution allows for the feasible simulation of large biomolecular systems, details such as anomer type, access of differing ring pucker conformations, and hydrogen bonding patterns are lost. Furthermore, inclusion of ring puckering parameters in CG systems, such as herapan sulphate chains, has been shown to reproduce key macroscopic hydrodynamic properties [326]. Due to these complexities, CG glycosylation was not explored in the simulations performed. As highlighted in Chapter 5, comparison between the interaction of glycosylated Ig domains and the large sugar headgroups of GM3 with results from atomistic simulations would be interesting to address in the future.

Accurate modelling of carbohydrates remains one of the biggest challenges facing the field of biomolecular simulation. In the spirit of developing a more accurate description of protein dynamics, the use of glycoprotein models will become increasingly relevant and require novel advances in experimental techniques [327, 328].

7.2.3 The MARTINI model and increasing membrane complexity

The addition of glycosylation sites is just one of the many ways in which the complexity of current biomolecular simulations can be increased to mimic environments close to those experienced *in vivo*. The complex and large scale biomolecular simulations performed in this work would not have been possible without the use of the MARTINI force field and the ability to model a mixed lipid membrane in particular. As highlighted in Chapter 6, a number of biomolecular interactions are omitted and typically higher order organisation of proteins within the membrane are dependent on a variety of factors such as lipid-mediated organisation, protein crowding, and cytoskeletal interactions. Work presented here focusses primarily on lipid-mediated protein organisation, however interactions with other membrane proteins, the cy-

toskeleton, and intracellular / extracellular proteins may play an important role in Nav channel organisation. There has been substantial progress in modelling the complexity of cell membranes in recent years, with development still ongoing [329]. It will be interesting to see how increasing model complexity influences the spatio-temporal dynamics of not only Nav channels, but VGICs more generally in the future.

7.2.4 Additional auxiliary subunits

Auxiliary subunits, such as β1-4 subunits, are not unique to Nav channels and in fact a number of ion channels that play a central role in the generation and propagation of the action potential are often found in association with auxiliary subunits *in vivo*. For example, in Cav channels transmembrane γ subunits disrupt Ca^{2+} influx to the cell through decreased channel density, increasing channel inactivation, and delayed recovery from inactivation [330]. In AMPA receptors (post-synaptic membrane proteins, critical in memory and learning), a subfamily of the Cav γ subunit named transmembrane AMPA receptor regulatory proteins (TARPs) influence channel gating [331] and are typically found in complex with the pore-forming dimer-of-dimers [332, 333]. In these receptors, TARP γ_2 (known commonly as stargazin) acts as a positive allosteric modulator in a variety of ways such as decelerating deactivation, accelerating recovery from desensitization, and boosting partial agonist efficacy [334, 335]. Interestingly, stargazin contains a conserved glutamic acid residue in its TMD, reflecting a similar conservation as found in Nav β subunits. The site of AMPA interaction is thought to occur in the extracellular loop regions of TARPs [336], however the potential contribution of the conserved glutamic acid residue in subunit interaction remains unclear. Additional auxiliary subunits in voltage-gated chloride (CLC) and various K^+ channels have also been shown to interact with auxiliary subunits. For example, CLC-2 channels are modulated by glialCAM subunits, single pass transmembrane proteins with an extracellular Ig domain [315, 316, 337, 338]. Single-pass transmem-

brane KCNE subunits interact with multiple Kv channels [339, 340], with a number of gene variants being involved in cardiac arrhythmias [340, 341]. In BK channels, β and γ subunits, both double-pass transmembrane proteins with large extracellular loops, alter various aspects of gating dynamics [342, 343].

These interactions are far from an exhaustive list and highlight that we are only just beginning to characterise large number of subunit interactions in ion channels. Unsurprisingly, given their involvement in a large number of cellular processes there is increasing evidence for ion channel auxiliary subunit roles in different diseases [8, 341, 344, 345]. A better understanding of how these subunits interact with their respective pore-forming subunit may help shed light on such diseases, as well as in a wider context for the treatment of various channelopathies.

7.3 Final remarks

Developments in both force fields and computational architectures are beginning to close the gap between experiment and simulation. In this thesis, the utilisation of both atomistic and CG methodologies, in a multiscale approach, have provided a starting point in our description of Nav subunit dynamics on the cell membrane. Our current understanding of protein - protein and protein - lipid interactions in the context of ion channel oligomerisation is an active area of research, one in which MD simulations will play a crucial role in the coming years.

Appendices

A

Mathematical background

A.1 Hamiltonian mechanics

An alternative formulation of classical mechanics was developed by W.R Hamilton, where the total energy of a system can be described by the Hamiltonian operator, H (or Hamiltonian for brevity).

$$H = T + V \tag{A.1}$$

Here, T and V represent the kinetic and potential energy of the system respectively. In contrast to Newtonian physics, the Hamiltonian provides a more general description of the system and instead works with variables $\mathbf{q}$ and $\mathbf{p}$ that define generalisable coordinates and momenta respectively. Since momentum is defined as $\mathbf{p} = m\mathbf{v}$ Equation A.1 may be rewritten as:

$$H(\mathbf{q}, \mathbf{p}) = T(\mathbf{p}) + V(\mathbf{q}) \tag{A.2}$$

$$= \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + V(\mathbf{q}_i) \tag{A.3}$$

In this reformulation of Newtonian physics the equations of motion now take the form:

$$\dot{\mathbf{q}}_i = \frac{\partial H}{\partial \mathbf{p}_i} \tag{A.4}$$

$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{q}_i} \tag{A.5}$$

Where $\dot{\mathbf{q}}_i$ and $\dot{\mathbf{p}}_i$ are the derivatives of the generalised coordinates and momenta respectively for particle i. It can also be shown that this formulation does conserve

energy of the system (i.e. such that $H = E$, and so $\frac{dH}{dt} = 0$).

$$\frac{d(H(\mathbf{q}(t), \mathbf{p}(t))}{dt} = \sum_{i=1}^{N} \left(\frac{\partial H}{\partial \mathbf{q}_i} \dot{\mathbf{q}}_i + \frac{\partial H}{\partial \mathbf{p}_i} \dot{\mathbf{p}}_i \right) \tag{A.6}$$

$$= \sum_{i=1}^{N} \left(\underbrace{\frac{\partial H}{\partial \mathbf{q}_i} \frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial H}{\partial \mathbf{p}_i} \frac{\partial H}{\partial \mathbf{q}_i}}_{=0} \right) \tag{A.7}$$

$$= 0 \tag{A.8}$$

Where in step A.6 $\rightarrow$ A.7 we have used the definitions in equations A.4 & A.5 and the chain rule.

A.2 The Verlet algorithm

The Verlet (and related) algorithm(s) are one of the most simple and powerful integrators used in MD simulations. The following derivation is based on the derivation outlined by Frenkel & Smit [346].

For reference the Taylor expansion is expressed as:

$$f(x) = \sum_{i=0}^{\infty} \frac{f^{(i)}(c)}{i!} (x - c)^i \tag{A.9}$$

The Taylor expansion of a particle coordinates ($\mathbf{r}$) at time t can be expressed as:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + f'(t)\Delta t + \frac{f''(t)}{2}\Delta t^2 + \frac{\Delta t^3}{3!}\dddot{\mathbf{r}} + \mathcal{O}(t^4) \tag{A.10}$$

similarly the choice of a negative time step, $-\Delta t$ gives:

$$\mathbf{r}(t - \Delta t) = \mathbf{r}(t) - f'(t)\Delta t + \frac{f''(t)}{2}\Delta t^2 - \frac{\Delta t^3}{3!}\dddot{\mathbf{r}} + \mathcal{O}(t^4) \tag{A.11}$$

Noting that $f'(t) = \mathbf{v}(t)$ and $f''(t) = \mathbf{a}(t)$, where $\mathbf{v}(t)$ and $\mathbf{a}(t)$ are velocity and acceleration respectively. Addition of equations A.10 and A.11 gives:

$$\mathbf{r}(t + \Delta t) + \mathbf{r}(t - \Delta t) = 2\mathbf{r}(t) + \frac{1}{2}\mathbf{a}(t)\Delta t^2 + \mathcal{O}(t^4) \tag{A.12}$$

$$\mathbf{r}(t + \Delta t) \approx 2\mathbf{r}(t) - \mathbf{r}(t - \Delta t) + \frac{1}{2}\mathbf{a}(t)\Delta t^2 \tag{A.13}$$

Reformulation of equation A.12 yields the velocity, so long as the history of the trajectory is known:

$$\mathbf{v}(t) = \frac{\mathbf{r}(t + \Delta t) - \mathbf{r}(t - \Delta t)}{2\Delta t} + \mathcal{O}(t^2) \tag{A.14}$$

A.3 Tait-Bryan angle formalism

The following brief description of Euler angles and the Tait-Bryan angle formalism is taken from H. Berendsen [347].

Rotational motion in rigid bodies is typically described by Euler angles. These angles (ϕ, θ, ψ) take the values through which a reference basis (X, Y, Z) is rotated to obtain a new basis named X′, Y′, and Z′ (Figure A.1). To map between the two coordinate systems a rotation matrix $\mathbf{R}$ is defined that transforms the coordinates of a particle, relative to its centre of mass (CoM), from the space-fixed → body-fixed system (Equation A.15) or the body-fixed → space-fixed (Equation A.16). The latter was used in the simulation methods described in Chapter 4, section 4.2.3.

$$\mathbf{r}_i = \mathbf{r}_{CoM} + \mathbf{R}^\top \mathbf{r_i}' \tag{A.15}$$

$$\mathbf{r}' = \mathbf{R}(\mathbf{r}_i - \mathbf{r}_{CoM}) \tag{A.16}$$

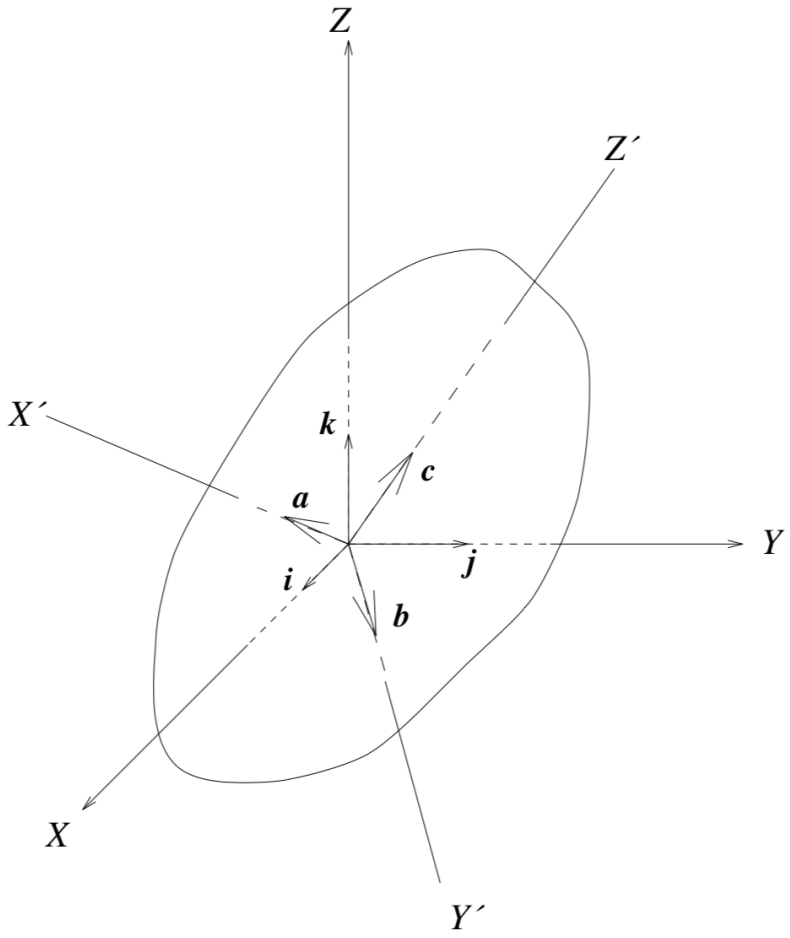
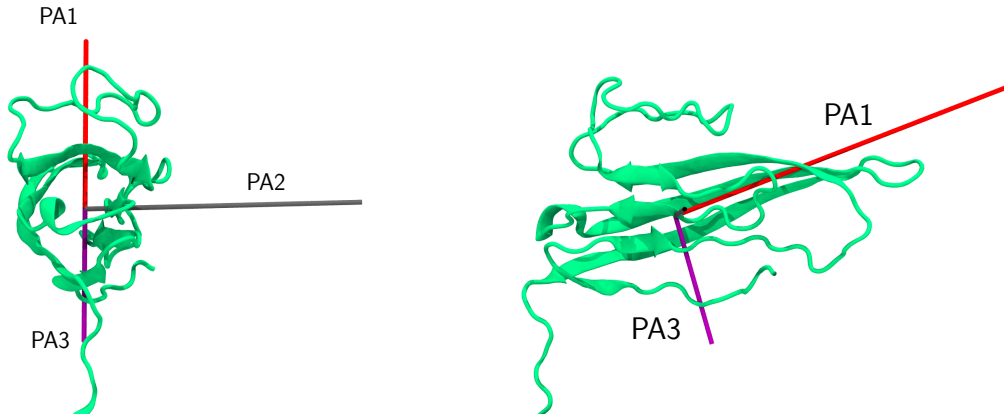


Figure A.1: A rigid body with a body-fixed coordinate system X′ Y′ Z′ and base vectors **a**, **b**, **c**, rotated with respect to the space-fixed coordinate system X Y Z with base vectors **i**,**j**, **k**. Image taken from H. Berendsen [347].

The PAs, calculated at each frame of a simulation within the Ig domain of the β subunit (Figure A.2), act as the basis of the body-fixed system whilst the unit vectors $\mathbf{e}_x$, $\mathbf{e}_y$, $\mathbf{e}_z$ act as the space-fixed system (where $\mathbf{e}_x = [1, 0, 0]$, $\mathbf{e}_y = [0, 1, 0]$ and $\mathbf{e}_z = [0, 0, 1]$).



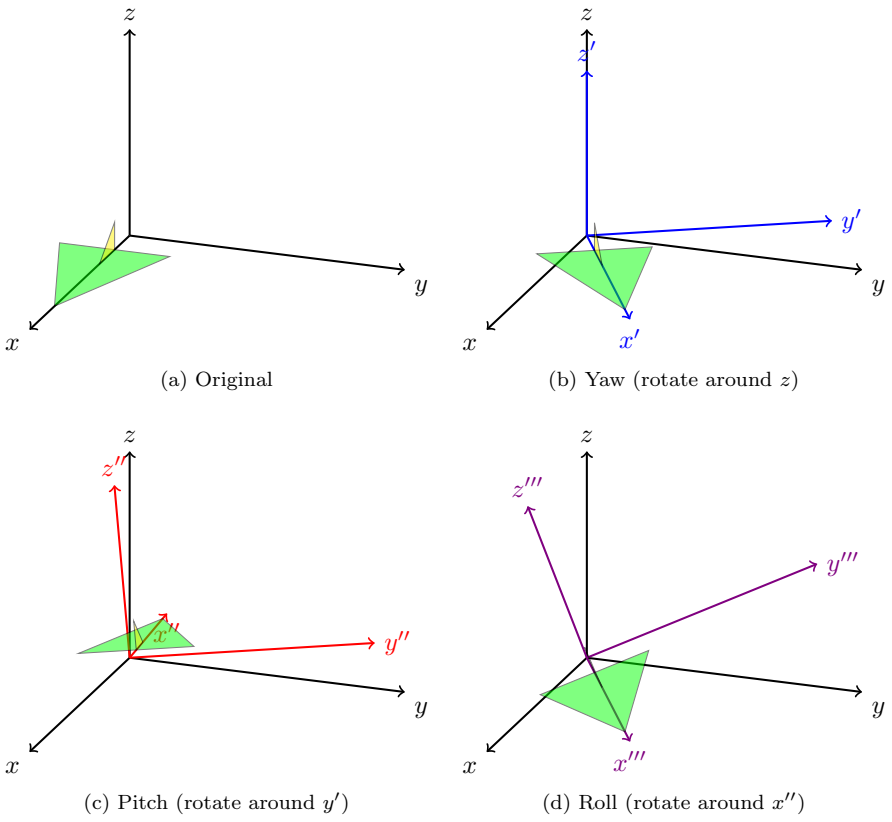
(a) (b)

Figure A.2: Principal axes calculated within the β subunit Ig domain. (a) End on view of the Ig domain with all three PAs visible. (b) A side on view of the Ig domain with PA1 and PA2 visible, PA2 is perpendicular to the plane formed by PA1 and PA3.

The rows of $\mathbf{R}$ (or columns of $\mathbf{R}^\top$) are the direction cosines between these axes, given by:

$$\mathbf{R} = \begin{pmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \\ c_x & c_y & c_z \end{pmatrix} = \begin{pmatrix} \mathbf{a}\cdot\mathbf{i} & \mathbf{a}\cdot\mathbf{j} & \mathbf{a}\cdot\mathbf{k} \\ \mathbf{b}\cdot\mathbf{i} & \mathbf{b}\cdot\mathbf{j} & \mathbf{b}\cdot\mathbf{k} \\ \mathbf{c}\cdot\mathbf{i} & \mathbf{c}\cdot\mathbf{j} & \mathbf{c}\cdot\mathbf{k} \end{pmatrix} \tag{A.17}$$

The Tait-Bryan angles are related to Euler angles. The key difference is that the rotations performed are done so using three distinct coordinate axes, that is, with each rotation the subsequent rotation is taken from the previous basis, shown visually in Figure A.3.



(a) Original (b) Yaw (rotate around z)

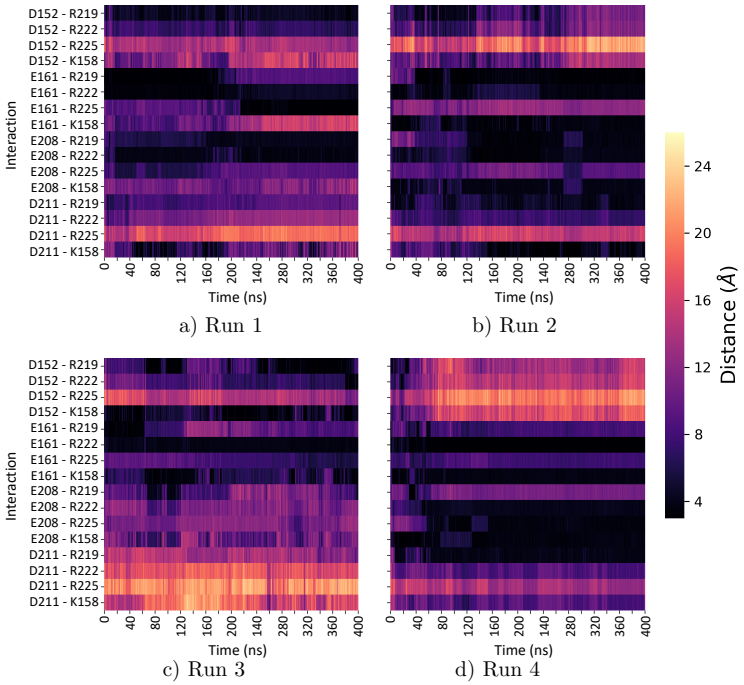
(c) Pitch (rotate around y') (d) Roll (rotate around x'')

Figure A.3: The general sequence of a rotation adopting the Tait-Bryan angle formalism. Often found in aerospace applications, x is taken as the pitch (i.e. of an aircraft) with the other angles commonly referred to as pitch and roll. (a) The starting frame with the long axis of the rigid body (indicated as toy aircraft model) aligned to x. (b) Rotation around z. (c) Rotation around y'. (d) Rotation around y''.

B

Supplementary material

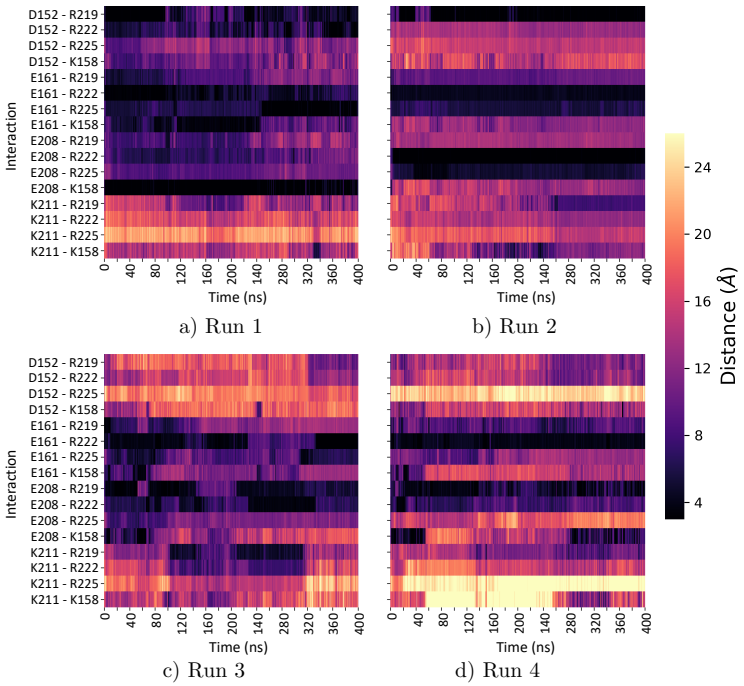
B.1 Supplementary material for Chapter 3



a) Run 1 b) Run 2

c) Run 3 d) Run 4

Figure B.1: Salt-bridge interactions present in VSDI of the T207S system.



a) Run 1 b) Run 2

c) Run 3 d) Run 4

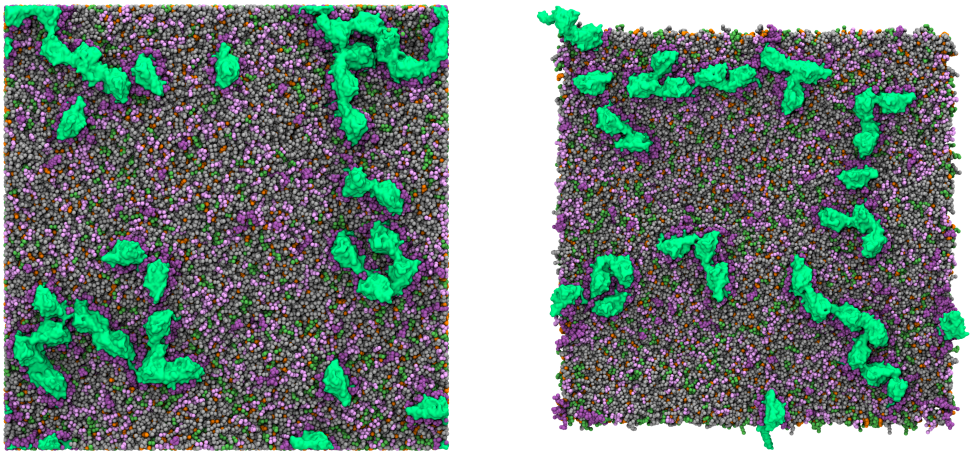
Figure B.2: Salt-bridge interactions present in VSDI of the double D211K - T207S system.

B.2 Supplementary material for Chapter 4



Figure B.3: β1 subunit R96 minimum distance from the membrane and RMSD over 200ns.

B.3 Supplementary material for Chapter 5



(a) Run 1 (b) Run 3

Figure B.4: β3 clustering in a CG mixed lipid membrane. (a) and (b) show runs 1 & 3 respectively, run 2 is shown in Chapter 5, Section 5.3.1, Figure 5.2.

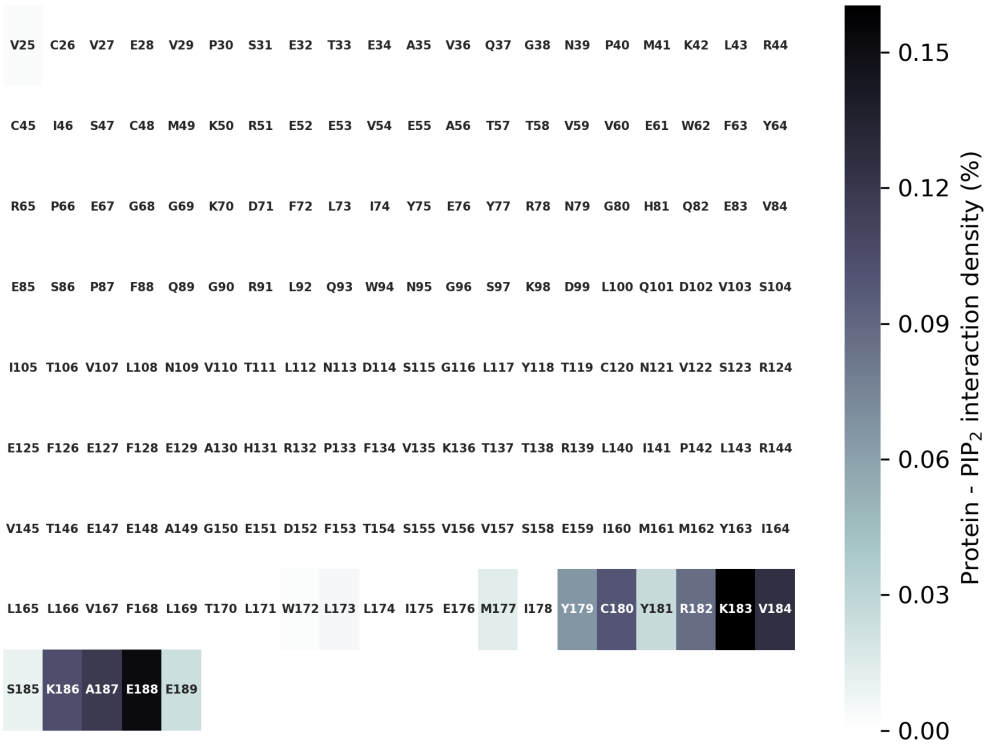
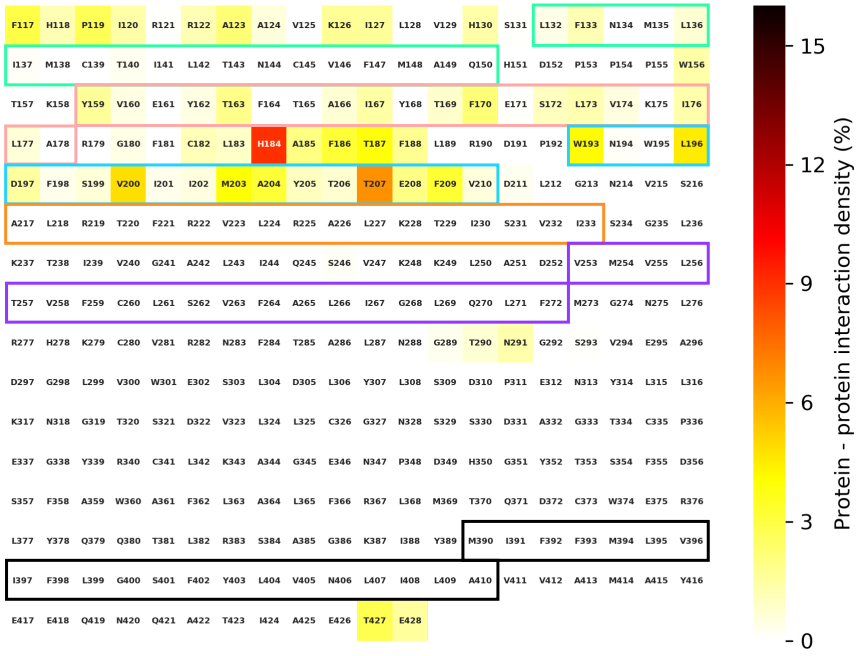
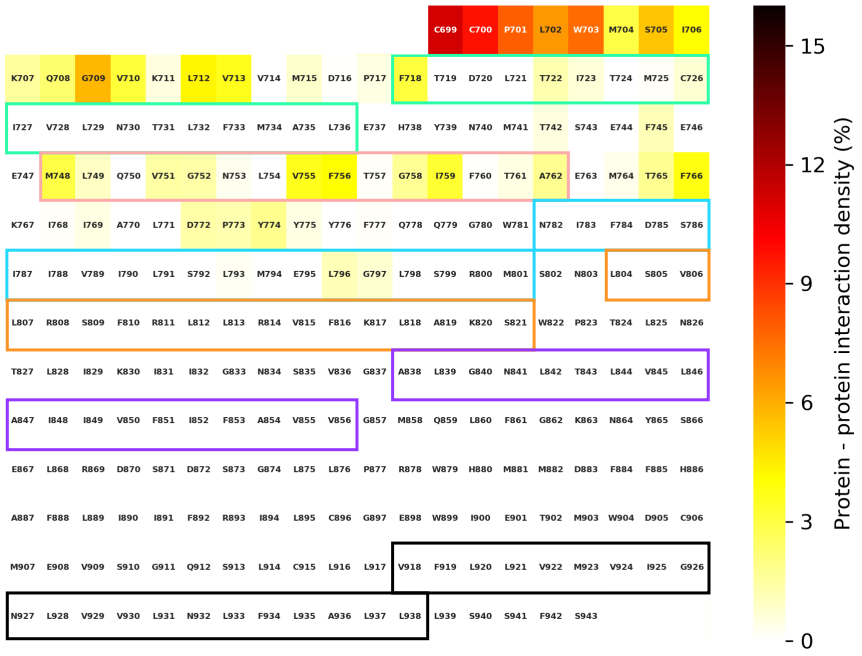


Figure B.5: Average β3 - PIP_2 interactions.

B.4 Supplementary material for Chapter 6

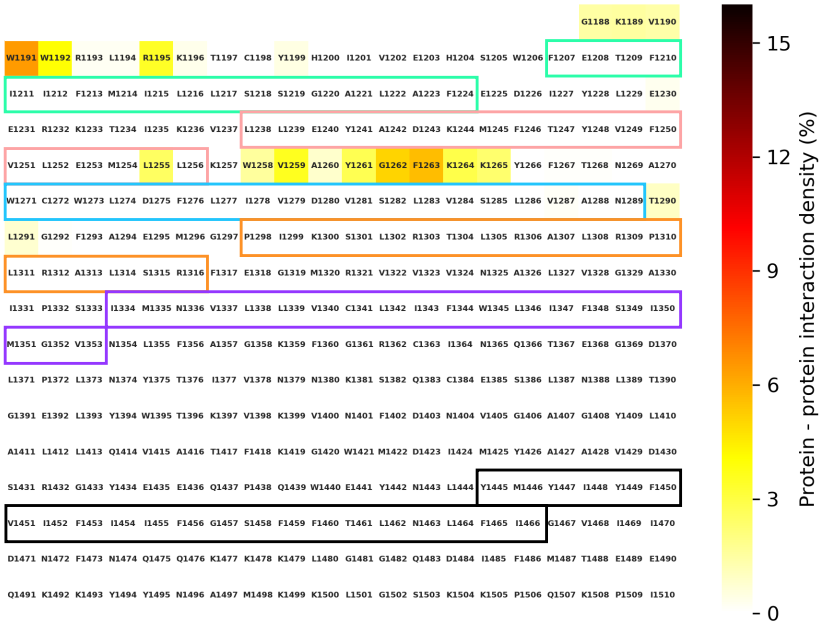


(a) DI

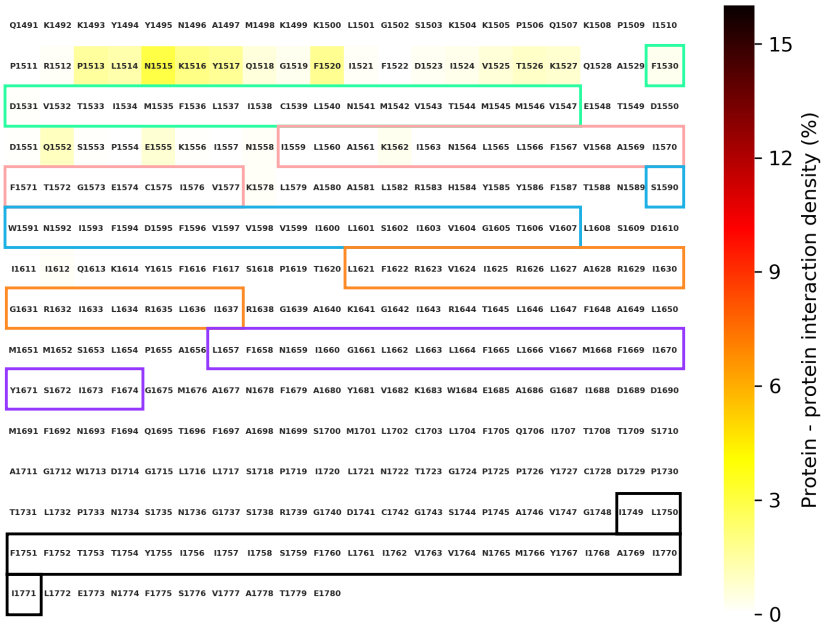


(b) DII

Figure B.6: $\beta 3$ contacts on the Nav1.5 α subunit. (a) and (b) show domains I & II respectively, the S1-6 helices in each domain are represented by boxes coloured green, pink, cyan, orange, purple, and black respectively.



(a) DIII



(b) DIV

Figure B.7: $\beta 3$ contacts on the Nav1.5 α subunit. (a) and (b) show domains III & IV respectively, the S1-6 helices in each domain are represented by boxes coloured green, pink, cyan, orange, purple, and black respectively.

B.5 Protein sequence alignments

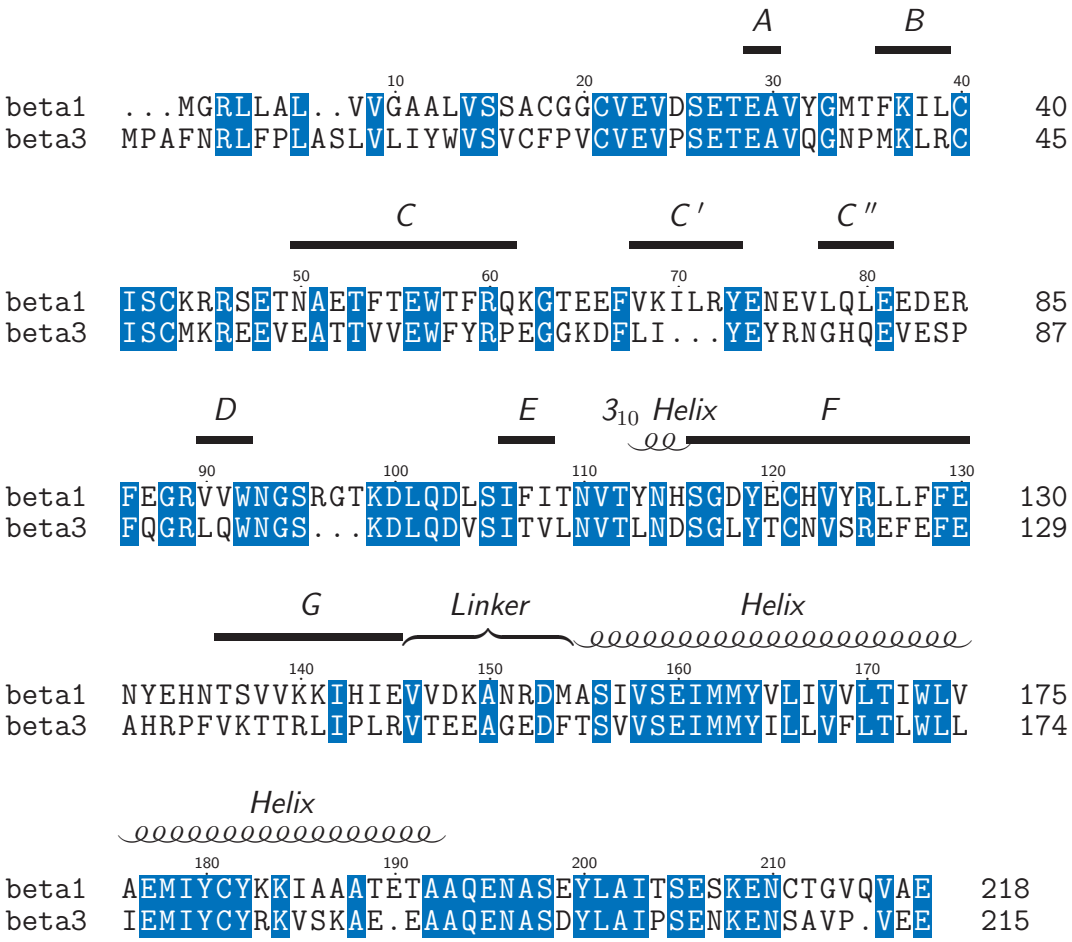


Figure B.8: Sequence alignment of Nav β1 and β3 subunits. Approximate location of strands in the Ig domain are indicated by black bars, in addition to the linker and helical regions. Numbering for β1 residues is shown above alignment.

References

[1] Q. Kuang, P. Purhonen, and H. Hebert, "Structure of potassium channels," *Cellular and Molecular Life Sciences*, vol. 72, p. 3677, 2015.

[2] C. A. Ahern, J. Payandeh, F. Bosmans, and B. Chanda, "The hitchhiker's guide to the voltage-gated sodium channel galaxy," *The Journal of General Physiology*, vol. 147, no. 1, pp. 1–24, 2016.

[3] Z. Yan, Q. Zhou, L. Wang, J. Wu, Y. Zhao, G. Huang, W. Peng, H. Shen, J. Lei, and N. Yan, "Structure of the Nav1.4-β1 Complex from Electric Eel," *Cell*, vol. 170, no. 3, pp. 470–482.e1, 2017.

[4] X. Pan, Z. Li, Q. Zhou, H. Shen, K. Wu, X. Huang, J. Chen, J. Zhang, X. Zhu, J. Lei, W. Xiong, H. Gong, B. Xiao, and N. Yan, "Structure of the human voltage-gated sodium channel Nav1.4 in complex with β1," *Science (New York, N.Y.)*, vol. 362, no. 6412, p. eaau2486, 2018.

[5] H. Shen, D. Liu, K. Wu, J. Lei, and N. Yan, "Structures of human Nav1.7 channel in complex with auxiliary subunits and animal toxins," *Science (New York, N.Y.)*, vol. 363, no. 6433, pp. 1303–1308, 2019.

[6] W. J. Brackenbury and L. L. Isom, "Na^+ Channel β Subunits: Overachievers of the Ion Channel Family," *Frontiers in Pharmacology*, vol. 2, pp. 1–11, 2011.

[7] S. Namadurai, N. R. Yereddi, F. S. Cusdin, C. L. H. Huang, D. Y. Chirgadze, and A. P. Jackson, "A new look at sodium channel β subunits," *Open biology*, vol. 5, no. 1, p. 140192, 2015.

[8] H. A. O'Malley and L. L. Isom, "Sodium channel β subunits: emerging targets in channelopathies," *Annual review of physiology*, vol. 77, pp. 481–504, 2015.

[9] C. Maffeo, S. Bhattacharya, J. Yoo, D. Wells, and A. Aksimentiev, "Modeling and Simulation of Ion Channels," *Chemical Reviews*, vol. 112, no. 12, pp. 6250–6284, 2012.

[10] A. L. Duncan, T. Reddy, H. Koldsø, J. Hélie, P. W. Fowler, M. Chavent, and M. S. P. Sansom, "Protein crowding and lipid complexity influence the nanoscale dynamic organization of ion channels in cell membranes," *Scientific Reports*, vol. 7, no. 1, p. 16647, 2017.

[11] T. Harayama and H. Riezman, "Understanding the diversity of membrane lipid composition," *Nature Reviews Molecular Cell Biology*, vol. 19, pp. 281 – 296, 2018.

[12] G. van Meer, D. R. Voelker, and G. W. Feigenson, "Membrane lipids: where they are and how they behave," *Nature Reviews Molecular Cell Biology*, vol. 9, no. 2, pp. 112–124, 2008.

[13] G. van Meer and A. I. P. M. de Kroon, "Lipid map of the mammalian cell," *J Cell Sci*, vol. 124, no. 1, pp. 5–8, 2011.

[14] Ü. Coskun and K. Simons, "Cell Membranes: The Lipid Perspective," *Structure*, vol. 19, no. 11, pp. 1543–1548, 2011.

[15] A. L. Hodgkin and A. F. Huxley, "A quantitative description of membrane current and its application to conduction and excitation in nerve," *The Journal of Physiology*, vol. 117, no. 4, pp. 500–544, 1952.

[16] B. Hille, *Ion channels of excitable membranes*. Sinauer, 2001.

[17] Y. Pirahanchi and N. R. Aeddula, *Physiology, Sodium Potassium Pump (Na^+ K^+ Pump)*. StatPearls Publishing, 2019.

[18] K. Lucas, "The "all or none" contraction of the amphibian skeletal muscle fibre," *The Journal of physiology*, vol. 38, no. 2-3, pp. 113–33, 1909.

[19] G. Meissner, "The structural basis of ryanodine receptor ion channel function," *The Journal of general physiology*, vol. 149, no. 12, pp. 1065–1089, 2017.

[20] H. Zeng and J. R. Sanes, "Neuronal cell-type classification: challenges, opportunities and the path forward," *Nature Reviews Neuroscience*, vol. 18, no. 9, pp. 530–546, 2017.

[21] A. F. Huxley and R. Stämpeli, "Evidence for saltatory conduction in peripheral myelinated nerve fibres," *The Journal of Physiology*, vol. 108, no. 3, pp. 315–339, 1949.

[22] R. Fitzhugh, "Computation of Impulse Initiation and Saltatory Conduction in a Myelinated Nerve Fiber," *Biophysical Journal*, vol. 2, no. 1, pp. 11–21, 1962.

[23] A. L. Harris, "Emerging issues of connexin channels: biophysics fills the gap," *Quarterly reviews of biophysics*, vol. 34, no. 3, pp. 325–472, 2001.

[24] J. Wang, S.-W. Ou, and Y.-J. Wang, "Distribution and function of voltage-gated sodium channels in the nervous system.," *Channels (Austin, Tex.)*, vol. 11, no. 6, pp. 534–554, 2017.

[25] H. Watanabe, T. T. Koopmann, S. Le Scouarnec, T. Yang, C. R. Ingram, J.-J. Schott, S. Demolombe, V. Probst, F. Anselme, D. Escande, A. C. P. Wiesfeld, A. Pfeufer, S. Kääb, H.-E. Wichmann, C. Hasdemir, Y. Aizawa, A. A. M. Wilde, D. M. Roden, and C. R. Bezzina, "Sodium channel β1 subunit mutations associated with Brugada syndrome and cardiac conduction disease in humans," *The Journal of clinical investigation*, vol. 118, no. 6, pp. 2260–8, 2008.

[26] D. Audenaert, L. Claes, B. Ceulemans, A. Löfgren, C. Van Broeckhoven, and P. De Jonghe, "A deletion in SCN1B is associated with febrile seizures and early-onset absence epilepsy," *Neurology*, vol. 61, no. 6, pp. 854–6, 2003.

[27] K. L. I. van Gassen, M. de Wit, M. van Kempen, W. Saskia van der Hel, P. C. van Rijen, A. P. Jackson, D. Lindhout, and P. N. E. de Graan, "Hippocampal Nav β3 expression in patients with temporal lobe epilepsy," *Epilepsia*, vol. 50, no. 4, pp. 957–962, 2009.

[28] J. J. Gargus, "Ion Channel Functional Candidate Genes in Multigenic Neuropsychiatric Disease," *Biological Psychiatry*, vol. 60, no. 2, pp. 177–185, 2006.

[29] B. S. Shah, E. B. Stevens, M. I. Gonzalez, S. Bramwell, R. D. Pinnock, K. Lee, and A. K. Dixon, "β3, a novel auxiliary subunit for the voltage-gated sodium channel, is expressed preferentially in sensory neurons and is upregulated in the chronic constriction injury model of neuropathic pain," *European Journal of Neuroscience*, vol. 12, no. 11, pp. 3985–3990, 2000.

[30] N. Takahashi, S. Kikuchi, Y. Dai, K. Kobayashi, T. Fukuoka, and K. Noguchi, "Expression of auxiliary β subunits of sodium channels in primary afferent neurons and the effect of nerve injury," *Neuroscience*, vol. 121, no. 2, pp. 441–450, 2003.

[31] W. S. Agnew, A. C. Moore, S. R. Levinson, and M. A. Raftery, "Identification of a large molecular weight peptide associated with a tetrodotoxin binding protein from the electroplax of Electrophoruselectricus," *Biochemical and Biophysical Research Communications*, vol. 92, no. 3, pp. 860–866, 1980.

[32] D. A. Beneski and W. A. Catterall, "Covalent labeling of protein components of the sodium channel with a photoactivable derivative of scorpion toxin," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 77, no. 1, pp. 639–43, 1980.

[33] W. A. Catterall, "The molecular basis of neuronal excitability," *Science*, vol. 223, no. 4637, pp. 653–661, 1984.

[34] W. A. Catterall, "From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels," *Neuron*, vol. 26, no. 1, pp. 13–25, 2000.

[35] F. H. Yu, V. Yarov-Yarovoy, G. A. Gutman, and W. A. Catterall, "Overview of molecular relationships in the voltage-gated ion channel superfamily," *Pharmacological reviews*, vol. 57, no. 4, pp. 387–95, 2005.

[36] M. Noda, S. Shimizu, T. Tanabe, T. Takai, T. Kayano, T. Ikeda, H. Takahashi, H. Nakayama, Y. Kanaoka, N. Minamino, K. Kangawa, H. Matsuo, M. A. Raftery, T. Hirose, S. Inayama, H. Hayashida, T. Miyata, and S. Numa, "Primary structure of Electrophorus electricus sodium channel deduced from cDNA sequence," *Nature*, vol. 312, no. 5990, pp. 121–127, 1984.

[37] J. Diss, S. Fraser, and M. Djamgoz, "Voltage-gated Na^+ channels: multiplicity of expression, plasticity, functional implications and pathophysiological aspects," *European Biophysics Journal*, vol. 33, no. 3, pp. 180–193, 2004.

[38] A. L. Goldin, R. L. Barchi, J. H. Caldwell, F. Hofmann, J. R. Howe, J. C. Hunter, R. G. Kallen, G. Mandel, M. H. Meisler, Y. B. Netter, M. Noda, M. M. Tamkun, S. G. Waxman, J. N. Wood, and W. A. Catterall, "Nomenclature of voltage-gated sodium channels," *Neuron*, vol. 28, no. 2, pp. 365–8, 2000.

[39] J. Satin, J. W. Kyle, M. Chen, P. Bell, L. L. Cribbs, H. A. Fozzard, and R. B. Rogart, "A mutant of TTX-resistant cardiac sodium channels with TTX-sensitive properties," *Science (New York, N.Y.)*, vol. 256, no. 5060, pp. 1202–5, 1992.

[40] L. Sivilotti, K. Okuse, A. Akopian, S. Moss, and J. Wood, "A single serine residue confers tetrodotoxin insensitivity on the rat sensory-neuron-specific sodium channel SNS," *FEBS Letters*, vol. 409, no. 1, pp. 49–52, 1997.

[41] I. Ogiwara, H. Miyamoto, N. Morita, N. Atapour, E. Mazaki, I. Inoue, T. Takeuchi, S. Itohara, Y. Yanagawa, K. Obata, T. Furuichi, T. K. Hensch, and K. Yamakawa, "Nav1.1 localizes to axons of parvalbumin-positive inhibitory interneurons: a circuit basis for epileptic seizures in mice carrying an Scn1a gene mutation," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 27, no. 22, pp. 5903–14, 2007.

[42] J. H. Caldwell, K. L. Schaller, R. S. Lasher, E. Peles, and S. R. Levinson, "Sodium channel Nav1.6 is localized at nodes of ranvier, dendrites, and synapses," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 10, pp. 5616–20, 2000.

[43] M. R. Kaplan, M. H. Cho, E. M. Ullian, L. L. Isom, S. R. Levinson, and B. A. Barres, "Differential control of clustering of the sodium channels Nav1.2 and Nav1.6 at developing CNS nodes of Ranvier," *Neuron*, vol. 30, no. 1, pp. 105–19, 2001.

[44] J. A. Lindia, M. G. Köhler, W. J. Martin, and C. Abbadie, "Relationship between sodium channel Nav1.3 expression and neuropathic pain behavior in rats," *Pain*, vol. 117, no. 1, pp. 145–153, 2005.

[45] T. R. Cummins, P. L. Sheets, and S. G. Waxman, "The roles of sodium channels in nociception: Implications for mechanisms of pain.," *Pain*, vol. 131, no. 3, pp. 243–57, 2007.

[46] J. S. Trimmer, S. S. Cooperman, S. A. Tomiko, J. Zhou, S. M. Crean, M. B. Boyle, R. G. Kalen, Z. Sheng, R. L. Barchi, F. J. Sigworth, R. H. Goodman, W. S. Agnew, and G. Mandel, "Primary structure and functional expression of a mammalian skeletal muscle sodium channel," *Neuron*, vol. 3, no. 1, pp. 33–49, 1989.

[47] M. B. Rook, M. M. Evers, M. A. Vos, and M. F. A. Bierhuizen, "Biology of cardiac sodium channel Nav1.5 expression," *Cardiovascular Research*, vol. 93, no. 1, pp. 12–23, 2012.

[48] D. M. Krzemien, K. L. Schaller, S. R. Levinson, and J. H. Caldwell, "Immunolocalization of sodium channel isoform NaCh6 in the nervous system," *The Journal of Comparative Neurology*, vol. 420, no. 1, pp. 70–83, 2000.

[49] A. Gasser, T. S.-Y. Ho, X. Cheng, K.-J. Chang, S. G. Waxman, M. N. Rasband, and S. D. Dib-Hajj, "An AnkyrinG-Binding Motif Is Necessary and Sufficient for Targeting Nav1.6 Sodium Channels to Axon Initial Segments and Nodes of Ranvier," *Journal of Neuroscience*, vol. 32, no. 21, pp. 7232–7243, 2012.

[50] J. J. Toledo-Aral, B. L. Moss, Z. J. He, A. G. Koszowski, T. Whisenand, S. R. Levinson, J. J. Wolf, I. Silos-Santiago, S. Halegoua, and G. Mandel, "Identification of PN1, a predominant voltage-dependent sodium channel expressed principally in peripheral neurons," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 4, pp. 1527–32, 1997.

[51] C. Han, J. Huang, and S. G. Waxman, "Sodium channel Nav1.8: Emerging links to human disease," *Neurology*, vol. 86, no. 5, pp. 473–83, 2016.

[52] S. D. Dib-Hajj, J. A. Black, and S. G. Waxman, "Nav1.9: a sodium channel linked to human pain," *Nature Reviews Neuroscience*, vol. 16, no. 9, pp. 511–519, 2015.

[53] M. Duménieu, M. Oulé, M. R. Kreutz, and J. Lopez-Rojas, "The Segregated Expression of Voltage-Gated Potassium and Sodium Channels in Neuronal Membranes: Functional Implications and Regulatory Mechanisms," *Frontiers in Cellular Neuroscience*, vol. 11, p. 115, 2017.

[54] J. M. Hull and L. L. Isom, "Voltage-gated sodium channel β subunits: The power outside the pore in brain development and disease," *Neuropharmacology*, vol. 132, pp. 43–57, 2018.

[55] G. A. Patino, W. J. Brackenbury, Y. Bao, L. F. Lopez-Santiago, H. A. O'Malley, C. Chen, J. D. Calhoun, R. G. Lafrenière, P. Cossette, G. A. Rouleau, and L. L. Isom, "Voltage-gated Na^+ channel β1B: a secreted cell adhesion molecule involved in human epilepsy," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 31, no. 41, pp. 14577–91, 2011.

[56] T. Zimmer and K. Benndorf, "The human heart and rat brain IIA Na^+ channels interact with different molecular regions of the β1 subunit," *The Journal of general physiology*, vol. 120, no. 6, pp. 887–95, 2002.

[57] E. J. Yu, S.-H. Ko, P. W. Lenkowski, A. Pance, M. K. Patel, and A. P. Jackson, "Distinct domains of the sodium channel β3-subunit modulate channel-gating kinetics and subcellular location," *The Biochemical journal*, vol. 392, no. Pt 3, pp. 519–26, 2005.

[58] C. Chen, V. Bharucha, Y. Chen, R. E. Westenbroek, A. Brown, J. D. Malhotra, D. Jones, C. Avery, P. J. Gillespie, K. A. Kazen-Gillespie, K. Kazarinova-Noyes, P. Shrager, T. L. Saunders, R. L. Macdonald, B. R. Ransom, T. Scheuer, W. A. Catterall, L. L. Isom, and L. L. Isom, "Reduced sodium channel density, altered voltage dependence of inactivation, and increased susceptibility to seizures in mice lacking sodium channel β2 subunits.," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 26, pp. 17072–7, 2002.

[59] L. F. Lopez-Santiago, M. Pertin, X. Morisod, C. Chen, S. Hong, J. Wiley, I. Decosterd, and L. L. Isom, "Sodium channel β2 subunits regulate tetrodotoxin-sensitive sodium channels in small dorsal root ganglion neurons and modulate the response to pain," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 26, no. 30, pp. 7984–94, 2006.

[60] L. Isom, D. Ragsdale, K. De Jongh, R. Westenbroek, B. Reber, T. Scheuer, and W. Catterall, "Structure and function of the β2 subunit of brain sodium channels, a transmembrane glycoprotein with a CAM motif," *Cell*, vol. 83, no. 3, pp. 433–442, 1995.

[61] C. Chen, J. D. Calhoun, Y. Zhang, L. Lopez-Santiago, N. Zhou, T. H. Davis, J. L. Salzer, and L. L. Isom, "Identification of the Cysteine Residue Responsible for Disulfide Linkage of Na^+ Channel α and β2 Subunits," *Journal of Biological Chemistry*, vol. 287, no. 46, pp. 39061–39069, 2012.

[62] J. Gilchrist, S. Das, F. Van Petegem, and F. Bosmans, "Crystallographic insights into sodium-channel modulation by the β4 subunit," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 110, no. 51, pp. E5016–24, 2013.

[63] K. A. McCormick, L. L. Isom, D. Ragsdale, D. Smith, T. Scheuer, and W. A. Catterall, "Molecular determinants of Na^+ channel function in the extracellular domain of the β1 subunit," *The Journal of biological chemistry*, vol. 273, no. 7, pp. 3954–62, 1998.

[64] J. Spampanato, "A Novel Epilepsy Mutation in the Sodium Channel SCN1A Identifies a Cytoplasmic Domain for Subunit Interaction," *Journal of Neuroscience*, vol. 24, no. 44, pp. 10022–10034, 2004.

[65] L. Meadows, J. D. Malhotra, A. Stetzer, L. L. Isom, and D. S. Ragsdale, "The intracellular segment of the sodium channel β1 subunit is required for its efficient association with the channel α subunit," *Journal of Neurochemistry*, vol. 76, no. 6, pp. 1871–1878, 2001.

[66] L. L. Isom and W. A. Catterall, "Na^+ channel subunits and Ig domains," *Nature*, vol. 383, no. 6598, pp. 307–308, 1996.

[67] L. L. Isom, "The role of sodium channels in cell adhesion," *Frontiers in bioscience: a journal and virtual library*, vol. 7, pp. 12–23, 2002.

[68] J. Srinivasan, M. Schachner, and W. A. Catterall, "Interaction of voltage-gated sodium channels with the extracellular matrix molecules tenascin-C and tenascin-R," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 95, no. 26, pp. 15753–7, 1998.

[69] J. D. Malhotra, K. Kazen-Gillespie, M. Hortsch, and L. L. Isom, "Sodium channel β subunits mediate homophilic cell adhesion and recruit ankyrin to points of cell-cell contact," *The Journal of biological chemistry*, vol. 275, no. 15, pp. 11383–8, 2000.

[70] D. P. McEwen and L. L. Isom, "Heterophilic interactions of sodium channel β1 subunits with axonal and glial cell adhesion molecules," *The Journal of biological chemistry*, vol. 279, no. 50, pp. 52744–52, 2004.

[71] C. F. Ratcliffe, R. E. Westenbroek, R. Curtis, and W. A. Catterall, "Sodium channel β1 and β3 subunits associate with neurofascin through their extracellular immunoglobulin-like domain," *The Journal of cell biology*, vol. 154, no. 2, pp. 427–34, 2001.

[72] D. Zhou, S. Lambert, P. L. Malen, S. Carpenter, L. M. Boland, and V. Bennett, "AnkyrinG is required for clustering of voltage-gated Na channels at axon initial segments and for normal action potential firing," *The Journal of cell biology*, vol. 143, no. 5, pp. 1295–304, 1998.

[73] C. Leterrier, N. Clerc, F. Rueda-Boroni, A. Montersino, B. Dargent, and F. Castets, "Ankyrin G Membrane Partners Drive the Establishment and Maintenance of the Axon Initial Segment," *Frontiers in Cellular Neuroscience*, vol. 11, p. 6, 2017.

[74] C. Chen, R. E. Westenbroek, X. Xu, C. A. Edwards, D. R. Sorenson, Y. Chen, D. P. McEwen, H. A. O'Malley, V. Bharucha, L. S. Meadows, G. A. Knudsen, A. Vilaythong, J. L. Noebels, T. L. Saunders, T. Scheuer, P. Shrager, W. A. Catterall, and L. L. Isom, "Mice Lacking Sodium Channel β1 Subunits Display Defects in Neuronal Excitability, Sodium Channel Expression, and Nodal Architecture," *Journal of Neuroscience*, vol. 24, no. 16, pp. 4030–4042, 2004.

[75] F. H. Yu, R. E. Westenbroek, I. Silos-Santiago, K. A. McCormick, D. Lawson, P. Ge, H. Ferriera, J. Lilly, P. S. DiStefano, W. A. Catterall, T. Scheuer, and R. Curtis, "Sodium channel β4, a new disulfide-linked auxiliary subunit with similarity to β2," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 23, no. 20, pp. 7577–85, 2003.

[76] K. Morgan, E. B. Stevens, B. Shah, P. J. Cox, A. K. Dixon, K. Lee, R. D. Pinnock, J. Hughes, P. J. Richardson, K. Mizuguchi, and A. P. Jackson, "β3: an additional auxiliary subunit of the voltage-sensitive sodium channel that modulates channel gating with distinct kinetics," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 97, no. 5, pp. 2308–13, 2000.

[77] B. S. Shah, E. B. Stevens, R. D. Pinnock, A. K. Dixon, and K. Lee, "Developmental expression of the novel voltage-gated sodium channel auxiliary subunit β3, in rat CNS," *The Journal of Physiology*, vol. 534, no. 3, pp. 763–776, 2001.

[78] S. A. Buffington and M. N. Rasband, "Na^+ Channel-Dependent Recruitment of Nav β4 to Axon Initial Segments and Nodes of Ranvier," *Journal of Neuroscience*, vol. 33, no. 14, pp. 6191–6202, 2013.

[79] H. M. Nguyen, H. Miyazaki, N. Hoshi, B. J. Smith, N. Nukina, A. L. Goldin, and K. G. Chandy, "Modulation of voltage-gated K^+ channels by the sodium channel β1 subunit," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 45, pp. 18577–82, 2012.

[80] K. A. Kazen-Gillespie, D. S. Ragsdale, M. R. D'Andrea, L. N. Mattei, K. E. Rogers, and L. L. Isom, "Cloning, localization, and functional expression of sodium channel β1a subunits," *The Journal of biological chemistry*, vol. 275, no. 2, pp. 1079–88, 2000.

[81] N. Qin, M. R. D'Andrea, M.-L. Lubin, N. Shafaee, E. E. Codd, and A. M. Correa, "Molecular cloning and functional expression of the human sodium channel β1b subunit, a novel splicing variant of the β1 subunit," *European Journal of Biochemistry*, vol. 270, no. 23, pp. 4762–4770, 2003.

[82] W. J. Brackenbury, J. D. Calhoun, C. Chen, H. Miyazaki, N. Nukina, F. Oyama, B. Ranscht, and L. L. Isom, "Functional reciprocity between Na^+ channel Nav1.6 and β1 subunits in the coordinated regulation of excitability and neurite outgrowth," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 5, pp. 2283–8, 2010.

[83] T. H. Davis, C. Chen, and L. L. Isom, "Sodium channel β1 subunits promote neurite outgrowth in cerebellar granule neurons," *The Journal of biological chemistry*, vol. 279, no. 49, pp. 51424–32, 2004.

[84] Y. Oh, S. Sashihara, J. A. Black, and S. G. Waxman, "Na^+ channel β1 subunit mRNA: differential expression in rat spinal sensory neurons," *Molecular Brain Research*, vol. 30, no. 2, pp. 357–361, 1995.

[85] J. Zhao, M. E. O'Leary, and M. Chahine, "Regulation of Nav1.6 and Nav1.8 peripheral nerve Na^+ channels by auxiliary β-subunits," *Journal of neurophysiology*, vol. 106, no. 2, pp. 608–19, 2011.

[86] M. Pertin, R.-R. Ji, T. Berta, A. J. Powell, L. Karchewski, S. N. Tate, L. L. Isom, C. J. Woolf, N. Gilliard, D. R. Spahn, and I. Decosterd, "Upregulation of the voltage-gated sodium channel $\beta 2$ subunit in neuropathic pain models: characterization of expression in injured and non-injured primary sensory neurons," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 25, no. 47, pp. 10970–80, 2005.

[87] Y. Qu, R. Curtis, D. Lawson, K. Gilbride, P. Ge, P. DiStefano, I. Silos-Santiago, W. Catterall, and T. Scheuer, "Differential Modulation of Sodium Channel Gating and Persistent Sodium Currents by the $\beta 1$, $\beta 2$, and $\beta 3$ Subunits," *Molecular and Cellular Neuroscience*, vol. 18, no. 5, pp. 570–580, 2001.

[88] S. K. G. Maier, R. E. Westenbroek, K. A. McCormick, R. Curtis, T. Scheuer, and W. A. Catterall, "Distinct subcellular localization of different sodium channel α and β subunits in single ventricular myocytes from mouse heart," *Circulation*, vol. 109, no. 11, pp. 1421–7, 2004.

[89] N. Gaborit, S. Le Bouter, V. Szuts, A. Varro, D. Escande, S. Nattel, and S. Demolombe, "Regional and tissue specific transcript signatures of ion channel genes in the non-diseased human heart," *The Journal of physiology*, vol. 582, no. Pt 2, pp. 675–93, 2007.

[90] S. G. Kaufmann, R. E. Westenbroek, A. H. Maass, V. Lange, A. Renner, E. Wischmeyer, A. Bonz, J. Muck, G. Ertl, W. A. Catterall, T. Scheuer, and S. K. G. Maier, "Distribution and function of sodium channel subtypes in human atrial myocardium," *Journal of molecular and cellular cardiology*, vol. 61, pp. 133–141, 2013.

[91] S. K. G. Maier, R. E. Westenbroek, T. T. Yamanushi, H. Dobrzynski, M. R. Boyett, W. A. Catterall, and T. Scheuer, "An unexpected requirement for brain-type sodium channels for control of heart rate in the mouse sinoatrial node," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, no. 6, pp. 3507–12, 2003.

[92] J. W. West, D. E. Patton, T. Scheuer, Y. Wang, A. L. Goldin, and W. A. Catterall, "A cluster of hydrophobic amino acid residues required for fast na^+-channel inactivation," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 89, no. 22, pp. 10910–4, 1992.

[93] C. A. Rohl, F. A. Boeckman, C. Baker, T. Scheuer, W. A. Catterall, and R. E. Klevit, "Solution Structure of the Sodium Channel Inactivation Gate," *Biochemistry*, vol. 38, no. 3, pp. 855–861, 1999.

[94] D. L. Capes, M. P. Goldschen-Ohm, M. Arcisio-Miranda, F. Bezanilla, and B. Chanda, "Domain IV voltage-sensor movement is both sufficient and rate limiting for fast inactivation in sodium channels," *The Journal of general physiology*, vol. 142, no. 2, pp. 101–12, 2013.

[95] A. F. Williams and A. N. Barclay, "The Immunoglobulin Superfamily - Domains for Cell Surface Recognition," *Annual Review of Immunology*, vol. 6, no. 1, pp. 381–405, 1988.

[96] D. K. Smith and H. Xue, "Sequence profiles of immunoglobulin and immunoglobulin-like domains," *Journal of Molecular Biology*, vol. 274, no. 4, pp. 530–545, 1997.

[97] V. Potapov, V. Sobolev, M. Edelman, A. Kister, and I. Gelfand, "Protein-Protein Recognition: Juxtaposition of Domain and Interface Cores in Immunoglobulins and Other Sandwich-like Proteins," *Journal of Molecular Biology*, vol. 342, no. 2, pp. 665–679, 2004.

[98] P. Bork, L. Holm, and C. Sander, "The Immunoglobulin Fold: Structural Classification, Sequence Patterns and Common Core," *Journal of Molecular Biology*, vol. 242, no. 4, pp. 309–320, 1994.

[99] S. Namadurai, D. Balasuriya, R. Rajappa, M. Wiemhöfer, K. Stott, J. Klingauf, J. M. Edwardson, D. Y. Chirgadze, and A. P. Jackson, "Crystal Structure and Molecular Imaging of the Nav Channel β3 Subunit Indicates a Trimeric Assembly," *The Journal of Biological Chemistry*, vol. 289, no. 15, pp. 10797–10811, 2014.

[100] J. Clatot, M. Hoshi, X. Wan, H. Liu, A. Jain, K. Shinlapawittayatorn, C. Marionneau, E. Ficker, T. Ha, and I. Deschênes, "Voltage-gated sodium channels assemble and gate as dimers," *Nature Communications*, vol. 8, no. 1, 2017.

[101] P. Hakim, N. Brice, R. Thresher, J. Lawrence, Y. Zhang, A. P. Jackson, A. A. Grace, and C. L.-H. Huang, "Scn3b knockout mice exhibit abnormal sinoatrial and cardiac conduction properties," *Acta physiologica (Oxford, England)*, vol. 198, no. 1, pp. 47–59, 2010.

[102] P. Hakim, I. S. Gurung, T. H. Pedersen, R. Thresher, N. Brice, J. Lawrence, A. A. Grace, and C. L.-H. Huang, "Scn3b knockout mice exhibit abnormal ventricular electrophysiological properties," *Progress in Biophysics and Molecular Biology*, vol. 98, no. 2-3, p. 251, 2008.

[103] S. Okata, S. Yuasa, T. Suzuki, S. Ito, N. Makita, T. Yoshida, M. Li, J. Kurokawa, T. Seki, T. Egashira, Y. Aizawa, M. Kodaira, C. Motoda, G. Yozu, M. Shimojima, N. Hayashiji, H. Hashimoto, Y. Kuroda, A. Tanaka, M. Murata, T. Aiba, W. Shimizu, M. Horie, K. Kamiya, T. Furukawa, and K. Fukuda, "Embryonic type Na^+ channel β-subunit, SCN3B masks the disease phenotype of Brugada syndrome," *Scientific Reports*, vol. 6, no. 1, p. 34198, 2016.

[104] Y. Shi, "A glimpse of structural biology through X-ray crystallography," *Cell*, vol. 159, no. 5, pp. 995–1014, 2014.

[105] E. G. Marklund and J. L. Benesch, "Weighing-up protein dynamics: the combination of native mass spectrometry and molecular dynamics simulations," *Current Opinion in Structural Biology*, vol. 54, pp. 50–58, 2019.

[106] E. Schrödinger, "An Undulatory Theory of the Mechanics of Atoms and Molecules," *Physical Review*, vol. 28, no. 6, pp. 1049–1070, 1926.

[107] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.

[108] G. Grff, H. Kalinowsky, and J. Traut, "A direct determination of the proton electron mass ratio," *Zeitschrift fr Physik A: Atoms and Nuclei*, vol. 297, no. 1, pp. 35–39, 1980.

[109] D. R. Hartree, "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, no. 1, pp. 111–132, 1928.

[110] D. Hartree and W. Hartree, "Self-consistent field, with exchange, for beryllium," *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, vol. 150, no. 869, pp. 9–33, 1935.

[111] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Physical Review*, vol. 136, no. 3B, pp. B864–B871, 1964.

[112] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Physical Review*, vol. 140, no. 4A, pp. A1133–A1138, 1965.

[113] D. C. Sherrill, "Frontiers in electronic structure theory," *The Journal of Chemical Physics*, vol. 132, no. 11, p. 110902, 2010.

[114] T. Helgaker, P. Jørgensen, and J. C. N. Q. H. . Olsen, *Molecular electronic-structure theory*. Chichester ; New York: Wiley, 2000.

[115] A. Warshel and M. Levitt, "Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme," *Journal of Molecular Biology*, vol. 103, no. 2, pp. 227–249, 1976.

[116] M. W. van der Kamp and A. J. Mulholland, "Combined Quantum Mechanics/Molecular Mechanics (QM/MM) Methods in Computational Enzymology," *Biochemistry*, vol. 52, no. 16, pp. 2708–2728, 2013.

[117] E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, and K. Schulten, "Discovery through the computational microscope," *Structure (London, England: 1993)*, vol. 17, no. 10, pp. 1295–306, 2009.

[118] R. O. Dror, R. M. Dirks, J. Grossman, H. Xu, and D. E. Shaw, "Biomolecular Simulation: A Computational Microscope for Molecular Biology," *Annual Review of Biophysics*, vol. 41, no. 1, pp. 429–452, 2012.

[119] M. O. Steinhauser and S. Hiermaier, "A review of computational methods in materials science: examples from shock-wave and polymer physics," *International journal of molecular sciences*, vol. 10, no. 12, pp. 5135–216, 2009.

[120] S. A. Hollingsworth and R. O. Dror, "Molecular Dynamics Simulation for All," *Neuron*, vol. 99, no. 6, pp. 1129–1143, 2018.

[121] J. A. McCammon, B. R. Gelin, and M. Karplus, "Dynamics of folded proteins," *Nature*, vol. 267, no. 5612, pp. 585–590, 1977.

[122] J. R. Perilla and K. Schulten, "Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations," *Nature Communications*, vol. 8, no. 1, p. 15959, 2017.

[123] M. S. Friedrichs, P. Eastman, V. Vaidyanathan, M. Houston, S. Legrand, A. L. Beberg, D. L. Ensign, C. M. Bruns, and V. S. Pande, "Accelerating molecular dynamic simulation on graphics processing units," *Journal of Computational Chemistry*, vol. 30, no. 6, pp. 864–872, 2009.

[124] L. C. Pierce, R. Salomon-Ferrer, C. Augusto F. de Oliveira, J. A. McCammon, and R. C. Walker, "Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics," *Journal of Chemical Theory and Computation*, vol. 8, no. 9, pp. 2997–3002, 2012.

[125] A. W. Götz, M. J. Williamson, D. Xu, D. Poole, S. Le Grand, and R. C. Walker, "Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born," *Journal of Chemical Theory and Computation*, vol. 8, no. 5, pp. 1542–1555, 2012.

[126] R. Salomon-Ferrer, A. W. Götz, D. Poole, S. Le Grand, and R. C. Walker, "Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald," *Journal of Chemical Theory and Computation*, vol. 9, no. 9, pp. 3878–3888, 2013.

[127] S. Lindert, D. Bucher, P. Eastman, V. Pande, and J. A. McCammon, "Accelerated Molecular Dynamics Simulations with the AMOEBA Polarizable Force Field on Graphics Processing Units," *Journal of Chemical Theory and Computation*, vol. 9, no. 11, pp. 4684–4691, 2013.

[128] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules," *Journal of the American Chemical Society*, vol. 117, no. 19, pp. 5179–5197, 1995.

[129] Robert Hooke, *Lectures de Potentia Restitutiva, Or of Spring Explaining the Power of Springing Bodies*. John Martyn, 1678.

[130] P. M. Morse, "Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels," *Physical Review*, vol. 34, no. 1, pp. 57–64, 1929.

[131] J. E. Jones, "On the Determination of Molecular Fields. I. From the Variation of the Viscosity of a Gas with Temperature," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 106, no. 738, pp. 441–462, 1924.

[132] C. I. Bayly, P. Cieplak, W. Cornell, and P. A. Kollman, "A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model," *The Journal of Physical Chemistry*, vol. 97, no. 40, pp. 10269–10280, 1993.

[133] P. Ren and J. W. Ponder, "Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations," *Journal of Computational Chemistry*, vol. 23, no. 16, pp. 1497–1506, 2002.

[134] P. Ren and J. W. Ponder, "Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation," *The Journal of Physical Chemistry B*, vol. 107, no. 24, pp. 5933–5947, 2003.

[135] P. Cieplak, F.-Y. Dupradeau, Y. Duan, and J. Wang, "Polarization effects in molecular mechanical force fields," *Journal of Physics: Condensed Matter*, vol. 21, no. 33, p. 333102, 2009.

[136] I. Newton, *Philosophiæ Naturalis Principia Mathematica*. Londini: Jussu Societatis Regiæ ac typis Josephi Streatii, 1687.

[137] L. Verlet, "Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules," *Physical Review*, vol. 159, no. 1, pp. 98–103, 1967.

[138] M. P. Allen and D. J. Tildesley in *Computer Simulation of Liquids*, p. 103, Oxford University Press, 2017.

[139] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters," *The Journal of Chemical Physics*, vol. 76, no. 1, pp. 637–649, 1982.

[140] M. P. Allen and D. J. Tildesley in *Computer Simulation of Liquids*, pp. 35–36, Oxford University Press, 2017.

[141] M. P. Allen and D. J. Tildesley in *Computer Simulation of Liquids*, pp. 39–40, Oxford University Press, 2017.

[142] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak, "Molecular dynamics with coupling to an external bath," *The Journal of Chemical Physics*, vol. 81, no. 8, pp. 3684–3690, 1984.

[143] T. Morishita, "Fluctuation formulas in molecular-dynamics simulations with the weak coupling heat bath," *The Journal of Chemical Physics*, vol. 113, no. 8, pp. 2976–2982, 2000.

[144] A. S. Lemak and N. K. Balabaev, "On The Berendsen Thermostat," *Molecular Simulation*, vol. 13, no. 3, pp. 177–187, 1994.

[145] E. Braun, S. M. Moosavi, and B. Smit, "Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited," *Journal of Chemical Theory and Computation*, vol. 14, no. 10, pp. 5262–5272, 2018.

[146] G. Bussi, D. Donadio, and M. Parrinello, "Canonical sampling through velocity rescaling," *The Journal of Chemical Physics*, vol. 126, no. 1, p. 014101, 2007.

[147] S. Nosé, "A unified formulation of the constant temperature molecular dynamics methods," *The Journal of Chemical Physics*, vol. 81, no. 1, pp. 511–519, 1984.

[148] W. G. Hoover, "Canonical dynamics: Equilibrium phase-space distributions," *Physical Review A*, vol. 31, no. 3, pp. 1695–1697, 1985.

[149] G. J. Martyna, M. L. Klein, and M. Tuckerman, "Nosé-Hoover chains: The canonical ensemble via continuous dynamics," *The Journal of Chemical Physics*, vol. 97, no. 4, pp. 2635–2643, 1992.

[150] M. Allen and D. J. Tildesley in *Computer Simulation of Liquids*, p. 131, Oxford University Press, 2017.

[151] H. J. C. Berendsen in *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*, pp. 199–201, Cambridge University Press, 2007.

[152] M. Parrinello and A. Rahman, "Crystal Structure and Pair Potentials: A Molecular-Dynamics Study," *Physical Review Letters*, vol. 45, no. 14, pp. 1196–1199, 1980.

[153] S. Nosé and M. Klein, "Constant pressure molecular dynamics for molecular systems," *Molecular Physics*, vol. 50, no. 5, pp. 1055–1076, 1983.

[154] M. P. Allen and D. J. Tildesley in *Computer Simulations of Liquids*, pp. 235–239, Oxford University Press, 2017.

[155] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, and M. Karplus, "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins," *The Journal of Physical Chemistry B*, vol. 102, no. 18, pp. 3586–3616, 1998.

[156] A. D. Mackerell, M. Feig, and C. L. Brooks, "Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations," *Journal of Computational Chemistry*, vol. 25, no. 11, pp. 1400–1415, 2004.

[157] N. Foloppe and A. D. MacKerell, Jr., "All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data," *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 86–104, 2000.

[158] A. D. MacKerell and N. K. Banavali, "All-atom empirical force field for nucleic acids: II. Application to molecular dynamics simulations of DNA and RNA in solution," *Journal of Computational Chemistry*, vol. 21, no. 2, pp. 105–120, 2000.

[159] D. H. de Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schäfer, X. Periole, D. P. Tieleman, and S. J. Marrink, "Improved Parameters for the Martini Coarse-Grained Protein Force Field," *Journal of Chemical Theory and Computation*, vol. 9, no. 1, pp. 687–697, 2013.

[160] L. Monticelli, S. K. Kandasamy, X. Periole, R. G. Larson, D. P. Tieleman, and S.-J. Marrink, "The MARTINI Coarse-Grained Force Field: Extension to Proteins," *Journal of Chemical Theory and Computation*, vol. 4, no. 5, pp. 819–834, 2008.

[161] S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman, and A. H. de Vries, "The MARTINI Force Field: Coarse Grained Model for Biomolecular Simulations," *The Journal of Physical Chemistry B*, vol. 111, no. 27, pp. 7812–7824, 2007.

[162] S. J. Marrink, A. H. de Vries, and A. E. Mark, "Coarse Grained Model for Semiquantitative Lipid Simulations," *The Journal of Physical Chemistry B*, vol. 108, no. 2, pp. 750–760, 2004.

[163] J. J. Uusitalo, H. I. Ingólfsson, P. Akhshi, D. P. Tieleman, and S. J. Marrink, "Martini Coarse-Grained Force Field: Extension to DNA," *Journal of Chemical Theory and Computation*, vol. 11, no. 8, pp. 3932–3945, 2015.

[164] J. J. Uusitalo, H. I. Ingólfsson, S. J. Marrink, and I. Faustino, "Martini Coarse-Grained Force Field: Extension to RNA," *Biophysical Journal*, vol. 113, no. 2, pp. 246–256, 2017.

[165] C. A. López, A. J. Rzepiela, A. H. de Vries, L. Dijkhuizen, P. H. Hünenberger, and S. J. Marrink, "Martini Coarse-Grained Force Field: Extension to Carbohydrates," *Journal of Chemical Theory and Computation*, vol. 5, no. 12, pp. 3195–3210, 2009.

[166] H. Lee, A. H. de Vries, S.-J. Marrink, and R. W. Pastor, "A Coarse-Grained Model for Polyethylene Oxide and Polyethylene Glycol: Conformation and Hydrodynamics," *The Journal of Physical Chemistry B*, vol. 113, no. 40, pp. 13186–13194, 2009.

[167] B. A. Wells and A. L. Chaffee, "Ewald Summation for Molecular Simulations," *Journal of Chemical Theory and Computation*, vol. 11, no. 8, pp. 3684–3695, 2015.

[168] T. Darden, D. York, and L. Pedersen, "Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems," *The Journal of Chemical Physics*, vol. 98, no. 12, pp. 10089–10092, 1993.

[169] P. P. Ewald, "Die Berechnung optischer und elektrostatischer Gitterpotentiale," *Annalen der Physik*, vol. 369, no. 3, pp. 253–287, 1921.

[170] M. Allen and D. Tildesley in *Computer Simulation of Liquids*, p. 219, Oxford University Press, 2017.

[171] J. Barker and R. Watts, "Monte Carlo studies of the dielectric properties of water-like models," *Molecular Physics*, vol. 26, no. 3, pp. 789–792, 1973.

[172] R. Watts, "Monte Carlo studies of liquid water," *Molecular Physics*, vol. 28, no. 4, pp. 1069–1083, 1974.

[173] M. P. Allen and D. J. Tildesley in *Computer Simulation of Liquids*, pp. 235–239, Oxford University Press, 2017.

[174] D. Frenkel and B. Smit in *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 545–550, Academic Press, Inc., 2nd ed., 2001.

[175] R. E. Amaro and A. J. Mulholland, "Multiscale methods in drug design bridge chemical and biological complexity in the search for cures," *Nature Reviews Chemistry*, vol. 2, no. 4, p. 0148, 2018.

[176] M. G. Saunders and G. A. Voth, "Coarse-Graining Methods for Computational Biology," *Annual Review of Biophysics*, vol. 42, no. 1, pp. 73–93, 2013.

[177] A. E. Hafner, J. Krausser, and A. Šarić, "Minimal coarse-grained models for molecular self-organisation in biology," *Current Opinion in Structural Biology*, vol. 58, pp. 43–52, 2019.

[178] S. C. L. Kamerlin, M. Haranczyk, and A. Warshel, "Progress in *Ab Initio* QM/MM Free-Energy Simulations of Electrostatic Energies in Proteins: Accelerated QM/MM Studies of pKa, Redox Reactions and Solvation Free Energies," *The Journal of Physical Chemistry B*, vol. 113, no. 5, pp. 1253–1272, 2009.

[179] K. E. Ranaghan, W. G. Morris, L. Masgrau, K. Senthilkumar, L. O. Johannissen, N. S. Scrutton, J. N. Harvey, F. R. Manby, and A. J. Mulholland, "Ab Initio QM/MM Modeling of the Rate-Limiting Proton Transfer Step in the Deamination of Tryptamine by Aromatic Amine Dehydrogenase," *The Journal of Physical Chemistry B*, vol. 121, no. 42, pp. 9785–9798, 2017.

[180] E. Lence, M. W. van der Kamp, C. González-Bello, and A. J. Mulholland, "QM/MM simulations identify the determinants of catalytic activity differences between type II dehydroquinase enzymes," *Organic & Biomolecular Chemistry*, vol. 16, no. 24, pp. 4443–4455, 2018.

[181] H. Koldsø and M. S. P. Sansom, "Organization and Dynamics of Receptor Proteins in a Plasma Membrane," *Journal of the American Chemical Society*, vol. 137, no. 46, pp. 14694–14704, 2015.

[182] A. K. Chaurasia, A. M. Rukangu, M. K. Philen, G. D. Seidel, and E. C. Freeman, "Evaluation of bending modulus of lipid bilayers using undulation and orientation analysis," *Physical Review E*, vol. 97, no. 3, p. 032421, 2018.

[183] M. Chavent, A. L. Duncan, P. Rassam, O. Birkholz, J. Hélie, T. Reddy, D. Beliaev, B. Hambly, J. Piehler, C. Kleanthous, and M. S. P. Sansom, "How nanoscale protein interactions determine the mesoscale dynamic organisation of bacterial outer membrane proteins," *Nature Communications*, vol. 9, no. 1, p. 2846, 2018.

[184] G. Guo, L. Zhang, and Y. Zhang, "Molecular dynamics study of the infiltration of lipid-wrapping C60 and polyhydroxylated single-walled nanotubes into lipid bilayers," *Frontiers of Physics*, vol. 10, no. 2, pp. 177–186, 2015.

[185] E. J. Wallace and M. S. P. Sansom, "Carbon Nanotube/Detergent Interactions via Coarse-Grained Molecular Dynamics," *Nano Letters*, vol. 7, no. 7, pp. 1923–1928, 2007.

[186] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, "The Protein Data Bank," *Nucleic Acids Research*, vol. 28, no. 1, pp. 235–242, 2000.

[187] S. K. Burley, H. M. Berman, C. Bhikadiya, C. Bi, L. Chen, L. DiCostanzo, C. Christie, K. Dalenberg, J. M. Duarte, S. Dutta, Z. Feng, S. Ghosh, D. S. Goodsell, R. K. Green, V. Guranović, D. Guzenko, B. P. Hudson, T. Kalro, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Periskova, A. Prlić, C. Randle, A. Rose, P. Rose, R. Sala, M. Sekharan, C. Shao, L. Tan, Y.-P. Tao, Y. Valasatava, M. Voigt, J. Westbrook, J. Woo, H. Yang, J. Young, M. Zhuravleva, and C. Zardecki, "RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy," *Nucleic Acids Research*, vol. 47, no. D1, pp. D464–D474, 2019.

[188] X. Periole, M. Cavalli, S.-J. Marrink, and M. A. Ceruso, "Combining an Elastic Network With a Coarse-Grained Molecular Force Field: Structure, Dynamics, and Intermolecular Recognition," *Journal of Chemical Theory and Computation*, vol. 5, no. 9, pp. 2531–2543, 2009.

[189] S. O. Yesylevskyy, L. V. Schäfer, D. Sengupta, and S. J. Marrink, "Polarizable Water Model for the Coarse-Grained MARTINI Force Field," *PLoS Computational Biology*, vol. 6, no. 6, p. e1000810, 2010.

[190] J. Michalowsky, L. V. Schäfer, C. Holm, and J. Smiatek, "A refined polarizable water model for the coarse-grained MARTINI force field with long-range electrostatic interactions," *The Journal of Chemical Physics*, vol. 146, no. 5, p. 054501, 2017.

[191] C. Arnarez, J. J. Uusitalo, M. F. Masman, H. I. Ingólfsson, D. H. de Jong, M. N. Melo, X. Periole, A. H. de Vries, and S. J. Marrink, "Dry Martini, a Coarse-Grained Force Field for Lipid Membrane Simulations with Implicit Solvent," *Journal of Chemical Theory and Computation*, vol. 11, no. 1, pp. 260–275, 2015.

[192] "PDB Statistics: Growth of Structures from 3DEM Experiments Released per Year." URL https://www.rcsb.org/stats/growth/em (Accessed on 01/11/2019).

[193] R. C. Edgar, "MUSCLE: a multiple sequence alignment method with reduced time and space complexity," *BMC Bioinformatics*, vol. 5, no. 1, p. 113, 2004.

[194] R. C. Edgar, "MUSCLE: multiple sequence alignment with high accuracy and high throughput," *Nucleic Acids Research*, vol. 32, no. 5, pp. 1792–1797, 2004.

[195] A. Fiser and A. Šali, "Modeller: Generation and Refinement of Homology-Based Protein Structure Models," *Methods in Enzymology*, vol. 374, pp. 461–491, 2003.

[196] M. Shen and A. Sali, "Statistical potential for assessment and prediction of protein structures," *Protein science: a publication of the Protein Society*, vol. 15, no. 11, pp. 2507–24, 2006.

[197] P. Benkert, S. C. E. Tosatto, and D. Schomburg, "QMEAN: A comprehensive scoring function for model quality assessment," *Proteins*, vol. 71, no. 1, pp. 261–277, 2008.

[198] F. Benedetti and C. Loison, "Comment to: Martini straight: Boosting performance using a shorter cutoff and GPUs by D.H. de Jong, S. Baoukina, H.I. Ingólfsson, and S.J.Marrink," *Computer Physics Communications*, vol. 228, pp. 146–151, 2018.

[199] A. Mancino, W. Glass, Y. Yan, P. Biggin, and D. Bowie, "Alternatively-spliced isoforms of the cardiac Nav1.5 sodium channel modify channel activation by distinct structural mechanisms," *unpublished, in prep.*

[200] T. Kiss, "Biogenesis of small nuclear RNPs.," *Journal of cell science*, vol. 117, no. Pt 25, pp. 5949–51, 2004.

[201] M. C. Wahl, C. L. Will, and R. Lührmann, "The Spliceosome: Design Principles of a Dynamic RNP Machine," *Cell*, vol. 136, no. 4, pp. 701–718, 2009.

[202] T. E. P. Consortium, "An integrated encyclopedia of DNA elements in the human genome," *Nature*, vol. 489, no. 7414, pp. 57–74, 2012.

[203] S. Djebali, C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J. Lagarde, W. Lin, F. Schlesinger, C. Xue, G. K. Marinov, J. Khatun, B. A. Williams, C. Zaleski, J. Rozowsky, M. Röder, F. Kokocinski, R. F. Abdelhamid, T. Alioto, I. Antoshechkin, M. T. Baer, N. S. Bar, P. Batut, K. Bell, I. Bell, S. Chakrabortty, X. Chen, J. Chrast, J. Curado, T. Derrien, J. Drenkow, E. Dumais, J. Dumais, R. Duttagupta, E. Falconnet, M. Fastuca, K. Fejes-Toth, P. Ferreira, S. Foissac, M. J. Fullwood, H. Gao, D. Gonzalez, A. Gordon, H. Gunawardena, C. Howald, S. Jha, R. Johnson, P. Kapranov, B. King, C. Kingswood, O. J. Luo, E. Park, K. Persaud, J. B. Preall, P. Ribeca, B. Risk, D. Robyr, M. Sammeth, L. Schaffer, L.-H. See, A. Shahab, J. Skancke, A. M. Suzuki, H. Takahashi, H. Tilgner, D. Trout, N. Walters, H. Wang, J. Wrobel, Y. Yu, X. Ruan, Y. Hayashizaki, J. Harrow,

M. Gerstein, T. Hubbard, A. Reymond, S. E. Antonarakis, G. Hannon, M. C. Giddings, Y. Ruan, B. Wold, P. Carninci, R. Guigó, and T. R. Gingeras, "Landscape of transcription in human cells," *Nature*, vol. 489, no. 7414, pp. 101–108, 2012.

[204] A. Kalsotra and T. A. Cooper, "Functional consequences of developmentally regulated alternative splicing," *Nature Reviews Genetics*, vol. 12, no. 10, pp. 715–729, 2011.

[205] S. Chaudhary, W. Khokhar, I. Jabre, A. S. N. Reddy, L. J. Byrne, C. M. Wilson, and N. H. Syed, "Alternative Splicing and Protein Diversity: Plants Versus Animals," *Frontiers in Plant Science*, vol. 10, p. 708, 2019.

[206] A. Chatelier, L. Dahllund, A. Eriksson, J. Krupp, and M. Chahine, "Biophysical Properties of Human Nav1.7 Splice Variants and Their Regulation by Protein Kinase A," *Journal of Neurophysiology*, vol. 99, no. 5, pp. 2241–2250, 2008.

[207] E. V. Fletcher, D. M. Kullmann, and S. Schorge, "Alternative Splicing Modulates Inactivation of Type 1 Voltage-gated Sodium Channels by Toggling an Amino Acid in the First S3-S4 Linker," *Journal of Biological Chemistry*, vol. 286, no. 42, pp. 36700–36708, 2011.

[208] A. Liavas, G. Lignani, and S. Schorge, "Conservation of alternative splicing in sodium channels reveals evolutionary focus on release from inactivation and structural insights into gating," *The Journal of Physiology*, vol. 595, no. 16, pp. 5671–5685, 2017.

[209] R. Onkal, J. H. Mattis, S. P. Fraser, J. K. J. Diss, D. Shao, K. Okuse, and M. B. A. Djamgoz, "Alternative splicing of Nav1.5: an electrophysiological comparison of 'neonatal' and 'adult' isoforms and critical involvement of a lysine residue," *Journal of Cellular Physiology*, vol. 216, no. 3, pp. 716–726, 2008.

[210] H. Shen, Q. Zhou, X. Pan, Z. Li, J. Wu, and N. Yan, "Structure of a eukaryotic voltage-gated sodium channel at near-atomic resolution," *Science*, vol. 355, no. 6328, p. eaal4326, 2017.

[211] The UniProt Consortium, "The Universal Protein Resource (UniProt)," *Nucleic Acids Research*, vol. 47, no. D1, pp. D506–D515, 2018.

[212] A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore, and T. Schwede, "SWISS-MODEL: homology modelling of protein structures and complexes," *Nucleic acids research*, vol. 46, no. W1, pp. W296–W303, 2018.

[213] Schrödinger LLC, "The PyMOL Molecular Graphics System," 2015.

[214] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, "GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers," *SoftwareX*, vol. 12, pp. 19–25, 2015.

[215] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, "Improved side-chain torsion potentials for the Amber ff99SB protein force field," *Proteins*, vol. 78, no. 8, pp. 1950–1958, 2010.

[216] J. P. M. Jämbeck and A. P. Lyubartsev, "Derivation and Systematic Validation of a Refined All-Atom Force Field for Phosphatidylcholine Lipids," *The Journal of Physical Chemistry B*, vol. 116, no. 10, pp. 3164–3179, 2012.

[217] J. P. M. Jämbeck and A. P. Lyubartsev, "An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes," *Journal of Chemical Theory and Computation*, vol. 8, no. 8, pp. 2938–2948, 2012.

[218] J. P. M. Jämbeck and A. P. Lyubartsev, "Another Piece of the Membrane Puzzle: Extending Slipids Further," *Journal of Chemical Theory and Computation*, vol. 9, no. 1, pp. 774–784, 2013.

[219] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, vol. 79, no. 2, pp. 926–935, 1983.

[220] C. Kandt, W. L. Ash, and D. Peter Tieleman, "Setting up and running molecular dynamics simulations of membrane proteins," *Methods*, vol. 41, no. 4, pp. 475–488, 2007.

[221] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," *Journal of Computational Chemistry*, vol. 18, no. 12, pp. 1463–1472, 1997.

[222] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations," *Journal of Computational Chemistry*, vol. 32, no. 10, pp. 2319–2327, 2011.

[223] R. Gowers, M. Linke, J. Barnoud, T. Reddy, M. Melo, S. Seyler, J. Domański, D. Dotson, S. Buchoux, I. Kenney, and O. Beckstein, "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations," in *Proceedings of the 15th Python in Science Conference*, pp. 98–105, 2016.

[224] W. Humphrey, A. Dalke, and K. Schulten, "VMD: visual molecular dynamics.," *Journal of molecular graphics*, vol. 14, no. 1, pp. 33–8, 27–8, 1996.

[225] M. Waskom, O. Botvinnik, D. O'Kane, P. Hobson, J. Ostblom, S. Lukauskas, D. C. Gemperline, T. Augspurger, Y. Halchenko, J. B. Cole, J. Warmenhoven, J. de Ruiter, C. Pye, S. Hoyer, J. Vanderplas, S. Villalba, G. Kunter, E. Quintero, P. Bachant, M. Martin, K. Meyer, A. Miles, Y. Ram, T. Brunner, T. Yarkoni, M. L. Williams, C. Evans, C. Fitzgerald, Brian, and A. Qalieh, "mwaskom/seaborn: v0.9.0 (July 2018)," 2018.

[226] J. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, no. 3, pp. 90–95, 2007.

[227] P. G. DeCaen, V. Yarov-Yarovoy, E. M. Sharp, T. Scheuer, and W. A. Catterall, "Sequential formation of ion pairs during activation of a sodium channel voltage sensor," *Proceedings of the National Academy of Sciences*, vol. 106, no. 52, pp. 22498–22503, 2009.

[228] L. Delemotte, M. A. Kasimova, M. L. Klein, M. Tarek, and V. Carnevale, "Free-energy landscape of ion-channel voltage-sensor domain activation," *Proceedings of the National Academy of Sciences*, vol. 112, no. 1, pp. 124–129, 2015.

[229] B. Chanda and F. Bezanilla, "Tracking Voltage-dependent Conformational Changes in Skeletal Muscle Sodium Channel during Activation," *The Journal of General Physiology*, vol. 120, no. 5, pp. 629–645, 2002.

[230] Z. Varga, W. Zhu, A. R. Schubert, J. L. Pardieck, A. Krumholz, E. J. Hsu, M. A. Zaydman, J. Cui, and J. R. Silva, "Direct Measurement of Cardiac Na^+ Channel Conformations Reveals Molecular Pathologies of Inherited Mutations," *Circulation: Arrhythmia and Electrophysiology*, vol. 8, no. 5, pp. 1228–1239, 2015.

[231] W. G. Glass, A. L. Duncan, and P. C. Biggin, "Computational investigation of voltage-gated sodium channel $\beta 3$ subunit dynamics," *Frontiers in Molecular Biosciences*, vol. 7, p. 40, 2020.

[232] G. A. Patino and L. L. Isom, "Electrophysiology and beyond: Multiple roles of Na^+ channel β subunits in development and disease," *Neuroscience Letters*, vol. 486, no. 2, pp. 53–59, 2010.

[233] W. Zhu, T. L. Voelker, Z. Varga, A. R. Schubert, J. M. Nerbonne, and J. R. Silva, "Mechanisms of noncovalent β subunit regulation of Nav channel gating," *The Journal of General Physiology*, vol. 149, no. 8, pp. 813–831, 2017.

[234] R. Weiss, R. Horn, and W. Catterall, "Functional differences between two classes of sodium channels in developing rat skeletal muscle," *Science*, vol. 233, no. 4761, pp. 361–364, 1986.

[235] A. H. Kanellopoulos, J. Koenig, H. Huang, M. Pyrski, Q. Millet, S. Lolignier, T. Morohashi, S. J. Gossage, M. Jay, J. E. Linley, G. Baskozos, B. M. Kessler, J. J. Cox, A. C. Dolphin, F. Zufall, J. N. Wood, and J. Zhao, "Mapping protein interactions of sodium channel Nav1.7 using epitopetagged genetargeted mice," *The EMBO Journal*, vol. 37, no. 3, pp. 427–445, 2018.

[236] G. Rougon and O. Hobert, "New insights into the diversity and function of neuronal immunoglbulin superfamily molecules," *Annual Review of Neuroscience*, vol. 26, no. 1, pp. 207–238, 2003.

[237] R. Danne, C. Poojari, H. Martinez-Seara, S. Rissanen, F. Lolicato, T. Róg, and I. Vattulainen, "*doGlycans* Tools for Preparing Carbohydrate Structures for Atomistic Simulations of Glycoproteins, Glycolipids, and Carbohydrate Polymers for GROMACS," *Journal of Chemical Information and Modeling*, vol. 57, no. 10, pp. 2401–2406, 2017.

[238] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, and R. J. Woods, "GLYCAM06: A generalizable biomolecular force field. Carbohydrates," *Journal of Computational Chemistry*, vol. 29, no. 4, pp. 622–655, 2008.

[239] M. Bansal, S. Kumart, and R. Velavan, "HELANAL: A Program to Characterize Helix Geometry in Proteins," *Journal of Biomolecular Structure and Dynamics*, vol. 17, no. 5, pp. 811–819, 2000.

[240] M. T. De Marothy and A. Elofsson, "Marginally hydrophobic transmembrane α-helices shaping membrane protein folding," *Protein science : a publication of the Protein Society*, vol. 24, no. 7, pp. 1057–74, 2015.

[241] S. Yoshida, M. Shibata, S. Yamamoto, M. Hagihara, N. Asai, M. Takahashi, S. Mizutani, T. Muramatsu, and K. Kadomatsu, "Homo-oligomer formation by basigin, an immunoglobulin superfamily member, via its N-terminal immunoglobulin domain," *European Journal of Biochemistry*, vol. 267, no. 14, pp. 4372–4380, 2000.

[242] J. U. Bowie, "Helix packing in membrane proteins," *Journal of Molecular Biology*, vol. 272, no. 5, pp. 780–789, 1997.

[243] A. Varki, "Biological roles of oligosaccharides: all of the theories are correct," *Glycobiology*, vol. 3, no. 2, pp. 97–130, 1993.

[244] R. A. Dwek, "Glycobiology: Toward Understanding the Function of Sugars," *Chemical Reviews*, vol. 96, no. 2, pp. 683–720, 1996.

[245] C. Reily, T. J. Stewart, M. B. Renfrow, and J. Novak, "Glycosylation in health and disease," *Nature Reviews Nephrology*, vol. 15, no. 6, pp. 346–366, 2019.

[246] R. G. Spiro, "Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds," *Glycobiology*, vol. 12, no. 4, pp. 43R–56R, 2002.

[247] H. S. Lee, Y. Qi, and W. Im, "Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study," *Scientific Reports*, vol. 5, no. 1, p. 8926, 2015.

[248] J. H. Lee, G. Ozorowski, and A. B. Ward, "Cryo-EM structure of a native, fully glycosylated, cleaved HIV-1 envelope trimer," *Science (New York, N.Y.)*, vol. 351, no. 6277, pp. 1043–8, 2016.

[249] B. L. Foley, M. B. Tessier, and R. J. Woods, "Carbohydrate force fields," *Wiley interdisciplinary reviews. Computational molecular science*, vol. 2, no. 4, pp. 652–697, 2012.

[250] X. Xiong, Z. Chen, B. P. Cossins, Z. Xu, Q. Shao, K. Ding, W. Zhu, and J. Shi, "Force fields and scoring functions for carbohydrate simulation," *Carbohydrate Research*, vol. 401, pp. 73–81, 2015.

[251] I. Alibay and R. A. Bryce, "Ring Puckering Landscapes of Glycosaminoglycan-Related Monosaccharides from Molecular Dynamics Simulations," *Journal of Chemical Information and Modeling*, vol. 59, no. 11, pp. 4729–4741, 2019.

[252] A. S. Hill, A. Nishino, K. Nakajo, G. Zhang, J. R. Fineman, M. E. Selzer, Y. Okamura, and E. C. Cooper, "Ion Channel Clustering at the Axon Initial Segment and Node of Ranvier Evolved Sequentially in Early Chordates," *PLoS Genetics*, vol. 4, no. 12, p. e1000317, 2008.

[253] J. Sajman, M. Trus, D. Atlas, and E. Sherman, "The L-type Voltage-Gated Calcium Channel co-localizes with Syntaxin 1A in nano-clusters at the plasma membrane," *Scientific Reports*, vol. 7, no. 1, p. 11350, 2017.

[254] E. Servili, M. Trus, D. Maayan, and D. Atlas, "β-Subunit of the voltage-gated Ca^{2+} channel Cav1.2 drives signaling to the nucleus via H-Ras," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 115, no. 37, pp. E8624–E8633, 2018.

[255] C. Gu and Y. Gu, "Clustering and Activity Tuning of Kv1 Channels in Myelinated Hippocampal Axons," *Journal of Biological Chemistry*, vol. 286, no. 29, pp. 25835–25847, 2011.

[256] S. Malin, J. Nerbonne, and M. M. Tamkun, "Delayed Rectifier K^+ Currents, IK, Are Encoded by Kv2 α-Subunits and Regulate Tonic Firing in Mammalian Sympathetic Neurons," *Journal of Neuroscience*, vol. 22, no. 23, pp. 10094–10105, 2013.

[257] J. Clatot, A. Ziyadeh-Isleem, S. Maugenre, I. Denjoy, H. Liu, G. Dilanian, S. N. Hatem, I. Deschênes, A. Coulombe, P. Guicheney, and N. Neyroud, "Dominant-negative effect of SCN5A N-terminal mutations through the interaction of Nav1.5 α-subunits," *Cardiovascular Research*, vol. 96, no. 1, pp. 53–63, 2012.

[258] L. L. Isom, K. S. De Jongh, and W. A. Catterall, "Auxiliary subunits of voltage-gated ion channels," *Neuron*, vol. 12, no. 6, pp. 1183–1194, 1994.

[259] M. Sheng and E. Kim, "Ion channel associated proteins," *Current Opinion in Neurobiology*, vol. 6, no. 5, pp. 602–608, 1996.

[260] J. L. Robertson, "The lipid bilayer membrane and its protein constituents," *The Journal of General Physiology*, vol. 150, no. 11, pp. 1472–1483, 2018.

[261] J. L. Sampaio, M. J. Gerl, C. Klose, C. S. Ejsing, H. Beug, K. Simons, and A. Shevchenko, "Membrane lipidome of an epithelial cell line," *Proceedings of the National Academy of Sciences*, vol. 108, no. 5, pp. 1903–1907, 2011.

[262] M. Sud, E. Fahy, D. Cotter, A. Brown, E. A. Dennis, C. K. Glass, A. H. Merrill, R. C. Murphy, C. R. H. Raetz, D. W. Russell, and S. Subramaniam, "LMSD: LIPID MAPS structure database," *Nucleic Acids Research*, vol. 35, no. Database, pp. D527–D532, 2007.

[263] D. Marquardt, B. Geier, and G. Pabst, "Asymmetric lipid membranes: towards more realistic model systems," *Membranes*, vol. 5, no. 2, pp. 180–96, 2015.

[264] P. F. Devaux and R. Morris, "Transmembrane Asymmetry and Lateral Domains in Biological Membranes," *Traffic*, vol. 5, no. 4, pp. 241–246, 2004.

[265] M. Mondal, B. Mesmin, S. Mukherjee, and F. R. Maxfield, "Sterols Are Mainly in the Cytoplasmic Leaflet of the Plasma Membrane and the Endocytic Recycling Compartment in CHO Cells," *Molecular Biology of the Cell*, vol. 20, no. 2, pp. 581–588, 2009.

[266] H. I. Ingólfsson, M. N. Melo, F. J. Van Eerden, C. Arnarez, C. A. Lopez, T. A. Wassenaar, X. Periole, A. H. De Vries, D. P. Tieleman, and S. J. Marrink, "Lipid organization of the plasma membrane," *Journal of the American Chemical Society*, vol. 136, no. 41, 2014.

[267] H. Koldsø, D. Shorthouse, J. Hélie, and M. S. P. Sansom, "Lipid Clustering Correlates with Membrane Curvature as Revealed by Molecular Simulations of Complex Lipid Bilayers," *PLOS Computational Biology*, vol. 10, no. 10, p. e1003911, 2014.

[268] H. Koldsø, T. Reddy, P. W. Fowler, A. L. Duncan, and M. S. P. Sansom, "Membrane Compartmentalization Reduces the Mobility of Lipids and Proteins within a Model Plasma Membrane," *The journal of physical chemistry. B*, vol. 120, no. 34, pp. 8873–8881, 2016.

[269] K. A. Marino, D. Prada-Gracia, D. Provasi, and M. Filizola, "Impact of Lipid Composition and Receptor Conformation on the Spatio-temporal Organization of μ-Opioid Receptors in a Multi-component Plasma Membrane Model," *PLOS Computational Biology*, vol. 12, no. 12, p. e1005240, 2016.

[270] S. J. Marrink, A. H. de Vries, and D. P. Tieleman, "Lipids on the move: Simulations of membrane pores, domains, stalks and curves," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1788, no. 1, pp. 149–168, 2009.

[271] R. C. Bernardi, M. C. R. Melo, and K. Schulten, "Enhanced sampling techniques in molecular dynamics simulations of biological systems," *Biochimica et biophysica acta*, vol. 1850, no. 5, pp. 872–877, 2015.

[272] T. A. Wassenaar, H. I. Ingólfsson, R. A. Böckmann, D. P. Tieleman, and S. J. Marrink, "Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations," *Journal of Chemical Theory and Computation*, vol. 11, no. 5, pp. 2144–2155, 2015.

[273] H. Berendsen, D. van der Spoel, and R. van Drunen, "GROMACS: A message-passing parallel molecular dynamics implementation," *Computer Physics Communications*, vol. 91, no. 1-3, pp. 43–56, 1995.

[274] E. Lindahl, B. Hess, and D. van der Spoel, "GROMACS 3.0: a package for molecular simulation and trajectory analysis," *Journal of Molecular Modeling*, vol. 7, no. 8, pp. 306–317, 2001.

[275] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, "GROMACS: Fast, flexible, and free," *Journal of Computational Chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005.

[276] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, "GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation," *Journal of Chemical Theory and Computation*, vol. 4, no. 3, pp. 435–447, 2008.

[277] S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl, "GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit," *Bioinformatics (Oxford, England)*, vol. 29, no. 7, pp. 845–54, 2013.

[278] P. Szilárd, M. J. Abraham, C. Kutzner, B. Hess, and E. Lindahl, "Tackling Exascale Software Challenges in Molecular Dynamics Simulations with GROMACS," 2015.

[279] A. Hagberg, D. Schilt, and P. Swart, "Exploring Network Structure, Dynamics, and Function using NetworkX," in *Exploring Network Structure, Dynamics, and Function using NetworkX in Proceedings of the 7th Python in Science conference* (G. Varoquaux, T. Vaught, and J. Millman, eds.), pp. 11–15, 2008.

[280] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, 2011.

[281] P. S. Schmalhorst, F. Deluweit, R. Scherrers, C.-P. Heisenberg, and M. Sikora, "Overcoming the Limitations of the MARTINI Force Field in Simulations of Polysaccharides," *Journal of Chemical Theory and Computation*, vol. 13, no. 10, pp. 5039–5053, 2017.

[282] K. Simons and E. Ikonen, "Functional rafts in cell membranes," *Nature*, vol. 387, no. 6633, pp. 569–572, 1997.

[283] H.-K. Wong, T. Sakurai, F. Oyama, K. Kaneko, K. Wada, H. Miyazaki, M. Kurosawa, B. De Strooper, P. Saftig, and N. Nukina, "β Subunits of voltage-gated sodium channels are novel substrates of β-site amyloid precursor protein-cleaving enzyme (BACE1) and γ-secretase," *The Journal of biological chemistry*, vol. 280, no. 24, pp. 23009–17, 2005.

[284] C. Yuan, C. Velázquez-Marrero, A. Bernardo, and S. N. Treistman, "Lipids Modulate the Increase of BK Channel Calcium Sensitivity by the β1 Subunit," *PLoS ONE*, vol. 9, no. 9, p. e107917, 2014.

[285] J. L. Salzer, "Polarized Domains of Myelinated Axons," *Neuron*, vol. 40, no. 2, pp. 297–318, 2003.

[286] S. A. Freeman, A. Desmazières, D. Fricker, C. Lubetzki, and N. Sol-Foulon, "Mechanisms of sodium channel clustering and its influence on axonal impulse conduction," *Cellular and molecular life sciences : CMLS*, vol. 73, no. 4, pp. 723–35, 2016.

[287] J. A. Black, N. Frézel, S. D. Dib-Hajj, and S. G. Waxman, "Expression of Nav1.7 in DRG Neurons Extends from Peripheral Terminals in the Skin to Central Preterminal Branches and Terminals in the Dorsal Horn," *Molecular Pain*, vol. 8, pp. 1744–8069–8–82, 2012.

[288] K.-J. Chang and M. N. Rasband, "Excitable Domains of Myelinated Nerves: Axon Initial Segments and Nodes of Ranvier," *Current Topics in Membranes*, vol. 72, pp. 159–192, 2013.

[289] D. Shao, K. Okuse, and M. B. Djamgoz, "Protein-protein interactions involving voltage-gated sodium channels: Post-translational regulation, intracellular trafficking and functional expression," *The International Journal of Biochemistry & Cell Biology*, vol. 41, no. 7, pp. 1471–1481, 2009.

[290] F. J. Berendt, K.-S. Park, and J. S. Trimmer, "Multisite phosphorylation of voltage-gated sodium channel α subunits from rat brain," *Journal of proteome research*, vol. 9, no. 4, pp. 1976–84, 2010.

[291] N. Makita, P. B. Bennett, and A. L. George, "Molecular determinants of beta 1 subunit-induced gating modulation in voltage-dependent Na^+ channels.," *The Journal of neuroscience : the official journal of the Society for Neuroscience*, vol. 16, no. 22, pp. 7117–27, 1996.

[292] Y. Qu, J. C. Rogers, S. F. Chen, K. A. McCormick, T. Scheuer, and W. A. Catterall, "Functional roles of the extracellular segments of the sodium channel α subunit in voltage-dependent gating and modulation by β1 subunits," *The Journal of biological chemistry*, vol. 274, no. 46, pp. 32647–54, 1999.

[293] J. S. Lowe, O. Palygin, N. Bhasin, T. J. Hund, P. A. Boyden, E. Shibata, M. E. Anderson, and P. J. Mohler, "Voltage-gated Nav channel targeting in the heart requires an ankyrin-G dependent cellular pathway," *The Journal of cell biology*, vol. 180, no. 1, pp. 173–86, 2008.

[294] D. Schmidt, Q.-X. Jiang, and R. MacKinnon, "Phospholipids and the origin of cationic gating charges in voltage sensors," *Nature*, vol. 444, no. 7120, pp. 775–779, 2006.

[295] Y. Xu, Y. Ramu, and Z. Lu, "Removal of phospho-head groups of membrane lipids immobilizes voltage sensors of K^+ channels," *Nature*, vol. 451, no. 7180, pp. 826–9, 2008.

[296] M. Milescu, F. Bosmans, S. Lee, A. A. Alabi, J. I. Kim, and K. J. Swartz, "Interactions between lipids and voltage sensor paddles detected with tarantula toxins," *Nature structural & molecular biology*, vol. 16, no. 10, pp. 1080–5, 2009.

[297] H. Zheng, W. Liu, L. Y. Anderson, and Q.-X. Jiang, "Lipid-dependent gating of a voltage-gated potassium channel," *Nature Communications*, vol. 2, no. 1, p. 250, 2011.

[298] Z. A. Sands and M. S. P. Sansom, "How does a voltage sensor interact with a lipid bilayer? Simulations of a potassium channel domain," *Structure (London, England: 1993)*, vol. 15, no. 2, pp. 235–44, 2007.

[299] M. Andersson, J. A. Freites, D. J. Tobias, and S. H. White, "Structural Dynamics of the S4 Voltage-Sensor Helix in Lipid Bilayers Lacking Phosphate Groups," *The Journal of Physical Chemistry B*, vol. 115, no. 27, pp. 8732–8738, 2011.

[300] L. Delemotte, M. Tarek, M. L. Klein, C. Amaral, and W. Treptow, "Intermediate states of the Kv1.2 voltage sensor from atomistic molecular dynamics simulations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 15, pp. 6109–14, 2011.

[301] D. W. Hilgemann, S. Feng, and C. Nasuhoglu, "The complex and intriguing lives of PIP_2 with ion channels and transporters," *Science's STKE: signal transduction knowledge environment*, vol. 2001, no. 111, p. re19, 2001.

[302] E. Sezgin, I. Levental, S. Mayor, and C. Eggeling, "The mystery of membrane organization: composition, regulation and roles of lipid rafts," *Nature Reviews Molecular Cell Biology*, vol. 18, no. 6, pp. 361–374, 2017.

[303] M. M. Lozano, Z. Liu, E. Sunnick, A. Janshoff, K. Kumar, and S. G. Boxer, "Colocalization of the Ganglioside GM1 and Cholesterol Detected by Secondary Ion Mass Spectrometry," *Journal of the American Chemical Society*, vol. 135, no. 15, pp. 5620–5630, 2013.

[304] T. Harder, P. Scheiffele, P. Verkade, and K. Simons, "Lipid domain structure of the plasma membrane revealed by patching of membrane components," *The Journal of cell biology*, vol. 141, no. 4, pp. 929–42, 1998.

[305] T. L. Yarbrough, T. Lu, H.-C. Lee, and E. F. Shibata, "Localization of Cardiac Sodium Channels in Caveolin-Rich Membrane Domains," *Circulation Research*, vol. 90, no. 4, pp. 443–449, 2002.

[306] A. Maguy, T. Hebert, and S. Nattel, "Involvement of lipid rafts and caveolae in cardiac ion channel function," *Cardiovascular Research*, vol. 69, no. 4, pp. 798–807, 2006.

[307] J. Rosenbluth, "Intramembranous particle distribution at the node of Ranvier and adjacent axolemma in myelinated axons of the frog brain," *Journal of Neurocytology*, vol. 5, no. 6, pp. 731–745, 1976.

[308] P. W. Fowler, J. Hélie, A. Duncan, M. Chavent, H. Koldsø, and M. S. P. Sansom, "Membrane stiffness is modified by integral membrane proteins," *Soft Matter*, vol. 12, no. 37, pp. 7792–7803, 2016.

[309] H. I. Ingólfsson, T. S. Carpenter, H. Bhatia, P.-T. Bremer, S. J. Marrink, and F. C. Lightstone, "Computational Lipidomics of the Neuronal Plasma Membrane," *Biophysical Journal*, vol. 113, no. 10, pp. 2271–2280, 2017.

[310] L. Iversen, S. Mathiasen, J. B. Larsen, and D. Stamou, "Membrane curvature bends the laws of physics and chemistry," *Nature Chemical Biology*, vol. 11, no. 11, pp. 822–825, 2015.

[311] G. Guigas and M. Weiss, "Effects of protein crowding on membrane systems," *Biochimica et Biophysica Acta (BBA) - Biomembranes*, vol. 1858, no. 10, pp. 2441–2450, 2016.

[312] M. R. Fernandez-Fernandez, D. Ruiz-Garcia, E. Martin-Solana, F. J. Chichon, J. L. Carrascosa, and J.-J. Fernandez, "3D electron tomography of brain tissue unveils distinct Golgi structures that sequester cytoplasmic contents in neurons," *Journal of cell science*, vol. 130, no. 1, pp. 83–89, 2017.

[313] P. Bassereau, R. Jin, T. Baumgart, M. Deserno, R. Dimova, V. A. Frolov, P. V. Bashkirov, H. Grubmüller, R. Jahn, H. J. Risselada, L. Johannes, M. M. Kozlov, R. Lipowsky, T. J. Pucadyil, W. F. Zeno, J. C. Stachowiak, D. Stamou, A. Breuer, L. Lauritsen, C. Simon, C. Sykes, G. A. Voth, and T. R. Weikl, "The 2018 biomembrane curvature and remodeling roadmap," *Journal of Physics D: Applied Physics*, vol. 51, no. 34, p. 343001, 2018.

[314] J. Körner, J. Meents, J.-P. Machtens, and A. Lampert, "β1 subunit stabilises sodium channel Nav1.7 against mechanical stress," *The Journal of Physiology*, 2018.

[315] M. Chung Moh, L. Hoon Lee, and S. Shen, "Cloning and characterization of hepaCAM, a novel Ig-like cell adhesion molecule suppressed in human hepatocellular carcinoma," *Journal of Hepatology*, vol. 42, no. 6, pp. 833–841, 2005.

[316] L. Favre-Kontula, A. Rolland, L. Bernasconi, M. Karmirantzou, C. Power, B. Antonsson, and U. Boschert, "GlialCAM, an immunoglobulin-like cell adhesion molecule is expressed in glial cells of the central nervous system," *Glia*, vol. 56, no. 6, pp. 633–645, 2008.

[317] B. C. Willis, D. Ponce-Balbuena, and J. Jalife, "Protein assemblies of sodium and inward rectifier potassium channels control cardiac excitability and arrhythmogenesis," *American journal of physiology. Heart and circulatory physiology*, vol. 308, no. 12, pp. H1463–73, 2015.

[318] R. G. Utrilla, P. Nieto-Marín, S. Alfayate, D. Tinaquero, M. Matamoros, M. Pérez-Hernández, S. Sacristán, L. Ondo, R. de Andrés, F. J. Díez-Guerra, J. Tamargo, E. Delpón, and R. Caballero, "Kir2.1-Nav1.5 Channel Complexes Are Differently Regulated than Kir2.1 and Nav1.5 Channels Alone," *Frontiers in Physiology*, vol. 8, 2017.

[319] J. Birch, D. Axford, J. Foadi, A. Meyer, A. Eckhardt, Y. Thielmann, and I. Moraes, "The fine art of integral membrane protein crystallisation," *Methods*, vol. 147, pp. 150–162, 2018.

[320] K. Murata and M. Wolf, "Cryo-electron microscopy for structural analysis of dynamic biological macromolecules," *Biochimica et Biophysica Acta (BBA) - General Subjects*, vol. 1862, no. 2, pp. 324–334, 2018.

[321] M. Igaev, C. Kutzner, L. V. Bock, A. C. Vaiana, and H. Grubmüller, "Automated cryo-EM structure refinement using correlation-driven molecular dynamics," *eLife*, vol. 8, 2019.

[322] J. Ognjenović, R. Grisshammer, and S. Subramaniam, "Frontiers in Cryo Electron Microscopy of Complex Macromolecular Assemblies," *Annual Review of Biomedical Engineering*, vol. 21, no. 1, pp. 395–415, 2019.

[323] R. D. Lins and P. H. Hünenberger, "A new GROMOS force field for hexopyranose-based carbohydrates," *Journal of Computational Chemistry*, vol. 26, no. 13, pp. 1400–1412, 2005.

[324] H. S. Hansen and P. H. Hünenberger, "A reoptimized GROMOS force field for hexopyranose-based carbohydrates accounting for the relative free energies of ring conformers, anomers, epimers, hydroxymethyl rotamers, and glycosidic linkage conformers," *Journal of Computational Chemistry*, vol. 32, no. 6, pp. 998–1032, 2011.

[325] S. Jo, T. Kim, V. G. Iyer, and W. Im, "CHARMM-GUI: A web-based graphical user interface for CHARMM," *Journal of Computational Chemistry*, vol. 29, no. 11, pp. 1859–1865, 2008.

[326] B. M. Sattelle, J. Shakeri, and A. Almond, "Does Microsecond Sugar Ring Flexing Encode 3D-Shape and Bioactivity in the Heparanome?," *Biomacromolecules*, vol. 14, no. 4, pp. 1149–1159, 2013.

[327] C. D. Castro, T. Klose, I. Speciale, R. Lanzetta, A. Molinaro, J. L. V. Etten, and M. G. Rossmann, "Structure of the chlorovirus PBCV-1 major capsid glycoprotein determined by combining crystallographic and carbohydrate molecular modeling approaches," *Proceedings of the National Academy of Sciences*, vol. 115, no. 1, pp. E44–E52, 2018.

[328] C. De Castro, G. A. Duncan, D. Garozzo, A. Molinaro, L. Sturiale, M. Tonetti, and J. L. Van Etten, "Biophysical Approaches to Solve the Structures of the Complex Glycan Shield of Chloroviruses," pp. 237–257, Springer, Singapore, 2018.

[329] S. J. Marrink, V. Corradi, P. C. Souza, H. I. Ingólfsson, D. P. Tieleman, and M. S. Sansom, "Computational Modeling of Realistic Cell Membranes," *Chemical Reviews*, vol. 119, no. 9, pp. 6184–6226, 2019.

[330] R.-S. Chen, T.-C. Deng, T. Garcia, Z. M. Sellers, and P. M. Best, "Calcium channel γ subunits: a functionally diverse protein family," *Cell Biochemistry and Biophysics*, vol. 47, no. 2, pp. 178–186, 2007.

[331] E. C. Twomey, M. V. Yelshanskaya, and A. I. Sobolevsky, "Structural and functional insights into transmembrane AMPA receptor regulatory protein complexes," *The Journal of general physiology*, 2019.

[332] Y. Zhao, S. Chen, C. Yoshioka, I. Baconguis, and E. Gouaux, "Architecture of fully occupied GluA2 AMPA receptor-TARP complex elucidated by cryo-EM," *Nature*, vol. 536, no. 7614, pp. 108–111, 2016.

[333] S. Chen, Y. Zhao, Y. Wang, M. Shekhar, E. Tajkhorshid, and E. Gouaux, "Activation and Desensitization Mechanism of AMPA Receptor-TARP Complex by Cryo-EM," *Cell*, vol. 170, no. 6, pp. 1234–1246.e14, 2017.

[334] S. Tomita, H. Adesnik, M. Sekiguchi, W. Zhang, K. Wada, J. R. Howe, R. A. Nicoll, and D. S. Bredt, "Stargazin modulates AMPA receptor gating and trafficking by distinct domains," *Nature*, vol. 435, no. 7045, pp. 1052–1058, 2005.

[335] A. D. Milstein, W. Zhou, S. Karimzadegan, D. S. Bredt, and R. A. Nicoll, "TARP Subtypes Differentially and Dose-Dependently Control Synaptic AMPA Receptor Gating," *Neuron*, vol. 55, no. 6, pp. 905–918, 2007.

[336] I. Riva, C. Eibl, R. Volkmer, A. L. Carbone, and A. J. Plested, "Control of AMPA receptor activity by the extracellular loops of auxiliary proteins," *eLife*, vol. 6, 2017.

[337] E. Jeworutzki, T. López-Hernández, X. Capdevila-Nortes, S. Sirisi, L. Bengtsson, M. Montolio, G. Zifarelli, T. Arnedo, C. Müller, U. Schulte, V. Nunes, A. Martínez, T. Jentsch, X. Gasull, M. Pusch, and R. Estévez, "GlialCAM, a Protein Defective in a Leukodystrophy, Serves as a ClC-2 Cl Channel Auxiliary Subunit," *Neuron*, vol. 73, no. 5, p. 951, 2012.

[338] S. Sirisi, X. Elorza-Vidal, T. Arnedo, M. Armand-Ugón, G. Callejo, X. Capdevila-Nortes, T. López-Hernández, U. Schulte, A. Barrallo-Gimeno, V. Nunes, X. Gasull, and R. Estévez, "Depolarization causes the formation of a ternary complex between GlialCAM, MLC1 and ClC-2 in astrocytes: implications in megalencephalic leukoencephalopathy," *Human Molecular Genetics*, vol. 26, no. 13, pp. 2436–2450, 2017.

[339] J.-P. David, J. I. Stas, N. Schmitt, and E. Bocksteins, "Auxiliary KCNE subunits modulate both homotetrameric Kv2.1 and heterotetrameric Kv2.1/Kv6.4 channels," *Scientific Reports*, vol. 5, no. 1, p. 12813, 2015.

[340] G. W. Abbott, "KCNE4 and KCNE5: K(+) channel regulation and cardiac arrhythmogenesis.," *Gene*, vol. 593, no. 2, pp. 249–60, 2016.

[341] S. M. Crump and G. W. Abbott, "Arrhythmogenic KCNE gene variants: current knowledge and future challenges," *Frontiers in Genetics*, vol. 5, p. 3, 2014.

[342] G. F. Contreras, A. Neely, O. Alvarez, C. Gonzalez, and R. Latorre, "Modulation of BK channel voltage gating by different auxiliary β subunits," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 109, no. 46, pp. 18991–6, 2012.

[343] Q. Li and J. Yan, "Modulation of BK Channel Function by Auxiliary β and γ Subunits," *International review of neurobiology*, vol. 128, pp. 51–90, 2016.

[344] E. S. A. Humphries and C. Dart, "Neuronal and Cardiovascular Potassium Channels as Therapeutic Drug Targets: Promise and Pitfalls," *Journal of biomolecular screening*, vol. 20, no. 9, pp. 1055–73, 2015.

[345] A. S. Haworth and W. J. Brackenbury, "Emerging roles for multifunctional ion channel auxiliary subunits in cancer," *Cell Calcium*, vol. 80, pp. 125–140, 2019.

[346] D. Frenkel and B. Smit in *Understanding Molecular Simulation: From Algorithms to Applications*, pp. 74–76, Academic Press, Inc., 2nd ed., 2001.

[347] H. J. C. Berendsen in *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*, pp. 406–413, Cambridge University Press, 2007.

THE AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE LICENSE TERMS AND CONDITIONS

Oct 09, 2020

This Agreement between William Glass ("You") and The American Association for the Advancement of Science ("The American Association for the Advancement of Science") consists of your license details and the terms and conditions provided by The American Association for the Advancement of Science and Copyright Clearance Center.

License Number	4918120468261
License date	Sep 29, 2020
Licensed Content Publisher	The American Association for the Advancement of Science
Licensed Content Publication	Science
Licensed Content Title	Structure of the human voltage-gated sodium channel $Na_v1.4$ in complex with β1
Licensed Content Author	Xiaojing Pan,Zhangqiang Li,Qiang Zhou,Huaizong Shen,Kun Wu,Xiaoshuang Huang,Jiaofeng Chen,Juanrong Zhang,Xuechen Zhu,Jianlin Lei,Wei Xiong,Haipeng Gong,Bailong Xiao,Nieng Yan
Licensed Content Date	Oct 19, 2018
Licensed Content Volume	362
Licensed Content Issue	6412

Volume number	362
Issue number	6412
Type of Use	Thesis / Dissertation
Requestor type	Scientist/individual at a research institution
Format	Print and electronic
Portion	Figure
Number of figures/tables	1
Title	Mr
Institution name	University of Oxford
Expected presentation date	Sep 2020
Portions	Figure 1F
Requestor Location	Mr. William Glass Balliol College Oxford, OX1 3BJ United Kingdom Attn: University of Oxford
Total	0.00 GBP

Terms and Conditions

American Association for the Advancement of Science TERMS AND CONDITIONS

Regarding your request, we are pleased to grant you non-exclusive, non-transferable permission, to republish the AAAS material identified above in your work identified above, subject to the terms and conditions herein. We must be contacted for permission for any uses other than those specifically identified in your request above.

The following credit line must be printed along with the AAAS material: "From [Full Reference Citation]. Reprinted with permission from AAAS."

All required credit lines and notices must be visible any time a user accesses any part of the AAAS material and must appear on any printed copies and authorized user might make.

This permission does not apply to figures / photos / artwork or any other content or materials included in your work that are credited to non-AAAS sources. If the requested material is sourced to or references non-AAAS sources, you must obtain authorization from that source as well before using that material. You agree to hold harmless and indemnify AAAS against any claims arising from your use of any content in your work that is credited to non-AAAS sources.

If the AAAS material covered by this permission was published in Science during the years 1974 - 1994, you must also obtain permission from the author, who may grant or withhold permission, and who may or may not charge a fee if permission is granted. See original article for author's address. This condition does not apply to news articles.

The AAAS material may not be modified or altered except that figures and tables may be modified with permission from the author. Author permission for any such changes must be secured prior to your use.

Whenever possible, we ask that electronic uses of the AAAS material permitted herein include a hyperlink to the original work on AAAS's website (hyperlink may be embedded in the reference citation).

AAAS material reproduced in your work identified herein must not account for more than 30% of the total contents of that work.

AAAS must publish the full paper prior to use of any text.

AAAS material must not imply any endorsement by the American Association for the Advancement of Science.

This permission is not valid for the use of the AAAS and/or Science logos.

AAAS makes no representations or warranties as to the accuracy of any information contained in the AAAS material covered by this permission, including any warranties of

merchantability or fitness for a particular purpose.

If permission fees for this use are waived, please note that AAAS reserves the right to charge for reproduction of this material in the future.

Permission is not valid unless payment is received within sixty (60) days of the issuance of this permission. If payment is not received within this time period then all rights granted herein shall be revoked and this permission will be considered null and void.

In the event of breach of any of the terms and conditions herein or any of CCC's Billing and Payment terms and conditions, all rights granted herein shall be revoked and this permission will be considered null and void.

AAAS reserves the right to terminate this permission and all rights granted herein at its discretion, for any purpose, at any time. In the event that AAAS elects to terminate this permission, you will have no further right to publish, publicly perform, publicly display, distribute or otherwise use any matter in which the AAAS content had been included, and all fees paid hereunder shall be fully refunded to you. Notification of termination will be sent to the contact information as supplied by you during the request process and termination shall be immediate upon sending the notice. Neither AAAS nor CCC shall be liable for any costs, expenses, or damages you may incur as a result of the termination of this permission, beyond the refund noted above.

This Permission may not be amended except by written document signed by both parties.

The terms above are applicable to all permissions granted for the use of AAAS material. Below you will find additional conditions that apply to your particular type of use.

FOR A THESIS OR DISSERTATION
If you are using figure(s)/table(s), permission is granted for use in print and electronic versions of your dissertation or thesis. A full text article may be used in print versions only of a dissertation or thesis.

Permission covers the distribution of your dissertation or thesis on demand by ProQuest / UMI, provided the AAAS material covered by this permission remains in situ.

If you are an Original Author on the AAAS article being reproduced, please refer to your License to Publish for rules on reproducing your paper in a dissertation or thesis.

FOR JOURNALS:
Permission covers both print and electronic versions of your journal article, however the AAAS material may not be used in any manner other than within the context of your article.

FOR BOOKS/TEXTBOOKS:
If this license is to reuse figures/tables, then permission is granted for non-exclusive world rights in all languages in both print and electronic formats (electronic formats are defined below).

If this license is to reuse a text excerpt or a full text article, then permission is granted for non-exclusive world rights in English only. You have the option of securing either print or electronic rights or both, but electronic rights are not automatically granted and do garner additional fees. Permission for translations of text excerpts or full text articles into other languages must be obtained separately.

Licenses granted for use of AAAS material in electronic format books/textbooks are valid only in cases where the electronic version is equivalent to or substitutes for the print version of the book/textbook. The AAAS material reproduced as permitted herein must remain in situ and must not be exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit), except in the case of permitted textbook companions as noted below.

You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

If your book is an academic textbook, permission covers the following companions to your textbook, provided such companions are distributed only in conjunction with your textbook at no additional cost to the user:

- Password-protected website
- Instructor's image CD/DVD and/or PowerPoint resource
- Student CD/DVD

All companions must contain instructions to users that the AAAS material may be used for non-commercial, classroom purposes only. Any other uses require the prior written permission from AAAS.

If your license is for the use of AAAS Figures/Tables, then the electronic rights granted herein permit use of the Licensed Material in any Custom Databases that you distribute the electronic versions of your textbook through, so long as the Licensed Material remains within the context of a chapter of the title identified in your request and cannot be downloaded by a user as an independent image file.

Rights also extend to copies/files of your Work (as described above) that you are required to provide for use by the visually and/or print disabled in compliance with state and federal laws.

This permission only covers a single edition of your work as identified in your request.

FOR NEWSLETTERS:
Permission covers print and/or electronic versions, provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit)

FOR ANNUAL REPORTS:
Permission covers print and electronic versions provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately (for example, if permission covers the use of a full text article, the article may not be offered for access or for purchase as a stand-alone unit)

FOR PROMOTIONAL/MARKETING USES:
Permission covers the use of AAAS material in promotional or marketing pieces such as information packets, media kits, product slide kits, brochures, or flyers limited to a single print run. The AAAS Material may not be used in any manner which implies endorsement or promotion by the American Association for the Advancement of Science (AAAS) or Science of any product or service. AAAS does not permit the reproduction of its name, logo or text on promotional literature.

If permission to use a full text article is permitted, The Science article covered by this permission must not be altered in any way. No additional printing may be set onto an article copy other than the copyright credit line required above. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to, the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto the article copies.

Additionally, article copies must be a freestanding part of any information package (i.e. media kit) into which they are inserted. They may not be physically attached to anything, such as an advertising insert, or have anything attached to them, such as a sample product. Article copies must be easily removable from any kits or informational packages in which they are used. The only exception is that article copies may be inserted into three-ring binders.

FOR CORPORATE INTERNAL USE:
The AAAS material covered by this permission may not be altered in any way. No additional printing may be set onto an article copy other than the required credit line. Any alterations must be approved in advance and in writing by AAAS. This includes, but is not limited to the placement of sponsorship identifiers, trademarks, logos, rubber stamping or self-adhesive stickers onto article copies.

If you are making article copies, copies are restricted to the number indicated in your request and must be distributed only to internal employees for internal use.

If you are using AAAS Material in Presentation Slides, the required credit line must be visible on the slide where the AAAS material will be reprinted

If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher." Access to any such CD, DVD, Flash Drive or Web page must be restricted to your organization's employees only.

FOR CME COURSE and SCIENTIFIC SOCIETY MEETINGS:
Permission is restricted to the particular Course, Seminar, Conference, or Meeting indicated in your request. If this license covers a text excerpt or a Full Text Article, access to the reprinted AAAS material must be restricted to attendees of your event only (if you have been granted electronic rights for use of a full text article on your website, your website must be password protected, or access restricted so that only attendees can access the content on your site).

If you are using AAAS Material on a CD, DVD, Flash Drive, or the World Wide Web, you must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written

permission from the publisher."

FOR POLICY REPORTS:
These rights are granted only to non-profit organizations and/or government agencies. Permission covers print and electronic versions of a report, provided the required credit line appears in both versions and provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately.

FOR CLASSROOM PHOTOCOPIES:
Permission covers distribution in print copy format only. Article copies must be freestanding and not part of a course pack. They may not be physically attached to anything or have anything attached to them.

FOR COURSEPACKS OR COURSE WEBSITES:
These rights cover use of the AAAS material in one class at one institution. Permission is valid only for a single semester after which the AAAS material must be removed from the Electronic Course website, unless new permission is obtained for an additional semester. If the material is to be distributed online, access must be restricted to students and instructors enrolled in that particular course by some means of password or access control.

FOR WEBSITES:
You must include the following notice in any electronic versions, either adjacent to the reprinted AAAS material or in the terms and conditions for use of your electronic products: "Readers may view, browse, and/or download material for temporary copying purposes only, provided these uses are for noncommercial personal purposes. Except as provided by law, this material may not be further reproduced, distributed, transmitted, modified, adapted, performed, displayed, published, or sold in whole or in part, without prior written permission from the publisher."

Permissions for the use of Full Text articles on third party websites are granted on a case by case basis and only in cases where access to the AAAS Material is restricted by some means of password or access control. Alternately, an E-Print may be purchased through our reprints department (brocheleau@rockwaterinc.com).

REGARDING FULL TEXT ARTICLE USE ON THE WORLD WIDE WEB IF YOU ARE AN 'ORIGINAL AUTHOR' OF A SCIENCE PAPER

If you chose "Original Author" as the Requestor Type, you are warranting that you are one of authors listed on the License Agreement as a "Licensed content author" or that you are acting on that author's behalf to use the Licensed content in a new work that one of the authors listed on the License Agreement as a "Licensed content author" has written.

Original Authors may post the 'Accepted Version' of their full text article on their personal or on their University website and not on any other website. The 'Accepted Version' is the version of the paper accepted for publication by AAAS including changes resulting from peer review but prior to AAAS's copy editing and production (in other words not the AAAS published version).

FOR MOVIES / FILM / TELEVISION:
Permission is granted to use, record, film, photograph, and/or tape the AAAS material in connection with your program/film and in any medium your program/film may be shown or heard, including but not limited to broadcast and cable television, radio, print, world wide web, and videocassette.

The required credit line should run in the program/film's end credits.

FOR MUSEUM EXHIBITIONS:
Permission is granted to use the AAAS material as part of a single exhibition for the duration of that exhibit. Permission for use of the material in promotional materials for the exhibit must be cleared separately with AAAS (please contact us at permissions@aaas.org).

FOR TRANSLATIONS:
Translation rights apply only to the language identified in your request summary above.

The following disclaimer must appear with your translation, on the first page of the article, after the credit line: "This translation is not an official translation by AAAS staff, nor is it endorsed by AAAS as accurate. In crucial matters, please refer to the official English-language version originally published by AAAS."

FOR USE ON A COVER:
Permission is granted to use the AAAS material on the cover of a journal issue, newsletter issue, book, textbook, or annual report in print and electronic formats provided the AAAS material reproduced as permitted herein remains in situ and is not exploited separately

By using the AAAS Material identified in your request, you agree to abide by all the terms and conditions herein.

Questions about these terms can be directed to the AAAS Permissions department permissions@aaas.org.

Other Terms and Conditions:

v 2

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

PARTIES:

1. **Cambridge University Press** [CompanyNumber] (Licensor); and
2. **William Glass** (Licensee).

Thank you for your recent permission request. Some permission requests for use of material published by the Licensor, such as this one, are now being facilitated by PLSclear.

Set out in this licence cover sheet (the **Licence Cover Sheet**) are the principal terms under which Licensor has agreed to license certain Licensed Material (as defined below) to Licensee. The terms in this Licence Cover Sheet are subject to the attached General Terms and Conditions, which together with this Licence Cover Sheet constitute the licence agreement (the **Licence**) between Licensor and Licensee as regards the Licensed Material. The terms set out in this Licence Cover Sheet take precedence over any conflicting provision in the General Terms and Conditions.

Free Of Charge Licence Terms

Licence Date:	29/09/2020
PLSclear Ref No:	42670

The Licensor

Company name:	Cambridge University Press
Address:	University Printing House Shaftesbury Road Cambridge CB2 8BS GB

The Licensee

Licensee Contact Name:	William Glass
Licensee Address:	

Licensed Material

title:	Simulating the Physical World Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics
ISBN:	9780521835275
publisher:	Cambridge University Press
Are you requesting permission to reuse the cover of the publication?	No
Figure number & title	15.1

Page numbers	406
Are you requesting permission to reuse your own work?	Yes

For Use In Licensee's Publication(s)

usage type	Book, Journal, Magazine or Academic Paper-Thesis / Dissertation
Will your dissertation be placed in an online repository?	Yes
Author	William Glass
Language	English
Title of dissertation/thesis	Multiscale Molecular Dynamics to Investigate Sodium Ion Channel Oligomerisation
University or institution	University of Oxford
Unlimited circulation?	Yes

Rights Granted

Exclusivity:	Non-Exclusive
Format:	Thesis / Dissertation
Language:	English
Territory:	World
Duration:	Lifetime of Licensee's Edition
Maximum Circulation:	Unlimited

GENERAL TERMS AND CONDITIONS

1. Definitions and Interpretation

1.1 Capitalised words and expressions in these General Terms and Conditions have the meanings given to them in the Licence Cover Sheet.

1.2 In this Licence any references (express or implied) to statutes or provisions are references to those statutes or provisions as amended or re-enacted from time to time. The term including will be construed as illustrative, without limiting the sense or scope of the words preceding it. A reference to in writing or written includes faxes and email. The singular includes the plural and vice versa.

2. Grant of Rights

2.1 The Licensor grants to Licensee the non-exclusive right to use the Licensed Material as specified in the Licence Cover Sheet.

2.2 The rights licensed to Licensee under this Licence do not include the right to use any third party copyright material incorporated in the Licensed Material. Licensee should check the Licensed Material carefully and seek permission for the use of any such third party copyright material from the relevant copyright owner(s).

2.3 Unless otherwise stated in the Licence Cover Sheet, the Licensed Material may be:

2.3.1 subjected to minor editing, including for the purposes of creating alternative formats to provide access for a beneficiary person (provided that any such editing does not amount to derogatory treatment); and/or

2.3.2 used for incidental promotional use (such as online retail providers' search facilities).

2.4 Save as expressly permitted in this Licence or as otherwise permitted by law, no use or modification of the Licensed Material may be made by Licensee without Licensor's prior written permission.

3. Copyright Notice and Acknowledgement

3.1 Licensee must ensure that the following notices and acknowledgements are reproduced prominently alongside each reproduction by Licensee of the Licensed Material:

3.1.1 the title and author of the Licensed Material;

3.1.2 the copyright notice included in the Licensed Material; and

3.1.3 the statement "Reproduced with permission of The Licensor through PLSclear."

4. Reversion of Rights

4.1 The rights licensed to Licensee under this Licence will terminate immediately and automatically upon the earliest of the following events to occur:

4.1.1 the Licensed Material not being used by Licensee within 18 months of the Licence Date;

4.1.2 expiry of the Licence Duration; or

4.1.3 the Maximum Circulation being reached.

5. Miscellaneous

5.1 By using the Licensed Material, Licensee will be deemed to have accepted all the terms and conditions contained in this Licence.

5.2 This Licence contains the entire understanding and agreement of the parties relating to its subject matter and supersedes in all respects any previous or other existing arrangements, agreements or understandings between the parties whether oral or written in relation to its subject matter.

5.3 Licensee may not assign this Licence or any of its rights or obligations hereunder to any third party without Licensor's prior written consent.

5.4 This Licence is governed by and shall be construed in accordance with the laws of England and Wales and the parties hereby irrevocably submit to the non-exclusive jurisdiction of the Courts of England and Wales as regards any claim, dispute or matter arising under or in relation to this Licence.