Elucidating the permeation mechanisms of the TRPV and TRPM ion channels using in silico electrophysiology

Ives, Callum Matthew

Award date: 2023

Licence: CC BY-NC-ND

Link to publication
Elucidating the permeation mechanisms of the TRPV and TRPM ion channels using in silico electrophysiology

Callum Matthew Ives

Submitted in satisfaction of the requirements for the Degree of Doctor of Philosophy

Primary Supervisor: Professor Ulrich Zachariae

Secondary Supervisor: Professor Geoffrey J. Barton

Division of Computational Biology

School of Life Sciences

University of Dundee

Scotland, UK

May, 2023
Declaration

I declare that this thesis is based on results obtained from research which I have personally carried out in the School of Life Sciences at the University of Dundee. I declare that the entire thesis is my own composition. Any work other than my own is clearly stated in the text and acknowledged with reference to any relevant investigators or contributors. This thesis has never been presented previously, in whole or in part, for the award of any higher degree. I have consulted all the references cited within the text of this thesis.

Signed:

Mr Callum Matthew Ives
Date: Tuesday 10th January 2023

I confirm that Callum Matthew Ives has carried out the research under my supervision and that they have fulfilled the conditions of the relevant Ordinance and Regulations of the University of Dundee, thereby qualifying them to submit this thesis in application for the degree of Doctor of Philosophy.

Signed:

Professor Ulrich Zachariae
Date: Tuesday 10th January 2023
# Table of Contents

**Publications** vii  

**Acknowledgements** viii  

**Abstract** x  

**Abbreviations** xi  

**List of Figures** xv  

**List of Tables** xix  

1 **Introduction** 1  
1.1 The role of ion channels in cellular physiology 1  
1.2 The transient receptor potential channel superfamily of cation ion channels 3  
1.3 TRPV: the vanilloid subfamily of TRP channels 6  
1.3.1 Structural features 6  
1.3.2 Implications in health and disease 7  
1.4 TRPM: the melastatin subfamily of TRP channels 8  
1.4.1 Structural features 9  
1.4.2 Implications in health and disease 10  
1.5 Project aims 11  

2 **Theory and methods** 13  
2.1 Core principles of molecular dynamics simulations 13  
2.2 Fixed point-charge forcefields 14  
2.3 Divalent cations in fixed point-charge forcefields 16  
2.4 Simulation protocol 21  
2.4.1 Thermodynamic ensembles, thermostats, and barostats 22  
2.4.2 *In silico* electrophysiology using molecular dynamics simulations 23
# TABLE OF CONTENTS

## 3 Ca\(^{2+}\)-selective permeation mechanism in the TRPV family of ion channels 27

3.1 Associated publications ........................................... 27

3.1.1 Authors’ Contributions ........................................... 27

3.2 Introduction ....................................................... 27

3.3 Methods ........................................................ 30

3.3.1 TRPV system construction ....................................... 30

3.3.2 Molecular dynamics simulations details ..................... 31

3.3.3 Simulation analysis ............................................. 33

3.4 Results .......................................................... 37

3.4.1 Continuous permeation of Ca\(^{2+}\) and Na\(^{+}\) in open-state TRPV5 and TRPV6 channels ............................................. 37

3.4.2 Pore cation binding sites and their preference for Ca\(^{2+}\)-binding ............................................. 39

3.4.3 A highly co-operative knock-on mechanism between three cation binding sites underpins selective Ca\(^{2+}\) permeation in TRPV channels ............................................. 45

3.4.4 Cation permeation in non-selective TRPV channels shows a lower degree of co-operativity ............................................. 51

3.4.5 Structural features distinguishing Ca\(^{2+}\)-selective from non-selective permeation ............................................. 54

3.4.6 Ca\(^{2+}\)-selective permeation is not strongly linked to the solvation states of permeation cations ............................................. 58

3.5 Discussion ....................................................... 60

## 4 Monovalent-selective permeation mechanism in the TRPM family of ion channels 63

4.1 Associated publications ........................................... 63

4.1.1 Authors’ Contributions ........................................... 63

4.2 Introduction ....................................................... 63

4.3 Methods ........................................................ 65

4.3.1 TRPM5 system construction ....................................... 65
# TABLE OF CONTENTS

4.4 Molecular dynamics simulations details ........................................... 66
4.4.1 Simulation analysis ................................................................. 69

4.5 Results ...................................................................................... 70
4.5.1 Cation conductance of the TRPM5 channel in di-cationic solutions 70
4.5.2 Low-voltage simulations in di-cationic solutions show the exclusive permeation of Na\(^+\) through TRPM5 ........................................ 71
4.5.3 The ion selectivity of TRPM5 is strongly voltage-dependent ....... 72
4.5.4 Mechanistic insights into ion permeation in TRPM5 from monocationic solutions ................................................................. 74
4.5.5 Solvation profiles of cations during channel permeation .......... 76
4.5.6 Selectivity for monovalent cations is based on permeation cooperativity between two binding sites ................................. 79
4.5.7 Why does the inner cavity form an attractive site for monovalent cations but a repulsive site for divalent cations? .................. 82

4.6 Discussion .................................................................................. 84

5 Comparison of the CompEL and EAEF techniques for in silico electrophysiology ................................................................. 87
5.1 Introduction ............................................................................... 87
5.2 Methods .................................................................................... 89
5.2.1 TRPV3 system construction ...................................................... 89
5.2.2 Molecular dynamics simulations details ................................. 89
5.2.3 Simulation analysis ................................................................. 90
5.3 Results ...................................................................................... 91
5.3.1 Case study: TRPV3 using the CHARMM36m force field ....... 91
5.3.2 Case study: TRPV3 using the AMBER19SB force field .......... 104
5.4 Discussion ................................................................................ 113

6 Concluding remarks ................................................................. 116
6.1 Summary of thesis ................................................................. 116
6.2 Future directions ........................................... 117

7 Bibliography .................................................. 120
Publications

Associated publications


Non-associated publications


Acknowledgements

This research was funded by the Medical Research Council as part of the Doctoral Training Programme in Quantitative and Interdisciplinary approaches to Biomedical Science [grant number MR/N013735/1].

Firstly, I would like to sincerely thank my supervisors, Prof Ulrich Zachariae and Prof Geoffrey Barton. To Geoff, a huge thank you for my introduction to the field of computational biology and for all your words of advice. To Ulrich, I wish to express my sincerest gratitude for giving me the opportunity to undertake this project. Your presence and guidance during my time in Dundee has provided me with a wealth of knowledge and skills that I will carry with me throughout my future career.

I would also like to say a massive thank you to past and present members of the Zachariae group, including Mr Neil Thomson, Mr Dominik Gurvič, Dr Lysbeth Antonides, Mr Alp Tegin Şahin, Mr Peter Ezzat, and Mr Tim Spankie. The positive and fun atmosphere within the group created an enjoyable culture to learn new skills and conduct exciting science.

Thank you to past and present PhD students within the Division of Computational Biology for creating such a fun place to work, including Dr Marcus Bage, Dr Maxim Tsenkov, and Mr Javier Sánchez Útges. I would also like to thank Dr Stuart MacGowan for his time and patience in introducing me to bioinformatics and computational biology when I first started in the Division.

I am grateful for all the hard work of the University of Dundee IT services for their management and maintenance of the High-Performance Computing Cluster. I would like to particularly thank Dr Thomas Walsh for all of his support.

I would like to thank my house mates during this PhD experience, Ms Hannah Donnelly and Mr Jack Hanna, for all of their support on days where simulations or code were proving particularly stubborn.
I am eternally grateful to my parents, grandparents, family, and friends for all of their support during this PhD experience. Your encouragement of me from a young age to explore my curiosity, and to follow it through to a university education, is what has led me here to this today.

Finally, I would like to say a huge thank you to Gemma, for her constant unwavering support and encouragement, without which this work would not have been possible. Your constant, joyful presence helped me through even the toughest days. I will never forget all of the help you provided throughout this time, i gcónaí.
Abstract

Transient Receptor Potential (TRP) channels are a superfamily of ion channels that have been implicated in the aetiology of a number of medical conditions. TRP channels conduct cations, and most members of this superfamily are described as being non-selective between cation types. However, a number of channels in the TRPV and TRPM subfamilies exhibit remarkable selectivity for Ca$^{2+}$ or monovalent ions, respectively. In this thesis, the atomistic mechanisms underpinning selective cation permeation in the TRPV and TRPM subfamilies were investigated using atomistic molecular dynamics simulations.

Specifically, it was found that the mechanism of Ca$^{2+}$-selectivity in the TRPV5 and TRPV6 channels is underpinned by a three site knock-on mechanism between adjacent ion binding sites; each of which possesses greater affinity for Ca$^{2+}$ binding over Na$^+$. Moreover, it was established that the mechanism of monovalent-selective permeation mechanism in the TRPM5 channel is based upon a two site knock-on mechanism for monovalent cations. By contrast, one of the two cation binding sites is abolished for Ca$^{2+}$ within the TRPM5 pore cavity, which explains the channel’s impermeability to Ca$^{2+}$ ions.

Finally, a systematic comparison was conducted between the two main methodologies used to create bio-mimetic, transmembrane voltages in atomistic molecular dynamics simulations on the basis of the TRPV3 channel. These comparisons formed an effort to identify any differences that should be taken into account when designing in silico electrophysiology experiments.

Overall, the research presented in this thesis substantially advances the understanding of the selective permeation mechanisms within two classes of ion channels of significant pharmaceutical interest.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABF</td>
<td>Adaptive Biasing Forces</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine Diphosphate</td>
</tr>
<tr>
<td>AMBER</td>
<td>Assisted Model Building with Energy Refinement</td>
</tr>
<tr>
<td>AMOEBA</td>
<td>Atomic Multipole Optimised Energetics for Biomolecular Applications</td>
</tr>
<tr>
<td>AMPA receptor</td>
<td>(\alpha)-amino-3-hydroxy-5-methyl-4-isoxazolepropionic Acid Receptor</td>
</tr>
<tr>
<td>ARD</td>
<td>Ankyrin Repeat Domain</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine Triphosphate</td>
</tr>
<tr>
<td>BCE</td>
<td>Before the Common Era</td>
</tr>
<tr>
<td>Ca(^{2+})</td>
<td>Calcium ion</td>
</tr>
<tr>
<td>CHARMM</td>
<td>Chemistry at Harvard Macromolecular Mechanics</td>
</tr>
<tr>
<td>Cl(^{-})</td>
<td>Chloride ion</td>
</tr>
<tr>
<td>CMAP</td>
<td>Correction Map</td>
</tr>
<tr>
<td>CompEL</td>
<td>Computational Electrophysiology</td>
</tr>
<tr>
<td>CPU</td>
<td>Central Processing Unit</td>
</tr>
<tr>
<td>Cryo-EM</td>
<td>Cryo-Electron Microscopy</td>
</tr>
<tr>
<td>CS</td>
<td>(Ca^{2+})-selective</td>
</tr>
<tr>
<td>CSM</td>
<td>Computed Structural Models</td>
</tr>
<tr>
<td>CSNB</td>
<td>Congenital Stationary Night Blindness</td>
</tr>
<tr>
<td>CTD</td>
<td>C-terminal domain</td>
</tr>
<tr>
<td>°C</td>
<td>Degrees Celsius</td>
</tr>
<tr>
<td>EAEF</td>
<td>External Applied Electric Field</td>
</tr>
<tr>
<td>EBI</td>
<td>European Bioinformatics Institute</td>
</tr>
<tr>
<td>ECCR</td>
<td>Electronic Continuum Correction with Rescaling</td>
</tr>
<tr>
<td>EPV</td>
<td>Extracellular Pore Vestibule</td>
</tr>
<tr>
<td>exSSI</td>
<td>Excess State-Specific Information</td>
</tr>
<tr>
<td>Fe(^{2+})</td>
<td>Iron ion</td>
</tr>
<tr>
<td>fs</td>
<td>Femtosecond</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphics Processing Unit</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>GROMACS</td>
<td>Groningen Machine for Chemical Simulations</td>
</tr>
<tr>
<td>GROMOS</td>
<td>Groningen Molecular Simulation</td>
</tr>
<tr>
<td>HMR</td>
<td>Hydrogen Mass Repartitioning</td>
</tr>
<tr>
<td>IDP</td>
<td>Intrinsically Disordered Protein</td>
</tr>
<tr>
<td>K</td>
<td>Degrees Kelvin</td>
</tr>
<tr>
<td>K⁺</td>
<td>Potassium ion</td>
</tr>
<tr>
<td>KcsA</td>
<td>K⁺ channel of Streptomyces A</td>
</tr>
<tr>
<td>kJ</td>
<td>Kilojoule</td>
</tr>
<tr>
<td>LG</td>
<td>Lower Gate</td>
</tr>
<tr>
<td>LINCS</td>
<td>Linear Constraint Solver</td>
</tr>
<tr>
<td>LJ</td>
<td>Lennard-Jones</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular Dynamics</td>
</tr>
<tr>
<td>Mg²⁺</td>
<td>Magnesium ion</td>
</tr>
<tr>
<td>MHR</td>
<td>Melastatin Homology Region</td>
</tr>
<tr>
<td>MM</td>
<td>Molecular Mechanics</td>
</tr>
<tr>
<td>mol</td>
<td>Mole</td>
</tr>
<tr>
<td>MSD</td>
<td>Macromolecular Structure Database</td>
</tr>
<tr>
<td>mV</td>
<td>Millivolt</td>
</tr>
<tr>
<td>Na⁺</td>
<td>Sodium ion</td>
</tr>
<tr>
<td>nAChR</td>
<td>Nicotinic Acetylcholine Receptor</td>
</tr>
<tr>
<td>nm</td>
<td>Nanometre</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>NPT ensemble</td>
<td>Isothermal-Isobaric Ensemble</td>
</tr>
<tr>
<td>ns</td>
<td>Nanosecond</td>
</tr>
<tr>
<td>NS</td>
<td>Non-selective</td>
</tr>
<tr>
<td>NVE ensemble</td>
<td>Canonical Ensemble</td>
</tr>
<tr>
<td>NVT ensemble</td>
<td>Micro-Canonical Ensemble</td>
</tr>
<tr>
<td>OPLS</td>
<td>Optimised Potentials for Liquid Simulations</td>
</tr>
<tr>
<td>PBC</td>
<td>Periodic Boundary Conditions</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>PCHPD</td>
<td>(4-phenylcyclohexyl)piperazine Derivative</td>
</tr>
<tr>
<td>PDBj</td>
<td>Protein Data Bank Japan</td>
</tr>
<tr>
<td>PH</td>
<td>Pore Helix</td>
</tr>
<tr>
<td>PI(4,5)P2</td>
<td>Phosphatidylinositol 4,5-bisphosphate</td>
</tr>
<tr>
<td>POPC</td>
<td>1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphocholine</td>
</tr>
<tr>
<td>ps</td>
<td>Picosecond</td>
</tr>
<tr>
<td>RCSB</td>
<td>Research Collaboratory for Structural Bioinformatics</td>
</tr>
<tr>
<td>RMSD</td>
<td>Root-Mean-Square-Deviation</td>
</tr>
<tr>
<td>RMSF</td>
<td>Root-Mean-Square-Fluctuation</td>
</tr>
<tr>
<td>RTX</td>
<td>Resiniferatoxin</td>
</tr>
<tr>
<td>RyR1</td>
<td>Ryanodine Receptor 1</td>
</tr>
<tr>
<td>SARS</td>
<td>Severe Acute Respiratory Syndrome</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard Error of Mean</td>
</tr>
<tr>
<td>SF</td>
<td>Selectivity Filter</td>
</tr>
<tr>
<td>sMD</td>
<td>Steered Molecular Dynamics</td>
</tr>
<tr>
<td>SSI</td>
<td>State-Specific Information</td>
</tr>
<tr>
<td>tr</td>
<td>Residence time</td>
</tr>
<tr>
<td>T2DM</td>
<td>Type II Diabetes Mellitus</td>
</tr>
<tr>
<td>TotCorr</td>
<td>Total Correlation</td>
</tr>
<tr>
<td>TRP</td>
<td>Transient Receptor Potential</td>
</tr>
<tr>
<td>TRPA</td>
<td>Transient Receptor Potential, Ankyrin</td>
</tr>
<tr>
<td>TRPC</td>
<td>Transient Receptor Potential, Canonical</td>
</tr>
<tr>
<td>TRPM</td>
<td>Transient Receptor Potential, Melastatin</td>
</tr>
<tr>
<td>TRPML</td>
<td>Transient Receptor Potential, Mucolipin</td>
</tr>
<tr>
<td>TRPN</td>
<td>Transient Receptor Potential, No Me chanoreceptor potential C</td>
</tr>
<tr>
<td>TRPP</td>
<td>Transient Receptor Potential, Polycystic</td>
</tr>
<tr>
<td>TRPS</td>
<td>Transient Receptor Potential, Soromelastatin</td>
</tr>
<tr>
<td>TRPV</td>
<td>Transient Receptor Potential, Vanilloid</td>
</tr>
<tr>
<td>TRPVL</td>
<td>Transient Receptor Potential, Vanilloid-like</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>TRPY</td>
<td>Transient Receptor Potential, Yeast</td>
</tr>
<tr>
<td>US</td>
<td>Umbrella Sampling</td>
</tr>
<tr>
<td>V</td>
<td>Volt</td>
</tr>
<tr>
<td>vdw</td>
<td>Van der Waals</td>
</tr>
<tr>
<td>VSLD</td>
<td>Voltage-Sensing Like Domain</td>
</tr>
<tr>
<td>wwPDB</td>
<td>worldwide Protein Data Bank</td>
</tr>
<tr>
<td>Zn$^{2+}$</td>
<td>Zinc ion</td>
</tr>
<tr>
<td>Å</td>
<td>Angstrom</td>
</tr>
<tr>
<td>µs</td>
<td>Microsecond</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Distribution of *Homo sapiens* drug targets by gene family and distribution of drugs targeting those gene families. ........................................... 3
1.2 Phylogenetic tree of TRP channels of *Homo sapiens*. .................... 4
1.3 Representative structures from each subfamily for the TRP superfamily. . 5
1.4 Domain organisation of a monomer of the TRPV subfamily of ion channels. 6
1.5 Domain organisation of a monomer of the TRPM subfamily of ion channels. 9
2.1 Schematic representing the electron cloud of an atom for fixed point-charge and polarisable force fields. ........................................... 18
2.2 2D representation of Periodic Boundary Conditions. ......................... 23
2.3 Schematic description of the external applied electric field (EAEF) methodology. ................................................................................. 25
2.4 Overview of the Computational Electrophysiology (CompEL) setup. .... 26
3.1 Structure of the truncated construct of TRPV5 of *Oryctolagus cuniculus* used within this work. .................................................... 29
3.2 Schematic of cation binding sites identified in the Ca$^{2+}$-selective TRPV5 channel. ............................................................... 40
3.3 Cation binding sites in TRPV5 identified by PENSA. ....................... 41
3.4 The effect of high voltage on Ca$^{2+}$-selectivity in simulations of TRPV5 in di-cationic solutions. ...................................................... 42
3.5 Pore architecture of TRPV channels from MD simulations, showing the average radius and hydrophobicity of the channel with respect to the relative z coordinate. ................................................. 43
3.6 W583 does not form a functionally important cation binding site in simulations of TRPV5. ............................................................ 44
3.7 Pore architecture of TRPV5 channels from MD simulations with protonated and deprotonated PI(4,5)P$_2$ molecules. ............................ 45
3.8 Occupancy probability and residence times of Ca$^{2+}$ and Na$^+$ cations from simulations of TRPV channels. ...................................... 46
3.9 Permeation state plots of permeating Ca\textsuperscript{2+} and Na\textsuperscript{+} cations through the Ca\textsuperscript{2+}-selective TRPV5 channel. 47

3.10 Excess state-specific information (ex\text{SSI}) between ion binding sites quantifies the degree of co-operativity in the knock-on mechanism of cation permeation in TRPV channels. 49

3.11 Permeation traces of the \( z \)-coordinate of permeating Ca\textsuperscript{2+} cations over time in a lower concentration mono-cationic solution of 25 mM CaCl\textsubscript{2}. 50

3.12 Permeation traces of the \( z \)-coordinate of permeating Ca\textsuperscript{2+} cations over time in a di-cationic solution of 75 mM CaCl\textsubscript{2} and 75 mM NaCl at a lower voltage of -205 mV. 51

3.13 The effect of similar binding site affinities on ex\text{SSI} generated from a model of two consecutive binding sites. 54

3.14 Total correlation of cation permeation between cation binding sites from simulations of TRPV channels. 55

3.15 Backbone dihedral angle distribution of SF residues in TRPV channels. 57

3.16 Multiple sequence alignment of the SF domains of TRPV sequences deposited in the Swiss-Prot database. 58

3.17 Architecture of the four-residue selectivity filter of the investigated TRPV channels. 59

3.18 Solvation state of permeating cations as they permeate through TRPV channels. 60

4.1 Structure and membrane voltage of CompEL simulations of TRPM5. 65

4.2 Overview of the structure of the TRPM5 channel of Danio rerio used in this work. 72

4.3 3D density map of Ca\textsuperscript{2+} cations around the EPV and SF of TRPM5. 73

4.4 Permeation traces of the \( z \)-coordinate of permeating cations over time through TRPM5. 75

4.5 Solvation and PMF profiles of Na\textsuperscript{+}, K\textsuperscript{+} and Ca\textsuperscript{2+} cations through the TRPM5 pore. 77
4.6 Negative logarithmic density profiles of permeating cations along the pore of the TRPM5 at different voltages. ........................................... 78

4.7 Solvation profiles of Na\(^+\) and Ca\(^{2+}\) cations through the TRPM5 pore in a mono-cationic solution. ......................................................... 80

4.8 Solvation profiles of Na\(^+\) and Ca\(^{2+}\) cations through the TRPM5 pore in a di-cationic solution. ......................................................... 81

4.9 SSI of cation transitions between binding sites and the average number of each cation within the pore of TRPM5. ......................................................... 83

4.10 Pore architecture of the monovalent-selective TRPM5 channel and the Ca\(^{2+}\)-selective TRPV5 channel from MD simulations. ...................... 85

5.1 Structure of the TRPV3 ion channel used to study in silico electrophysiology methodologies. ................................................................. 88

5.2 Current-voltage (I-V) plot of TRPV3 in 150 mM NaCl at transmembrane voltages produced using either CompEL or EAEF. ...................... 92

5.3 Calculated conductances from MD simulations of ion permeation through TRPV3 under negative and positive voltages generated by CompEL or EAEF. ................................................................. 93

5.4 Negative logarithmic density profiles for Na\(^+\) ions traversing the pore of TRPV3 under a high transmembrane voltage. ...................... 95

5.5 Negative logarithmic density profiles for Na\(^+\) ions traversing the pore of TRPV3 under a medium transmembrane voltage. ...................... 96

5.6 Negative logarithmic density profiles for Na\(^+\) ions traversing the pore of TRPV3 under a low transmembrane voltage. ...................... 97

5.7 Distribution of RMSD of the TRPV3 protein, pore domain, and ARD. ........ 99

5.8 RMSF of TRPV3 during simulations with a high voltage. ...................... 102

5.9 Average difference between RMSF values of TRPV3 during simulations under both negative and positive voltages, generated by CompEL or EAEF techniques. ................................................................. 103
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.10</td>
<td>Comparison of the relative movement of the ARD region of TRPV3 under a range of transmembrane voltages.</td>
<td>105</td>
</tr>
<tr>
<td>5.11</td>
<td>Calculated conductances from MD simulations of ion permeation through TRPV3 under negative and positive voltages generated by CompEL or EAEF using the AMBER force field.</td>
<td>106</td>
</tr>
<tr>
<td>5.12</td>
<td>Negative logarithmic density profiles for Na$^+$ ions through the pore of TRPV3 under a high transmembrane voltage using the AMBER19SB force field.</td>
<td>108</td>
</tr>
<tr>
<td>5.13</td>
<td>Distribution of RMSD of the TRPV3 protein, pore domain, and ARD from simulations using the AMBER19SB force field.</td>
<td>110</td>
</tr>
<tr>
<td>5.14</td>
<td>RMSF of TRPV3 during simulations using the AMBER19SB force field with a high voltage.</td>
<td>112</td>
</tr>
<tr>
<td>5.15</td>
<td>Comparison of the relative movement of the ARD region of TRPV3 under a range of transmembrane voltages using the AMBER19SB force field.</td>
<td>113</td>
</tr>
</tbody>
</table>
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary of protein constructs used in this study.</td>
</tr>
<tr>
<td>3.2</td>
<td>Summary of simulation details of the Ca$^{2+}$-selective TRPV5 channel.</td>
</tr>
<tr>
<td>3.3</td>
<td>Summary of simulation details of the Ca$^{2+}$-selective TRPV6 channel.</td>
</tr>
<tr>
<td>3.4</td>
<td>Summary of simulation details of the non-selective TRPV3 channel.</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary of simulation details of the non-selective TRPV2 channel.</td>
</tr>
<tr>
<td>3.6</td>
<td>Summary of simulation details of additional control simulations of the Ca$^{2+}$-selective TRPV5 channel.</td>
</tr>
<tr>
<td>3.7</td>
<td>Average time to permeate through the TRPV pore, as defined by the $z$ position between binding sites A and C.</td>
</tr>
<tr>
<td>3.8</td>
<td>Calculated conductances from MD simulations of ion permeation in Ca$^{2+}$-selective TRPV channels.</td>
</tr>
<tr>
<td>3.9</td>
<td>Selectivity ratios of Ca$^{2+}$ and Na$^{+}$ permeation events from simulations of TRPV channels in a di-cationic solution.</td>
</tr>
<tr>
<td>3.10</td>
<td>Calculated $exSSI$ and $exSSI_{norm}$ values of cation transition from binding sites in MD simulations of TRPV channels.</td>
</tr>
<tr>
<td>3.11</td>
<td>Calculated conductances from MD simulations of ion permeation in non-selective TRPV channels.</td>
</tr>
<tr>
<td>3.12</td>
<td>Root mean square fluctuation (RMSF) of the backbone of SF residues of TRPV channels from MD simulations.</td>
</tr>
<tr>
<td>4.1</td>
<td>Summary of CompEL simulation details of the TRPM5 channel.</td>
</tr>
<tr>
<td>4.2</td>
<td>Summary of external applied field simulation details of the TRPM5 channel.</td>
</tr>
<tr>
<td>4.3</td>
<td>Calculated conductances and selectivities from CompEL simulations of ion permeation in the TRPM5 channel.</td>
</tr>
<tr>
<td>4.4</td>
<td>Calculated conductances from applied field simulations of ion permeation in the TRPM5 channel.</td>
</tr>
<tr>
<td>5.2</td>
<td>Summary of calculated voltages from simulations performed in this study utilising CompEL or EAEFs.</td>
</tr>
</tbody>
</table>
5.3 Average RMSD values from simulations of TRPV3 under transmembrane voltages. ................................................................. 100
5.4 Average movement of the ARD domain of TRPV3 under transmembrane voltages. ................................................................. 104
5.5 Average RMSD values from simulations of TRPV3 under transmembrane voltages using the AMBER19SB force field. ................. 111
5.6 Average movement of the ARD domain of TRPV3 under transmembrane voltages using the AMBER19SB force field. ................. 111
Introduction

1.1 The role of ion channels in cellular physiology

A key capability of unicellular and multicellular organisms is the ability to communicate between cells. This process can be achieved in many ways, including the translocation of ions across cellular membranes, mediated by ion channels. For example, the translocation of signalling ions such as Ca\(^{2+}\) [8], and electrical impulses carried by Na\(^{+}\), Ca\(^{2+}\), K\(^{+}\), and Cl\(^{-}\) ions [9]. Ion channels are pore-forming, integral membrane proteins that allow ions to permeate through a pore down their electrochemical gradients.

The significance of ion channels within cellular physiology has long been established. One of the key findings was the determination of the ionic basis of action potentials by Hodgkin and Huxley in 1952 [10]. Despite many researchers having long hypothesised that transmembrane ion flux was dependent on conduction pathways within specific membrane proteins, this could not be directly demonstrated until 1976. At this time, Neher and Sakmann developed the patch clamp technique, which they used to conduct single-channel electrophysiological experiments on the nicotinic acetylcholine receptor (nAChR) from frog muscle fibres [11]. Since this seminal work, many ion channels have been identified, cloned, and characterised across all phyla of life.

A central theme in cellular physiology is the Nernst potential for a certain ionic species \((E_i)\). \(E_i\) is the transmembrane voltage at which the net flow of ion \(i\) will be zero, as the transmembrane concentration gradient is balanced by the transmembrane electric potential. Therefore, \(E_i\) is dependent on the established concentration gradient for a given ion by active transporters, and the uneven distribution of impermanent organic anions. The \(E_i\) of an ion can be calculated using Equation 1.1, where \(R\) is the gas constant, \(T\) is the absolute temperature, \(z\) is the valence of ion \(i\), \(F\) is the Faraday constant, and \([i]_{\text{extra}}\) and \([i]_{\text{intra}}\) are the extra- and intracellular concentrations of ion \(i\), respectively.

\[
E_i = \frac{RT}{zF} \ln \left( \frac{[i]_{\text{extra}}}{[i]_{\text{intra}}} \right)
\] (1.1)

The transmembrane voltage of a cell depends on a combination of which and how...
many ion channels are open at a specific time. For example, if K\(^+\) channels are the predominantly open channel type, then the transmembrane voltage will tend towards \(E_K\) (\(\sim -80\) mV). Alternatively, if Na\(^+\) channels are the predominantly open channel type, then the transmembrane voltage will tend towards \(E_{Na}\) (\(\sim +70\) mV). In excitable cells, the transmembrane voltage will tend towards \(E_K/-80\) mV in resting conditions, and tend towards \(E_{Na}/+70\) mV at the action potential peak.

The physiological implications of any ion channel relies on two main properties, i) gating and ii) ion selectivity. The gating of an ion channel refers to its activation (opening) or deactivation (closing) of the channel pore. When an ion channel is in an open state, it is described as being conductive, allowing ionic species to travel through the channel and cross the membrane that the ion channel is embedded in. Conversely, an ion channel in a closed state is described as being in a non-conducting and impermeable state. The gating of an ion channel can be triggered and modulated by several factors and forces, including transmembrane voltage, temperature, pH, ligand-binding, and mechanical activation.

The selectivity of an ion channel describes its ability to form an efficient permeation pathway for certain ionic species, whilst discriminating against other undesired ionic species. The selectivity of an ion channel is a major factor of its physiological role, and is one of the key constituents in conferring the specific physiological role to an ion channel within an organism. The degree of ion selectivity can vary greatly, from a channel simply being selective for cations over anions, to a channel being selective for specific ionic species over other ionic species of the same charge and valency. The selectivity of ion channels is even more remarkable when one considers the rate of ion transport achieved in addition to this selectivity, up to the order of \(10^8\) ions per second [12].

Ion channels represent the second most populous target for small molecule drugs within the ChEMBL database [13, 14, 15] (Figure 1.1). Drugs that target ion channels modulate the properties of gating and selectivity. For example, the local anaesthetic lidocaine has been shown to stabilise the inactivated state of \(NaV1\) channels, which is a property exploited to treat pain [16].

Consequently, a detailed, atomistic understanding of the permeation mechanisms of
1 INTRODUCTION

ion transport through ion channels is vital to be able to understand i) how abnormal permeation mechanisms can manifest in clinical conditions, and ii) how to mechanistically overcome these issues through the rational design of channel-interacting drugs.

1.2 The transient receptor potential channel superfamily of cation ion channels

The transient receptor potential (TRP) family of ion channels was first identified in 1969 when Cosens and Manning discovered a mutant strain of *Drosophila melanogaster* that exhibited blindness in the presence of constant, bright light [17]. This mutant *D. melanogaster* was named *trp*, and cloning of the mutated *trp* gene identified the first member of the TRP superfamily of ion channels.

Since the initial work by Cosens and Manning, the known TRP family has grown substantially to become a superfamily of polymodal signal-detecting ion channels with a number of constituent members across multiple kingdoms of life. Based upon sequence homology, six subfamilies of the TRP superfamily have been identified within the *Mammalia* class. The nomenclature of these subfamilies is defined by the suffix of the subfamily. In mammalian TRP channels, these subfamilies are TRPC (Canonical), TRPV (Vaniloid), TRPM (Melastatin), TRPA (Ankyrin), TRPML (Mucolipin), and TRPP (Polycystic) [18]. In *Homo sapiens*, there have been 26 TRP channels identified to date (Figure 1.2).

In addition to the subfamilies identified in *Mammalia*, further TRP subfamilies have
1 INTRODUCTION

Figure 1.2: Phylogenetic tree of TRP channels of *Homo sapiens*. Sequences were obtained from manually-curated Swiss-Prot sequences from the UniProtKB database [19, 20]. The average distance tree was calculated in Jalview [21] using the BLOSUM62 substitution matrix.

been identified in other classes and phyla of the *Eukarya* kingdom. These include the TRPN [22], TRPY [23, 24, 25], TRPVL [26], and TRPS [27] subfamilies.

TRP channels assemble primarily as homotetramers to form functional ion channels. A conserved structural feature across all TRP channels is the presence of six transmembrane helices (S1-S6) per subunit, forming two distinct transmembrane domains; a four-helix bundle comprising of helices S1-S4 forming the voltage-sensing like domain (VS LD), and the pore-forming domain consisting of helices S5 and S6 [18]. It should be noted that the S1-S4 regions are referred to as voltage-sensing like, as the existence and/or mechanism of voltage gating in TRP channels is a contested subject [28]. As well as these transmembranous helices, TRP channels also have a conserved amphipathic TRP helix that is located parallel to the surface of the intracellular leaflet of the membrane [18]. This TRP helix directly links the pore-lining S6 helix and the cytoplasmic domains, and has been proposed to be a key structure in channel gating. In addition to this conserved transmembrane architecture, members of the TRP superfamily display highly diverse extramembrane loops and N- and C-terminal domains between the different subfamilies [29] (Figure 1.3).

Due to their size and nature as membrane proteins, it has proven historically diffi-
cult to determine the 3D structure of TRP channels at an atomic resolution using X-ray crystallography and NMR methodologies. However, since the dawn of the so-called “Cryo-EM resolution revolution” [30], the number of TRP channel structures determined has increased dramatically. This has provided researchers with a wealth of structural information on TRP channel structure-function relationships, including structures of TRP channels with numerous ligands bound and the channel in various, specific gating states.

Figure 1.3: Representative structures from each subfamily for the TRP superfamily. The structures depicted are TRPV2 (PDB ID: 6OO3), TRPA1 (PDB ID: 3J9P), TRPM2 (PDB ID: 6MIX), TRPC3 (PDB ID: 6CUD), TRPN (PDB ID: 5VKQ), TRPML1 (PDB ID: 5WJ5), and TRPP2 (PDB ID: 5T4D). Figure reprinted from Koivisto et al. [31].

TRP channels are gated open by a particularly wide range of stimuli, which include temperature, small molecules, transmembrane voltage changes, and mechanical cues [32, 33, 34]. Recently, the 2021 Nobel Prize in Physiology or Medicine was awarded to Professor David Julius for his work on temperature gating in TRP channels [35], specifically the heat-sensing TRPV1 channel [32, 33, 8] and the cold-sensing TRPM8 channel [36].

TRP channels are described as cation non-selective ion channels, i.e., they selectively permeate cations over anions, but are largely non-selective between different cationic species and valencies. However, there are a few exceptions to this rule within the TRP superfamily. In particular, the TRPV5 and TRPV6 channels exhibit high selectivity for Ca\(^{2+}\) (P\(_{Ca}/P_{Na}\) ≈ 100:1) [37, 38]. Additionally, the TRPM4 and TRPM5 channels exhibit high selectivity for monovalent cations and impermeability to divalent cations [39]. Finally, the TRPM6 and TRPM7 channels are permeable to Mg\(^{2+}\) and Zn\(^{2+}\) cations [40], whilst the remaining TRP family members are not.

The work conducted within this thesis focuses specifically on two of these TRP sub-
families, specifically the TRPV and TRPM subfamilies.

1.3 TRPV: the vanilloid subfamily of TRP channels

TRPV channels are arguably the most studied subfamily of the TRP superfamily. The TRPV subfamily consists of six members (TRPV1-6), with TRPV1 the most well studied amongst these. TRPV1 is the primary molecular transducer of heat sensation in Homo sapiens [32, 33, 8] and is activated by capsaicin, the irritating compound found in chilli peppers.

1.3.1 Structural features

As with all TRP channels, TRPV channels consist of six transmembrane helices organised in a tetrameric assembly (Figure 1.4). The selectivity filter (SF) of TRPV channels is four residues long, with the sequence of this SF differing between TRPV1-4 and TRPV5-6 (discussed further in Chapter 3).

![Figure 1.4: Domain organisation of a monomer of the TRPV subfamily of ion channels. Figure reprinted from Kärki and Tojkander [41].](image)

The N-terminus of each TRPV channel contains six ankyrin-repeat domains (ARDs) per chain (Figure 1.4). ARDs consist of ~33 amino acid long repeats, each comprising a pair of anti-parallel α-helices [42], organised in a right-handed fold. These adjacent ankyrin repeats are almost parallel to one another, with a 2-3° counter-clockwise offset. This results in the ARD of TRPV channels having a small left-handed twist [42].

ARD domains have been identified in all phyla, and have been shown to play a large number of biological roles. In particular, ARD domains have been implicated in protein-
protein and protein-substrate interactions [43, 44, 45]. The precise role of ARDs within TRPV channels remains unclear. ARDs have been implicated in pore assembly and channel gating through ligand interactions and in interactions with the cytoskeleton [46]. Despite the uncertainty of their role in TRPV channel function, mutations within the ARD of TRPV channels can have devastating consequences to health, as demonstrated, e.g., in the TRPV4 channel [47, 48] (discussed in Chapter 1.3.2).

Prior to the determination of full-length TRP channel structures, the structure of isolated ARD domains of TRPV channels were determined via X-ray crystallography. To date, there are 16 individual structures for the ARD domain of TRPV channels [49, 50, 51, 52, 53, 54, 55, 56, 57, 58].

1.3.2 Implications in health and disease

As described previously, the TRPV1 channel is the primary detector of noxious heat thermosensation [32, 33, 8]. trpv1⁻/⁻ mice models displayed reduced heat hyperalgesia after inflammation [34]. TRPV1 antagonists were therefore investigated a potential analgesic targets. Paradoxically, TRPV1 agonists have also been reported to have analgesic properties. For example, dilute capsaicin creams are commonly used as an over-the-counter pain relief. Moreover, trials have also been conducted using resiniferatoxin (RTX) in patients with intractable cancer pain. RTX is a high affinity activator of TRPV1, and has been used to create permanent analgesia in those suffering from debilitating pain [59, 60]. In addition to a pain relief target, TRPV1 antagonists have also been used as targets for the development of therapeutics for pruritus [61] and cough [62].

Conditions including pulmonary edema, acute lung injury, and sepsis all result in excess fluid in the lung. Consequently, endothelial cell dysfunction is a key pathological mechanism in these conditions. TRPV4 has been shown to be a vital osmosensor [63, 64, 65]. TRPV4-selective antagonists have been demonstrated to prevent an increase in fluid in the lungs in models representing heart failure [66].

The TRPV6 channel has also been implicated in cancer pathology [67, 68], with increased TRPV6 expression in prostate, colon, breast, thyroid, and ovary carcinomas [69,
1 INTRODUCTION

Of these, the most extensive studies of TRPV6 in cancer have been related to prostate cancer. Consequently, TRPV6 antagonists have emerged as a promising anti-cancer therapeutic. A peptide inhibitor of TRPV6 (SOR-C13) has recently completed a Phase I clinical trial [73], and been shown to reduce growth in ovarian and prostate cancers in a number of animal models [74, 75]. In addition to peptide inhibitors, (4-phenylcyclohexyl)-piperazine derivatives (PCHPDs) have also been shown to be potent and selective TRPV6 inhibitors, with structural data explaining the inhibition mechanism of PCHPDs [76].

There are multiple heritable diseases associated with mutations occurring in trpv genes. Mutations in trpv3 have been associated with Olmsted diseases [77] and focal palmoplantar keratoderma [78]. Furthermore, mutations in the trpv4 gene have been shown to result in a variety of conditions, including scapulopereoneal muscular atrophy [47], Charcot-Marie-Tooth disease [54], and autosomal dominant brachyolmia [79]. In addition to these, TRPV4 has also been implicated in the aetiology of several autosomal dominant skeletal dysplasias [48]: metatropic dysplasia Kozlowski type of spondylometaphyseal dysplasia, Maroteaux type of spondyloepiphysyal dysplasia, and parastrempmatic dysplasia.

1.4 TRPM: the melastatin subfamily of TRP channels

TRPM channels form the largest subfamily within the TRP superfamily with eight subfamily members (TRPM1-8). They are Ca\(^{2+}\)-activated ion channels. TRPM channels are ubiquitously expressed in Homo sapiens tissue. In particular: TRPM1 is most expressed in the brain, TRPM2 in the brain and bone marrow, TRPM3 in the brain and pituitary, TRPM4 in the intestine and pancreas, TRPM5 in the intestine, pancreas, and prostate, TRPM6 in the intestine and brain, TRPM7 in adipose tissue, and TRPM8 in the prostate and liver [80]. Amongst the most well-known TRPM channels is the TRPM8 channel. This channel is the primary molecular transducer of cold thermosensation in Homo sapiens [36, 81].
1 INTRODUCTION

1.4.1 Structural features

TRPM channels have the same conserved transmembrane domain that is present in all TRP channels; a tetrameric assembly of six transmembrane helices and the amphipathic TRP helix. Within the VSLD of TRPM channels, a Ca\(^{2+}\) binding site is present between residues of the S2, S3, and TRP helices (Ca\(_{TMD}\) binding site). Binding of Ca\(^{2+}\) at this site has been implicated to underpin Ca\(^{2+}\)-dependent activation of TRPM channels, as demonstrated in structures of the TRPM5 channel in distinct activation states [82]. TRPM channels possess a shorter SF than that seen in TRPV channels, with a three-residue long SF.

![Diagram of TRPM channel](image.png)

Figure 1.5: Domain organisation of a monomer of the TRPM subfamily of ion channels. The C-terminal domain differs among TRPM subfamily members. Figure reprinted from Huang et al. [83].

The N- and C-termini of TRPM channels contain large cytoplasmic domains. The N-terminus forms four characteristic TRPM homology regions (MHRs), which are often grouped into the MHR1/2 and MHR3/4 domains [83] (Figure 1.5). These MHR domains have been shown to be ligand binding sites for multiple ligand types in a number of TRPM channels. For example, TRPM2 is activated by adenosine diphosphate ribose (ADPR) in the presence of Ca\(^{2+}\). Structures of TRPM2 of *Danio rerio* and of *Homo sapiens*
showed that ADPR binds to the MHR1/2 domain, and mutation of the binding residues abolished ADPR/Ca\(^{2+}\)-induced channel activation [84, 85]. In addition, ligands have been shown to bind at the interface between the MHR1/2 and MHR3/4 domains. In TRPM4, adenosine triphosphate (ATP) inhibits Ca\(^{2+}\)-induced currents [86]. However, structures of the ATP-bound TRPM4 structure of *Mus musculus* identified an ATP binding site formed by aromatic residues at the interface of the MHR1/2 domain MHR3/4 domains [87].

The C-termini of TRPM channels form a coiled-coil “rib” and “pole” structure, which has been implicated to play an important role in subunit assembly and in ligand binding [83]. The C-terminal domain (CTD) of TRPM channels forms the ligand binding sites for nucleotides and interaction partners. For example in TRPM2 channels, the CTD contains a NUDT9-H domain, homologous to the ADP-ribose pyrophosphate NUDT9 domain [84, 88, 89, 85]. This chimerisation of an ion channel and enzyme domain has led to TRPM channels sometimes being referred to as “chanzymes”. Additionally, the TRPM6 and TRPM7 channels contain a functional serine/threonine-specific kinase [90], belonging to an atypical family of eukaryotic \(\alpha\)-kinases [91].

### 1.4.2 Implications in health and disease

Due to their expression and physiological significance, TRPM channels have been implicated in a number of pathological conditions [92, 93].

As described previously, TRPM8 is the main detector of cold thermosensation [36, 81]. Genetic deletion of the *trpm8* gene in mice resulted in hypersensitivity following exposure to cold [94]. Therefore, TRPM8 is a promising analgesic target. In fact, TRPM8 is activated by menthol [95], which has been used as a topical analgesic since Ancient Egyptian times (as mentioned in the Ebers papyrus from 1550 BCE) [96]. However, it should be noted that menthol also modulates several other ion channels, including TRPV3 [97] and TRPA1 [98]. TRPM8 antagonists are an active target pursued by pharmaceutical companies for a novel class of analgesic drugs [92].

Due to its expression within the pancreatic \(\beta\)-cells and role in insulin secretion, TRPM5 has been proposed as a potential therapeutic target in the treatment of obesity and Type II
INTRODUCTION

Diabetes Mellitus (T2DM) [99]. This proposal is based on TRPM5 potentiation stimulating pancreatic β-cells to increase insulin secretion.

Mutations in trpm genes have also been reported to result in a number of heritable diseases. Mutations in the trpm1 gene can result in the development of congenital stationary night blindness (CSNB) [100, 101], a condition where patients have visual difficulties from birth in dim-light situations (known as nyctalopia). Furthermore, mutations in the trpm4 gene have been implicated as a cause of familial heart block type I [102], a serious cardiovascular condition. Finally, mutations in the trpm6 gene cause familial hypomagnesia and secondary hypocalcemia [103, 104].

1.5 Project aims

As described previously, one of the key characteristics of an ion channel is its ion selectivity; its ability to form an efficient permeation pathway for a specific ionic species, and to discriminate against other undesired ionic species. In the research conducted within this thesis I set out to investigate the molecular determinants underpinning the Ca\(^{2+}\)-selective permeation mechanism of the TRPV5 and TRPV6 channels (Chapter 3), and the monovalent-selective cation permeation mechanisms of the TRPM5 channel (Chapter 4).

In recent years, molecular dynamics (MD) simulations have been successfully employed to shed light on ion channel function and the mode of action of channel-acting drugs in atomistic detail, for instance on K\(^{+}\) channels [105, 106], Na\(^{+}\) channels [107, 108], Cl\(^{-}\) channels [109, 110], and ligand-gated ion channels [111]. This methodology utilises computer simulations to model the dynamics and interactions of atoms within a molecular system. The benefit of MD over other classical biophysical or structural methodologies is that it allows us to study a system with atomistic detail, on timescales not typically accessible to other techniques. The high degree of molecular detail provided by MD simulations has led to it being referred to as a computational microscope among researchers and the literature [112, 113, 114].

Therefore, to address the proposed aims, I decided to utilise a computational-based
approach, conducting *in silico* electrophysiology simulations of the TRPV and TRPM channels. In these simulations, a continuous flow of permeating cations was achieved by creating a bio-mimetic, transmembrane electrochemical potential, by one of two methodologies, namely the Computational Electrophysiology (CompEL) method and the external applied electric field (EAEF) method. I additionally conducted a systematic comparison of these two methodologies, to identify any differences in using these methods to model ion permeation through an ion channel (Chapter 5). In total, this thesis reports the results from simulations with an aggregated simulation time of 58.26 µs that produced 8,040 permeation events across TRP channels, providing a basis for investigating the selectivity and permeation mechanisms in these channels with very high statistical power.
2 Theory and methods

The specific details of the simulation approaches used are described thoroughly in each of the respective Chapters. In this Chapter, I will give an overview on the theory underpinning MD simulations, with a particular emphasis on topics pertinent to in silico electrophysiology.

2.1 Core principles of molecular dynamics simulations

A key principle of molecular dynamics is the concept of the ergodic hypothesis, one of the key foundations in Boltzmann’s formulation of statistical mechanics [115]. The ergodic hypothesis describes how a trajectory will sample all possible states of a system, if given sufficient time. If the ergodic hypothesis is assumed to be true for the system, then if one were to study a single system for a long enough period of time, the result would be the same as if one studied a large group (ensemble) of systems at a single point. This is often expressed as saying that the time average of the microscopic dynamics ($\langle A \rangle_{time}$) is equal to the ensemble average of the macroscopic dynamics ($\langle A \rangle_{ensemble}$), as described in Equation 2.1.

$$\langle A \rangle_{ensemble} = \langle A \rangle_{time} \tag{2.1}$$

Therefore, the ergodic hypothesis makes it possible to study the thermodynamic properties of a system from its trajectory over time. If one assumes that the nucleus of an atom follows the Newtonian laws of motion, the trajectory of an atom can be obtained through the integration of Newton’s second law (Equation 2.2) [116]:

$$-\frac{dU}{dr_i} = m_i \frac{d^2r_i}{dt^2} \tag{2.2}$$

Where $U$ is equal to the potential energy of the system as a function of all atomic coordinates ($r_i$), and $m_i$ is the mass of atom $i$. The derivative of the potential energy of $U$ is the force ($F_i$) acting on atom $i$ at position $r_i$, and determines its acceleration and thus its velocity and the position of atom $i$ after a given timestep, $dt$. 

13
The value of this timestep has to be smaller than the fastest motion of the system. Typically, this is the bond vibrations formed by hydrogen atoms which occur on the femtosecond timescale. It is possible to increase the timestep, however, by using methodologies such as the LINCS algorithm [117] and approaches like hydrogen mass-repartitioning (HMR) [118, 119]; which are both employed throughout this thesis. These algorithms serve to constrain or abolish the fast oscillations of covalently bonded hydrogen atoms, which are irrelevant for the function of the biomolecules under investigation.

The deterministic nature of MD simulations allows us to compute the state of the system at any time point if we know the position and velocity of all the particles in the system. Several algorithms have been developed for the efficient integration of Equation 2.2. In this thesis, the leap-frog algorithm was used for the integration of the equations of motion [120].

2.2 Fixed point-charge forcefields

In MD, the potential energy function described in Equation 2.2 is dependent not only on the atomic Cartesian coordinates, but also on the geometric and energetic properties of inter-particle interactions. The combination of parameters and mathematical functions that describe these interactions are commonly referred to as a "force field" [121]. This force field is a mathematical expression that describes the dependence of the energy of a system on the coordinates of its constituent particles and their interactions [122]. The first force fields began to be used in the 1960s, with the aim of using them with the molecular mechanics (MM) method to predict molecular structures, vibrational spectra, and enthalpies of molecules. However, since this time the topic of force field development has evolved to be able to facilitate the simulation of increasingly complicated molecular systems. Examples of several popular force fields include the CHARMM [123], AMBER [124], GROMOS [125], and OPLS [126] force fields.

\[
E_{\text{system}} = E_{\text{bonded}} + E_{\text{non-bonded}}
\]  

(2.3)

The majority of this work was conducted using the CHARMM (Chemistry at Harvard
Macromolecular Mechanics) forcefield, specifically the CHARMM36m force field [127].
As with all force fields, the energy function of the CHARMM forcefield is dependent on the combination of bonded and non-bonded energies (Equation 2.3). The bonded and non-bonded terms of the CHARMM36m potential energy function are given by Equation 2.4 and Equation 2.5, respectively [128].

\[
E_{\text{bonded}} = \sum_{\text{bonds}} K_b (b - b_0)^2 + \sum_{\text{angles}} K_\theta (\theta - \theta_0)^2 + \sum_{\text{impropers}} K_\phi (\phi - \phi_0)^2 + \sum_{\text{dihedrals}} K_{\phi,n} (1 + \cos(n\phi - \delta_n)) + \sum_{\text{Urey–Bradley}} K_{UB} (r_{1,3} - r_{1,3,0})^2 + \sum_{\text{residues}} u_{\text{CMAP}} (\phi, \psi)
\]

The bonded terms of the potential energy function (Equation 2.4) describes the covalent connectivity patterns of atoms within the molecule, whereby \(b\), \(\theta\), \(\phi\), and \(\phi\) correspond to the bond length, valence angle, improper dihedral angle, and dihedral angles, respectively. The terms \(b_0\), \(\theta_0\), and \(\phi_0\) are the respective equilibrium terms, and \(K_b\), \(K_\theta\), and \(K_\phi\) are the respective spring constants of harmonic potentials. The cosine expansion describing the potential around dihedral angles is further supplemented with a Urey-Bradley term. This is a harmonic potential that handles the two terminal atoms of a torsion angle with a separate quadratic term dependent on the atom-atom distance. The bonded term is further supplemented with a “correction map” (CMAP), which is a term applied to the backbone \(\Phi\) and \(\Psi\) dihedral angles of proteins. This CMAP improves the secondary structure propensities of the protein.

\[
E_{\text{non-bonded}} = \sum_{\text{NBpairs}} \varepsilon_{ij} \left[ \left( \frac{r_{ij}}{r_{ij}^\text{min}} \right)^{12} - 2 \left( \frac{r_{ij}}{r_{ij}^\text{min}} \right)^{6} \right] + \sum_{\text{NBpairs}} \frac{q_i q_j}{4\pi \varepsilon_0 \| \mathbf{r}_i - \mathbf{r}_j \|}
\]

The non-bonded term of the potential energy function (Equation 2.5) describes the electrostatic and van der Waals (vdW) interactions within the system. In this \(q_i\) and \(q_j\) are the respective partial atomic charges of atoms \(i\) and \(j\). The vdW term is treated by a
Lennard-Jones (LJ) 6-12 potential, whereby $\varepsilon_{ij}$ is the well depth, $r_{ij}^{\text{min}}$ is the radius, and $\|r_i - r_j\|$ is the distance between atoms $i$ and $j$. For context, a LJ potential models the soft attractive and hard repulsive forces between neutral atoms, based upon their relative proximity [129]. This potential is made up of the Pauli repulsion (with an exponent of 12) and the attractive London dispersion interaction (given an exponent of 6).

In addition to the CHARMM36m force field, some simulations conducted and reported in Chapter 5 utilised an AMBER (Assisted Model Building with Energy Refinement) force field, specifically the AMBER19SB force field [127]. The AMBER family of force fields contain some differences to the previously described CHARMM forcefield. These differences are present in both the bonded and non-bonded portion of the energy function [121].

The first difference in respect of bonded interactions is that in the AMBER forcefield, improper dihedral angles contribute to the energy function via the dihedral term. However, in the CHARMM force fields a separate quadratic term for improper dihedrals is included (Equation 2.4). Secondly, the CHARMM force field incorporates a Urey-Bradley term for dihedral angles, whereas the AMBER force field does not [121].

Moreover, in the non-bonded interactions of the force field, differences are present in the handling of 1,4 non-bonded interactions; atoms A and D in the dihedral formed by A-B-C-D. In the AMBER force field, the 1,4-LJ interactions are scaled by $\frac{1}{2}$ and Coulomb interactions by $\frac{1}{12}$. The application of scaling constants is very common in force fields, and other scaling factors are applied in the aforementioned OLPS and GROMOS force fields. However, in the CHARMM force field, a far simpler approach is taken. Here, 1,4 non-bonded interactions are not scaled by default, with special 1,4-LJ parameters applied to a select few atom type pairs [121].

### 2.3 Divalent cations in fixed point-charge forcefields

As described previously, the equation for non-bonded interactions contains a term describing the electrostatic interactions of the system (Equation 2.5), based upon the partial atomic charges of an atom. However, this assignment of a fixed partial charge does not
account for the effects of polarisation, and hence these force fields are described as additive or non-polarisable. Despite neglecting the effects of polarisation, these force fields have been shown to accurately reproduce the energies and dynamics when used to study a molecular system, as emphasised by their popularity and use.

However, there are circumstances where the lack of inclusion of polarisation results in deviation from the experimentally derived energies. This issue is particularly pertinent for multivalent ionic species, which poses a particular problem when one considers the biological significance of these ionic species. For example, Ca\textsuperscript{2+} ions play a vital role in signalling pathways [130], Mg\textsuperscript{2+} is a mineral nutrient that plays an essential role in the biological activation of adenosine triphosphate (ATP) [131], Zn\textsuperscript{2+} is a key co-factor for a number enzymes such as carbonic anhydrase [132], and Fe\textsuperscript{2+} in the haem group of haemoglobin is necessary for oxygen transport through the bloodstream [133]. As discussed in Chapter 1, TRP channels also conduct multivalent cations, predominantly Ca\textsuperscript{2+}. Furthermore, the TRPV5 and TRPV6 channels are Ca\textsuperscript{2+}-selective, and the TRPM6 and TRPM7 channel are Mg\textsuperscript{2+}- and Zn\textsuperscript{2+}-permeable.

In simulations of these divalent cations in non-polarisable force fields, the protein-cation binding energies are overestimated due to the neglect of ionic polarisation effects [134]. This issue is a particular problem for MD simulations investigating permeation mechanisms, whereby these overestimated binding energies result in cations binding strongly to the protein and not permeating through the channels.

Consequently, there has been a huge amount of effort within the field to overcome these issues caused by the neglect of polarisation. One such approach is to incorporate terms for electronic polarisation within the non-bonded term of the potential energy function, in so called polarisable forcefields. Methods to explicitly model polarisation can be grouped into three main approaches (Figure 2.1): fluctuation charge models, induced point dipole models, and the Drude oscillator model.

In the fluctuating charge model, the polarisation is based upon the mechanism of atomic partial charges seen in Equation 2.5, however the magnitude of the individual partial charges can change during the simulation. This is achieved by assigning virtual
masses to each charge, and handling them as an additional degree of freedom in the equations of motion [136]. Despite this approach being attractive due to the relatively small additional computational cost, the major disadvantage is the difficulty in representing polarisation that does not occur in the direction of the existing bonds.

In the induced point dipole model, the polarisation is based upon the mechanism of fixed atomic partial charges seen in Equation 2.5, with the addition of inducible point dipoles that are determined by the electric field of the site [137]. This approach has the benefit of being relatively simple to parameterise, however is challenging to implement within an existing MD simulation engine. One example of an induced point dipole based force field is the AMBER FF02 force field which is implementable in the AMBER simulation package [138, 139].

In the Drude oscillator model, the polarisation of an atom is modelled via the addition of a Drude particle, attached to the core of the atom by a harmonic spring [140, 141]. The polarisability of a given atom ($\alpha$) is then described by the distribution of the total charge on the atom ($q$). The charges of the Drude particle ($q_D$) and the core atom ($q_A$) are such that $q = q_A + q_D$. The atomic polarisability ($\alpha$) and the induced atomic dipole ($\mu$) can therefore be described as in Equations 2.6 and 2.7, respectively, where $k_D$ is the force constant of the Drude-atom harmonic bond (1000 kcal mol$^{-1}$ Å$^{-2}$ for the Drude-2013 forcefield [140]) and $E$ is an electric field.
\[
\alpha = \frac{q_D^2}{k_D} \tag{2.6}
\]

\[
\mu = \frac{q_D^2 E}{k_D} \tag{2.7}
\]

In addition to these three approaches utilising point charges, there have also been force fields developed that use multipole moments. These multipole moments are anisotropic, and so naturally capture any non-spherical components of atomic charge density. This approach has been used to design the AMOEBA (Atomic Multipole Optimised Energetics for Biomolecular Applications) force field [142, 143, 144].

Despite the promise of polarisable force fields, there are still a number of limitations to them [145]. As has been demonstrated in the development of non-polarisable force fields, weaknesses in force field development can only be identified when they are used beyond the simple simulations used for their parameterisation, and if they are utilised widely by the scientific community to truly test these force fields. It is this constant cycle of improvements that has resulted in the multiple versions of non-polarisable force fields. For example, the CHARMM force field for protein simulations has included CHARMM19 [146], CHARMM22 [123], CHARMM22/CMAP [147], CHARMM36 [148], and CHARMM36m [127]. Therefore, the relative infancy of polarisable force fields compared to their non-polarisable force fields increases the likelihood of unforeseen shortcomings in their parameterisation.

An additional obstacle to the use of polarisable forcefields in the reputation of these force fields to be “slow”; that is for them to take a larger amount of central processing unit (CPU) hours to achieve the same simulated time. Although the speed of simulations with polarisable force fields has greatly improved over the past few years, they are still slower than non-polarisable force fields due to the additional components of their functional forms.

For this reason, there have been attempts within the field to modify the individual parameters for divalent cations within non-polarisable forcefields to improve their accu-
racy. One exemplar approach is electronic continuum correction with rescaling (ECCR), which involves re-scaling of the ionic charges by the inverse of the electronic part of the water dielectric constant (i.e. a factor of 0.75) and an adjustment of the ionic radii. This approach has been shown to reproduce the experimental structure of aqueous CaCl$_2$ obtained from neutron scattering [134], and improved accuracy in binding free energies of Ca$^{2+}$ cations with the EF-hand loops of calmodulin [149].

An alternative, and particularly interesting approach, is to distribute the charge and vdW centres of the ion to multiple sites, via a so-called multi-site model. Recently, Zhang et al. published a new multi-site Ca$^{2+}$ model specifically optimised for Ca$^{2+}$-protein and Ca$^{2+}$-water interactions [150]. This model correctly replicated experimental values for the hydration free energy and the number of coordinated water molecules in the first solvation shell, and showed Ca$^{2+}$-protein binding energies comparable to the quantum mechanic and polarisable models. The multi-site Ca$^{2+}$ model has previously been used to investigate ion permeation of Ca$^{2+}$ in a range of channels, including the type-1 ryanodine receptor [150, 151], AMPA receptors [152], and recently the E protein of SARS-CoV-2 [3]. Please note however, that at the start of this project only Zhang et al. [150] had been published, and that the studies described in [151, 152, 3] have been published whilst this project has been underway.

Therefore, for the work in this thesis I opted to use non-polarisable force fields, rather than polarisable force fields, due to their extensive use and development, and the quicker simulation time. To overcome the described shortcomings of these non-polarisable, fixed point-charge forcefields with divalent cations, I utilised the multi-site Ca$^{2+}$ model of Zhang et al. [150]. This approach was preferred to ECCR due to its demonstrated success in being applied to model Ca$^{2+}$ permeation in an ion channel. Therefore, I was able to model Ca$^{2+}$ permeation events in TRP channels, and elucidate selectivity for or against permeation of this ion type. The potential use of polarisable force fields or ECCR in future research of the permeation mechanisms in TRP channels is discussed in Chapter 6.
2.4 Simulation protocol

The precise set-up of a simulation system and the parameters set for a simulation have a large impact on the resultant output of the simulation. The precise settings used throughout the system are described in Chapters 3.3, 4.3, and 5.2. Here, I describe some of the theory and principles behind some of the parameters and methodologies used in the MD simulations within this thesis.

The starting point for atomistic molecular simulations of biological systems is to build the simulation system. As described in Equations 2.2, to calculate the movement of a particle \( i \), we need to know the initial coordinates of the particles. In the case of simulating biological molecules, such as proteins and nucleic acids, these initial coordinates correspond to the 3D structure of the biomolecules. Experimentally-determined protein and nucleic acids structures are published in the publicly available worldwide Protein Data Bank (wwPDB) repository [153]; a consortium operated with the Research Collaboratory for Structural Bioinformatics (RCSB), the Macromolecular Structure Database (MSD) at the European Bioinformatics Institute (EBI), and the Protein Data Bank Japan (PDBj) at the Institute for Protein Research in Osaka University as the database custodians. This database contains \( \sim 200,000 \) individual, structural models at the time of writing. These structures have been determined by a number of structural biological techniques, including X-ray crystallography, cryo-electron microscopy (cryo-EM), and nuclear magnetic resonance (NMR). These techniques have had a major impact in biophysical research, as exemplified by 15 Nobel Prize awards associated with X-ray crystallographic structures [154, 155], and the 2017 Nobel Prize in Chemistry being awarded “for developing cryo-electron microscopy for the high-resolution structure determination of biomolecules in solution” [156]. In recent years, the improvement in structure prediction tools, such as AlphaFold2 [157] and RoseTTAFold [158], have resulted in the determination of computed structural models (CSMs) of increasing accuracy. The RCSB PDB now offers access to \( \sim 1 \) million CSMs from the AlphaFoldDB and RoseTTAFold model archive. All simulations conducted within this thesis utilised publicly-available, experimentally-determined structures deposited within the wwPDB.
2.4.1 Thermodynamic ensembles, thermostats, and barostats

In Chapter 2.1, the concept of statistical ensembles of a system was introduced, and how the ergodic hypothesis allows us to derive the properties of a true thermodynamic system from the laws of classical and statistical mechanics. For this purpose, we can consider different systems with different degrees of coupling to their surroundings as different ensembles. Specifically, in MD simulations there are three main ensembles to be considered: the canonical ensemble (also referred to as the $NVE$ ensemble), the micro-canonical ensemble (also referred to a the $NVT$ ensemble), and the isothermal-isobaric ensemble (also referred to as the $NPT$ ensemble).

All simulations conducted within this thesis utilise the NPT ensemble, that is that the ensemble maintains a constant total number of particles ($N$), a constant pressure ($P$), and a constant temperature ($T$). This ensemble is particularly pertinent for calculating the Gibbs free energy of a system; the thermodynamic quantity that describes the maximum amount of work a system can do at constant pressure and temperature.

To maintain the temperature and pressure of a system, MD simulations use thermostat and barostat algorithms, respectively. In the NPT ensemble, these algorithms maintain the temperature and pressure of the system. Throughout this thesis, the Nosé-Hoover thermostat [159] was used to maintain the temperature of the system at 310 K. This thermostat algorithm maintains the system temperature by coupling particles within the simulations system to a virtual temperature bath. Furthermore, the Parrinello-Rahman barostat was used to maintain the pressure at 1 bar [160]. This barostat algorithm modulates the volume of the system to maintain a target pressure.

In order to maintain the key NPT ensemble, it is necessary to keep a constant number of atoms within the system. However, the addition of walls or barrier around the simulation system would lead to artificial interactions with a wall. One popular way to deal with the finite size of a system is through the application of periodic boundary conditions (PBC) [161]. In this approach, the atoms at the edge of a unit cell interact with the atoms at the opposite side of the unit cell, and if an atom leaves the unit cell it will be replaced by its periodic image from the opposite side (Figure 2.2). Therefore, the incorporation of
PBC conditions allows for the approximation of an infinite system by simulating a finite unit cell.

2.4.2 *In silico* electrophysiology using molecular dynamics simulations

The major driving forces for ionic permeation through ion channels are the transmembrane voltage and the concentration gradients across the membrane. To be able to model ion permeation in MD simulations, it is therefore necessary to reproduce these driving forces, principally the transmembrane voltage. To date, there are two main methodologies to model these for *in silico* electrophysiological experiments: an external applied electric field (EAEF) and the Computational Electrophysiology (CompEL) protocol.

In the EAEF methodology, an external force is applied to all charged atoms of the system [162, 163]. This generates a linear potential across the unit cell of the simulation system. The bulk aqueous salt solution of the system will naturally self-orient itself to reduce the magnitude of the electric field as much as possible, in turn producing its own
non-uniform reaction field. However, the protein-membrane system is not able to re-orient itself to negate the electric field. Therefore, the resulting potential across the system is the sum of the applied field and the reaction field (Figure 2.3). Furthermore, this reaction potential also means that despite the non-periodic nature of the resulting potential, the force experienced by charged species across the periodic boundary is continuous.

The magnitude of the voltage produced is shown in Equation 2.8, where \( V \) is the resulting voltage, \( E \) is the magnitude of the electric field, and \( L \) is the unit cell length in the direction of the applied field.

\[
V = E \times L
\]  (2.8)

However, this approach may become problematic in the case of simulations of membrane proteins with large cytoplasmic domains. Unlike the bulk solution, these domains cannot freely re-orient themselves to counteract the applied field and therefore may be under an artefactual force.

On the other hand, the CompEL methodology produces a transmembrane voltage by using small ionic imbalances (\( \Delta q \)) [164, 165]. In this approach, a simulation system consisting of a single bilayer is duplicated, producing a "stacked" double bilayer system (Figure 2.4). This results in the formation of two separated bulk solution compartments; compartment A and compartment B. The number of molecules in each of these compartments for each ionic species can then be specified by the user. By producing a small \( \Delta q \) between these compartments, the transmembrane potential (\( \Delta U \)) of the system can be calculated according to the membrane capacitance, as shown in Equation 2.9.

\[
\Delta U = \frac{\Delta q}{C}
\]  (2.9)

The CompEL protocol determines the number of the ionic species at regular, short intervals in each of these compartments during the course of the simulation. If it is found that the numbers in each compartment differ from the reference value specified, automated water-ion exchanges are carried out to maintain the desired \( \Delta q \), and thus the same \( \Delta U \).
minimise disturbance of the protein-membrane system, particles for ion-water exchanges are selected from those located near the central planes of each compartment.

Despite the CompEL methodology requiring a double, stacked system it is only marginally more computationally expensive compared to the EAEF approach. This is because despite it being necessary for the system to be twice as big, it is possible to investigate permeation events through the channels embedded in each of the two ion channels. Therefore, the only additional computational cost is as a result of the regular determination of the number of ionic species in each compartment, and water-ion exchanges when necessary. Furthermore, the stacked nature of the CompEL approach allows a user to have each of
Figure 2.4: Overview of the Computational Electrophysiology (CompEL) setup. Schematic of a parallel (A) and anti-parallel (B) simulation system consisting of two membranes (yellow), two ion channels (green), water (blue/grey), cations (purple), and anions (green). A snapshot from the parallel simulation setup of the KcsA channel (C) used by Köpfer et al. [105] and the resulting potential (D). Figure reprinted from Kutzner et al. [165].

the ion channels within the system either experience the same or opposite voltage polarities. As demonstrated in Figure 2.4, if the protein-membrane systems are parallel to one another, then each channel will experience different voltage polarities (i.e. one will be under negative voltage and the other under positive voltage). On the other hand, if the protein-membrane system are anti-parallel and mirrored, then each of the channels will experience the same voltage polarity.

In this thesis, both of these methodologies have been used extensively. In Chapter 3 the EAEF method is used, in Chapter 4 the EAEF and the CompEL methods are used, and in Chapter 5 a direct comparison between the methodologies is conducted.
3 Ca\(^{2+}\)-selective permeation mechanism in the TRPV family of ion channels

3.1 Associated publications

The work in this chapter has been published in the following publications:


3.1.1 Authors’ Contributions

CMI and UZ conceived the idea and designed the computational study, CMI conducted the simulations, CMI analysed the simulation data, CMI and NJT developed the use of SSI for knock-on permeation, NJT analysed the SSI data, UZ supervised the work, CMI and UZ wrote the manuscript with contributions from NJT, and all authors edited and reviewed the manuscript.

3.2 Introduction

The significance of Ca\(^{2+}\) in cellular function was first recognised by Sydney Ringer in 1883, who demonstrated that low concentrations of calcium were required for the contraction of cardiac muscle [166]. Ca\(^{2+}\) is now recognised as a versatile signalling agent, with cellular Ca\(^{2+}\) concentrations impacting a broad array of physiological processes ranging from cell proliferation to cell suicide [8, 167, 168, 169]. However, the cytoplasmic concentration of Ca\(^{2+}\) ions is usually kept low due to cytotoxic consequences [170]. Therefore, the controlled opening of channels in cellular and organellar membranes is one of the required mechanisms to allow the influx of this ion from the exoplasm and internal storage compartments into the cytoplasm; this subsequently initiates the Ca\(^{2+}\) signalling cascade. The question of how Ca\(^{2+}\) channels selectively permeate Ca\(^{2+}\) in low concentrations over vastly more abundant Na\(^{+}\) ions, and yet conduct them at high rates, has been a
A key example of ion channels that mediate Ca\(^{2+}\) permeation across the cytoplasmic membrane is the TRP channel superfamily. In their open state, these polymodal signal-detecting TRP channels allow the transmembrane flux of cations down their electrochemical gradient, thereby increasing the intracellular Ca\(^{2+}\) and Na\(^{+}\) concentration [173]. The malfunction of TRP channels underlies a wide range of diseases, and they are therefore of immense biomedical importance, serving as drug targets for a variety of existing and candidate drugs [92].

Whilst they are all cation-selective, most TRP channels electrophysiologically characterised to date show only limited discrimination between cation types, as well as between divalent and monovalent cations. However, the TRPV5 and TRPV6 channels are unique due to their high selectivity for Ca\(^{2+}\) cations over Na\(^{+}\) cations (\(P_{Ca}/P_{Na} \approx 100:1\) from reversal potential measurements in combination with ion substitution) [37, 38]. Phylogenetic analysis has demonstrated that the TRPV5 and TRPV6 channels of vertebrates originated from an ancestral TRPV5/6 gene, which then diverged to form TRPV5 and TRPV6 from a duplication event after speciation [174]. Both of these channels are constitutively active due to basal levels of phosphatidylinositol 4,5-bisphosphate (PI(4,5)P\(_2\)) in the cellular membrane, and play a key role in Ca\(^{2+}\) homeostasis in the body [175]. Despite their characteristic Ca\(^{2+}\)-selectivity, both channels have been shown to permeate monovalent cations such as Na\(^{+}\) when divalent cations are absent [176, 37, 177, 178].

By contrast, the remaining members of the TRPV subfamily, TRPV1-4, conduct both Ca\(^{2+}\) and Na\(^{+}\) cations, even in the presence of high Ca\(^{2+}\) concentrations, although they are still slightly Ca\(^{2+}\)-selective, with a permeability ratio \(P_{Ca}/P_{Na} \sim 10:1\) [39]. These channels gate in response to a number of stimuli, including raised temperature – in particular the archetypal member TRPV1 [32, 33, 34], which has led to TRPV1-4 being referred to as thermoTRPV channels – as well as endogenous and exogenous ligands.

In this chapter, I set out to elucidate the molecular basis of Ca\(^{2+}\)-selectivity and permeation in the TRPV channel subfamily. I conducted atomistic MD simulations of TRPV channels under transmembrane voltage, and compared the cation permeation mechanism
observed in the Ca$^{2+}$-selective TRPV5 and TRPV6 channels to the permeation mechanism in two exemplar non-selective TRPV channels, TRPV2 and TRPV3. In total, I observed 2,851 full ion traversals from 17.25 µs of MD simulations, allowing us to decipher the permeation mechanisms and principles of ion selectivity in the TRPV family with statistical power. My findings suggest that ion conduction in TRPV channels proceeds via a co-operative knock-on mechanism involving multiple ion binding sites. The degree of co-operativity in ion permeation, linking the multiple binding sites, determines the degree of ion selectivity in the channels.

Figure 3.1: Structure of the truncated construct of TRPV5 of *Oryctolagus cuniculus* used within this work, from the extracellular side (*top-left*) and in the plane of the lipid bilayer (*bottom-left*). In this study, the pore is defined as the region between the constrictions of the channel, namely the top residue of the SF (referred to as the α-position of the SF) and the hydrophobic lower gate (*right*).
Table 3.1: Summary of protein constructs used in this study.

<table>
<thead>
<tr>
<th>PDB ID</th>
<th>Orthologue</th>
<th>Residue range</th>
<th>Upper-gate residue</th>
<th>Lower-gate residue</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV2</td>
<td>6BO4</td>
<td>Rattus norvegicus</td>
<td>327 to 691</td>
<td>E609</td>
<td>I642</td>
</tr>
<tr>
<td>TRPV3</td>
<td>6PVP</td>
<td>Mus musculus</td>
<td>375 to 745</td>
<td>D641</td>
<td>I674</td>
</tr>
<tr>
<td>TRPV5</td>
<td>6DMU</td>
<td>Oryctolagus cuniculus</td>
<td>262 to 639 (and PI(4,5)P₂)</td>
<td>D542</td>
<td>I575</td>
</tr>
<tr>
<td>TRPV6</td>
<td>6BO8</td>
<td>Homo sapiens</td>
<td>262 to 638</td>
<td>D542</td>
<td>I575</td>
</tr>
</tbody>
</table>

3.3 Methods

3.3.1 TRPV system construction

Truncated TRPV simulation systems consisting of the membrane-domain of the channels were constructed as described in Table 3.1. The systems were built using the CHARMM-GUI server [179]. The charged N- and C-terminal residues were neutralised by capping with acetyl (ACE) and N-methylamide (CT3) groups, respectively, and all missing non-terminal residues were modelled [180]. In the case of the TRPV5 system, the parameters for PI(4,5)P₂ were generated using the CHARMM General Force Field (CGenFF) [181] through the ligand reader and modeller in CHARMM-GUI [182].

The structures were aligned in the membrane using the PPM server [187], inserted into a 1-palmitoyl-2-oleoyl-sn-glycerol-3-phosphocholine (POPC) bilayer of 150 x 150 Å size using the CHARMM-GUI membrane builder [188, 189], and then solvated. Ions were added using GROMACS 2020.2 [190, 191] to neutralise any system charges and add ions to a concentration of either 150 mM NaCl, 150 mM CaCl₂, or a mixture of 75 mM NaCl and 75 mM CaCl₂. In the case of simulations containing Ca²⁺, the standard CHARMM36m Ca²⁺ molecules were then replaced with the multi-site Ca²⁺ of Zhang et al. [150]. HMR of the system was used to allow the use of 4-fs time steps in simulations of NaCl solutions [118, 119]. The multi-site Ca²⁺ model used for simulations of CaCl₂ however is incompatible with a 4-fs time step, and therefore any simulations including Ca²⁺ cations were performed with HMR but at a time step of 2-fs. The protein was restrained in the open-state by applying harmonic restraints on the α-carbon atoms of the lower gate residues (see Table 3.1).
3.3.2 Molecular dynamics simulations details

All simulations were performed using GROMACS 2020.2 [190, 191] and the CHARMM36m force field for the proteins, lipids, and ions [127]. The TIP3P water model was used to model solvent molecules [192]. The system was minimised and equilibrated using the suggested equilibration inputs from CHARMM-GUI [193]. In brief, the system was equilibrated using the NPT ensemble for a total time of 1.85 ns with the force constraints on the system components being gradually released over six equilibration steps. The systems were then further equilibrated by performing a 15 ns simulation with no electric field applied. To prevent closing of the lower-gate of the pore, harmonic restraints were applied to maintain the distance between the $\alpha$-carbon atoms of the lower gate residues of each respective chain (Table 3.1). To drive ion permeation, an external applied electric field was applied by using the method of Aksimentiev et al. [162] to production simulations with an $E_0$ of -0.03 V nm$^{-1}$; this resulted in a transmembrane voltage of $\sim$410 mV with negative polarity in the intracellular region. The temperature was maintained at 310 K using the Nosé-Hoover thermostat [159] and the pressure was maintained semi-isotropically at 1 bar using the Parrinello-Rahman barostat [160]. Periodic boundary conditions were used throughout the simulations. Long-range electrostatic interactions were modelled using the particle-mesh Ewald method [194] with a cut-off of 12 Å. The LINCS algorithm [117] was used to constrain bonds with hydrogen atoms. All individual simulations were 250 ns long and repeated five times for each system, as summarised in Tables 3.2, 3.3, 3.4, 3.5, and 3.6.
### 3 CA\(^{2+}\)-SELECTIVE PERMEATION MECHANISM IN THE TRPV FAMILY OF ION CHANNELS

#### Table 3.2: Summary of simulation details of the Ca\(^{2+}\)-selective TRPV5 channel.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPV5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>6DMU (262-639 PI(4,5)P2)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
<tr>
<td>Ligand</td>
<td>PI(4,5)P2 (CGenFF)</td>
</tr>
<tr>
<td>Ion</td>
<td>150 mM CaCl(_2) 291 Ca(^{2+}) (Zhang et al.) 574 Cl(^-) (CHARMM36m)</td>
</tr>
<tr>
<td>Independent simulations</td>
<td>5</td>
</tr>
<tr>
<td>Total simulation time (µs)</td>
<td>1.25</td>
</tr>
<tr>
<td>Estimated voltage (mV)</td>
<td>-410</td>
</tr>
<tr>
<td>Total permeations</td>
<td>85</td>
</tr>
</tbody>
</table>

#### Table 3.3: Summary of simulation details of the Ca\(^{2+}\)-selective TRPV6 channel.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPV6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>6BO8 (262-638)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
<tr>
<td>Ligand</td>
<td>-</td>
</tr>
<tr>
<td>Ion</td>
<td>150 mM CaCl(_2) 279 Ca(^{2+}) (Zhang et al.) 562 Cl(^-) (CHARMM36m)</td>
</tr>
<tr>
<td>Independent simulations</td>
<td>5</td>
</tr>
<tr>
<td>Total simulation time (µs)</td>
<td>1.25</td>
</tr>
<tr>
<td>Estimated voltage (mV)</td>
<td>-410</td>
</tr>
<tr>
<td>Total permeations</td>
<td>189</td>
</tr>
</tbody>
</table>

#### Table 3.4: Summary of simulation details of the non-selective TRPV3 channel.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPV3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>6PVP (375-745)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
<tr>
<td>Ligand</td>
<td>-</td>
</tr>
<tr>
<td>Ion</td>
<td>150 mM CaCl(_2) 291 Ca(^{2+}) (Zhang et al.) 582 Cl(^-) (CHARMM36m)</td>
</tr>
<tr>
<td>Independent simulations</td>
<td>5</td>
</tr>
<tr>
<td>Total simulation time (µs)</td>
<td>1.25</td>
</tr>
<tr>
<td>Estimated voltage (mV)</td>
<td>-410</td>
</tr>
<tr>
<td>Total permeations</td>
<td>941</td>
</tr>
</tbody>
</table>
Table 3.5: Summary of simulation details of the non-selective TRPV2 channel.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPV2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>6BO4 (327-691)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
<tr>
<td>Ligand</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ion</th>
<th>150 mM CaCl$_2$</th>
<th>150 mM NaCl</th>
<th>75 mM CaCl$_2$ + 75mm NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>253 Ca$^{2+}$ (Zhang et al.)</td>
<td>253 Na$^+$ (CHARMM36m)</td>
<td>127 Ca$^{2+}$ (Zhang et al.)</td>
</tr>
<tr>
<td></td>
<td>526 Cl$^-$ (CHARMM36m)</td>
<td>273 Cl$^-$ (CHARMM36m)</td>
<td>127 Na$^+$ (CHARMM36m)</td>
</tr>
<tr>
<td></td>
<td>401 Cl$^-$ (CHARMM36m)</td>
<td></td>
<td>401 Cl$^-$ (CHARMM36m)</td>
</tr>
</tbody>
</table>

| Independent simulations | 5 | 5 | 5 |
| Total simulation time (µs) | 1.25 | 1.25 | 1.25 |
| Estimated voltage (mV) | -410 | -410 | -410 |
| Total permeations | 60 | 59 | 72 |

Table 3.6: Summary of simulation details of additional control simulations of the Ca$^{2+}$-selective TRPV5 channel.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPV5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>6DMU(262-639PI(4,5)P2)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
<tr>
<td>Ligand</td>
<td>PI(4,5)P2 (CGenFF) PI(4,5)P2 deprotonated (CGenFF)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ion</th>
<th>75 mM CaCl$_2$ + 75mm NaCl</th>
<th>25 mM CaCl$_2$</th>
<th>150 mM CaCl$_2$</th>
<th>150 mM NaCl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>143 Ca$^{2+}$ (Zhang et al.)</td>
<td>52 Ca$^{2+}$ (Zhang et al.)</td>
<td>291 Ca$^{2+}$ (CHARMM36m)</td>
<td>311 Na$^+$ (CHARMM36m)</td>
</tr>
<tr>
<td></td>
<td>151 Na$^+$ (CHARMM36m)</td>
<td>96 Cl$^-$ (CHARMM36m)</td>
<td>574 Cl$^-$ (CHARMM36m)</td>
<td>287 Cl$^-$ (CHARMM36m)</td>
</tr>
</tbody>
</table>

| Independent simulations | 3 | 3 | 3 | 3 |
| Total simulation time (µs) | 0.75 | 0.3 | 0.75 | 0.45 |
| Estimated voltage (mV) | -205 | -410 | -410 | -410 |
| Total permeations | 22 | 2 | 0 | 95 |

### 3.3.3 Simulation analysis

Analysis of MD trajectory data was performed using in-house written Python scripts, utilising GROMACS modules [190, 191], the SciPy library of tools [195, 196, 197, 198], and MDAnalysis [199, 200]. Analysis of the pore architecture was performed using CHAP [201]. All plots were generated in Python using Matplotlib [202] and Seaborn [203].

**Calculating conductance and selectivity from in silico electrophysiology experiments**

The conductance of the channels ($G_{\text{channel}}$) was calculated according to Equation 3.1, where $N_p$ is the number of permeation events, $Q_{\text{ion}}$ is the charge of the permeating ion in Coulomb, $t_{\text{traj}}$ is the length of the trajectory, and $V_m$ is the transmembrane voltage. The mean conductance and standard error were calculated from overlapping 50 ns windows of the trajectory.

$$G_{\text{channel}} = \frac{N_p \times Q_{\text{ion}}}{t_{\text{traj}} \times V_m} \quad (3.1)$$
3 CA2+-SELECTIVE PERMEATION MECHANISM IN THE TRPV FAMILY OF ION CHANNELS

The selectivity ($P_{Ca}/P_{Na}$) was calculated as the ratio between the total sum of Ca2+ permeation events and the total sum of Na+ permeation events, from five-fold replicated 250 ns simulations in di-cationic solutions with 75 mM CaCl2 and 75 mM NaCl.

Identification of cation binding sites from MD simulations of TRPV channels

Cation binding sites were identified by plotting timeseries of each permeating ion with respect to their position along the pore axis. To further validate this, the Featurizer function of PENSA [204] was used to identify the 12 most occupied ion binding sites, as determined by 3D density maxima within 7 Å of the protein. This analysis was performed on a trajectory of concatenated five-fold replicated 250 ns simulations with a 200 ps time-step from mono-cationic simulations.

Calculating ion occupancy probabilities and residence times in the identified cation binding sites

The ion occupancy ($O_{ion}$) of the identified cation binding sites was calculated by dividing the number of frames ($N_{occupied}$), in which an ion’s centre of geometry is within 3.5 Å of the centre of geometry of the ion co-ordinating binding site atoms by the number of frames in the time window ($N_{frames}$). These atoms for the respective binding sites were: the carboxylate oxygen atoms of the $\alpha$-position residue of the SF for binding site A, the carbonyl oxygen atoms of the $\beta$- and $\gamma$-position residues of the SF for binding site B, and the terminal carbon atoms of the hydrocarbon sidechain of the isoleucine of the hydrophobic lower gate (I575 in TRPV5) and the amide oxygen atoms of neighbouring asparagine residue (N572 in TRPV5). A cut-off distance of 3.5 Å was chosen based upon the maximum reported distance for calcium-oxygen interactions [205, 206, 207]. The mean ion occupancies and standard error were calculated from non-overlapping 50 ns windows of the five-fold replicated 250 ns simulation trajectories with a 20 ps time-step.

The ion residence times ($t_r$) were calculated by averaging the amount of time an individual ion was located within 3.5 Å of the centre of geometry of the ion co-ordinating binding site atoms. The mean $t_r$ and standard error were calculated from five-fold repli-
Characterising permeation co-operativity through mutual information using SSI from PENSA

To characterise the level of co-operativity in the knock-on permeation mechanisms in TRPV channels, I used PENSA to calculate the state-specific information (SSI) shared between discrete state transitions in the occupancy distributions of each binding site [5, 204]. A timeseries distribution with a timestep of 20 ps for each binding site was obtained, whereby for each frame, if an ion occupied the binding site then this ion’s atom ID number was recorded, whereas if the binding site was unoccupied, an ID of -1 was recorded. The ID numbers were discrete, and changes between ID numbers in each binding site therefore represent discrete state transitions. By quantifying the mutual information shared between changes to the ID numbers in each site, I was able to determine whether ion transitions at one site were coupled to transitions at another during a 20 ps time interval. From this I concluded whether cations are “knocking” each other, or dissociation occurred independently from one another. The time interval was iteratively optimised to keep noise and finite sampling effects to a minimum (see below). It was found that both were smallest when an interval of 20 ps was used.

Similar to McClendon et al., it was observed that finite sampling resulted in independent distributions sharing mutual information [208, 209]. To overcome this, a statistical threshold was calculated for each simulation via randomly permuted copies of the original data. Random permutations of the original data maintained marginal probabilities for binding site occupation in each simulation while at the same time quantifying the effect of finite sampling on the measurement of state-specific information. State-specific information was then calculated between two independently permuted versions of the occupancy distribution for the minimum entropy binding site. Since the upper bound of mutual information between two variables is equal to the lowest entropy of those variables, the binding site corresponding to the lowest entropy was used for obtaining the threshold. This ensured that the portion of SSI which could be attributed to random noise...
between any two binding sites was always less than or equal to the SSI. This measurement was repeated 1,000 times in order to resolve a Gaussian distribution from which the 99% confidence threshold was obtained. This threshold was subtracted from the measured values to resolve excess mutual information, or excess SSI (exSSI), shared in discrete state transitions. As it is not possible to transfer negative information, negative exSSI values were corrected to a value of 0. The maximum SSI value was also derived, representing a theoretical upper limit for the information that can be shared between two binding sites, where exSSI\textsubscript{max} is given by subtracting the random threshold from the minimum entropy of the two binding sites in question.

\[
exSSI(A,B)_{\text{max}} = \min(H(A),H(B)) - \text{threshold}(A,B) \tag{3.2}
\]

To quantify the interdependence of all three ion binding sites within the TRPV pores, the total correlation (TotCorr) was obtained using Equation 3.3, where \(H(A), H(B), H(C)\) represent the entropy of binding sites \(A, B\) and \(C\), respectively, and \(H(A,B,C)\) the joint entropy of binding sites \(A,B\) and \(C\).

\[
\text{TotCorr} = H(A) + H(B) + H(C) - H(A,B,C) \tag{3.3}
\]

Characterising the architecture of the selectivity filter of TRPV channels

To determine the area formed between residues in the SF, the area of the quadrilateral between the adjacent chains was calculated on the \(X\) and \(Y\) axes. For this, I used the carboxylate oxygen atoms of the adjacent chains for the \(\alpha\)-position residue, and the carbonyl oxygen atoms for the \(\beta\)-, \(\gamma\)-, and \(\delta\)-position residues.

To quantify the SF areas from my MD simulations, the mean and standard error were calculated from non-overlapping 50 ns windows of the five-fold replicated 250 ns trajectories with a 200 ps time-step. Furthermore, the SF areas of all TRPV structures deposited in the PDB, available as of 4\textsuperscript{th} February 2022, were determined. A total of 101 structures were analysed, with non-tetrameric structures or structures without all the atoms of interest modelled not included. The mean and standard error of the mean were calculated for
3.4 Results

3.4.1 Continuous permeation of Ca\(^{2+}\) and Na\(^{+}\) in open-state TRPV5 and TRPV6 channels

I performed MD simulations of the pore domain of open-state TRPV5 [185] and TRPV6 [186] channels embedded in POPC lipid bilayers under transmembrane voltage (∼-410 mV). The aqueous solutions contained either 150 mM CaCl\(_2\) or 150 mM NaCl (herein referred to as mono-cationic solutions), or a mixture consisting of 75 mM CaCl\(_2\) and 75 mM NaCl (herein referred to as di-cationic solutions). All simulations performed with Ca\(^{2+}\) ions utilised the multi-site Ca\(^{2+}\) model developed by Zhang et al. [150], unless otherwise stated. In both the mono-cationic and the di-cationic solutions, the applied voltage drove a continuous flow of permeating ions through all investigated open-state TRPV channels. Overall, I recorded 433 complete inward channel crossings for Ca\(^{2+}\) and 417 for Na\(^{+}\) in simulations of the Ca\(^{2+}\)-selective TRPV channels.

In TRPV5 and TRPV6, Ca\(^{2+}\) ions traversed the entire pore within average time spans of 28.4 ± 3.9 ns (TRPV5) and 12.0 ± 1.0 ns (TRPV6) (Table 3.7). This greater permeation time of Ca\(^{2+}\) cations is likely a result of their greater affinity for the cation binding sites. However, the difference in permeation time is less than one would expect based on the residency times (Figure 3.8). For example, the average time for a Ca\(^{2+}\) cation to permeate through TRPV5 is ∼1.5-fold slower than Na\(^{+}\) permeation, whereas the residency time of cations at binding site A of TRPV5 would suggest a 12-fold difference. This supports the concept of co-operativity between successive binding sites increasing the unbinding rate of ions from binding sites, as proposed by Hille [172].

The calculated Ca\(^{2+}\) and Na\(^{+}\) conductances from my simulations are shown in (Table 3.8). The considerable Na\(^{+}\) conductances I observed agree with the experimental finding that the highly Ca\(^{2+}\)-selective TRPV channels conduct Na\(^{+}\) well in the absence of Ca\(^{2+}\) [176, 37, 177, 178]. Notably, these conductances are in quantitative agreement with
3 CA$^{2+}$-SELECTIVE PERMEATION MECHANISM IN THE TRPV FAMILY OF ION CHANNELS

Table 3.7: Average time to permeate through the TRPV pore, as defined by the $z$ position between binding sites A and C. The mean permeation time and standard error of the mean were calculated from five-fold replicated 250 ns simulations in mono-cationic 150 mM CaCl$_2$ or 150 mM NaCl. The average permeation time was defined as the time taken between the cation binding at binding site A, and dissociating from binding site C into the bulk solution.

<table>
<thead>
<tr>
<th>Permeation time (ns)</th>
<th>Ca$^{2+}$</th>
<th>Na$^+$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV2</td>
<td>28.4 ± 7.6</td>
<td>7.0 ± 1.8</td>
</tr>
<tr>
<td>TRPV3</td>
<td>6.3 ± 1.2</td>
<td>2.8 ± 0.2</td>
</tr>
<tr>
<td>TRPV5</td>
<td>28.4 ± 3.9</td>
<td>18.2 ± 3.8</td>
</tr>
<tr>
<td>TRPV6</td>
<td>12.0 ± 1.0</td>
<td>12.1 ± 1.5</td>
</tr>
</tbody>
</table>

Published values measured for Na$^+$ in vitro [210, 38].

By contrast, control simulations of TRPV5 using the default CHARMM36m force field parameters for Ca$^{2+}$, but otherwise identical conditions, did not exhibit ion permeation; instead, the Ca$^{2+}$ ions remained tightly bound to the protein ion binding sites for the entire course of the simulations. This observation is reflective of the shortcomings of standard parameters for divalent cations in fixed-point charge force fields and highlights the improved accuracy of multi-site Ca$^{2+}$ models in simulating divalent cation permeation and reproducing in vitro conductances. In addition, no Cl$^-$ anions were observed to permeate TRPV channels in any of my simulations.

Table 3.8: Calculated conductances from MD simulations of ion permeation in Ca$^{2+}$-selective TRPV channels. Mean inward conductances and standard error of the mean (SEM) were calculated from overlapping 50 ns windows from five-fold replicated 250 ns simulations. Permeation of mono-cationic Ca$^{2+}$ or Na$^+$ cations was simulated in a solution of 150 mM CaCl$_2$ or 150 mM NaCl, respectively. Permeation of a di-cationic mixture of Ca$^{2+}$ and Na$^+$ cations was investigated in a solution of 75 mM CaCl$_2$ and 75 mM NaCl.

<table>
<thead>
<tr>
<th>Conductance (pS)</th>
<th>Ca$^{2+}$</th>
<th>Na$^+$</th>
<th>Ca$^{2+}$ &amp; Na$^+$</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV5</td>
<td>53 ± 7</td>
<td>49 ± 6</td>
<td>92 ± 10</td>
<td>59 ± 6 (Na$^+$) [210]</td>
</tr>
<tr>
<td>TRPV6</td>
<td>117 ± 12</td>
<td>61 ± 6</td>
<td>29 ± 6</td>
<td>58 ± 4 (Na$^+$) [38]</td>
</tr>
</tbody>
</table>

I note that the $P_{Ca}/P_{Na}$ values obtained from my simulations overall show lower Ca$^{2+}$ selectivity than the reported literature values (Table 3.9). I surmised that this might be, at least partially, due to the higher voltages used in my simulations to enhance the sampling rate. Supplementary simulations performed at a lower voltage demonstrated that, indeed, the selectivity for Ca$^{2+}$ increases with lower voltages across the membrane (Figure 3.4).
Table 3.9: Selectivity ratios of Ca\(^{2+}\) and Na\(^{+}\) permeation events from simulations of TRPV channels in a di-cationic solution.

<table>
<thead>
<tr>
<th>Channel</th>
<th>(P_{\text{Ca}}/P_{\text{Na}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV2</td>
<td>5.0</td>
</tr>
<tr>
<td>TRPV3</td>
<td>1.7</td>
</tr>
<tr>
<td>TRPV5</td>
<td>2.6</td>
</tr>
<tr>
<td>TRPV6</td>
<td>4.4</td>
</tr>
</tbody>
</table>

Below a voltage threshold of \(\sim 205\) mV, however, the sampling of permeation events in the simulations became very poor, such that I was not able to reliably probe the precise voltage range used in the experiments.

Additionally, I can not rule out the contribution of force field inaccuracies. My simulations with Na\(^{+}\) show a remarkable agreement between experimentally recorded and simulated conductance. Even though the Ca\(^{2+}\) model I used has been carefully parameterised [150], modelling divalent cations is a far from trivial task and this is the first multi-site Ca\(^{2+}\) model with which simulations of ion channel current have become possible. It can therefore not be excluded that further iterations of model refinement may eventually be required to not only reflect experimental solvation free energies and protein affinity [150], but also accurately reproduce experimental conductance values.

### 3.4.2 Pore cation binding sites and their preference for Ca\(^{2+}\)-binding

Prior to the determination of the atomic structures of Ca\(^{2+}\) channels and the development of channel-permeable models for Ca\(^{2+}\) ions, it had been suggested from experimental observations that Ca\(^{2+}\) channels may obtain their selectivity through competition, \textit{i.e.} by divalent cations, such as Ca\(^{2+}\), binding more tightly to their ion binding sites than monovalent cations, such as Na\(^{+}\) [171, 172].

By analysing the individual traces of permeating Ca\(^{2+}\) cations along the pore axis \(z\) of TRPV5 and TRPV6 over time, I identified three cation binding sites inside the channels (Figure 3.2). I refer to these cation binding sites as sites A, B and C, viewed from the extracellular entrance of the channel SF to the hydrophobic lower gate. The three cation binding sites were further confirmed by 3D density analysis using PENSA [204] (Figure 3.3). The PENSA analysis also identified a number of cation binding sites outside of
the pore within the extracellular loops of both TRPV5 and TRPV6 (Figure 3.3), in line
with the previous suggestion that TRPV6 contains negatively charged "recruitment sites"
that funnel cations towards the entrance of the pore [211, 212].

![Figure 3.2: Schematic of cation binding sites identified in the Ca\(^{2+}\)-selective TRPV5 channel. Permeation traces of the \(z\)-coordinate of permeating Ca\(^{2+}\) cations over time established that cations are bound at three regions within the pore (left). The location of the residues constituting these three binding sites in Ca\(^{2+}\)-selective TRPV channels is shown on the structure of TRPV5 (right). Please note, only cations that fully permeate through the pore within the 250 ns simulation are shown in the plot (left).](image)

Of the three binding sites I observed, binding site A is formed by the carboxylate oxygen atoms of the ring of acidic residues at the SF entrance (referred to here as the SF \(\alpha\)-position); binding site B is formed by the carbonyl oxygen atoms of the bottom two SF residues (SF \(\gamma\)- and \(\delta\)-positions); and binding site C is formed jointly by the hydrophobic gate consisting of a ring of isoleucine residues (I575 in TRPV5), and by the amide oxygen atoms of the neighbouring asparagine residues (N572 in TRPV5) near the cytoplasmic exit of the pore (Figure 3.2). The location of these binding sites coincides with constrictions in the pore profile, as determined using CHAP [201] (Figure 3.5). The distance between binding sites A and B is \(~5\ \text{Å}\), and that between binding sites B and C is \(~14\ \text{Å}\). I note that Hughes et al. [185] reported a further constriction below the hydrophobic gate (binding site C) formed by W583 in the TRPV5 structure, and an analogous constriction at W583 can be observed in TRPV6 [186]. However, my simulations do not suggest that the side chains of W583 constitute a functionally important ion binding site, as shown in Figure 3.6.
3 CA\textsuperscript{2+}.SELECTIVE PERMEATION MECHANISM IN THE TRPV FAMILY OF ION CHANNELS

Figure 3.3: Cation binding sites in TRPV5 identified by PENSA. 3D density maxima of Ca\textsuperscript{2+} cations within 7 Å of the protein was analysed to identify 12 cation binding sites shown as pseudo-atoms (left). This analysis identified binding sites A (red), B (orange), and C (yellow) and several other "recruitment sites" (purple). The location of these recruitment sites (purple) help to attract cations and funnel them towards the pore entrance (top right and top left). My MD simulations suggested that the following residues act as recruitment sites: E522, D525, T528, F531, S532, E535, Y547, and Y549.

Additionally, the cryo-EM structure utilised of TRPV5 of O. cuniculus was solved to a resolution of 4 Å and resolved a PI(4,5)P\textsubscript{2} molecule bound to each chain of the homotetrameric structure [185]. At such a resolution, it is not possible to accurately model the protonation state of the PI(4,5)P\textsubscript{2} molecule. Hughes et al. report that PI(4,5)P\textsubscript{2} binding induces conformational changes related to channel activity. As my simulation protocol included harmonic restraints on the lower gate of the protein to maintain the open state of the structure, the protonation state of the PI(4,5)P\textsubscript{2} would be inconsequential to the stability of the pore during my MD simulations. To confirm this, I performed three-fold 150 ns simulations of TRPV5 with deprotonated PI(4,5)P\textsubscript{2} molecules (molecule charge = -4), and compared the pore architecture to the production simulations of five-fold 250 ns simulations of TRPV5 with protonated PI(4,5)P\textsubscript{2} molecules. As can be seen in Figure 3.7, my simulations show no significant difference in the pore architecture dependent on the protonation state of the PI(4,5)P\textsubscript{2} in TRPV5 simulations. This demonstrates that my
Figure 3.4: The effect of high voltage on Ca\textsuperscript{2+}-selectivity in simulations of TRPV5 in di-cationic solutions. Simulations performed at a lower voltage of $\sim$205 mV (left) resulted in an increased occupancy probability of Ca\textsuperscript{2+} cations (orange) and a reduced occupancy probability of Na\textsuperscript{+} cations (blue) compared to simulations at a higher voltage of $\sim$410 mV (centre). This resulted in increased Ca\textsuperscript{2+}-selectivity in lower voltage simulations, as summarised in the $P_{Ca}/P_{Na}$ value (right).

simulation protocol and use of harmonic restraints on the lower gate are able to reliably constrain the protein structures in their open-state conformations.
Figure 3.5: Pore architecture of TRPV channels from MD simulations, showing the average radius and hydrophobicity of the channel with respect to the relative z coordinate, obtained using CHAP [201]. The mean radius or hydrophobicity (black) and standard deviation (grey) were calculated from concatenated trajectories of five-fold replicated 250 ns simulations in 150 mM CaCl$_2$ with a 200 ps time-step. The shaded grey region represents the standard deviation. The average position of binding sites A, B, and C are shown as shaded red, orange, and yellow regions, respectively. The dashed line in the pore radius plots indicates the radius of a dehydrated Ca$^{2+}$ ion.

In mono-cationic Ca$^{2+}$ solutions, the three binding sites showed Ca$^{2+}$ occupancy probabilities of 0.69 ± 0.05, 0.67 ± 0.04, and 0.57 ± 0.03 in TRPV5, and 0.43 ± 0.04, 0.54 ± 0.05, and 0.29 ± 0.02 in TRPV6 (from A to C, respectively; Figure 3.8). In mono-cationic Na$^+$ solutions, similar occupancies were observed. However, the Na$^+$ residence times ($t_r$) at the three binding sites were markedly lower than those observed for Ca$^{2+}$, with ratios of $t_r$(Ca$^{2+}$) : $t_r$(Na$^+$) varying between ~35:1 and ~3:1 (Figure 3.8). These residence times suggest that Ca$^{2+}$ ions have a greater affinity for these binding sites than Na$^+$. 

This observation was further substantiated when the occupancy of binding sites in di-cationic solutions was analysed, in which Ca$^{2+}$ and Na$^+$ cations are competing for the binding sites. Binding sites A, B and C in the Ca$^{2+}$-selective channels all showed high occupancies with Ca$^{2+}$ in the mixed cationic solutions (Figure 3.8). Across all the replicate simulations, I recorded average Ca$^{2+}$ occupancies of 0.51 ± 0.05, 0.48 ± 0.06, and 0.37 ± 0.03 for binding sites A, B and C in TRPV5, respectively, and of 0.68 ± 0.04, 0.61 ± 0.04, and 0.40 ± 0.04 for binding sites A, B and C in TRPV6 (Figure 3.8). By contrast, the Na$^+$ occupancy of each of the three binding sites under these conditions was found to be below 0.07, both in TRPV5 and TRPV6; that is, the ratio between Ca$^{2+}$ and
Figure 3.6: W583 does not form a functionally important cation binding site in simulations of TRPV5 in 150 mM CaCl$_2$. W583 is located on the S6 helix, below the hydrophobic lower gate (centre). Analysis of the occupancy probability (left) and the residence time (right) showed that the constriction formed by W583 does not coordinate Ca$^{2+}$ cations as efficiently as binding sites A, B, and C in my simulations.

Na$^+$ occupancy varies between ~85:1 and ~7:1 (Figure 3.8). These values indicate a free energy difference of between 11.5 kJ mol$^{-1}$ and 5.0 kJ mol$^{-1}$ for the preferential binding of Ca$^{2+}$ over Na$^+$. 

Effect of \(P(4, 5)P_2\) protonation on TRPV5 system with harmonic restraints

![Pore radius of protonated](image1)

![Pore hydrophobicity of protonated](image2)

![Pore radius of deprotonated](image3)

![Pore hydrophobicity of deprotonated](image4)

Figure 3.7: Pore architecture of TRPV5 channels from MD simulations with protonated (top) and deprotonated (bottom) \(P(4,5)P_2\) molecules. The plots show the mean radius and hydrophobicity of the channel with respect to the relative \(z\) coordinate. The shaded grey region represents the standard deviation. The average position of binding sites A, B, and C are shown as shaded red, orange, and yellow regions, respectively.

### 3.4.3 A highly co-operative knock-on mechanism between three cation binding sites underpins selective \(\text{Ca}^{2+}\) permeation in TRPV channels

The observed increased affinity for \(\text{Ca}^{2+}\) cations at the pore binding sites compared to \(\text{Na}^+\) means that, in a mixed solution, \(\text{Ca}^{2+}\) will preferentially occupy these binding sites; however, this also implies that \(\text{Ca}^{2+}\) ions face a greater energy barrier when they dissociate from the binding sites. In mono-cationic solutions, this would result in a greatly reduced \(\text{Ca}^{2+}\) conductance with respect to \(\text{Na}^+\). For instance, based upon the observed residence times in mono-cationic solutions, I would expect an approximately 12-fold reduced \(\text{Ca}^{2+}\) unbinding rate compared to \(\text{Na}^+\) for binding site A in TRPV5. However, a much reduced \(\text{Ca}^{2+}\) conductance is neither observed in my simulations nor in the experimental literature.
Figure 3.8: Occupancy probability (left) and residence times (right) of Ca\textsuperscript{2+} (orange columns) and Na\textsuperscript{+} (blue columns) cations from simulations of ion permeation in both mono-cationic and di-cationic ion solutions. The plots show the occupancy probability of all proteins in this study, including Ca\textsuperscript{2+}-selective (CS) and non-selective (NS) TRPV channels. The mean occupancy probability and SEM were calculated from non-overlapping 50 ns windows from five-fold replicated 250 ns simulations; the mean residence time and SEM from five-fold replicated 250 ns simulations. The location of the residues constituting binding sites A, B, and C in Ca\textsuperscript{2+}-selective TRPV channels is shown on the structure of TRPV5 (centre).

Due to the divalent charge of Ca\textsuperscript{2+}, increasing the affinity to cation binding sites, this dichotomy had previously been suggested to exist, and it was hypothesised that this paradox could be resolved by assuming co-operativity between successive unbinding events such as in a knock-on mechanism [171, 172].

In the classic knock-on mechanism, which for example underpins K\textsuperscript{+} channel function, ions transition into and out of multiple ion binding sites in a highly correlated fashion [213, 214, 215, 216, 105]. For example, early experiments by Hodgkin and Keynes and later flux-ratio measurements established that 3–3.4 K\textsuperscript{+} ions moved in lockstep with each other during permeation through K\textsuperscript{+} channels [217, 218].

For each permeating ion in a simulation of TRPV5, Figure 3.9 shows the association and dissociation of Ca\textsuperscript{2+} and Na\textsuperscript{+} ions at binding sites A, B and C from top to bottom as colour code (bound to A, red; bound to B, orange; bound to C, yellow; transiting within the pore but not bound to a binding site, blue; located in extracellular solvent, dark grey; located in intracellular solvent, light grey). As can be seen for Ca\textsuperscript{2+} in TRPV5 for example (Figure 3.9 left), the plot demonstrates that (i) permeating Ca\textsuperscript{2+} ions spend the vast majority of their time within the pore at the three binding sites (reflected in the scarcity
3  CA\textsuperscript{2+} -SELECTIVE PERMEATION MECHANISM IN THE TRPV FAMILY OF ION CHANNELS

of blue boxes vs. red, orange and yellow), (ii) dual and triple occupancy of the three sites, A, B, and C, with Ca\textsuperscript{2+} is frequently observed (horizontal slices across plot: triple occupancy is observed in 27.2\% of the simulation frames, dual occupancy in 49.7\%), and (iii) transitions between states show a high degree of correlation, i.e., the ions frequently move in concert into and out of their respective binding sites (horizontal slices; binding state transitions). By contrast, during Na\textsuperscript{+} permeation (Figure 3.9 centre), the ions are predominantly transiting across the pore without occupying particular binding sites for extended time spans (blue, on average 53\% of the traversal time for each ion).

![Figure 3.9: Permeation state plots of permeating Ca\textsuperscript{2+} (left) and Na\textsuperscript{+} (centre) cations through the Ca\textsuperscript{2+}-selective TRPV5. Permeation state plots shows the state of each permeating cation (columns) at a given time point (rows) by assigning a state to the ions to indicate whether the cation is bound to a binding site, transitioning between binding sites, or in the bulk solution: bound to A, red; bound to B, orange; bound to C, yellow; transiting within the pore but not bound to a binding site, blue; located in extracellular solvent, dark grey; located in intracellular solvent, light grey. Comparison of permeation state plots for Ca\textsuperscript{2+} and Na\textsuperscript{+} cations show that Ca\textsuperscript{2+} permeation proceeds in a well-ordered manner with three Ca\textsuperscript{2+} cations within the pore and knocking adjacent cations to the next binding site. Na\textsuperscript{+} permeation on the other hand is far less ordered, with regularly more than three Na\textsuperscript{+} cations within the pore at a given time. Each plot (left and centre) shows exemplars from a single 250 ns simulation of TRPV5 performed in mono-cationic 150 mM CaCl\textsubscript{2} or 150 mM NaCl, respectively. The structure of TRPV5 shows the colours used in the permeation state plots and the location of the residues constituting the three binding sites (right). Please note, only cations that fully permeate through the pore within the 250 ns simulation are shown in the plots, whereas the binding site occupancies reported above reflect both permeating and non-permeating ions (left and right).

In order to go beyond visual inspection of the trajectories and to assess the co-operativity of ion permeation in a quantitative way, a new approach based on mutual information was developed, taking into account the ”state” of each ion binding site. To achieve this, I assigned a specific binding state (unoccupied, or occupied with a specific ion) to binding sites A, B, and C and used the recently developed approach, state-specific information
on pairs of adjacent sites to quantify the degree of coupling between ion binding transitions at each of these sites (see Methods). This analysis yields a coefficient quantifying the co-operativity between ion binding and unbinding at neighbouring or more distant binding sites, where a greater coefficient signifies a higher degree of coupling. This coupling suggests that when an ion transitions from one site it is more likely there is a transition at the other. To correct for the non-zero mutual information that samples of two completely independent variables can display due to finite-size effects, the approach of McClendon et al. [208] and Pethel et al. [209] was used to yield excess mutual information, or excess SSI (exSSI). Additionally, a theoretical upper limit was determined for the maximum mutual information that can be shared between two binding sites by using the minimum state entropy among the two sites. Note that this quantity represents an absolute upper limit; reaching it would require both binding sites to exhibit idealised simultaneous states and state transitions throughout the entire simulated time.

The SSI analysis showed that in the Ca\(^{2+}\)-selective TRPV channels under the simulated conditions, TRPV5 and TRPV6, a high level of information above noise is shared between the transition of ions into and out of binding sites A and B, respectively, both for the permeation of Ca\(^{2+}\) and Na\(^{+}\) (exSSI between 0.8–1.6 bits, Table 3.10 and Figure 3.10). This suggests that the ion binding and unbinding processes at each of these binding sites are coupled to one another, constituting a knock-on mechanism at relatively short range. I observed 3-4 water molecules on average between cations bound at binding sites A and B during knock-on, demonstrating a "soft" knock-on mechanism to be in place, as opposed to the "direct" knock-on mechanism between dehydrated K\(^{+}\) ions in K\(^{+}\)-selective cation channels [105]. As detailed further below, my simulations indicate that only a moderate level of ion desolvation occurs in the SF of the studied TRPV channels, such that the hydration shell of the permeating ions remains largely intact during knock-on.

Similarly, the transitions of ions into and out of binding sites B and C show a large degree of correlation for both Ca\(^{2+}\) and Na\(^{+}\) (Figure 3.10). In the case of binding sites B and C however, this requires a remote knock-on mechanism to be in operation, since these sites are \(~\sim 14 \text{ Å} \) apart. The concept of a remote knock-on event was first proposed by Tind-
Table 3.10: Calculated \( \text{exSSI} \) and \( \text{exSSI}_{\text{norm}} \) values of cation transition from binding sites in MD simulations of TRPV channels. The mean \( \text{exSSI} \) or \( \text{exSSI}_{\text{norm}} \), and standard error of the mean, were calculated from five-fold replicated 250 ns simulations in mono-cationic solutions of 150 mM CaCl\(_2\) or 150 mM NaCl.

<table>
<thead>
<tr>
<th>Cation</th>
<th>( \text{exSSI}(A, B) )</th>
<th>( \text{exSSI}(A, B)_{\text{norm}} )</th>
<th>( \text{exSSI}(B, C) )</th>
<th>( \text{exSSI}(B, C)_{\text{norm}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Ca}^{2+} )</td>
<td>0.15 ± 0.06</td>
<td>0.12 ± 0.05</td>
<td>0.88 ± 0.13</td>
<td>0.44 ± 0.03</td>
</tr>
<tr>
<td>( \text{Na}^+ )</td>
<td>0.26 ± 0.13</td>
<td>0.18 ± 0.06</td>
<td>0.13 ± 0.03</td>
<td>0.18 ± 0.02</td>
</tr>
<tr>
<td>TRPV3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Ca}^{2+} )</td>
<td>0.34 ± 0.09</td>
<td>0.17 ± 0.04</td>
<td>0.96 ± 0.17</td>
<td>0.32 ± 0.05</td>
</tr>
<tr>
<td>( \text{Na}^+ )</td>
<td>0.05 ± 0.04</td>
<td>0.03 ± 0.02</td>
<td>0.26 ± 0.16</td>
<td>0.10 ± 0.06</td>
</tr>
<tr>
<td>TRPV5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Ca}^{2+} )</td>
<td>1.25 ± 0.15</td>
<td>0.50 ± 0.05</td>
<td>1.01 ± 0.25</td>
<td>0.40 ± 0.06</td>
</tr>
<tr>
<td>( \text{Na}^+ )</td>
<td>1.05 ± 0.31</td>
<td>0.30 ± 0.06</td>
<td>0.84 ± 0.17</td>
<td>0.28 ± 0.04</td>
</tr>
<tr>
<td>TRPV6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \text{Ca}^{2+} )</td>
<td>0.81 ± 0.16</td>
<td>0.31 ± 0.03</td>
<td>0.53 ± 0.25</td>
<td>0.27 ± 0.04</td>
</tr>
<tr>
<td>( \text{Na}^+ )</td>
<td>1.55 ± 0.18</td>
<td>0.39 ± 0.03</td>
<td>0.88 ± 0.10</td>
<td>0.32 ± 0.03</td>
</tr>
</tbody>
</table>

jong et al. based upon Brownian dynamics simulations [219], and observed by Zhang et al. in atomistic MD simulations of Ca\(^{2+}\) permeation in the RyR1 channel [150]. Our SSI analysis suggests that the degree of co-operativity in the remote knock-on mechanism between binding sites B and C (Figure 3.10 right) is slightly smaller than the co-operativity in the closer knock-on mechanism between binding sites A and B (Figure 3.10 left).

Figure 3.10: Excess state-specific information (exSSI) between ion binding sites quantifies the degree of co-operativity in the knock-on mechanism of cation permeation in TRPV channels. The mean exSSI and SEM between transitions from binding sites A and B (left) and binding sites B and C (right) are shown for Ca\(^{2+}\) (orange) and Na\(^+\) (blue) cations in mono-cationic solutions. For each exSSI, the mean maximum exSSI\(_{\text{max}}\) and standard error is also shown (grey), and the exSSI\(_{\text{norm}}\) is reported in Table 3.10.

To further investigate this cation co-operativity, I also investigated the effect of reduced cation concentrations and transmembrane voltages. \textit{In vitro} electrophysiology experiments of TRPV5 and TRPV6 in the literature have been conducted with a maximum Ca\(^{2+}\) concentration of 30 mM [37, 220], and concentrations of up to 30 mM Ca\(^{2+}\) were
used in studies of other TRPV channels [221]. Moreover, many electrophysiological recordings on TRPV channels were performed at Na\(^+\) concentrations of 150 mM [176]. I therefore believe that the concentrations I probed in my simulations do not deviate too strongly from realistic concentrations used under experimental conditions. However, to confirm that the co-operativity and knock-on mechanism observed was not as a consequence of the higher cation concentrations, I also performed simulations of TRPV5 in a lower Ca\(^{2+}\) concentration of 25 mM CaCl\(_2\), rather than 150 mM CaCl\(_2\), which is well within the range of concentrations used in *in vitro* experiments. As can be seen in Figure 3.11, the knock-on mechanism between Ca\(^{2+}\) ions is preserved at a lower Ca\(^{2+}\) concentration. Specifically, two three-site knock-on permeation events can be seen at \(~50\) and \(~78\) ns.

![Permeation traces of the z-coordinate of permeating Ca\(^{2+}\) cations over time in a lower concentration mono-cationic solution of 25 mM CaCl\(_2\).](image)

Figure 3.11: Permeation traces of the z-coordinate of permeating Ca\(^{2+}\) cations over time in a lower concentration mono-cationic solution of 25 mM CaCl\(_2\).
Furthermore, permeation traces from the lower voltage simulations of TRPV5 in a di-cationic solution (as used for Figure 3.4), showed that a clear and distinct knock-on mechanism between permeating Ca\(^{2+}\) cations is present in lower voltage simulations. This confirms that the observed knock-on mechanism and cation co-operativity is not a result of the increased voltage used in these simulations as compared to experiment.

Figure 3.12: Permeation traces of the \(\varepsilon\)-coordinate of permeating Ca\(^{2+}\) cations over time in a di-cationic solution of 75 mM CaCl\(_2\) and 75 mM NaCl at a lower voltage of -205 mV.

### 3.4.4 Cation permeation in non-selective TRPV channels shows a lower degree of co-operativity

To determine if the remaining, non-selective TRPV channels showed a different permeation mechanism, I next performed simulations of the open-state TRPV2 [183] and TRPV3 [184] channels using the same simulation approach as described for the Ca\(^{2+}\)-selective TRPV channels. These simulations of non-selective TRPV channels also showed continuous ion permeation, with cation conductances, again, in good agreement with published conductance values measured \textit{in vitro} (Table 3.11). Overall, I recorded 706 complete inward channel crossings for Ca\(^{2+}\) and 1,176 for Na\(^+\) from simulations of the
The occupancy of binding sites B and C in the TRPV2 and TRPV3 systems showed no clear difference to the Ca\textsuperscript{2+}-selective channels. By contrast, the occupancy of binding site A was markedly reduced for both Na\textsuperscript{+} and Ca\textsuperscript{2+} ions in the mono-cationic solutions (Figure 3.8). This suggests that cations are less well coordinated at binding site A in the non-selective TRPV channels, leading to lower affinity binding in the mono-cationic solutions. All binding sites, however, exhibited a preference for binding Ca\textsuperscript{2+} in the di-cationic solutions.

Table 3.11: Calculated conductances from MD simulations of ion permeation in non-selective TRPV channels. Mean inward conductances and standard error of mean (SEM) were calculated from overlapping 50 ns windows from five-fold replicated 250 ns simulations. Mono-cationic solutions at 150 mM concentration; di-cationic mixture of Ca\textsuperscript{2+} and Na\textsuperscript{+} at a concentration of 75 mM each.

<table>
<thead>
<tr>
<th></th>
<th>Conductance (pS)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ca\textsuperscript{2+}</td>
</tr>
<tr>
<td>TRPV2</td>
<td>40 ± 4</td>
</tr>
<tr>
<td>TRPV3</td>
<td>196 ± 10</td>
</tr>
</tbody>
</table>

I was therefore curious if the three-site knock-on mechanism described previously for Ca\textsuperscript{2+}-selective TRPV channels is also at play in the non-selective TRPV channels. Our SSI analysis confirmed that the co-operativity between binding sites B and C in the non-selective TRPV channels was comparable to those calculated for the Ca\textsuperscript{2+}-selective TRPV channels (Table 3.10 and Figure 3.10). However, the correlation between ion binding transitions at binding sites A and B was substantially reduced in both of the non-selective TRPV systems (Figure 3.10). The nearly complete absence of co-operativity from binding sites A and B demonstrates that a knock-on mechanism is not occurring between these two sites in the non-selective TRPV channels. Instead, my findings suggest that cation permeation in the non-selective TRPV channels occurs via a two-site knock-on mechanism between binding sites B and C.

Since the ion occupancy observed at binding site A is reduced in the case of the non-selective TRPV channels, it is plausible that this lower affinity also impacts on the coupling between transitions at binding sites A and B. To test this notion further, the relationship between affinity differences of a pair of binding sites and the knock-on co-operativity...
was tested systematically by using a toy model with two energy wells (binding sites) possessing a range of different depths (affinities). This toy model consisted of binding site occupancy data generated for two binding sites, binding sites $\varepsilon$ and $\zeta$, for the equivalent of a 100 ns simulation with the ion occupation state of each binding site reported every 20 ps, as in the analysis of real simulation trajectories. In this simple model, the occupancy probability of binding site $\varepsilon$ was 0.9, and the occupancy probability of binding site $\zeta$ was adjusted to investigate the effect of differing binding site affinities. The ratio of occupancy probabilities between $\varepsilon:\zeta$ was: 1:1, 1:0.75, 1:0.5, 1:0.25, 1:0.1, 1:0.01. As shown in Figure 3.13, there is a linear relationship between the affinity difference and the observed exSSI. This demonstrates that the diminished affinity of binding site A is likely to be the major reason for the loss of co-operativity in the SF of the non-selective TRPV channels. Our results show that similar binding affinity is a necessary but not sufficient condition for a high degree of co-operativity between two cation binding sites.

Based on the SSI approach to quantify mutual information in binding and unbinding events at different ion binding sites, and using the concept of total correlation to evaluate the overall co-operativity in a system across all coupled events, I next calculated the total correlation of ion permeation for all the TRP channels investigated. The reduction in the number of correlated knock-on sites within the non-selective TRP channels can be distinctly seen when comparing the total correlation for each of the non-selective and Ca$^{2+}$-selective TRPV channels (Figure 3.14).
Figure 3.13: The effect of similar binding site affinities on exSSI generated from a model of two consecutive binding sites. When the binding site affinity of binding site $\xi$ was reduced relative to binding site $\zeta$, the exSSI also decreased in a linear fashion.

The ion binding sites in non-selective TRPV channels preferentially bind Ca$^{2+}$ over Na$^+$, as in Ca$^{2+}$-selective TRPV channels, which explains why the non-selective TRPV channels in fact show slight Ca$^{2+}$-selectivity ($P_{Ca}/P_{Na} \sim 10:1$ [39]). However, my data suggests that the reduced level of coordination at binding site A, and especially the effect on the three-site co-operativity it imparts together with sites B and C, reduces the Ca$^{2+}$-selectivity from $P_{Ca}/P_{Na} \sim 100:1$ seen in TRPV5 and TRPV6, and in this way ultimately determines the difference between Ca$^{2+}$-selective and non-selective permeation.

3.4.5 Structural features distinguishing Ca$^{2+}$-selective from non-selective permeation

To understand why cation coordination at binding site A is weakened in non-selective TRPV channels, decoupling its co-operativity, I investigated the area formed between the
Figure 3.14: Total correlation of cation permeation between cation binding sites from simulations of TRPV channels. Comparison of the total correlation showed that the Ca\(^{2+}\)-selective TRPV channels have greater total correlation than non-selective TRPV channels (centre; Ca\(^{2+}\), orange bars; Na\(^{+}\), blue bars). This greater total correlation in Ca\(^{2+}\)-selective TRPV channels is a consequence of a knock-on mechanism between three binding sites (left). However, in non-selective TRPV channels, cation co-ordination at binding site A is reduced, resulting in reduced co-operativity, and a two binding site knock-on mechanism between binding sites B and C only (right).

four subunits of the TRPV channel for each residue in the selectivity filter. The crosssectional area formed by the carboxylate oxygen atoms at the SF \(\alpha\)-position is larger in the non-selective TRPV channels than in the Ca\(^{2+}\)-selective TRPV channels (Figure 3.17). Interestingly, despite the increased average area formed by the carboxylate oxygen atoms at the \(\alpha\)-position residue in the selectivity filter, the average area formed by the carbonyl oxygen atoms at both the \(\gamma\)- and \(\delta\)-positions are smaller in non-selective TRPV channels than in Ca\(^{2+}\)-selective TRPV channels (Figure 3.17). These differences in selectivity architecture were further confirmed using the pore profile calculated using CHAP [201] (Figure 3.5). Our simulations showed no differences in the flexibility of selectivity filters between Ca\(^{2+}\)-selective and non-selective TRPV channels as determined by RMSF calculations of the backbone atoms (Table 3.12). In contrast, analysis of the backbone dihedral angles showed some differences in the angle distribution between Ca\(^{2+}\) and non-selective TRPV channels (Figure 3.15). The most obvious difference was in the \(\beta\)-position of the SF, where non-selective TRPV channels showed wider distribution of the \(\phi\) angle, and principally on the \(\psi\) angle. Comparison of the amino acid composition of the SF of TRPV channels shows that the \(\beta\)-position residue in Ca\(^{2+}\)-selective TRPV channels is
usually an isoleucine residue; however, in non-selective TRPV channels this β-residue is a much smaller glycine residue (Figure 3.17). As the sidechain of the β-position residue faces into hydrophobic pocket between the SF and pore-helix, the hydrophobic sidechain of isoleucine is involved in hydrophobic interactions with this region, increasing the backbone stability of the β-position in Ca²⁺-selective TRPV channels. On the other hand, the glycine residue at the β-position in the non-selective TRPV channels will have greater backbone flexibility. Furthermore, the increased flexibility at this β-position will affect the adjacent α- and γ-positions, explaining why these two positions show differences in their backbone dihedral angle distributions between Ca²⁺-selective and non-selective TRPV channels, but no difference in the δ-position.

To expand the geometric analysis to all available TRPV channel structures, I also calculated the average area formed by selectivity filter residues from all available TRPV structures deposited in the Protein Data Bank (PDB) (Figure 3.17). Analysis of these static structures showed the same two trends observed within my MD simulations: (1) Ca²⁺-selective TRPV channels clearly have a smaller average area formed by the carboxylate oxygen atoms of the SF α-position; and (2), non-selective TRPV channels have a slightly narrower constriction at the SF γ- and δ-position residues. Our MD simulations suggested that the wider opening at the α-position leads to a weaker cation binding interaction at binding site A and cation coordination, which, in turn, decouples binding site A from the co-operative knock-on mechanism that underpins permeation in the Ca²⁺-selective TRPV channels.

The non-selective channels possess a narrower constriction formed by the side chains and carbonyl groups of the γ- and δ-position residues forming binding site B, whereas

Table 3.12: Root mean square fluctuation (RMSF) of the backbone of SF residues of TRPV channels from MD simulations. The mean RMSF and standard error of the mean for each residue was calculated from five-fold replicated 250 ns simulations of each channel in 150 mM CaCl₂.
the Ca\textsuperscript{2+}-selective TRPV channels exhibit an even narrower constriction at binding site A. Our simulations suggest that the greater occupancy of binding site A is a result of the more confined geometry of this charged binding site in the Ca\textsuperscript{2+}-selective TRPV channels. This, in turn, leads to a higher occupancy of the uncharged binding site B in the Ca\textsuperscript{2+}-selective TRPV channels (see Figure 3.8), as binding site B receives cations from the adjacent binding site A via the knock-on mechanism, rather than having to bind them from bulk solution, as would be the case in the non-selective channels. Thus, despite the constriction not being as narrow, binding site B has a higher occupancy in the Ca\textsuperscript{2+}-selective than in the non-selective TRPV channels.
3.4.6 Ca$^{2+}$-selective permeation is not strongly linked to the solvation states of permeation cations

A previously reported mechanism of how cation selectivity can be achieved is by desolvation of permeating cations. Differences in desolvation energies between cationic species provide a thermodynamic penalty that can be more favourable for the permeation of one cationic species over another. Such a mechanism has been reported to underpin K$^+$-selectivity over Na$^+$ in K$^+$ channels [105, 106], for example. Alternative mechanisms suggested to yield K$^+$ selectivity are based on protein flexibility, especially the plasticity of the K$^+$ channel SF [224], and the number of stacked ion binding sites in the SF [225], where the reduction from four to three stacked K$^+$ binding sites has been shown to abolish K$^+$ selectivity. By contrast, Na$^+$ selectivity has been suggested to rely chiefly on a "snug fit" coordination of the Na$^+$ ion in the SFs of eukaryotic Na$_V$ channels and the preservation of its solvation shell while permeating bacterial Na$_V$ channels [226].

To investigate whether a high degree of desolvation was a dominant factor in Ca$^{2+}$-selectivity in TRPV channels, I determined the number of oxygen atoms within a 3 Å
radius of the cations, representing their first solvation shell (Figure 3.18). In the bulk solution, both Ca\textsuperscript{2+} and Na\textsuperscript{+} cations showed their expected water co-ordination number of 7 and 5.6, respectively. As the cations entered the pore, I saw a small degree of partial dehydration of permeating cations at the SF (Figure 3.18). In particular, the carboxylate oxygen atoms of the acidic residue at the entrance of the SF co-ordinated an incoming cation, with these displacing up to $\sim$2 co-ordinated water molecules from the first solvation shell of the cation. I also observe some desolvation around, or below, binding site C. However, no major differences in the solvation shell of permeating Ca\textsuperscript{2+} or Na\textsuperscript{+} cations, or indeed between the Ca\textsuperscript{2+}-selective and non-selective TRPV channels were observed. The finding that the degree of desolvation does not differ substantially between Ca\textsuperscript{2+}-selective and non-selective channels indicates that their selectivity is not exclusively based on a mechanism of ion dehydration. The largest level of desolvation is seen at ion binding site A in the TRPV5 channel for Ca\textsuperscript{2+}. Since the overall desolvation penalty for Ca\textsuperscript{2+} is larger than for Na\textsuperscript{+} [227], one would expect this to lead to the preferential binding of Na\textsuperscript{+} at this site, whereas the opposite is observed (Figure 3.8), further arguing against dehydration as
3 CA$^{2+}$-SELECTIVE PERMEATION MECHANISM IN THE TRPV FAMILY OF ION CHANNELS

major selectivity mechanism in TRPV channels.

![Figure 3.18: Solvation state of permeating cations as they permeate through TRPV channels. The mean number of oxygen atoms of water molecules (blue), the number of oxygen atoms of protein residues (red), and total number of any oxygen atoms (black) within 3 Å of each permeating cation are plotted. The curves were smoothed using a Gaussian filter with a sigma value of 3.]

3.5 Discussion

Our simulations showed three main cation binding sites in the permeation pathway of TRPV channels, which I term binding sites A, B and C. Binding sites A and B are formed by the carboxylate oxygen atoms of the $\alpha$-position residue of the SF and by the carbonyl oxygen atoms of the $\gamma$- and $\delta$-position residues of the SF, respectively. Binding site C is located just above the hydrophobic lower gate of the pore, and is formed by the isoleucine residues of the lower gate and amide oxygen atoms of the neighbouring asparagine residues.

The identification of these cation binding sites in my MD simulations is in agreement with previously published TRPV structures and other MD simulations. In their crystal structure of the *Rattus norvegicus* TRPV6, Saotome *et al.* identified three regions of electron density within the channel pore which they interpreted as cation-binding sites [211]. It should be noted that there are some minor differences in the residues forming binding site C to the binding site reported here, likely due to the structure of TRPV6 of Saotome *et al.* being in the closed-state [211], rather than the open-state structure of
McGoldrick et al. used in the present study [186]. During the transition from a closed-state to the open-state, TRP channels undergo a rotation about the pore-forming S6 helix which changes the pore-facing residues [186].

Moreover, numerous structures of TRPV channels deposited within the PDB are resolved with cations bound at one of the cation binding sites identified in my MD simulations. These include structures from several orthologues of the TRPV channels simulated in this study, as well as of the TRPV1 and TRPV4 channels that were not simulated in this study. For example, structures of the closed-state TRPV1 channel of Rattus norvegicus [228] and the open-state TRPV4 channel of Homo sapiens in complex with 4α-PDD [229], both model cations bound at binding site B. This observation further validates the existence of the identified cation binding sites, as well as their conservation among TRPV channels and perhaps the wider TRP superfamily.

In addition to structural data, Sakipov et al. performed equilibrium MD simulations of Ca\(^{2+}\) movement in the closed-state structure of Rattus norvegicus TRPV6 and identified cation binding sites in the SF [212]. The binding sites from this work are also generally in agreement with my simulations results and the aforementioned structural data. However, Sakipov et al. reported that their simulation data identified two Ca\(^{2+}\) cations residing at binding site A, with adjacent chains each occupying an ion. In accordance with previous X-ray crystallographic data, two ions associated to binding site A are not seen in a major population of my simulation ensembles, which I attribute to the use of an optimised multi-site model for Ca\(^{2+}\) ions in the present study [150] and increased sampling for improved statistical analysis.

In mono-cationic solutions, I observed a high probability of at least two of the three binding sites (A, B and C) being simultaneously occupied with either Na\(^{+}\) or Ca\(^{2+}\), respectively, with mostly insignificant differences between the occupancy values for Na\(^{+}\) and Ca\(^{2+}\) at each individual binding site. However, the residence times were markedly reduced at all sites for Na\(^{+}\) ions, and overall a tendency towards weaker binding for either ion at binding site A at the extracellular SF entrance was observed.

In mixed, di-cationic solutions of Na\(^{+}\) and Ca\(^{2+}\), by contrast, Ca\(^{2+}\) ions strongly out-
competed Na⁺ ions for association at all pore binding sites, both in the Ca²⁺-selective and non-selective TRPV channels. Therefore, according to my data, all the binding sites display a much greater affinity for Ca²⁺. This gave rise to the question of how permeation efficiency for conducting Ca²⁺ ions is achieved in these channels and how this relates to their varying degrees of selectivity, since higher affinity binding is usually expected to lead to slower permeation rates.

By ensuring co-operativity between binding/unbinding events at multiple binding sites, permeation rates can be enhanced [171, 172]. I hypothesised that the level of co-operativity between sites A, B and C in the pore could underpin the difference between highly and less Ca²⁺-selective TRPV channels. I therefore developed a novel method to quantify co-operativity during ion permeation in pores with multiple ion binding sites based on mutual information and total correlation measures using the SSI approach [5]. I anticipate that this method will be similarly useful for the study of permeation mechanisms and the basis of selectivity in other channels. Our analysis showed that there is a substantial degree of co-operativity for Ca²⁺ permeation between binding sites B and C across all TRPV channels, whereas a clear distinction exists between the co-operativity between binding sites A and B in Ca²⁺-selective and non-selective TRPV channels. In the non-selective TRPV channels, binding site A is decoupled from binding site B. By contrast, binding sites A and B are even more strongly coupled than binding sites B and C in the case of the Ca²⁺-selective channels. I suggest that this marked difference in co-operativity mechanistically explains the different levels of Ca²⁺ selectivity in TRPV channels.
4 Monovalent-selective permeation mechanism in the TRPM family of ion channels

4.1 Associated publications

The work in this chapter has been published in the following publication:


4.1.1 Authors’ Contributions

CMI and UZ conceived the idea and designed the computational study, CMI conducted the simulations, CMI analysed the simulation data, NJT analysed the SSI data, UZ supervised the work, CMI and UZ wrote the manuscript with contributions from NJT, and all authors edited and reviewed the manuscript.

4.2 Introduction

The translocation of ions across cellular and organellar membranes via ion channels is essential to ensure cellular ionic homeostasis and provides a key pathway of intra- and intercellular communication. Ion channels catalyse the permeation of ions across the membrane up to an order of $10^6$ ions per second, while at the same time often displaying strict selectivity for particular ionic species [230]. The transient receptor potential (TRP) superfamily of ion channels comprises a large group of cation-selective channels that are implicated in a wide range of physiological processes [173, 231]. Due to their physiological importance, TRP channels are associated with a large number of pathological conditions [232], including in the aetiology of several rare, genetic conditions. Therefore, many members of the TRP channel superfamily constitute major pharmaceutical target proteins [92, 31].

63
Within the TRP channel superfamily, TRPM channels form the largest subfamily, consisting of eight members (TRPM1-8) [233, 234]. TRPM channels assemble as homotetramers, in which each subunit provides six transmembrane helices (S1-S6), a cytosolic N-terminus domain composed of four melastatin homology regions, and a cytosolic C-terminus coiled-coil domain [18, 29]. In keeping with most members of the TRP superfamily, TRPM channels are described as being cation non-selective, that is, they conduct cations but do not differentiate substantially between cationic species. However, in the TRPM subfamily, TRPM4 and TRPM5 are exceptions to this observation, since both channels are selective for monovalent cations and impermeable to divalent cations [39]. TRPM4 and TRPM5 are therefore the only members of the wider TRP superfamily to display selectivity for monovalent cations.

Although TRPM4 and TRPM5 are close homologues as they share a high degree of sequence homology and have similar biophysical characteristics, there are some variations in their activation mechanisms. For example, while both channels are activated by intracellular Ca\textsuperscript{2+} concentrations, TRPM5 is approximately 20-fold more sensitive to Ca\textsuperscript{2+} than TRPM4 [86]. TRPM5 signalling has been implicated in the taste of sweet, bitter, and umami tastes in type II taste bud cells [235, 236], and in the secretion of insulin by pancreatic \(\beta\)-cells [237, 238]. Consequently, TRPM5 is a potential drug target for a number of conditions, including metabolic conditions such as type II diabetes mellitus [99]. Several structures of the TRPM4 and TRPM5 channel have been published to date, however an open-state structure has only been solved for TRPM5 [82].

In the present work, I set out to characterise the cation permeation mechanism of the TRPM5 channel, focusing in particular on the basis for its monovalent cation selectivity. I conducted atomistic MD simulations and \textit{in silico} electrophysiology of the open-state structure of \textit{Danio rerio} TRPM5 [82], using Na\textsuperscript{+}, K\textsuperscript{+}, and Ca\textsuperscript{2+} ions as examples for monovalent and divalent biological ions. In total, I recorded 568 ion permeation events within 17.25 \(\mu\)s of aggregated time in \textit{in silico} electrophysiology simulations. The findings suggest a new mechanism of selectivity that is based on an extra binding site for monovalent cations in the channel’s cavity, rather than on differentiating ions in the se-
lectivity filter. Conduction of monovalent ions is therefore a synergistic process incorpo-
rating co-operativity between multiple binding sites, whereas the conduction of divalent
ions is inefficient in comparison.

Figure 4.1: Structure and membrane voltage of CompEL simulations of TRPM5. a) Snapshot of
the CompEL system showing the TRPM5 pore domain of Danio rerio used in this study inserted
into a double bilayer simulation system in an anti-parallel fashion so that both proteins experi-
ence identical voltage polarity. Cations within the aqueous compartments are shown as spheres
(orange: calcium; blue: sodium), highlighting the 9:1 ion concentration gradient between the
compartments. b) The CompEL charge differences applied across the aqueous compartments (∆q)
resulted in transmembrane voltages of ∼ -50 mV, -130 mV, -380 mV, and -610 mV in addition to
the concentration gradient. c) Average pore radius of TRPM5 along the pore axis from MD simu-
lations. The regions in grey shade represent the average positions of the major pore constrictions
in TRPM5, formed by Q906 (upper gate) and I966 (lower gate). The dashed line indicates the
radius of a completely dehydrated Ca\textsuperscript{2+} ion for comparison.

4.3 Methods

4.3.1 TRPM5 system construction

A truncated TRPM5 simulation system consisting of the membrane-domain of the chan-
nel was constructed by using residues 698-1020, including the N-acetyl-β-D-glucosamine
of the glycosylated N921 residue, of the Danio rerio TRPM5 structure [82]. The bound
Ca\textsuperscript{2+} cations occupying the Ca\textsubscript{TMD} binding site, which have been proposed to be im-
plicated in Ca\textsuperscript{2+}-dependent activation of TRPM5, was also modelled. The system was
built using the CHARMM-GUI server [179]. The charged N- and C-terminal residues
were neutralised by capping with acetylated (ACE) and N-methylamidated (CT3) groups, respectively. All missing non-terminal residues were modelled [180].

The structure was aligned in the membrane using the PPM server [187], and inserted into a POPC bilayer of 160 x 160 Å size using the CHARMM-GUI membrane builder [188, 189], and then solvated. Ions were added using GROMACS 2020.2 [190, 191] to neutralise any system charges and add ions to a concentration of either 150 mM NaCl, 150 mM KCl, 150 mM CaCl$_2$ (referred to as mono-cationic solutions), or a mixture of 75 mM NaCl and 75 mM CaCl$_2$ (referred to as di-cationic solutions). In the case of simulations containing Ca$^{2+}$, the standard CHARMM36m Ca$^{2+}$ ions were then replaced with the multi-site Ca$^{2+}$ of Zhang et al. [150].

4.4 Molecular dynamics simulations details

All simulations were performed using GROMACS 2020.2 [190, 191] or GROMACS 2022 [239], and the CHARMM36m force field for the proteins, lipids, and ions [127]. The TIP3P water model was used to model solvent molecules [192]. The system was minimised and equilibrated using the suggested equilibration inputs from CHARMM-GUI [193]. In brief, the system was equilibrated using the NPT ensemble for a total time of 1.85 ns with the force constraints on the system components being gradually released over six equilibration steps. The systems were then further equilibrated by performing a 15 ns simulation with no electric field applied. To prevent closing of the lower-gate of the pore, harmonic restraints were applied to maintain the distance between the α-carbon atoms of the lower gate I966 residue of each respective chain. The temperature was maintained at 310 K using the Nose-Hoover thermostat [159] and the pressure was maintained semi-isotropically at 1 bar using the Parrinello-Rahman barostat [160]. Periodic boundary conditions were used throughout the simulations. Long-range electrostatic interactions were modelled using the particle-mesh Ewald method [194] with a cut-off of 12 Å. The LINCS algorithm [117] was used to constrain bonds with hydrogen atoms. Hydrogen mass re-partitioning (HMR) of the system was used to allow the use of 4-fs time steps in simulations of NaCl solutions. The multi-site Ca$^{2+}$ model used for simulations
of CaCl$_2$ however is incompatible with a 4-fs time step, and therefore any simulations including Ca$^{2+}$ cations were performed with HMR but at a time step of 2-fs.

A summary of all simulations can be found in more detail in Tables 4.1 and 4.2.

Table 4.1: Summary of CompEL simulation details of the TRPM5 channel. All simulations were conducted in a di-cationic solution of 75 mM NaCl and 75 mM CaCl$_2$. In all simulations, the Ca$^{2+}$ cations occupying the Ca$_{TMD}$ were modelled, and remained bound for the duration of the simulations.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPM5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>7MBS (698-1020)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
<tr>
<td><strong>In silico electrophysiology methodology</strong></td>
<td>CompEL (anti-parallel with a 9:1 concentration gradient)</td>
</tr>
</tbody>
</table>
| **Ion** | 75 mM NaCl + 75mm CaCl$_2$
266 Na$^+$ (CHARMM36m)
274 Ca$^{2+}$ (Zhang et al.)
814 Cl$^-$ (CHARMM36m) |
| Independent simulations | 3 3 3 3 |
| Total simulation time (us) | 1.5 1.5 1.5 1.5 |
| Total aggregated simulation time (us) | 3 3 3 3 |
| **Ionic ratios between compartments (Extracellular : Intracellular)** | 239 : 27 Na$^+$
681 : 133 Cl$^-$
239 : 27 Na$^+$
680 : 134 Cl$^-$
239 : 27 Na$^+$
677 : 137 Cl$^-$
239 : 27 Na$^+$
673 : 141 Cl$^-$ |
| Estimated voltage (mV) | -50 -130 -380 -610 |
| **Permeation events** | 15 Na$^+$
0 Ca$^{2+}$
0 Cl$^-$ | 18 Na$^+$
0 Ca$^{2+}$
0 Cl$^-$ | 32 Na$^+$
19 Ca$^{2+}$
1 Cl$^-$ | 115 Na$^+$
168 Ca$^{2+}$
6 Cl$^-$ |
| Total number of permeation events | 15 18 52 289 |

CompEL simulations

The CompEL protocol [164, 165] of GROMACS was employed to create a transmembrane voltage and to drive ion permeation. I performed this using an anti-parallel double bilayer system, so that both channels experienced the same voltage polarity, with negative polarity in the intracellular region. Simulations were performed in a di-cationic solution of 75 mM NaCl and 75 mM CaCl$_2$ with a range of ionic imbalances (Δq), resulting in simulations with voltages of $\sim$ -50 mV, -130 mV, -380 mV, and -610 mV. To further drive cation permeation, a neutral ion concentration gradient of 9:1 between the extracellular and intracellular solutions was also created (Figure 4.1). All CompEL simulations were 500 ns long and repeated three times for each system, resulting in an aggregated simula-
Table 4.2: Summary of external applied field simulation details of the TRPM5 channel. All simulations were conducted in a mono-cationic solution of either 150 mM NaCl, 150 mM KCl, or 150 mM CaCl$_2$. In all simulations, the Ca$^{2+}$ cations occupying the Ca$_{TMD}$ were modelled, and remained bound for the duration of the simulations.

<table>
<thead>
<tr>
<th>Protein</th>
<th>TRPM5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>7MBS (698-1020)</td>
</tr>
<tr>
<td>Force field</td>
<td>CHARMM36m</td>
</tr>
<tr>
<td>Water</td>
<td>TIP3P</td>
</tr>
</tbody>
</table>

In silico electrophysiology methodology

<table>
<thead>
<tr>
<th>Ion</th>
<th>150 mM NaCl</th>
<th>150 mM KCl</th>
<th>150 mM CaCl$_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>267 Na$^+$ (CHARMM36m)</td>
<td>267 K$^+$ (CHARMM36m)</td>
<td>271 Ca$^{2+}$ (Zhang et al.)</td>
</tr>
<tr>
<td></td>
<td>275 Cl$^-$ (CHARMM36m)</td>
<td>275 Cl$^-$ (CHARMM36m)</td>
<td>542 Cl$^-$ (CHARMM36m)</td>
</tr>
</tbody>
</table>

| Independent simulations | 3 | 3 | 3 |
| Total simulation time (us) | 0.75 | 0.75 | 0.75 |
| Estimated voltage (mV) | -340 | -340 | -340 |
| Permeation events | 83 Na$^+$ | 34 K$^+$ | 54 Ca$^{2+}$ |
| Total number of permeation events | 83 | 34 | 54 |

External applied field simulations

In addition to CompEL simulations, simulations were also performed in a mono-cationic simulation of 150 mM NaCl, 150 mM KCl, 150 mM CaCl$_2$, using the method of Aksimentiev et al. to produce an external electric field [162]. These simulations had an $E_0$ of either -0.03, -0.0175, or -0.0044 V nm$^{-1}$, resulting in a transmembrane voltage of $\sim$340 mV, 200 mV, or 50 mV, respectively, with negative polarity in the intracellular region. All applied field simulations were 250 ns long and repeated three times for each system.
4.4.1 Simulation analysis

Analysis of MD trajectory data was performed using in-house written Python scripts, utilising GROMACS modules [190, 191], the SciPy library of tools [195, 196, 197, 198], and MDAnalysis [199, 200]. Analysis of the pore architecture was performed using CHAP [201]. All plots were generated in Python using Matplotlib [202] and Seaborn [203]. All MD inputs and analysis scripts used for this study are deposited in a public GitHub repository, available at: https://github.com/cmives/Na_selectivity_mechanism_of_TRPM_channels.

Calculating conductance and selectivity from in silico electrophysiology experiments

The conductance of the channels ($C_{ion}$) was calculated according to Equation 4.1, where $N_p$ is the number of permeation events, $Q_{ion}$ is the charge of the permeating ion in Coulomb, $t_{traj}$ is the length of the trajectory, and $V_{tm}$ is the transmembrane voltage. The mean conductance and standard error were calculated from overlapping 50 ns windows of the trajectory.

$$C_{ion} = \frac{N_p \times Q_{ion}}{t_{traj} \times V_{tm}}$$  (4.1)

The selectivity ($P_{Na}/P_{Ca}$) from the di-cationic CompEL simulations was calculated as the ratio between the total sum of Na$^{2+}$ permeation events and the total sum of Ca$^{2+}$ permeation events across all simulations.

Identification of cation binding sites from MD simulations of TRPV channels

Cation binding sites were identified by plotting timeseries of each permeating ion with respect to their position along the pore axis. To further validate these positions, a 3D density mesh was generated for cations within 10 Å of the protein. This analysis was performed on a trajectory of concatenated three-fold replicated 500 ns simulations from mono-cationic simulations with a voltage of $\sim$ -50 mV produced by the CompEL method.

Characterising permeation co-operativity through mutual information using SSI from PENSA

To characterise the level of co-operativity in the knock-on permeation mechanisms of
the TRPM5 channel, PENSA was used to calculate the state-specific information (SSI) shared between discrete state transitions in the occupancy distributions of both of the binding sites [5, 240]. The methodology used is the same as that described in detail in Chapter 3.3.

In brief, a timeseries distribution with a timestep of 20 ps for each binding site was obtained, whereby for each frame, if an ion occupied the binding site then this ion’s atom ID number was recorded, whereas if the binding site was unoccupied, an ID of -1 was recorded. This was then quantified by mutual-information based SSI whether ion transitions at one site were coupled to ion transitions at the adjacent site. To account for statistical noise that can arise from even independent distributions [208, 209], a statistical threshold was calculated which was then subtracted from the measured values to resolve the excess mutual information, or excess SSI (exSSI).

4.5 Results

4.5.1 Cation conductance of the TRPM5 channel in di-cationic solutions

*In silico* simulations of *Danio rerio* TRPM5 [82] embedded in a POPC lipid bilayer were performed, with a di-cationic ionic solution of 135 mM NaCl and 135 mM CaCl$_2$ in the central dense aqueous compartment, and 15 mM NaCl and 15 mM CaCl$_2$ in the outer diluted aqueous compartments (Figure 4.1). An anti-parallel CompEL setup was used to produce a bio-mimetic transmembrane voltage of $\sim$ -50 mV, as well as higher voltages of -130 mV, -380 mV and -610 mV to increase the number of permeation events and improve the statistics of the analyses (Figure 4.1). In addition to the membrane voltages, the 9:1 ion concentration gradient between the middle and the outside bulk compartment also drove ion permeation.

Our simulations under ion and voltage gradients showed a continuous flow of permeating ions, resulting in a total of 374 permeation events across all investigated simulation conditions. Despite the ion gradient providing an additional driving force for permeation alongside the voltage, the calculated conductances from the *in silico* electrophysiology simulations, in a range between 7 and 38 pS (Table 4.3), are generally in good agreement
4 MONOVALENT-SELECTIVE PERMEATION MECHANISM IN THE TRPM FAMILY OF ION CHANNELS

with the published conductance values of 23–25 pS from in vitro electrophysiology in NaCl based solutions [241, 242].

4.5.2 Low-voltage simulations in di-cationic solutions show the exclusive permeation of Na\(^+\) through TRPM5

At the lowest voltages of \(~\sim\) -50 mV and -130 mV, complete Na\(^+\)-selectivity in mixed solutions was observed, with no recorded Ca\(^{2+}\) permeation during an accumulated simulation time of 1.5 \(\mu\)s, while in the same time span, 15 Na\(^+\) ions traverse the TRPM5 pore, in accordance with its general conductance level (Table 4.3).

Analysis of the pore architecture of TRPM5 showed no major conformational changes throughout the course of the simulations. The TRPM5 pore possesses two main constrictions: an upper constriction formed by the sidechains of Q906 of the three-residue selectivity filter (SF), and a lower constriction formed by the sidechains of I966 of the lower gate (Figure 4.1). A minor constriction can also be observed \(~\sim 13\) \(\AA\) above the SF, in the extracellular pore vestibule (EPV) (Figure 4.2). This constriction is formed by the loop between the pore helix (PH) and the S6 helix.

In simulations at -50 mV and -130 mV, Na\(^+\) cations first enter the EPV region of the TRPM5 pore, where they show a broad association. Permeating Na\(^+\) cations then traverse the SF rapidly, and enter the pore cavity. They spend a substantial amount of time occupying the cavity before passing through the lower gate and exiting the pore at the intracellular face.

As opposed to monovalent Na\(^+\), Ca\(^{2+}\) ions did not readily enter the inner pore of TRPM5 during the course of the simulations. Ca\(^{2+}\) cations chiefly occupied the EPV region at the extracellular entrance (see Figure 4.2). 3D density maps of Na\(^+\) and Ca\(^{2+}\) ions further confirmed this observation (Figure 4.3). The maps show substantial Ca\(^{2+}\) density in the EPV, particularly near the acidic residues on the loop between the PH and the S6 helix, namely: E910, E911, D919, D920, D925, and E928. It was also observed that Ca\(^{2+}\) ions occasionally migrate from the EPV toward the pore, however they are blocked from entering the cavity at the SF, particularly at the constriction formed by
4.5.3 The ion selectivity of TRPM5 is strongly voltage-dependent

As the membrane voltage is increased, the Na$^+$-selectivity ($P_{Na}/P_{Ca}$) in the simulations is diminished (Table 4.3). At a voltage of both $\sim -50$ mV and $\sim -130$ mV, complete Na$^+$-selectivity was observed, with no Ca$^{2+}$ permeation events in any of the simulations. At a voltage of $\sim -380$ mV, the in silico electrophysiology simulations continue to display slightly Na$^+$-selective permeation; however, when the voltage is further increased to $\sim -610$ mV, Na$^+$-selectivity is lost. Furthermore, higher voltage simulations also yielded a small number of Cl$^-$ permeation events, with anions permeating through to the extracellular solution.

Our findings indicate a strongly voltage-dependent energy surface for permeation.
4 MONOVALENT-SELECTIVE PERMEATION MECHANISM IN THE TRPM FAMILY OF ION CHANNELS

Figure 4.3: 3D density map of Ca\textsuperscript{2+} cations around the EPV and SF of TRPM5. The density of Ca\textsuperscript{2+} ions was calculated from concatenated trajectories of TRPM5 in a di-cationic solution under a transmembrane voltage of \( \sim -50 \) mV generated by the CompEL method. A major density maxima is seen within the EPV, where Ca\textsuperscript{2+} binds and associates. Occasionally, a Ca\textsuperscript{2+} cation approaches the SF, however is not able to traverse past through the SF at bio-mimetic voltages. The sidechains of Q906 of the SF and I966 of the lower gate are shown as sticks for clarity (grey).

across the TRPM5 pore, suggesting relatively weak cation binding sites within the pore domain. These observations are reminiscent of the effects seen for high voltage simulations of the related TRPV channel subfamily, in which Ca\textsuperscript{2+}-selectivity for TRPV5 and TRPV6 was reduced at higher-voltage regimes.

In order to explore the ion permeation dynamics in TRPM5 and their underlying energetic properties further, I aimed to enhance the sampling of both Na\textsuperscript{+} and Ca\textsuperscript{2+} permeation events, while at the same time maintain a Na\textsuperscript{+}-selective voltage range. Thus a voltage of \( \sim -340 \) mV was selected as a compromise for studying permeation in mono-cationic solutions to ensure a sufficient number of traversals of both Ca\textsuperscript{2+} and Na\textsuperscript{+} while remaining within the monovalent-selective regime.
Table 4.3: Calculated conductances and selectivities from CompEL simulations of ion permeation in the TRPM5 channel. Mean inward conductances and standard error of the mean (SEM) were calculated from overlapping 50 ns windows from three-fold replicated 500 ns simulations of an anti-parallel double bilayer system. Mean selectivity ratios of Na\(^+\) and Ca\(^{2+}\) permeation events and SEM were calculated from three-fold replicated 500 ns simulations. The raw number of permeation events for each cation is displayed in brackets below the respective conductance value.

<table>
<thead>
<tr>
<th>Conc. gradient</th>
<th>Voltage (mV)</th>
<th>Conductance (pS)</th>
<th>(P_{Na}/P_{Ca})</th>
</tr>
</thead>
<tbody>
<tr>
<td>135 mM : 15 mM</td>
<td>50</td>
<td>16 ± 3.1 (15)</td>
<td>16 ± 3.1</td>
</tr>
<tr>
<td></td>
<td>130</td>
<td>7 ± 1.3 (18)</td>
<td>7 ± 1.3</td>
</tr>
<tr>
<td></td>
<td>380</td>
<td>4 ± 0.5 (32)</td>
<td>9 ± 1.0</td>
</tr>
<tr>
<td></td>
<td>610</td>
<td>10 ± 0.9 (115)</td>
<td>38 ± 3.1</td>
</tr>
</tbody>
</table>

4.5.4 Mechanistic insights into ion permeation in TRPM5 from mono-cationic solutions

I next conducted in silico electrophysiology simulations using an EAEF of ∼ -340 mV across the membrane to investigate the permeation mechanism of Na\(^+\) and Ca\(^{2+}\) ions in mono-cationic solutions, as well as K\(^+\) ions, with sufficient sampling efficiency (Table 4.4). As shown in Figure 4.4, there is a clear difference between the behaviour of monovalent cations in the channel, such as Na\(^+\) and K\(^+\) ions, and the divalent Ca\(^{2+}\) ions.

Whereas Na\(^+\) and K\(^+\) ions occupy the central cavity of the channel for most of the simulated time, any permeating Ca\(^{2+}\) ions traverse the inner cavity rapidly, not showing any apparent immobilisation within the cavity. Despite occupying the cavity for extended periods of time, Na\(^+\) and K\(^+\) ions do not seem to bind to a particular binding site or residue within the cavity, but explore nearly the entire cavity volume before they permeate to the intracellular side.

Looking at the density of ions along the pore axis, and using the negative logarithmic density as an estimate for the underlying free energy profile at the non-equilibrium permeation conditions under membrane voltage, it can be seen that the cavity region forms a shallow, wide energy minimum for the monovalent cations, whereas Ca\(^{2+}\) ions experience
Figure 4.4: Exemplar permeation traces of the $z$-coordinate of permeating cations over time. The permeation traces of Na$^+$ (blue, top), K$^+$ (purple, middle), and Ca$^{2+}$ (orange, bottom) are plotted from simulations performed in a mono-cationic solution with an EAEF. The shaded grey regions represent the average position of the pore constrictions formed by Q906 (upper) and I966 (lower). Simulation data is from simulations performed in a mono-cationic solution with an EAEF producing a voltage of $\sim$ -340 mV. Please note, only cations that fully permeate through the pore within the 250 ns simulation are shown in the plot.

a small energy barrier in the same region (Figure 4.5). In contrast, both monovalent and divalent ions show binding to a shallow binding site, corresponding to the EPV. Both ion types experience a slight energy barrier to permeation near the intracellular channel exit.

Furthermore, additional simulations using an EAEF of differing magnitudes were con-
Table 4.4: Calculated conductances from applied field simulations of ion permeation in the TRPM5 channel. Mean inward conductances and standard error of the mean (SEM) were calculated from overlapping 50 ns windows from three-fold replicated 500 ns simulations of a single bilayer system. The raw number of permeation events for each cation is displayed in brackets below the respective conductance value.

<table>
<thead>
<tr>
<th>Voltage (mV)</th>
<th>Ion solution</th>
<th>Conductance (pS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-50</td>
<td>150 mM NaCl</td>
<td>19 ± 6.7 (4)</td>
</tr>
<tr>
<td></td>
<td>150 mM CaCl₂</td>
<td>0 ± 0.0 (0)</td>
</tr>
<tr>
<td>-200</td>
<td>150 mM NaCl</td>
<td>17 ± 3.1 (15)</td>
</tr>
<tr>
<td></td>
<td>150 mM CaCl₂</td>
<td>9 ± 3.8 (4)</td>
</tr>
<tr>
<td>-340</td>
<td>150 mM NaCl</td>
<td>52 ± 5.6 (83)</td>
</tr>
<tr>
<td></td>
<td>150 mM KCl</td>
<td>21 ± 4.2 (34)</td>
</tr>
<tr>
<td></td>
<td>150 mM CaCl₂</td>
<td>85 ± 5.0 (54)</td>
</tr>
</tbody>
</table>

ducted. Notably, whereas the main features of the ion density and free energy estimates occur across all tested voltages, Ca²⁺ is increasingly excluded from the cavity, and is no longer able to enter into the cavity at the lowest voltage of -50 mV during the time span of the simulations (Figure 4.6). This shows again that the ion selectivity of TRPM5 is voltage-dependent. I next aimed to elucidate the molecular foundations of this behaviour and selectivity in general.

### 4.5.5 Solvation profiles of cations during channel permeation

To probe if cation desolvation played a part in yielding selective permeation, especially at the narrowest constriction near Q906 of the SF, the number of water oxygen atoms within a 3 Å radius of cations was calculated, representing their first solvation shell (Figure 4.5). The desolvation of permeating ions has previously been reported as one important mechanism to ensure ion selectivity in some channels. Differences in the desolvation energies of permeating ions provide a thermodynamic penalty which can underpin more favourable
Figure 4.5: Solvation and PMF profiles of Na$^+$ (blue), K$^+$ (purple) and Ca$^{2+}$ (orange) cations through the TRPM5 pore. These simulations were performed in a mono-cationic solution, with an external applied field used to produce a transmembrane voltage of $\sim$ -340 mV. 

a) The mean number of oxygen atoms of water molecules (solid line) and of any oxygen atoms of any molecule (dashed line) within 3 Å of each permeating cation is plotted. 

b) Negative logarithmic density profiles of permeating cations along the pore of the TRPM5. The logarithmic ion densities represent quasi-free energies (with a nominal unit of kT). Density minima reflect stably bound ions (i.e. binding sites), while maxima indicate barriers between the binding sites. The location of the pore constrictions formed by Q906 (upper) and I966 (lower) are represented as grey regions. Both plots have been smoothed using a Gaussian filter with a sigma value of 2.

permeation of an ionic species over another. This mechanism of ion selectivity by desolvation is suggested to be particularly pertinent in K$^+$ channels [105, 106].

In the bulk solution of the simulated systems, Na$^+$, K$^+$, and Ca$^{2+}$ ions show the expected water coordination number of their solvation shells. As both Na$^+$ and K$^+$ ions enter the pore of TRPM5, they are partially desolvated by Q906, the side chain displacing 1–2
Figure 4.6: Negative logarithmic density profiles of permeating cations along the pore of the TRPM5 at different voltages. These simulations were performed in a mono-cationic solution, with an EAEF used to produce transmembrane voltages of $\sim -50$ mV (top), $\sim -200$ mV (centre), and $\sim -340$ mV (bottom). The logarithmic ion densities represent quasi-free energies (with a nominal unit of kT). The location of the pore constrictions formed by Q906 (upper) and 1966 (lower) are represented as grey regions. Both plots have been smoothed using a Gaussian filter with a sigma value of 2.
water molecules from the first solvation shell of the ion. After traversing the constriction at Q906, these monovalent ions are then re-solvated in the pore cavity, before again being partially de-solvated at the hydrophobic lower gate formed by I966. By contrast, the rapidly permeating Ca$^{2+}$ ions do not show any significant desolvation when they cross the SF or hydrophobic lower gate of TRPM5.

Furthermore, the solvation profiles of permeating cations show an additional region of differing desolvation within the EPV highlighted previously (Figure 4.2). In this region, Ca$^{2+}$ and K$^{+}$ ions are partially desolvated, indicating closer interactions with the acidic residues in the EPV region. By contrast, Na$^{+}$ cations do not show any desolvation there.

Similar solvation profiles were observed for permeating cations in both the EAEF simulations in a mono-cationic solution (Figure 4.7), and the CompEL simulations in a di-cationic solution (Figure 4.8), at a range of voltage magnitudes.

Summarising, these data suggest that ion desolvation in the SF is not a major factor in achieving selectivity for monovalent cations. Since both monovalent and divalent cations occupy the EPV, filtering for monovalent ions must occur later in the permeation pathway. However, Ca$^{2+}$ ions are not desolvated when they reach the inner cavity.

As the energetic penalty for desolvating Ca$^{2+}$ is far larger than for Na$^{+}$ and K$^{+}$ [227], the observed desolvation profile cannot explain the deselection of Ca$^{2+}$ ions for further traversal of the cavity and pore.

### 4.5.6 Selectivity for monovalent cations is based on permeation co-operativity between two binding sites

I hypothesised that the permeation mechanism of monovalent ions, such as Na$^{+}$, may be more efficient than that of Ca$^{2+}$ ions due to the presence of an additional binding site for monovalent cations in the internal channel cavity. In previous work, a mutual-information based assessment of the permeation co-operativity across multiple ion binding sites in channels was developed, termed state-specific information (SSI; [5, 240]). In brief, SSI quantifies the probability that a state change in one binding site, *i.e.* a change from binding an ion to being vacant upon ion permeation, is correlated to a similar state change in a
Figure 4.7: Solvation profiles of Na\(^+\) (blue) and Ca\(^{2+}\) (orange) cations through the TRPM5 pore. These simulations were performed in a mono-cationic solution, with an EAEF used to produce transmembrane voltages of \(~-50\) mV (top), \(~-200\) mV (centre), and \(~-340\) mV (bottom). The mean number of oxygen atoms of water molecules (solid line) and of any oxygen atoms of any molecule (dashed line) within 3 Å of each permeating cation is plotted. The location of the pore constrictions formed by Q906 (upper) and I966 (lower) are represented as grey regions. All plots have been smoothed using a Gaussian filter with a sigma value of 2.
Figure 4.8: Solvation profiles of Na\(^+\) (blue) and Ca\(^{2+}\) (orange) cations through the TRPM5 pore. These simulations were performed in a di-cationic solution, with the CompEL methodology used to produce transmembrane voltages of \(~ -50\) mV (top), \(~ -130\) mV (second from top), \(~ -380\) mV (second from bottom), and \(~ -610\) mV (bottom). The mean number of oxygen atoms of water molecules (solid line) and of any oxygen atoms of any molecule (dashed line) within 3 Å of each permeating cation is plotted. The location of the pore constrictions formed by Q906 (upper) and I966 (lower) are represented as grey regions. All plots have been smoothed using a Gaussian filter with a sigma value of 2.
second binding site, in which case the unbinding events are coupled to one another.

This SSI approach was applied to ion conduction in TRPM5, as done in Chapter 3 to study TRPV channels, focusing on the pair of binding sites at the EPV and the channel cavity (see Figure 4.2). These two binding sites are shallow but locate directly to the main pore axis; in addition, they show moderate to high occupancy with monovalent cations, respectively (Figure 4.9). By using SSI, it was found that both Na$^+$ and K$^+$ ions display a high level of correlation between binding and unbinding at the two successive sites, whereas the permeation of Ca$^{2+}$ ions shows only a low degree of correlation (Figure 4.9). This means that a distant knock-on mechanism is in operation between incoming monovalent ions, which bind transiently at the SF as well as over substantial time spans within the cavity. In other words, the cavity serves as a reservoir, more likely releasing a Na$^+$ or K$^+$ ion to the cytoplasm when a further monovalent cation approaches and binds to the SF of TRPM5. By contrast, Ca$^{2+}$ ions permeate on their own, driven solely by the transmembrane electric field. These results suggest that this mechanistic difference is what ultimately underlies the monovalent-ion selectivity in TRPM5, since the efficiency of monovalent conduction is optimised by co-operative permeation across two channel binding sites. However, this monovalent selectivity is abolished when the driving force for Ca$^{2+}$ permeation (i.e. the transmembrane electric field) exceed a certain threshold (i.e. at high voltages).

4.5.7 Why does the inner cavity form an attractive site for monovalent cations but a repulsive site for divalent cations?

The presence of a water-filled internal cavity is a conserved feature among most cation channels, which serves to maintain a high degree of ion hydration despite its location in the centre of the lipid bilayer. As in other channels, the major permeating species, Na$^+$ and K$^+$ ions, are re-hydrated and transiently captured in the cavity following their permeation through the SF. At higher voltages, Ca$^{2+}$ ions can enter and traverse the cavity but do not show any deviation from their bulk hydration number during this process. At lower voltages, by contrast, Ca$^{2+}$ ions are excluded from entering the cavity.
Figure 4.9: State-specific information (SSI) of cation transitions between binding sites and the average number of each cation within the pore of TRPM5. **a)** Excess state-specific information \((\text{exSSI})\) between the EPV and pore cavity ion binding sites quantifies the degree of co-operativity in the permeation mechanisms of \(\text{Na}^+\) (blue), \(\text{K}^+\) (purple), and \(\text{Ca}^{2+}\) (orange) ions. The mean exSSI and SEM between transitions from the two binding sites were calculated from simulations performed in mono-cationic simulations with an EAEF producing a voltage of \(\sim \sim 340\) mV. **b)** The mean probability for the number of each cationic species within the pore was calculated from non-overlapping 50 ns windows from three-fold replicated 250 ns simulations, with the pore defined as the region between the constrictions formed by Q906 and I966.

Therefore, the difference between the pore and cavity properties of the highly \(\text{Ca}^{2+}\)-selective TRP channel, TRPV5, and the monovalent-selective TRPM5 was investigated. As shown in Figure 4.10, the general features of the pore are preserved with a constriction at the extracellular SF, a wider internal cavity region, and a second constriction at the intracellular gate. TRPM5 has a markedly shorter SF, while its cavity is wider than that of TRPV5. However, there is a substantial difference in the pore lining of the two TRP channels. Whereas the TRPV5 SF constitutes a strongly hydrophilic region, TRPM5 does not display increased hydrophilicity within its SF. The transition from the SF to the cavity is slightly hydrophobic in TRPM5, while it is a hydrophilic region in TRPV5. There are no differences between the hydrophobicity of the two channels at the intracellular gates.
Therefore, the monovalent-selectivity of TRPM5 was attributed to the raised hydrophobicity of its SF and upper portion of its inner cavity regions. In particular, the transition between the SF and the cavity in TRPM5 is lined by large hydrophobic Phe residues at the bottom of the SF (F904). The energetic penalty to place monovalent cations within a hydrophobic region is much smaller than for divalent cations. As observed in simulations under increased voltage, the protein matrix does not form favourable interactions with Ca$^{2+}$ ions in this area, which could serve to replace water molecules in their hydration shell. Accordingly, this region is identical to the position along the pore axis at which Ca$^{2+}$ ions are repelled back into the extracellular space at lower voltages. At supra-physiological voltages, however, the increased force acting on the divalent Ca$^{2+}$ ions is sufficient to drag some of the ions across the SF and inner cavity along with their complete hydration shell. The permeating Ca$^{2+}$ ions do not display any noticeable interaction with protein residues during this process, which indicates that this effect is mainly driven by the strong electric field across the membrane under these conditions.

4.6 Discussion

The calculated conductances and ion selectivities observed in the \textit{in silico} electrophysiology simulations of TRPM5 are in good agreement with published \textit{in vitro} values. In particular, in simulations of an equimolar mixture of NaCl and CaCl$_2$ under bio-mimetic voltages, the simulations model the impermeability of TRPM5 to divalent cations. This selectivity is even more remarkable when one considers that this selectivity was captured with a 1:1 ratio of Na$^+$.Ca$^{2+}$ ions in these simulations, compared to the \textit{in vivo} ratio of $\sim$112:1 [243]. To my best knowledge, the simulations reported herein represent the first \textit{in silico} study of the ion permeation mechanism in any TRPM channel.

However, when the voltage of these simulations is increased to supra-physiological voltages there is a distinct loss in the selectivity of TRPM5. A similar effect was previously described in simulations of TRPV channels in Chapter 3. These findings show how the increased forces acting on divalent cations at these increased voltages are sufficient to enable them to traverse energetically unfavourable regions of the pore. As a result,
Figure 4.10: Pore architecture of the monovalent-selective TRPM5 channel (green) and the Ca^{2+}-selective TRPV5 channel (cyan) from MD simulations. The average pore radius (a) and hydrophobic profile (b) for each channel was calculated using CHAP [201]. The standard deviation is shown as shaded regions. The profile of the TRPV5 was generated from simulation data reported in Chapter 3.

these findings pose a crucial question for deciding on a voltage for *in silico* electrophysiology experiments, revealing subtleties in the trade-off between the improved sampling of increased conductances and the correct recovery of ionic selectivity.

Our simulations identify two shallow, broad ion binding sites for monovalent cations in the TRPM5 channel; one within the EPV above the SF, and a second within the pore cavity, whereas divalent cations do not interact favourably within the cavity. Application of this methodology identified that a knock-on mechanism is occurring between monovalent ions within this region. This co-operative mechanism between ion binding and unbinding events at adjacent binding sites facilitates enhanced permeation rates [172].

On the other hand, divalent cations permeate on their own, driven solely by the transmembrane voltage. At physiological voltages, this is not sufficient to permit permeation. I attribute the energetic unfavourability of Ca^{2+} ions within the TRPM5 pore due to in-
creased hydrophobicity of the SF and upper portion of the pore cavity.

Moreover, numerous closed-state structures of TRPM4, a close homologue of TRPM5, have been published within the PDB. Several of these structures include Na\(^+\) cations that have been modelled within the pore cavity [87, 244]. Consequently, these structures (PDB IDs: 6BCJ, 6BCL, and 6BWI) suggest that the presence of a broad monovalent cation binding site is a conserved feature among TRPM4 and TRPM5 channels.

As additional open-state structures of other TRPM channels and different orthologues are solved and published, it will be possible to further test this permeation mechanism. Additional *in silico* electrophysiological experiments of TRPM4, the other monovalent-selective TRPM channel, and of non-selective TRPM channels will facilitate a better understanding of cation permeation across the TRPM family.

In conclusion, I report that monovalent cations permeate through the pore of TRPM5 via a co-operative, distant knock-on mechanism between a binding site in the EPV and a binding site in the pore cavity. The channel, however, is impermeable to divalent cations at physiological voltages, as they experience an energetic barrier within the pore cavity, thus disrupting the co-operative permeation mechanism. I suggest that this difference in co-operativity mechanistically explains the monovalent cation-selectivity of the TRPM5 channel.
5 Comparison of the CompEL and EAEF techniques for \textit{in silico} electrophysiology

5.1 Introduction

As demonstrated in Chapters 1 and 2, MD simulations are a powerful tool for investigating the mechanisms of ion permeation through an ion channel. The ability of this approach to provide atomistic insights into these processes makes them a very attractive computational methodology. Continuous ion permeation can be achieved in these simulations by producing a transmembrane voltage, mimetic to the membrane potential present in biological cells. To date, there are two main methodologies used to produce these voltages, namely the EAEF and CompEL techniques, as discussed in Chapter 2.4.2. The use of the EAEF methodology was used to study the permeation mechanisms of TRPV channels and TRPM channels in Chapters 3 and 4, respectively, and the CompEL methodology was also used to model the monovalent-selectivity of TRPM5 at physiological transmembrane voltages in Chapter 4.

However, with both methodologies demonstrating the great potential of \textit{in silico} electrophysiology, it poses the question of which methodology is preferable when designing experiments to elucidate ion permeation mechanisms using MD simulations. A comparison between the two methods was previously conducted by Melcr \textit{et al}., finding no difference between the structural parameters of model membranes [245]. However, this study was carried out on impermeable membranes, and did not include any ion channels within the system. Therefore, it remains unclear if these two methods are equivalent when being used to study ion permeation through ion channels. This question is particularly interesting for ion channels possessing large soluble domains. As the EAEF technique creates additional forces across the entire system, it was a point of interest whether this would affect the dynamics of these domains more than with the CompEL technique.

In this chapter, a systematic comparison of the CompEL and EAEF techniques for modelling ion permeation through an ion channel was conducted, using a number of different simulation systems as case studies. MD simulations of the wild-type structure
of TRPV3 of *Mus musculus* [184] were conducted at a range of voltages using both CompEL and EAEF. Additionally, simulations were conducted of TRPV3 both with and without harmonic restraints to keep the lower-gate of the pore open. Furthermore, simulations were performed using both the CHARMM36m [127] and AMBER19SB [246] force fields. Therefore, this comparison represents a broad-scoped comparison of *in silico* electrophysiology of TRPV3 using the CompEL and EAEF techniques at differing (i) voltage amplitudes, (ii) voltage polarities, (iii) protein restraints, and (iv) force fields. In total, these *in silico* electrophysiology experiments produced 4,628 permeation events from an aggregated simulation time of 23.76 µs.

This work, to my best knowledge, represents the first effort to identify any differences in the ion permeation and channel structure in MD simulations using these two methodologies. The ultimate goal of this work is to help provide guides for the biophysical community when conducting *in silico* electrophysiology, and to ascertain if there are circumstances where one methodology is more appropriate than the other.

Figure 5.1: Structure of the TRPV3 ion channel used to study *in silico* electrophysiology methodologies. The open-state structure of the cation non-selective TRPV3 of *Mus musculus* (PDB: 6PVP) is shown from the extracellular side (a) and from the plane of the membrane (b). The pore-domain (residues 433-707, *yellow*) and ankyrin-repeat domain (ARD, residues 118-363, *cyan*) are indicated by colouring.
5 COMPARISON OF THE COMPEL AND EAEF TECHNIQUES FOR IN SILICO ELECTROPHYSIOLOGY

5.2 Methods

5.2.1 TRPV3 system construction

A variety of TRPV3 systems were created using the full-length, open-state of *Mus musculus* TRPV3 of Singh et al. [184]. The system was built using the CHARMM-GUI server [179]. The charged N- and C-terminal residues were neutralised by capping with acetylated (ACE) and N-methylamidated (CT3) groups, respectively. All missing non-terminal residues were modelled [180]. The structures were aligned in the membrane using the PPM server [187], and inserted into a POPC bilayer of 180 x 180 Å size using the CHARMM-GUI membrane builder [188, 189], and then solvated. Ions were added to neutralise the system charges and bring the bulk solution to a concentration of 150 mM NaCl. HMR was used in the system to allow the use of a 4-fs time step.

5.2.2 Molecular dynamics simulations details

All simulations were carried out using GROMACS 2022 [239]. Simulations of TRPV3 were performed using the CHARMM36m force field for the proteins, lipids, and ions [127], and the TIP3P water model was used to model solvent molecules [192]. Additional simulations of TRPV3 were performed using the AMBER19SB force field for proteins and ions [246], the SLipids force field for lipids [247]; in these simulations, the OPC water model was used to model solvent molecules [248]. For simulations using AMBER19SB, the OPC water model was chosen as it is the recommended water model to be used with this force field [246].

The system was minimised and equilibrated using the suggested equilibration inputs from CHARMM-GUI [193]. In brief, the system was equilibrated in the NPT ensemble for a total time of 1.85 ns with restraints on the system components being gradually released over six equilibration steps. The systems were then further equilibrated by performing a 15 ns simulation with no electric field applied. The temperature was maintained at 310 K using the Nose-Hoover thermostat [159] and the pressure was maintained semi-isotropically at 1 bar using the Parrinello-Rahman barostat [160]. Periodic boundary conditions were used throughout the simulations. Long-range electrostatic interactions were
Table 5.1: Summary of simulations performed in the comparison of the CompEL and EAEF techniques.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Force field</th>
<th>Construct</th>
<th>Ion solution</th>
<th>Transmembrane voltage methodology</th>
<th>Voltage</th>
<th>Simulation duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV3</td>
<td>CHARMM36m</td>
<td>6PVP (+ harmonic restraints)</td>
<td>150 mM NaCl and TIP3P water</td>
<td>CompEL</td>
<td>High</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Medium</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Low</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EAEF</td>
<td>High</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Medium</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Low</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td>AMBER19SB</td>
<td>6PVP (- harmonic restraints)</td>
<td>150 NaCl and TIP3P water</td>
<td>CompEL</td>
<td>High</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Low</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>EAEF</td>
<td>High</td>
<td>3 x 330 ns</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Low</td>
<td>3 x 330 ns</td>
</tr>
</tbody>
</table>

modelled using the particle-mesh Ewald method [194] with a cut-off of 12 Å in simulations using CHARMM36m and a cut-off of 9 Å in simulations using AMBER19SB. The LINCS algorithm [117] was used to constrain bonds with hydrogen atoms in simulations using CHARMM36m and to constrain all bonds in simulations using AMBER19SB. A summary of all simulations is shown in Table 5.1.

Simulations of these systems were performed using either the CompEL methodology [164, 165] or an external electric field [162]. CompEL simulations were conducted using a parallel double bilayer system, so that both channels experienced opposite voltage polarities. CompEL simulations were performed at several ionic imbalances (Δq), and EAEF simulations at several E₀ values, to produce a range of voltages (Table 5.2). This resulted in a number of simulations with approximately equal transmembrane voltages, generated by the two different in silico electrophysiology techniques.

5.2.3 Simulation analysis

Analysis of MD trajectory data was performed using in-house written Python scripts, utilising GROMACS modules [190, 191], the SciPy library of tools [195, 196, 197, 198], and MDAnalysis [199, 200]. All plots were generated in Python using Matplotlib [202]
Table 5.2: Summary of calculated voltages from simulations performed in this study utilising CompEL or EAEFs.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Voltage</th>
<th>Transmembrane voltage methodology</th>
<th>Voltage (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRPV3</td>
<td>High</td>
<td>CompEL (Δq = 14)</td>
<td>± 860</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EAEF (E₀ = ± 0.0482 V nm⁻¹)</td>
<td>± 760</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>CompEL (Δq = 7)</td>
<td>± 440</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EAEF (E₀ = ± 0.0322 V nm⁻¹)</td>
<td>± 500</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>CompEL (Δq = 3)</td>
<td>± 180</td>
</tr>
<tr>
<td></td>
<td></td>
<td>EAEF (E₀ = ± 0.0161 V nm⁻¹)</td>
<td>± 250</td>
</tr>
</tbody>
</table>

and Seaborn [203].

5.3 Results

5.3.1 Case study: TRPV3 using the CHARMM36m force field

Continuous cation permeation was observed in all simulations conducted of full-length TRPV3 with harmonic restraints acting on the α-carbon atoms of I674 of the lower gate. These were applied to keep the protein in an open-state. In total, these simulations across a range of different voltage magnitudes and polarities resulted in 3,990 permeation events. From these permeation events, the relationship between the current-voltage characteristic was evaluated by plotting I-V curves, as seen in Figure 5.2. From this plot it can be deduced the current of TRPV3 is greater under negative voltages, and that a slightly larger current is produced by simulations using the EAEF protocol, rather than CompEL. This is further emphasised when assessing the calculated conductances from the two methodologies (Figure 5.3), where it is evident that simulations using EAEF produced a greater conductance than those using CompEL, at both negative and positive voltages.

MD simulations allow one to obtain an understanding of ion permeation on a molec-
Figure 5.2: Current-voltage (I-V) plot of TRPV3 in 150 mM NaCl at transmembrane voltages produced using either CompEL (green) or EAEF (purple). The relationship between the voltage applied and the recorded current indicates an inward rectification.

At the molecular level, and to gain insight that may be difficult to achieve from other biophysical techniques. For example, MD simulations allow us to investigate the free energy surface for a permeating cation. Specifically, we can explore how the free energy changes as a function of a chosen coordinate, e.g. the axis of a pore. This is referred to as the potential of mean force (PMF). The PMF can be calculated from MD simulations using a variety of methodologies, such as steered MD (sMD) [249], adaptive biasing forces (ABF) [250], and umbrella sampling (US) [251]. However, in MD simulations with sufficient sampling, the negative logarithmic density of a species of particles, such as ions, along a reaction coordinate can be used to calculate the PMF. In the case of a system under voltage, this protocol can also be used to estimate quasi-PMF profiles under non-equilibrium conditions. This approach has previously been used to describe the free energy surfaces of KcsA [106] and the RyR1 channel [150].
5 COMPARISON OF THE COMPEL AND EAEF TECHNIQUES FOR IN SILICO ELECTROPHYSIOLOGY

Figure 5.3: Calculated conductances from MD simulations of ion permeation through TRPV3 under negative (left) and positive voltages (right) generated by CompEL (green) or EAEF (purple). Mean inward conductances and SEM were calculated from overlapping 30 ns windows from three-fold replicated 330 ns simulations.

These pseudo-PMF profiles (herein simply referred to as PMF profiles) show similar shapes for Na\(^+\) between the CompEL and EAEF techniques under high (Figure 5.4), medium (Figure 5.5), and low (Figure 5.6) transmembrane voltages. Looking firstly at the quasi-free energy of Na\(^+\) under negative voltage, energy wells are present in the pore of TRPV3 at binding site B (∼11 Å), a minor broad peak within the cavity (∼6-3 Å), and at binding site C (∼-10 Å). These free energy wells and barriers are recapitulated in simulations with transmembrane voltages produced by either CompEL or EAEF.

Similarly, the PMF profiles of Na\(^+\) cations under positive voltages also show remarkable similarity between the two methodologies. Interestingly, comparison between the PMF profiles under voltages of different polarities also highlights some differences in the free energy of the pore; these mainly occur just inside the pore at both the lower and upper gates. At the lower gate (which would be encountered first by cations under a positive voltage), a large free energy barrier can be observed, as opposed to the free energy pro-
file under negative voltages. The magnitude of this free energy barrier may underpin the lower conductance of TRPV3 under positive voltages, as observed in Figure 5.2.
Figure 5.4: Negative logarithmic density profiles for Na\(^+\) ions traversing the pore of TRPV3 under a high transmembrane voltage. The profiles of TRPV3 under negative (dashdotted line) and positive (dotted line) voltages generated by either the CompEL (green) and EAEF (purple) techniques are plotted. The logarithmic ion densities represent quasi-free energy profiles (with a unit of kT). Density minima reflect stably bound ions (i.e. binding sites), while maxima indicate barriers between the binding sites. The location of the cation binding sites discussed in Chapter 3 are represented as grey regions. The plots have been smoothed using a Gaussian filter with a sigma value of 3.
Figure 5.5: Negative logarithmic density profiles for Na\textsuperscript{+} ions traversing the pore of TRPV3 under a medium transmembrane voltage. The profiles of TRPV3 under negative (dashdotted line) and positive (dotted line) voltages generated by either the CompEL (green) and EAEF (purple) techniques are plotted. The logarithmic ion densities represent quasi-free energies (with a nominal unit of kT). Density minima reflect stably bound ions (i.e. binding sites), while maxima indicate barriers between the binding sites. The location of the cation binding sites discussed in Chapter 3 are represented as grey regions. The plots have been smoothed using a Gaussian filter with a sigma value of 3.
Figure 5.6: Negative logarithmic density profiles for Na\(^{+}\) ions traversing the pore of TRPV3 under a low transmembrane voltage. The profiles of TRPV3 under negative (dashdotted line) and positive (dotted line) voltages generated by either the CompEL (green) and EAEF (purple) techniques are plotted. The logarithmic ion densities represent quasi-free energies (with a nominal unit of kT). Density minima reflect stably bound ions (\textit{i.e.} binding sites), while maxima indicate barriers between the binding sites. The location of the cation binding sites discussed in Chapter 3 are represented as grey regions. The plots have been smoothed using a Gaussian filter with a sigma value of 3.

In addition to analyses of conductance and quasi-free energies of permeating ions, an investigation was carried out to assess the difference in the structure of TRPV3 under transmembrane voltages from the two methodologies. One measure of the change of
protein structure during a simulation is the root-mean-square deviation (RMSD).

In this study, the RMSD of the backbone atoms was used as a measure of changes in the tertiary structure of the protein, relative to the starting structure of the simulation. In particular, the distributions of RMSDs over the course of the simulations were compared. Figure 5.7 contains violin plots for systems of different voltage and restraints, comparing the RMSD distributions of the backbone atoms of the entire protein, as well as focusing on the pore and ARD domains (as defined in Figure 5.1). The distribution of RMSD from the simulations shows that the range of values, and thus the degree of structural change, demonstrates no obvious differences between simulations conducted using the CompEL or EAEF methodology. This observation is true regardless of the voltage polarity, and regardless of whether harmonic restraints are applied or not. Small differences in the distribution can be observed in certain pairs of simulations (for example, in simulations under a low, negative voltage using position restraints). However, these differences show no trend and are more likely resulting from stochasticity, rather than from a systematic difference. The mean and standard deviation from these RMSD distributions are summarised in Table 5.3. This table indicates that the average RMSD values and the standard deviation are similar between CompEL and the EAEF methodologies.

Figure 5.7 also compares the RMSD of the whole protein with the pore domain and ARD. The pore domain of the channel shows lower RMSD values than the ARD domain. This is not unexpected, however, as the pore domain in embedded in and constrained by the POPC bilayer used in this simulation. The ARD region on the other hand is a cytoplasmic domain, and therefore has far more flexibility. The dynamics of the ARD are discussed later in this Chapter.
Figure 5.7: Distribution of RMSD of the TRPV3 protein (top), pore domain (centre), and ARD (bottom). Violin plots compare the distribution at a range of voltages produced by either CompEL (green) or EAEF (purple), at either negative (left) and positive (right) voltages. A structural guide to the domains of TRPV3 can be found in Figure 5.1.
In addition to the RMSD, the related root-mean-square-fluctuation (RMSF) was also used to compare the structural flexibility of the systems during the simulations. Whereas the RMSD quantifies the average positional differences of an entire structure (or sub-selection of the structure) over time, the RMSF describes the average fluctuation of a particular residue over time. The RMSF of particle $i$ is calculated according to Equation 5.1, where $x_i$ is the coordinates of particle $i$ and $\langle x_i \rangle$ is the ensemble average position of $i$.

$$RMSF_i = \sqrt{\left\langle (x_i - \langle x_i \rangle)^2 \right\rangle}$$  \hspace{1cm} (5.1)
Comparison of the RMSF profiles of TRPV3 with harmonic restraints on the lower gate under a high transmembrane voltage showed similar plots between the EAEF and CompEL simulations, under both negative and positive voltages (Figure 5.8). A large degree of fluctuation can be observed in the unstructured N-terminal (residues 85-115) and C-terminal (residues 707-745) regions of the protein. In addition, further peaks in the RMSF are apparent at the region corresponding to finger 3 loop of the ARD (residues 239-263). This highly flexible region has been shown to be important in protein-protein interactions in TRPV3, and mutation of residues in this region can lead to altered channel activity and pharmacology [53]. Indeed, the ARD of TRPV channels has long been a topic of interest in the field, with implications of their role in tetramerisation [252, 253], thermosensation [254], and channelopathies [255]. Furthermore, the loops between the S1 and S2 helices (residues 461-480) and the S5 and PH helices (residues 608-624) are also shown as having high RMSF values. All of the aforementioned regions displaying higher RMSF values, and thus greater flexibility, are equally identified in simulations under both negative and positive voltages, generated by both the EAEF and CompEL methodologies.

To better compare the RMSF profiles of simulations conducted using EAEF or CompEL, the difference in the RMSF values between the two methodologies was then calculated for all simulations conducted using the CHARMM force field (Table 5.1). This plot showed that the average RMSF difference between the two methodologies is minimal, with the difference usually $\sim 0.2 \text{ Å}$ and within the standard error of the mean of a difference of $0 \text{ Å}$ (Figure 5.9). However, a difference of $\sim 1.7 \text{ Å}$ can be observed for the most distal residues of the N-terminal region. This indicates that the unstructured N-terminus displays more flexibility in simulations with a transmembrane voltage generated by the EAEF compared to those generated using CompEL. Additional simulations will be necessary to ascertain if unstructured regions and intrinsically disordered proteins (IDPs) are consistently and systematically more flexible in EAEF simulations.

As described previously, the ARD of TRPV channels is implicated in a wide array of mechanisms contributing to proper channel functioning. The dynamics of this domain are particularly pertinent in the homologous TRPN protein of Drosophila melanogaster,
where compression of the ARD underpins the channels mechanosensitivity, acting as a spring that transfers forces to the pore, via a so-called "push-to-open" mechanism [256].

To investigate the dynamics of the ARD of TRPV3 under voltage, the movement of the domain was calculated, both along the pore axis ($z$-plane) and radially from the pore axis ($xy$-plane) relative to the initial structure. Distributions of this movement again show broadly similar dynamics for simulations conducted using EAEF or CompEL (Figure 5.10). Under both negative and positive voltage, simulations conducted using both methodologies show movement along both the $z$ and $xy$ planes, with this movement occurring in both
 directions. By comparing the mean distance moved along these planes (Table 5.4), it can be seen that the average distance moved for both simulations is approximately about 0 Å. Again, differences can be observed between individual pairs of simulations, however these differences are spurious, with no clear trend between simulations conducted using either the CompEL or EAEF methods. This becomes evident when one looks at the combined data average (Table 5.4).

Figure 5.9: Average difference between RMSF values of TRPV3 during simulations under both negative (left) and positive (right) voltages, generated by CompEL or EAEF techniques. The mean RMSF is plotted across all simulations performed using the CHARMM force field, with the standard error of the mean plotted as a shaded region. A positive difference indicates a greater RMSF in CompEL simulations, and a negative difference indicates a greater RMSF in EAEF simulations.
## COMPARISON OF THE COMPEL AND EAEF TECHNIQUES FOR IN SILICO ELECTROPHYSIOLOGY

Table 5.4: Average movement of the ARD domain of TRPV3 under transmembrane voltages. The mean and standard deviation were calculated from three-fold replicated 330 ns simulations. The mean and standard deviation from a combination of all simulations systems of that voltage polarity are also shown.

<table>
<thead>
<tr>
<th>Voltage</th>
<th>Distance along pore axis (Å)</th>
<th>Radial distance from pore axis (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CompEL</td>
<td>EAEF</td>
</tr>
<tr>
<td><strong>Negative</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High (restr)</td>
<td>1.2 ± 2.3</td>
<td>1.1 ± 1.8</td>
</tr>
<tr>
<td>Medium (restr)</td>
<td>0.5 ± 2.1</td>
<td>0.9 ± 1.6</td>
</tr>
<tr>
<td>Low (restr)</td>
<td>0.9 ± 2.3</td>
<td>0.8 ± 2.0</td>
</tr>
<tr>
<td>High (unrestr)</td>
<td>1.1 ± 2.6</td>
<td>0.4 ± 1.4</td>
</tr>
<tr>
<td>Low (unrestr)</td>
<td>0.3 ± 3.0</td>
<td>1.1 ± 1.3</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td><strong>0.8 ± 0.3</strong></td>
<td><strong>0.9 ± 0.3</strong></td>
</tr>
<tr>
<td><strong>Positive</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High (restr)</td>
<td>1.2 ± 2.2</td>
<td>0.3 ± 2.3</td>
</tr>
<tr>
<td>Medium (restr)</td>
<td>1.2 ± 2.7</td>
<td>0.9 ± 1.1</td>
</tr>
<tr>
<td>Low (restr)</td>
<td>1.5 ± 2.5</td>
<td>1.3 ± 1.4</td>
</tr>
<tr>
<td>High (unrestr)</td>
<td>0.6 ± 3.0</td>
<td>0.5 ± 1.3</td>
</tr>
<tr>
<td>Low (unrestr)</td>
<td>-0.1 ± 3.7</td>
<td>0.5 ± 1.9</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td><strong>0.9 ± 0.6</strong></td>
<td><strong>0.7 ± 0.5</strong></td>
</tr>
</tbody>
</table>

### 5.3.2 Case study: TRPV3 using the AMBER19SB force field

Due to differences in their parameterisation (as discussed in Chapter 2.2), simulations conducted with different force fields can result in small differences in the resulting ensembles. Therefore, the investigation of force field effects is an important question in the biophysics community. For example, the effect of different force fields has been investigated for simulations of the multidrug efflux protein P-Glycoprotein [257], Cytochrome P450 [258], and intrinsically disordered proteins [259]. For this reason, comparative simulations of TRPV3 using the AMBER19SB force field [246] were also performed. These simulations were conducted under a high transmembrane voltage and with harmonic restraints on the lower-gate to maintain the channel in an open-state (Table 5.1).
Figure 5.10: Comparison of the relative movement of the ARD region of TRPV3 under a range of transmembrane voltages. The distance of the ARD from the centre of the TRPV3 pore in both the $z$ (top) and $xy$ (bottom) planes is plotted. Violin plots compare the distribution at a range of voltages produced by either CompEL (green) or EAEF (purple), at either negative (left) and positive (right) voltages.
The simulations using the AMBER force field also resulted in continuous cation permeation through the TRPV3, producing 350 permeation events. Similarly to the calculated conductances from the CHARMM force field, the conductances from the AMBER forcefield show an increased value for EAEF simulations compared to the CompEL simulations. It should be noted, however, that the conductances from these simulations are approximately ~ 7x smaller than those calculated from the CHARMM force field. The root-cause of this difference however remains unclear at the present time, and additional simulations should be conducted to ascertain if this reduction is meaningful or stochastic.

The PMF profiles resulting from these simulations showed no notable difference dependent on the methodology used to generate the transmembrane voltage (Figure 5.12), as observed with the CHARMM force field. The PMF profiles under negative voltage captured free energy wells corresponding to binding site B (\( \sim 11 \, \text{Å} \)) and binding site C (\( \sim -10 \, \text{Å} \)). Interestingly, the minor broad peak within the cavity observed in Figure 5.4, is not observed in simulations conducted with the AMBER force field. The cause of this observation is currently unclear, however.
Furthermore, the PMF profiles of Na\textsuperscript{+} cations under positive voltages also showed no differences between the different *in silico* electrophysiology methodologies. However, these PMF profiles also showed differences compared to similar simulations conducted using the CHARMM force field. Firstly, the free energy barrier observed just inside the pore at the lower gate is $\sim 1$ kT smaller in the simulations using the AMBER force field compared to the CHARMM force field. Additionally, the region within the pore cavity spanning from $\sim -4$ to $4$ Å shows a shift of $\sim -1.75$ kT in simulations using the AMBER force field, indicating this region being more energetically favourable compared with the CHARMM force field. Therefore, the reduced Na\textsuperscript{+} conductance recorded in AMBER simulations may be as a result of greater Na\textsuperscript{+} binding, as demonstrated by the increased free energy within the pore cavity.
Figure 5.12: Negative logarithmic density profiles for Na\(^+\) ions through the pore of TRPV3 under a high transmembrane voltage using the AMBER19SB force field. The profiles of TRPV3 under negative (dashdotted line) and positive (dotted line) voltages generated by either the CompEL (green) and EAEF (purple) techniques are plotted. The logarithmic ion densities represent quasi-free energies (with a nominal unit of kT). Density minima reflect stably bound ions (i.e. binding sites), while maxima indicate barriers between the binding sites. The location of the cation binding sites discussed in Chapter 3 are represented as grey regions. The plots have been smoothed using a Gaussian filter with a sigma value of 3.

In addition, the effect of the two different methodologies on protein structure were investigated with the AMBER force field. The distribution of the RMSD showed no major differences between the two methodologies (Figure 5.13). It can be observed that under a
negative voltage, the median RMSD is slightly greater in simulations utilising the Com-
pEL method. However, under a positive voltage, the median RMSD is slightly greater in simulations utilising the EAEF method. This can also been seen when comparing the mean RMSD values, as displayed in Table 5.5. However, these mean differences are small (\(\sim 0.3 \text{ Å} \)), with significant overlap of the standard deviation. It is therefore likely that this difference is a result of the spurious differences observed within pairs of simulations, as seen with the CHARMM forcefield.
Figure 5.13: Distribution of RMSD of the TRPV3 protein (top), pore domain (centre), and ARD (bottom) from simulations using the AMBER19SB force field. Violin plots compare the distribution at a range of voltages produced by either CompEL (green) or EAEF (purple), at either negative (left) and positive (right) voltages. A structural guide to the domains of TRPV3 can be found in Figure 5.1.
Table 5.5: Average RMSD values from simulations of TRPV3 under transmembrane voltages using the AMBER19SB force field. The mean RMSD and standard deviation were calculated from three-fold replicated 330 ns simulations.

<table>
<thead>
<tr>
<th>Voltage</th>
<th>RMSD (Å)</th>
<th>CompEL</th>
<th>EAEF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Negative</td>
<td>High (restr)</td>
<td>4.3 ± 0.6</td>
<td>4.1 ± 0.6</td>
</tr>
<tr>
<td>Positive</td>
<td>High (restr)</td>
<td>4.1 ± 0.9</td>
<td>4.4 ± 0.7</td>
</tr>
</tbody>
</table>

Similarly, the RMSF profiles of TRPV3 showed similar shapes for both methodologies, under both a negative and positive voltage (Figure 5.14). These traces indicate the same regions of increased flexibility as highlighted by simulations using the CHARMM force field, principally: the N- and C-termini, the finger 3 loop of the ARD, the loop between the S1 and S2 helices, and the loop between the S5 and PH helices.

Finally, the movement of the ARD domain along the z and xy plane was evaluated during the course of simulations with the AMBER forcefield. Comparison of the distribution (Figure 5.15) and mean averages (Table 5.6) indicated no differences between simulations conducted dependent on the use of the CompEL or EAEF techniques.

Table 5.6: Average movement of the ARD domain of TRPV3 under transmembrane voltages using the AMBER19SB force field. The mean and standard deviation were calculated from three-fold replicated 330 ns simulations.

<table>
<thead>
<tr>
<th>Voltage</th>
<th>Distance along pore axis (Å)</th>
<th>Radial distance from pore axis (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CompEL</td>
<td>EAEF</td>
</tr>
<tr>
<td>Negative</td>
<td>High (restr)</td>
<td>-0.6 ± 1.9</td>
</tr>
<tr>
<td>Positive</td>
<td>High (restr)</td>
<td>-0.6 ± 2.1</td>
</tr>
</tbody>
</table>
Figure 5.14: RMSF of TRPV3 during simulations using the AMBER19SB force field with a high voltage. The RMSF was compared under both negative (left) and positive (right) voltages, generated by both the CompEL (top, green) and EAEF (bottom, purple) techniques. The mean RMSF is plotted, with the standard deviation plotted as a shaded region.
5 COMPARISON OF THE COMPEL AND EAEF TECHNIQUES FOR IN SILICO ELECTROPHYSIOLOGY

Figure 5.15: Comparison of the relative movement of the ARD region of TRPV3 under a range of transmembrane voltages using the AMBER19SB force field. The distance of the ARD from the centre of the TRPV3 pore in both the $z$ (top) and $xy$ (bottom) planes is plotted. Violin plots compare the distribution at a range of voltages produced by either CompEL (green) or EAEF (purple), at either negative (left) and positive (right) voltages.

5.4 Discussion

The parameters and settings used to conduct MD simulations have an explicit effect on the resulting experimental outputs. Therefore, careful thought must be applied when setting up an MD simulation to ensure that the simulation will optimally model the desired system.

For systems in which the aim is to model the voltage present across biological membranes, two main methodologies are currently available; through ionic imbalances using the CompEL technique, or through a static electric field using the EAEF technique. Previous studies have described that there is no difference between the effects of these two methods on the membrane component of the system [245]. However, this study did not incorporate any ion channels or other proteins within the system. Therefore, it has remained unclear if these two methods would be equivalent for ion channels, particularly those with
large soluble domains. In this Chapter, I have presented a systematic comparison of these two methodologies, considering various measurements, such as the conductance, PMF profile of Na\(^+\) cations, and structural changes of the protein.

Analysis of the conductance from these simulations has indicated that there is a difference in the conductance generated from in silico electrophysiology experiments depending on the methodology used; specifically that simulations using EAEF produces a greater current than CompEL. Conversely, analysis of the PMF profile and protein structure displays no substantial difference between simulations conducted of the TRPV3 channel using either method. These seemingly opposing conclusions were observed in simulations conducted using both the CompEL and EAEF techniques.

One hypothesis to explain this opposition is based upon how the two methodologies generate their respective voltages, as described in Chapter 2.4.2. In CompEL, an ionic imbalance creates a voltage across the membrane of the system. On the other hand, in EAEF simulations, an external electric field is applied across the entire system in a given direction. Despite this, the bulk aqueous salt solution of a system will self-orient itself to reduce the magnitude of the electric field (see Figure 2.3). As a result, in both methods an ion should only feel the potential in proximity to the membrane.

However, in Chapter 2.4.2 it was described how the cytoplasmic domains of proteins cannot freely re-orient to counteract the applied field, and that this may lead to an artefactual force. It remains possible that the cytoplasmic domains of TRPV3 disrupt the reaction potential of the electric field, meaning that charged atoms can feel the transmembrane voltage when further from the membrane than in CompEL simulations. Therefore, in simulations utilising an EAEF, charged atoms will be attracted towards the pore more than those utilising CompEL, and will likely result in a greater number of permeation events. However, the effect of the electric field will be equal on the ion channel itself, explaining the lack of observed differences in protein structure.

To probe this hypothesis, additional simulations could be conducted on a range of ion channels with cytoplasmic domains of differing sizes. Alternatively, a protein design approach could be taken, where truncations are made to the TRPV3 cytoplasmic domains
to reduce their presence in the bulk solution. If this hypothesis were correct, one would expect that the greater the cytoplasmic domains, the greater the disparity between the two methodologies. Consequently, the accuracy of the EAEF technique would be dependent on the size of the cytoplasmic domains of an ion channel.

Whilst the stated similarities and differences have been observed for TRPV3, it will also be necessarily to increase the scope of this study by including other proteins to be able to make conclusive statements. In particular, it would be optimal for these additional systems to include ion channels of differing families and characteristics of the already included TRPV3 channel. For example, simulations conducted using channels of varying cation/anion selectivities, varying sizes, and varying stoichiometries.

Furthermore, this Chapter has highlighted some of the interesting differences arising from the use of different force fields. These were particularly pertinent in the calculated conductances and the PMF profiles of Na\(^+\) cations through the TRPV3 pore. Further analyses will need to be performed to ascertain the route cause of these differences, i.e. whether they result from the ion parameters directly or indirectly from the different water models (TIP3P for CHARMM36m and OPC for AMBER19SB).
6 Concluding remarks

6.1 Summary of thesis

The aim of this thesis was to elucidate the permeation and selectivity mechanisms of the TRPV and TRPM subfamilies of the TRP superfamily. To achieve this, atomistic MD simulations were conducted of open-state structures of TRP channels. In silico electrophysiology experiments were performed by creating bio-mimetic transmembrane voltages, using either the CompEL or EAEF methodologies.

In Chapter 3, the Ca$^{2+}$-selective permeation mechanism of the TRPV5 and TRPV6 channels was determined. Three cation binding sites were identified within the pore, each of which displayed greater affinity for Ca$^{2+}$ binding over Na$^+$. Ca$^{2+}$ permeation proceeded via a knock-on mechanism between these three cation binding sites within the channel pore, therefore imparting selectivity for Ca$^{2+}$ over Na$^+$. A comparison of the non-selective TRPV2 and TRPV3 channels found that these channels contain one less binding site than the Ca$^{2+}$-selective TRPV5 and TRPV6 channels. A novel application of mutual information, termed SSI, between ion binding and unbinding events at consecutive binding sites enabled the quantification of the degree of knock-on taking place in ion permeation. This showed that the level of Ca$^{2+}$ selectivity in TRPV channels is determined by the co-operativity, or coupling, between the transitions at three pore cation binding sites. The Ca$^{2+}$-selective TRPV channels display a highly correlated three-site knock-on, whereas the cation binding site at the extracellular entrance is decoupled from the mechanism in the non-selective TRPV channels, which reduces the overall preference for Ca$^{2+}$ permeation.

In Chapter 4, the monovalent-selective permeation mechanism of the TRPM5 channel was investigated. It was found that two broad binding sites for monovalent cations (Na$^+$ and K$^+$) were present in the TRPM5 channel, within the extracellular pore vestibule and within the pore cavity. Using SSI, it was shown that monovalent cation permeation proceeds via a co-operative knock-on mechanism between these binding sites. On the other hand, divalent cations experience an energetic barrier within the pore cavity. There-
fore, the co-operative permeation mechanism of TRPM5 is absent for divalent cations. Comparison between the pore architecture of the monovalent-selective TRPM5 channel and the aforementioned Ca$^{2+}$-selective TRPV5 channel suggested that this divalent cation energetic barrier is as a consequence of the more hydrophobic nature of the TRPM5 pore.

Finally, in Chapter 5 a systematic comparison was undertaken between the CompEL and EAEF methods; both of which are used to create a transmembrane voltage to drive ion permeation in MD simulations. In this work, TRPV3 was used as a model system, and comparative simulations were conducted at a range of voltages produced by both in silico electrophysiology protocols. No difference in protein structure was observed between these simulations. However, it was observed that the EAEF method consistently produced greater Na$^+$ conductances than those from the CompEL method.

### 6.2 Future directions

The research presented in this thesis has made a substantial contribution to the understanding of the permeation and selectivity mechanisms in TRP channels. For proposed future areas of research I will group Chapters 3 and 4 into one area of focus, and evaluate Chapter 5 as a separate area of focus.

To further develop the studies on the permeation mechanisms of TRPV and TRPM channels, the scope of the current simulations performed should be expanded. Firstly, additional simulations could be performed to encompass other, as yet untested, channels of the TRPV and TRPM subfamilies. This would be particularly pertinent for the work conducted in Chapter 4 where simulations focused on a singular channel, rather than a selection as in Chapter 3. By increasing the number of channels simulated, the confidence in the conclusions drawn from these studies will be increased. Additionally, as one would expect the selectivity mechanisms of a channel to be the same regardless of its species, simulations could also be performed in different orthologues. This would test and possibly confirm the evolutionary conservation of the deduced mechanisms.

In addition to expanding the work conducted in Chapters 3 and 4 by simulating additional TRP channel structures, the studies could also be expanded by performing addi-
tional simulations with different force fields. As described in Chapter 2.2, different force fields can produce different results in simulations due to differences in their parameterisation. Therefore, additional simulations could be performed utilising an alternative force field to further probe the conclusions drawn in this thesis, and to rule out these results being force field-dependent.

Furthermore, as discussed in Chapter 2.3 there are known issues with the parameters of divalent cations in non-polarisable, point-charge force fields. Whilst the multi-site Ca\(^{2+}\) model of Zhang et al. has proven highly successful in this thesis and other published work, it remains possible that further optimisation of this model may be required. For example, in Chapter 3 the in silico \(P_{Ca}/P_{Na}\) of TRPV5 and TRPV6 was far lower than the in vitro electrophysiology values reported within the literature. It remains possible that this may be due to an underestimation in Ca\(^{2+}\)-protein binding energies. Consequently, future work could aim to repeat the work of Chapters 3 and 4 using an alternative method to correct for Ca\(^{2+}\) polarisation, such as ECCR. Furthermore, simulations could be conducted using polarisable force fields, to evaluate the permeation mechanisms observed with these force fields.

In this thesis, co-operative knock-on mechanisms of ion permeation were quantified using SSI. This novel application of mutual information was used heavily in Chapters 3 and 4 to compare the degree of knock-on between different TRP channels and different ionic species. However, further work is needed to benchmark this methodology. For example, the knock-on permeation observed in the SF of K\(^{+}\)-selective channels is arguably the classic and most well known example of a knock-on ion permeation mechanism. Therefore, it would be advantageous to calculate the SSI of permeation in this channel, and to compare these results with those already obtained for TRP channels.

Chapter 5 of this thesis presented the preliminary experiments in a systematic study of the two principle in silico electrophysiology methods. Although the data presented in this chapter includes 23.76 \(\mu s\) of simulation data and 4,628 permeation events, all of this data was generated from a single exemplar protein. In the future it would therefore be necessary to extend this study to incorporate other proteins. Ideally, these other proteins would
encompass a wide range of channel architectures and selectivities. This would allow conclusions to be drawn on the overall differences between the CompEL and EAEF methods. Additionally, simulations should be conducted on ion channels with cytoplasmic domains of differing sizes to probe the observed differences in conductances generated between the two methods.

The described future directions would further the successes already reported within this thesis, and would continue the progress towards a universal understanding of cation permeation and selectivity mechanisms within the TRPM and TRPV subfamilies, and the wider TRP superfamily.
# Bibliography


[109] K. A. McKiernan, A. K. Koster, M. Maduke, and V. S. Pande. “Dynamical model of the CLC-2 ion channel reveals conformational changes associated with selectivity-


[151] C. Liu, A. Zhang, N. Yan, and C. Song. “Atomistic details of charge/space com-
petition in the Ca2+selectivity of ryanodine receptors”. In: *Journal of Physical

Plested, P. Carloni, and H. Sun. “Mechanism of calcium permeation in a glutamate
receptor ion channel”. In: *ChemRxiv* (2022). DOI: 10.26434/chemrxiv-2022-
x73hl-v2.

data bank”. In: *Nature Structural and Molecular Biology* 10 (2003), pp. 980–980.


[155] *All Nobel prizes*. URL: https://www.nobelprize.org/prizes/lists/all-
nobel-prizes/.

prizes/chemistry/2017/summary/.

[157] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tun-

J. Wang, Q. Cong, L. N. Kinch, R. D. Schaeffer, et al. “Accurate prediction of
protein structures and interactions using a three-track neural network”. In: *Science*


new molecular dynamics method”. In: *Journal of Applied Physics* 52 (1981),


parallelism from laptops to supercomputers”. In: *SoftwareX* 1-2 (2015), pp. 19–25. DOI: 10.1016/j.softx.2015.06.001.


