

A bond graph approach to integrative biophysical modelling

Michael Pan
ORCID: 0000-0002-8978-7350

Submitted in total fulfilment of the requirements of the degree
Doctor of Philosophy

June 2019

Department of Biomedical Engineering,
Melbourne School of Engineering,
The University of Melbourne

Abstract

A major goal of systems biology is to develop comprehensive, multi-scale mathematical models of physiological systems that integrate biological knowledge from the scale of molecules to the scale of tissues and organs. Models on this scale hold great potential in advancing our knowledge of biology and medicine, but they have yet to be achieved in complex biological systems. It is widely acknowledged that constructing large-scale models requires the reuse and integration of existing models; however, model integration is currently challenging because many existing models violate the conservation laws of physics, especially conservation of energy. It is therefore highly desirable to express models in a framework that respects the laws of physics and thermodynamics. Bond graphs are an energy-based modelling framework, initially developed for use in multi-physics engineering systems to help derive equations consistent with the laws of physics. More recently, bond graphs have been applied to the field of biology where they have helped in making models physically and thermodynamically consistent. While bond graphs provide several advantages for large-scale modelling such as thermodynamic consistency and hierarchical modelling, they have yet to be applied to large-scale dynamic models of biological systems.

This thesis aims to develop methods based on the bond graph framework to facilitate model reuse and integration. These methods are demonstrated by applying them to biomolecular systems within the cardiac cell. Firstly, bond graphs are applied to membrane transporters, demonstrating that bond graphs can be used to correct thermodynamic inconsistencies within existing models. Secondly, independently developed models of ion channels and transporters are coupled into a model of cardiac electrophysiology, showing that bond graphs can be used to systematically explain the issues of drift and non-unique steady states that affect many existing models. Finally, a generalised method for simplifying models of enzyme kinetics is developed and used to facilitate the development of simple, thermodynamically consistent models of enzymes that are easily incorporated into larger models.

Declaration

This is to certify that:

- this thesis comprises only my original work towards the PhD except where indicated in the preface;
- due acknowledgement has been made in the text to all other material used; and
- this thesis is fewer than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices

Michael Pan

26 June 2019

Preface

The following chapters contain material from published works that have resulted from the research presented in this thesis:

- Chapter 2 comprises a slightly modified version of the author accepted manuscript for the following publication that was published in the *Journal of Theoretical Biology* on 21 November 2019:

Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Crampin, E.J., 2019. A thermodynamic framework for modelling membrane transporters. *Journal of Theoretical Biology.* 481, 10–23. doi: 10.1016/j.jtbi.2018.09.034

The chapter also contains work from the following manuscript that was submitted for publication in *Physiome* on 2 November 2017:

Pan, M., Gawthrop, P.J., Cursons, J., Tran, K., Crampin, E.J., 2017. The cardiac Na^+/K^+ ATPase: An updated, thermodynamically consistent model. *Physiome.*

Preprint available at arXiv:1711.00989

- Chapter 3 comprises a slightly modified version of the author accepted manuscript for the following peer-reviewed publication that was published on 1 June 2018:

Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Crampin, E.J., 2018. Bond graph modelling of the cardiac action potential: implications for drift and non-unique steady states. *Proc. R. Soc. A* 474, 20180106. doi: 10.1098/rspa.2018.0106

Chapter 4 contains unpublished material not submitted for publication. I acknowledge the contribution of Dr. Peter Cudmore, who helped develop the software used to generate the results for this chapter.

The research presented in this thesis was funded by an Australian Government Research Training Program Scholarship.

Acknowledgements

I am fortunate to have received the assistance and kindness of many others in crafting this body of work. It is only through the guidance of my supervisors Prof. Edmund Crampin, Dr. Joseph Cursons and Prof. Peter Gawthrop that I was able to turn a budding scientific interest into a thesis. They have been unwavering in their support and generous with their time and advice, which has allowed me to navigate the complexities and nuances of academic research. I am particularly grateful for the late nights and weekends they spent revising drafts of papers and chapters, as well as their ability to always see the value in my work even when I couldn't.

Edmund was the first to inform me of this project, and I am eternally grateful for his agreement to supervise me. Further, he took interest in my development, provided me with astute feedback to steer my work in the right direction and funded me towards the end of my degree. Joe's empathy and pragmatism have been instrumental in moving my work forward. He went to great lengths to regularly scrutinise my notes in detail and walked across campus to meet with me even after we moved into separate buildings. I thank Peter for bringing his technical expertise to my research, listening to my doubts about my work and for providing me with opportunities to be involved with his research. Thanks to my supervisors, my PhD experience has been largely enjoyable and rewarding.

In the initial stages of my project, I found discussions with Kenneth Tran very useful in helping me come to grips with the modelling of transporters and cardiac cells. I also thank my committee chair Jonathan Manton for his encouragement and for ensuring that my project stayed on track.

I am grateful for the support provided by both past and present members of the Systems Biology Laboratory: Daniel Hurley, Vijay Rajagopal, Greg Bass, Ivo Siekmann, Hilary Hunt, Claire Miller, Matt Faria, Stuart Johnston, David Ladd, Agne Tilunaite and Pete Cudmore – who also developed software tools that drastically improved the productivity of my work. They have been a source of endless advice and discussion, and more importantly, they have been as friendly and fun as they are motivated.

I thank my colleagues at the Department of Biomedical Engineering and the mathematical biologists at the School of Mathematics and Statistics for their company over the years, particularly to my office mates Shouryadipta Ghosh, Jess Crawshaw and Punya Alahakoon Mudiyanselage.

During the past few years, I spent countless hours of gossip, games and walks with my friends Darren, Ewe Jin, Winnie, Nat, David, Keith, Chris, Ken, Mandy, Naomi, Marina, Felicia, Annette, Mourad and Melinda. They have been genuinely understanding, encouraging and inspiring. It is in no small part due to them that I have remained sane over the years, and I thank them for reminding me of the life I have outside of work.

Finally, my deepest gratitude goes to my parents for their love and care, making sure I was healthy every day, celebrating each of my successes along the way, and for sparking my initial interest in becoming a scientist, engineer and lifelong learner.

Contents

1	Introduction	1
1.1	Systems biology and the Physiome	1
1.2	Mathematical modelling	2
1.3	Model integration	10
1.4	The case for an energetic framework	13
1.5	Bond graph modelling	14
1.6	Conclusions	26
2	A thermodynamic framework for modelling membrane transporters	29
2.1	Introduction	29
2.2	Hypothetical models	31
2.3	Thermodynamic models of cardiac cell transporters	42
2.4	Discussion	49
2.5	Conclusion	52
2.A	TF components in biochemical networks	52
2.B	Converting existing kinetic models to bond graphs	53
2.C	Modifications to the Terkildsen et al. model of the Na^+/K^+ ATPase . . .	56
2.D	Stoichiometry and parameters for SERCA	59
2.E	Stoichiometry and parameters for Na^+/K^+ ATPase	62

3 Bond graph modelling of the cardiac action potential	67
3.1 Introduction	67
3.2 Methods	69
3.3 Results	83
3.4 Discussion	90
3.5 Conclusion	96
3.A Fitting ion channel parameters	96
3.B Channel-specific modelling issues	100
3.C Ion transporters	103
3.D Ca^{2+} buffering	105
3.E Bond graph parameters	106
3.F Charge conserved moiety	107
4 Generalised model reduction of enzyme kinetics	109
4.1 Introduction	109
4.2 Methods	111
4.3 Results	119
4.4 Discussion	132
4.5 Conclusion	136
4.6 Future work	136
4.A Model reduction	137
4.B Fitting energetic parameters to kinetic data	144
4.C Fitting experimental data for Na^+/K^+ ATPase	145
5 Conclusions	149
5.1 Summary of research	149
5.2 Implications and further work	150

List of Figures

1.1	Cardiac cell physiology	7
1.2	A simple bond graph example representing an RLC circuit.	15
1.3	Bond graph modules.	17
1.4	Bond graph representations of biochemical systems.	22
2.1	A simple enzyme cycle	32
2.2	Transport of a species coupled to an energy supply	38
2.3	Transport of a charged species	40
2.4	Kinetic and bond graph representations of the cardiac SERCA model by Tran et al.	43
2.5	Simulation of the SERCA pump	44
2.6	The cardiac Na^+/K^+ ATPase model	46
2.7	Fit of the cardiac Na^+/K^+ ATPase model to current-voltage measurements	47
2.8	Fit of the cardiac Na^+/K^+ ATPase model to metabolite dependence data	48
2.9	Simulation of the Na^+/K^+ ATPase	49
2.10	Bond graph of the reaction $2A \rightleftharpoons B$	52
3.1	Cardiac action potential model	71
3.2	Conceptual representations of physical systems	73
3.3	The bond graph model of a plateau K^+ channel	77
3.4	Comparison of I-V curves between the Luo-Rudy and bond graph models	78
3.5	Channel states of a Na^+ channel	80

3.6	Fits for bond graph parameters against corresponding gating equations from the Luo-Rudy model	81
3.7	A simulation of the cardiac action potential using a bond graph model . .	84
3.8	Effect of stimulus type and variable ion concentrations on drift in the bond graph model	88
3.9	Effect of initial conditions on steady-state behaviour	89
3.10	Fitting the <i>f</i> -gate of the L-type Ca^{2+} channel	103
3.11	The bond graph model of NCX	104
3.12	Fit of NCX model to data	105
4.1	A reversible Michaelis-Menten model	112
4.2	Simulations of the Michaelis-Menten model	119
4.3	Variation of the reversible Michaelis-Menten model with the amount of substrate	120
4.4	A four-state enzyme cycle	122
4.5	The cycling rate of the four-state enzyme model	123
4.6	Fit of reduced versions of the four-state model to synthetic data generated from the full model	125
4.7	Fits of all possible enzymatic mechanisms to synthetic data generated from the four-state model	127
4.8	Fits of possible models of the Na^+/K^+ ATPase to data	128
4.9	Selected mechanisms from fitting potential models to data for the Na^+/K^+ ATPase	130
4.10	Fits of selected Na^+/K^+ ATPase mechanisms to data	131
4.11	Reduction of bond graph models of enzymes	134
4.12	Generalised model of full enzyme kinetics	139
4.13	Fast timescale of enzyme kinetic models	140
4.14	Slow timescale of enzyme kinetic models	141

List of Tables

1.1	Bond graph components and mathematical relationships.	19
1.2	Summary of effort and flow variables.	20
2.1	Forward stoichiometric matrix for the Tran et al. (2009) SERCA model	59
2.2	Reverse stoichiometric matrix for the Tran et al. (2009) SERCA model	59
2.3	Parameters for the Tran SERCA model	60
2.4	Parameters for the bond graph model of the SERCA pump	61
2.5	Forward stoichiometric matrix for the Terkildsen et al. (2007) model of Na ⁺ /K ⁺ ATPase	62
2.6	Reverse stoichiometric matrix for the Terkildsen et al. (2007) model of Na ⁺ /K ⁺ ATPase	63
2.7	Kinetic parameters for the updated cardiac Na ⁺ /K ⁺ ATPase model	64
2.8	Parameters for the bond graph version of the updated cardiac Na ⁺ /K ⁺ ATPase model	65
3.1	Conserved moieties associated with chemostat selection	86
3.2	The values of conserved moieties and chemostats under different initial conditions	91
3.3	Permeabilities of the GHK equations used for the bond graph model	97
3.4	Gate transition parameters	99
3.5	Summary of conditions used to simulate f-gate for fitting parameters	102
3.6	Amounts of each ion channel per cell	106

4.1	Energetic parameters estimated using synthetic data for the four-state enzyme model	124
4.2	Simulation conditions for Figure 4.10	132
4.3	The binding steps used to generate candidate models for the Na^+/K^+ ATPase	146

List of Symbols and Acronyms

Bond graph symbols

0	Equal effort junction.
1	Equal flow junction.
C	Charge store (e.g. chemical species, capacitor, tank).
I	Momentum store (e.g. inductor, inertia).
R	Dissipative component (e.g. resistor, pipe).
Re	Chemical reaction.
Re_GHK	A modulated Re component used to describe the Goldman-Hodgkin-Katz equation.
Se	Source of effort.
SS	Source-sensor components, representing external connections.
TF	Transformer.

Nomenclature

A^f	Forward affinity; the total chemical potential of the reactants of a reaction [J/mol].
A^r	Reverse affinity; the total chemical potential of the products of a reaction [J/mol].
A	Reaction affinity [J/mol].
C	Capacitance [F].
e	Effort.
e_0	Total amount of enzyme.
Exp	Element-wise exponential.
F	Faraday's constant (96485 C/mol).
f	Flow.
G	The left nullspace matrix of the stoichiometric matrix that defines the conserved moieties of a biochemical system.
I	Current [A].
K_d	Dissociation constant.
K_s	The thermodynamic constant of a species s [mol ⁻¹].
K	A vector containing the thermodynamic constants of all species in a biochemical network.
K	A diagonal matrix containing the species thermodynamic constants of a biochemical network.
k_r^+	Forward kinetic constant of reaction r .
k_r^-	Reverse kinetic constant of reaction r .
k^+	A vector containing all the forward kinetic constants of a biochemical network.
k^-	A vector containing all the reverse kinetic constants of a biochemical network.
Ln	Element-wise logarithm.
N	Stoichiometric matrix.
N^f	Forward stoichiometric matrix.
N^r	Reverse stoichiometric matrix.

R	Universal gas constant (8.314 J/mol/K).
T	Temperature [K].
v	The rate of a reaction [mol/s].
V	A vector of the reaction velocities in a biochemical network, except in Chapter 3 where it is the membrane potential.
\mathbf{V}	A vector of the reaction velocities in a biochemical network (Chapter 3).
v_{cyc}	The reaction rate of an enzyme, normalised by the total amount of enzyme [s^{-1}].
V_m	Membrane potential.
x	Molar amount [mol].
X	A vector containing the molar amounts of each species in a biochemical network, or the state vector in more generalised systems.
\mathbf{X}	A vector containing the molar amounts of each species in a biochemical network (Chapter 1).
X^{cd}	A vector containing the amounts of the dynamically varying species.
z	A unit of charge.
ΔG	Gibb's free energy of reaction [J/mol].
ΔG^0	Standard free energy of reaction under conditions where all concentrations are set to 1M [J/mol].
κ_r	The rate constant of a reaction r [mol/s].
$\boldsymbol{\kappa}$	A vector containing the rate constants of all reactions in a biochemical network.
$\boldsymbol{\kappa}$	A diagonal matrix containing the reaction rate constants of a biochemical network.
μ	Chemical potential [J/mol].

Acronyms

ADP	adenosine diphosphate.
AMP	adenosine monophosphate.
AP	action potential.
APD	action potential duration.
ATP	adenosine triphosphate.
BG	bond graph.
GHK	Goldman-Hodgkin-Katz.
NCX	sodium-calcium exchanger.
Pi	inorganic phosphate.
SERCA	sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase.
SR	sarcoplasmic reticulum.

Chapter 1

Introduction

1.1 Systems biology and the Physiome

For most of the 20th century, the study of biology was dominated by a reductionist approach that involved cataloguing the details of every biological molecule and studying them in isolation (Hartwell et al., 1999; Kohl et al., 2010). While this approach is useful for many purposes, it is increasingly appreciated that biological function arises from complex interactions that can only be studied in terms of systems rather than separate processes (Cornish-Bowden et al., 2004; Kohl et al., 2010). Therefore, the effect of a gene or protein on phenotype is best understood when also considering both the wider biological network and the environment (Gonçalves et al., 2013; Kohl et al., 2010; Cornish-Bowden et al., 2004). The field of biology is becoming increasingly data-intensive, and the amount of data generated from biological experiments is increasing at a pace that exceeds our ability to understand it. Despite the vast amount of data gathered from high-throughput experiments (including genomics, transcriptomics and proteomics), we are yet to understand how biomolecular components work together to give rise to physiological function (Beard et al., 2005; Cornish-Bowden et al., 2004). As a result, the field of systems biology arose out of the need to understand biological systems as complex systems rather than a collection of isolated components. Qualitative reasoning alone is often insufficient for analysing complex systems, thus mathematical models are essential tools for systems biology (Noble, 2002; Gonçalves et al., 2013).

While mathematical models were initially formulated as simplifications of the biological systems they represent, there is a growing push to produce realistic models that account for all processes that occur within cells and tissues (Babtie and Stumpf, 2017). A major goal for systems biology, undertaken in the Physiome and Virtual Physiological Human Projects, is to develop a comprehensive quantitative model of the entire human body (the “physiome model”) that integrates all biological knowledge from the scale of molecules to entire organisms (Bassingthwaite, 2000; Crampin et al., 2004a; Hunter and Borg, 2003). Such a model could prove revolutionary for how we approach biology and medicine. Experimentally, *in silico* testing of a physiome model could generate new hypotheses for experiments (Hunter and Borg, 2003; Babtie and Stumpf, 2017). Clinically, a physiome model could be used to link the effects of drugs at the biomolecular level to tissue- and

organ-level observations, allowing modelling of biomolecular systems to be integrated with medical imaging techniques to tailor treatments for specific individuals (Hunter and Borg, 2003; Hunter, 2016).

With the complexity of biological systems, it is imperative that we build large-scale models by reusing existing models validated in the literature (Smith et al., 2007; Mallavarapu et al., 2009). For instance, the field of cardiac cell modelling has frequently reused and integrated independently developed models to advance from models of electrophysiology to fully integrated models that incorporate electrophysiology, signalling, contraction and metabolism (Luo and Rudy, 1994a; Crampin and Smith, 2006; Tran et al., 2015). Another example of successful model reuse is the whole-cell model of *Mycoplasma genitalium*, in which individual models were coupled together to form a comprehensive model accounting for every biochemical process within the cell (Karr et al., 2012). However, the task of model integration for physiome modelling remains challenging as models of biological processes can span across multiple temporal and spatial scales (Nickerson et al., 2016). Currently, the reuse of computational models is complicated by issues such as the inflexibility of software implementations (Mallavarapu et al., 2009; Neal et al., 2014), lack of a common software framework (Nickerson et al., 2016) and difficulty in dealing with models expressed under different formalisms (Machado et al., 2011). To deal with these issues, part of the Physiome Project has been directed at creating software standards that facilitate the reuse and integration of existing models (Nickerson et al., 2016). However, as discussed later in § 1.3.2, there has been a relative lack of attention directed towards the equally important physical issues.

In Section 1.2, I summarise the utilities that mathematical models can provide to physiology, and briefly outline modelling techniques that are presently used in biochemistry and cardiac cell physiology, two areas of study in which mathematical models are frequently coupled together. In Section 1.3, I then outline the progress the community has made towards effective model integration as well as unresolved limitations. In Section 1.4, I describe how an energetic framework could potentially resolve these limitations. A specific energetic framework, the bond graph, is introduced in Section 1.5 which outlines the basic theory of bond graph modelling and summarises how bond graphs have been applied to physiological systems. Finally, in Section 1.6, I justify the aim of my thesis, which is to use bond graphs to facilitate model reuse and integration, following which I provide an outline of the thesis.

1.2 Mathematical modelling

Biological systems are physical systems. Therefore, like mechanical, electrical and chemical systems, a solid mathematical and theoretical underpinning is required to make sense of quantitative data generated from biological systems (Omholt and Hunter, 2016). Mathematical models are useful tools for driving mechanistic understanding and predictive analysis, and this is exemplified in the Michaelis-Menten model and the Hodgkin-Huxley model, which are widely used in both experimental and theoretical studies. The Michaelis-Menten model is commonly used in biochemistry to explain measurements from enzyme assays (Schnell and Maini, 2003; Keener and Sneyd, 2009),

while the Hodgkin-Huxley model provided the first quantitative explanation of how action potentials in excitable membranes arise (Hodgkin and Huxley, 1952). It is interesting to note that in both cases, mechanistic assumptions encoded within these models had not been elucidated at the time the models were first proposed but were later observed experimentally (Gunawardena, 2014; Bezanilla and Armstrong, 1977; Rudy and Silva, 2006). Thus, mathematical models can be used to complement experimental techniques in understanding biological systems.

As discussed in Bailey (1998) and Beard et al. (2005), mathematical models of physiological systems aim to achieve any combination of the following five goals: (1) “to organize disparate information into a coherent whole”; (2) “to think (and calculate) logically about what components and interactions are important in a complex system”; (3) “to simulate, predict, and optimize procedures, experiments, and therapies”; (4) “to disprove hypotheses and to define improved hypotheses”; and (5) “to understand the essential qualitative features”. In order to rigorously test assumptions and hypotheses, mathematical models require one to quantitatively encode their assumptions about the system in question (Mallavarapu et al., 2009; Finkelstein et al., 2004). Therefore, it is important to note that while models may incorporate our best understanding about a system, they are ultimately incorrect, and in accordance with goal 4, should be seen as vehicles for helping us understand biology rather than the final goal (Bassingthwaighe, 2000; Babtie and Stumpf, 2017; Kohl et al., 2010). Thus models can only be seen as good or bad in the context of a purpose, and the purpose of a model dictates how it is formulated and the amount of detail it contains (Kohl et al., 2010; Voit, 2017).

Biological systems may be modelled using empirical or biophysical approaches. While the following paragraphs contrast these two approaches, it should be noted that they are complimentary and both are essential for systems biology (Villaverde and Banga, 2014). Empirical (or data-driven) approaches are commonly used for modelling complex systems where little is known about specific interactions. Examples of these include partial least squares regression (Helland, 1990), correlation networks (Batushansky et al., 2016; Weckwerth, 2003), mutual information networks (Meyer et al., 2007) and machine learning (Baker et al., 2018). While empirical approaches generally neglect the system dynamics, frameworks such as S-systems enable the empirical prediction of dynamics – although in the case of S-systems, this comes at the expense of losing stoichiometric constraints (Beard et al., 2004b; Voit, 2017). Because empirical approaches do not model potentially complicated interactions between components, their primary strength is that they are computationally efficient, allowing them to model large biological networks. However, the lack of a mechanistic underpinning is also a limitation as empirical models only provide insights on the specific datasets they were trained on, and fail to extrapolate beyond that data (Birtwistle et al., 2013). As a result, they are often inconsistent with each other and fail when coupled together (Beard et al., 2005; Smith et al., 2007).

In contrast to empirical models, biophysical (or mechanistic) models are built upon the fundamental laws of physics (Smith et al., 2007). These models have been applied to a wide range of physiological systems, including biochemical reaction networks (Beard and Qian, 2008), muscle mechanics (Zahalak, 1990), tumour growth (Tracqui, 2009) and neural activity (Hasenstaub et al., 2010). Because physical conservation laws often relate to the dynamics of a system, differential equations are a commonly used framework for biophysical modelling (Keener and Sneyd, 2009). An advantage of a

biophysical description is that parameters have physical interpretations, aiding intuitive understanding (Forger et al., 2007). Compared to empirical models, biophysical models describe processes in more detail, but at the expense of computational complexity, which can impede parameter estimation (Gonçalves et al., 2013). However, because biophysical models explicitly encode physical approximations and assumptions about a system, they provide a framework for addressing potential issues in the process of model integration (Beard et al., 2005; Smith et al., 2007). As a result, biophysical models are more likely than empirical models to remain predictive once coupled (Smith et al., 2007; Niederer et al., 2017).

Biochemical systems and cardiac cells are two systems in which models are frequently coupled together. In the rest of this section, I briefly review how these systems are modelled.

1.2.1 Biochemical reaction networks

Biochemical networks are a collection of chemical species and the reactions the species are involved in. In many cases, the rates of each reaction follow the law of mass action, where the rate of reaction is proportional to the amount of each substrate. More specifically, if $\{S_i\}$ represents the set of (indexed) species of a network, for a reversible reaction

$$\sum_{i \in E} \nu_i^f S_i \rightleftharpoons \sum_{j \in P} \nu_j^r S_j \quad (1.1)$$

with reactants $\{S_i : i \in E\}$, products $\{S_i : i \in P\}$ and stoichiometric coefficients ν_i^f and ν_j^r , the rate of the forward reaction is

$$v^+ = k^+ \prod_{i \in E} x_i^{\nu_i^f} \quad (1.2)$$

and the rate of the reverse reaction is

$$v^- = k^- \prod_{j \in P} x_j^{\nu_j^r} \quad (1.3)$$

where x_i [mol or M] is the amount of S_i (Keener and Sneyd, 2009; Ederer and Gilles, 2007). The total rate of reaction v [mol/s or M/s] is therefore

$$v = v^+ - v^- = k^+ \prod_{i \in E} x_i^{\nu_i^f} - k^- \prod_{j \in P} x_j^{\nu_j^r} \quad (1.4)$$

While irreversible reactions can be modelled by excluding either the forward or reverse reaction, the principle of microscopic reversibility states that any physical process can in principle be reversed. Thus, in reality, all reactions are reversible (Polettini and Esposito, 2014).

General biochemical systems can be modelled by summing together contributions of the reactions to each of the species. A convenient form for expressing this relationship is

$$\frac{d\mathbf{X}}{dt} = NV(\mathbf{X}) \quad (1.5)$$

where \mathbf{X} is a vector of concentrations (or molar amounts) of each species and $V(\mathbf{X})$ is a vector containing the reaction rates described by Eq. 1.4 (van der Schaft et al., 2013; Palsson, 2006). N is the stoichiometric matrix that relates the fluxes to the rates of change of each species. For example, the reaction scheme



can be represented using

$$\mathbf{X} = \begin{bmatrix} x_A \\ x_B \\ x_C \end{bmatrix}, \quad N = \begin{bmatrix} -1 & 0 & 1 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}, \quad V = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} k_1^+ x_A - k_1^- x_B \\ k_2^+ x_B - k_2^- x_C \\ k_3^+ x_C - k_3^- x_A \end{bmatrix} \quad (1.7)$$

where k_1^+ , k_1^- , k_2^+ , k_2^- , k_3^+ and k_3^- are the kinetic parameters associated with the reactions. Similarly, the reaction scheme



can be represented using

$$\mathbf{X} = \begin{bmatrix} x_A \\ x_B \\ x_E \\ x_C \end{bmatrix}, \quad N = \begin{bmatrix} -2 & 0 \\ 0 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}, \quad V = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} k_1^+ x_A^2 x_E - k_1^- x_C \\ k_2^+ x_C - k_2^- x_B x_E \end{bmatrix} \quad (1.9)$$

The stoichiometric representation in Eq. 1.5 is useful because it enables systems-level analyses using established matrix methods. For example, if we define G as the left nullspace matrix of N , which contains a basis for the kernel of N^T in its rows, we find that $G\dot{\mathbf{X}} = GNV = 0$. Thus, the entries of $G\mathbf{X}$ are pools of species that do not change over time and are therefore known as conserved moieties (Palsson, 2006; Klipp et al., 2009). Conservation of adenosine (ATP + ADP + AMP) is an example of a conserved moiety that is often observed in metabolic systems (Haraldsdóttir and Fleming, 2016; Schuster and Hilgetag, 1995). In the example described in Eqs. 1.8–1.9, a possible left nullspace matrix is

$$G = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{bmatrix} \quad (1.10)$$

The first row of G describes conservation of enzyme ($x_E + x_C = \text{constant}$), while the second row describes conservation of substrate ($x_A + 2x_B + 2x_C = \text{constant}$).

Similarly, the right nullspace matrix H can be defined as a matrix containing a basis for the kernel of N in its columns. Then if $V = Hc$ for some column vector c , $\dot{\mathbf{X}} = NHc = 0$. Hence H defines the space of reaction fluxes that do not affect the concentrations of species, which are known as cycles (Palsson, 2006; Klipp et al., 2009). Note that in an isolated biochemical system, all steady-state cycles must be zero because reactions with finite flux dissipate energy, and systems with non-zero steady-state fluxes would generate heat without a source of energy, violating the laws of thermodynamics. Therefore, the

right nullspace matrix also defines a set of constraints (the Wegscheider conditions, or detailed balance) that the kinetic parameters must satisfy to be thermodynamically consistent (Ederer and Gilles, 2007; Beard et al., 2004a; Liebermeister et al., 2010). A shortcoming of the kinetic representation is that parameters cannot be adjusted independently without precluding the possibility of violating the laws of thermodynamics; for example, the parameters in the reaction scheme in Eq. 1.6 must satisfy the constraint

$$\frac{k_1^+ k_2^+ k_3^+}{k_1^- k_2^- k_3^-} = 1 \quad (1.11)$$

in order to be thermodynamically consistent. As will be discussed in § 1.5.5, an alternative energetic formulation avoids this issue.

1.2.2 Cardiac cell physiology

The physiological function of cardiac cells depends upon an organised set of processes that span the electrical, chemical and mechanical domains (Figure 1.1). Because of the multi-physical nature of these processes as well as the complex feedback loops involved, the cardiac cell has become the target of several modelling studies, leading to numerous iterations between models and experiment (Fink et al., 2011; Noble and Rudy, 2001). The electrical signal initiating the series of events leading to contraction, known as the action potential (Figure 1.1B, top panel), travels along the plasma membrane and causes voltage-gated L-type Ca^{2+} channels to open (Noble and Rudy, 2001). This causes a small and localised increase in intracellular Ca^{2+} . While not sufficient to substantially change the overall Ca^{2+} concentration within the cell, the small amount of Ca^{2+} that moves into the cell binds to ryanodine receptors (RyRs) that are present on the sarcoplasmic reticulum membrane and localised near L-type Ca^{2+} channels. Ca^{2+} opens ryanodine receptors, triggering the release of large amounts of Ca^{2+} from the sarcoplasmic reticulum into the cytoplasm (Figure 1.1B, middle panel) – known as calcium-induced calcium release (Fearnley et al., 2011). While Ca^{2+} release through ryanodine receptors is required for contraction, these signals can also be modulated by inositol trisphosphate receptors (IP3Rs) that trigger signalling pathways leading to cardiac hypertrophy (Nakayama et al., 2010). Protein kinase A (PKA) is another regulator of Ca^{2+} release. It can phosphorylate both L-type Ca^{2+} channels and ryanodine receptors, increasing their open probabilities and making myocytes more contractile (Fearnley et al., 2011). Once released into the cytoplasm, Ca^{2+} activates contraction by binding to troponin, allowing actin and myosin filaments to slide across each other to generate tension within the cell (Figure 1.1B, bottom panel). ATP drives the process of cross-bridge cycling that causes myosin filaments to pull actin filaments inwards (Keener and Sneyd, 2009; Fearnley et al., 2011), and cross-bridge cycling is the major contributor to energy consumption in cardiac cells.

During muscle relaxation, the ions that have traversed membranes during the contraction phase must be returned to their original compartments. This is achieved through various membrane transport processes (Figure 1.1A). Two major active transporters that drive this process are the sarcoplasmic/endoplasmic reticulum Ca^{2+} -ATPase (SERCA) and the Na^+/K^+ ATPase. SERCA is primarily responsible for returning Ca^{2+} from the cytosol to the sarcoplasmic reticulum, whereas the Na^+/K^+ ATPase is used to return Na^+

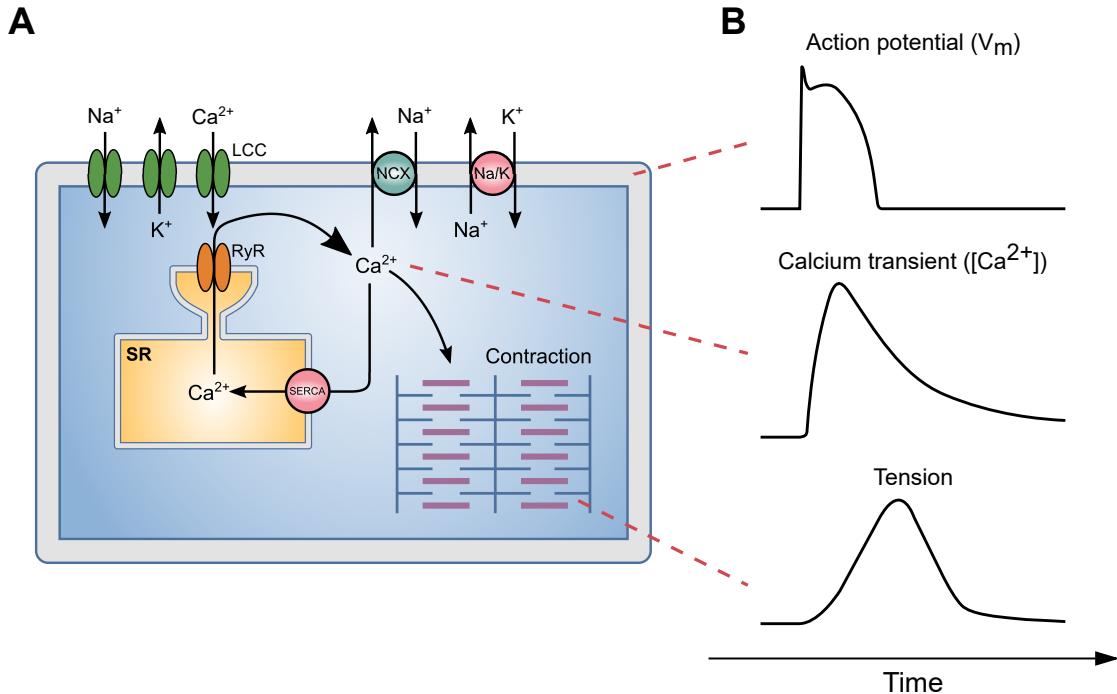


Figure 1.1: Cardiac cell physiology. (A) Schematic of the cardiac cell; (B) The components of cardiac physiology give rise to a wide range of measurable signals, including the membrane potential that arises from ionic currents (top), the intracellular Ca²⁺ transient arising from the movement of Ca²⁺ between compartments (middle), and the tension arising from the activation of contraction by Ca²⁺ (bottom). Abbreviations: SR, sarcoplasmic reticulum; LCC, L-type Ca²⁺ channel; Na/K, Na⁺/K⁺ ATPase; SERCA, sarcoplasmic/endoplasmic Ca²⁺-ATPase; RyR, ryanodine receptor; NCX, sodium-calcium exchanger; V_m , membrane potential.

and K⁺ to the extracellular and intracellular sides of the plasma membrane respectively. The Na⁺ gradient generated by the Na⁺/K⁺ ATPase is also used to drive the Na⁺-Ca²⁺ exchanger (NCX), which moves Ca²⁺ from the cytosol to the extracellular space (Fearnley et al., 2011; Noble and Rudy, 2001).

Numerous models exist for specific subsystems within the cardiac cell. Models of cardiac electrophysiology have been developed for different animal species to explain electrophysiological measurements on cardiac membrane (Noble and Rudy, 2001). In these models, the membrane is modelled as a linear capacitor. The rate of accumulation of charge across the membrane is the sum of the contribution of all ionic currents across the membrane, therefore the membrane potential V_m [V] can be expressed using the differential equation

$$\frac{dV_m}{dt} = -\frac{1}{C_m} \sum_j I_j \quad (1.12)$$

where C_m [$\mu\text{F}/\text{cm}^2$] is the membrane capacitance and I_j [$\mu\text{A}/\text{cm}^2$] are the transmembrane currents that arise from the movement of ions across the sarcolemma. In cardiac cells, commonly modelled currents include the fast Na⁺ current I_{Na} , time-independent K⁺

current I_{K1} , time-dependent K⁺ current I_K , plateau K⁺ current I_{Kp} and L-type Ca²⁺ current I_{CaL} (Luo and Rudy, 1994a; Rudy and Silva, 2006).

The mathematical expressions describing I_j may take on different forms of varying complexity. The simplest are algebraic relationships, for example, the linear current-voltage relationship

$$I_j = \bar{g}_j(V_m - E_j) \quad (1.13)$$

where \bar{g}_j [mS/cm²] is the channel conductance, V_m is the membrane potential in mM and E_j [mV] is the Nernst potential where the channel is at equilibrium (Keener and Sneyd, 2009). The Nernst potential can be expressed in terms of the ion concentrations as

$$E_j = \frac{RT}{z_j F} \ln \left(\frac{[s_j]_e}{[s_j]_i} \right) \quad (1.14)$$

where $R = 8314$ J/kmol/K is the universal gas constant, T [K] is the absolute temperature, $F = 96485$ C/mol is Faraday's constant, s_j is the ion transported by the channel, $[s_j]_a$ is the concentration of s_j on the intracellular ($a = i$) or extracellular ($a = e$) side and z_j is the charge of s_j . Another frequently used algebraic relationship is the nonlinear Goldman-Hodgkin-Katz (GHK) equation:

$$I_j = P_j \frac{z_j^2 F^2}{RT} V_m \left(\frac{[s_j]_i - [s_j]_e e^{-z_j F V_m / RT}}{1 - e^{-z_j F V_m / RT}} \right) \quad (1.15)$$

where P_j [cm/s] is the permeability of the ion channel (Keener and Sneyd, 2009). A limitation of algebraic relationships is that they do not incorporate any dynamics associated with gating processes that open and close the channel. In models of cardiac electrophysiology, gating is most often incorporated using a gating term o :

$$I_j = o \cdot I_{\max} \quad (1.16)$$

where I_{\max} is the maximum current that may take on an algebraic relationship such as that described by Eq. 1.15 or Eq. 1.16 (Rudy and Silva, 2006; Luo and Rudy, 1994a). The open probability o is a function of channel-specific gating variables; for example, a simplistic model of a Na⁺ channel may use $o = m^3 h$, where m and h are the gating variables that evolve according to the differential equation

$$\frac{dw}{dt} = \alpha_w(V)(1 - w) - \beta_w(V)w \quad (1.17)$$

where for $w \in \{m, h\}$, $\alpha_w(V)$ is the opening transition rate and $\beta_w(V)$ is the closing transition rate. This follows the Hodgkin-Huxley formalism, where the gating processes m and h are assumed to be independent of each other. However, more complex Markov state models can be used to account for the dynamics of gating processes that cannot be decomposed into independent processes (Rudy and Silva, 2006).

The first models of cardiac electrophysiology were based on the Hodgkin-Huxley model of the squid giant axon (Noble and Rudy, 2001). To maintain the plateau of the action potential in an energy-efficient manner, two K⁺ currents were required: a regular channel and a delayed rectifying current (Noble, 1962). Later models included the L-type

Ca^{2+} channel, enabling the investigation of intracellular Ca^{2+} handling as well as the influence of Ca^{2+} concentration on the action potential (McAllister et al., 1975; Beeler and Reuter, 1977). While initial models of cardiac electrophysiology contained fixed ion concentrations, more recent models examined variation in ion concentrations over time. However, to maintain ion concentrations in these models, ion transport processes such as the Na^+/K^+ ATPase and NCX needed to be included. As a result, these dynamic models were far more detailed than the models preceding them (DiFrancesco and Noble, 1985; Luo and Rudy, 1994a). Since then, models of cardiac electrophysiology have been developed in greater detail and for a wide range of animals (Rasmusson et al., 1990; Shaw and Rudy, 1997; Faber and Rudy, 2000; Pandit et al., 2001; Hund and Rudy, 2004; Aslanidi et al., 2009). An increasing availability of data from undiseased human cardiomyocytes has enabled the development of human cardiomyocyte models (ten Tusscher et al., 2004; O'Hara et al., 2011).

In addition to models of cardiac electrophysiology, models of Ca^{2+} release have been used to reproduce Ca^{2+} transients seen in experiments (Hinch et al., 2004; Greenstein et al., 2006). Models of cardiac mechanics attempt to reconstruct measured force-velocity relationships within cardiac tissue, and these can range from phenomenological models (Hill, 1938) to more biophysical models that attempt to describe the transduction of chemical to mechanical energy (Huxley, 1957; Rice et al., 2008; Tran et al., 2010).

Despite the detail in which individual processes within the cardiac cell have been modelled, fully integrated models of cardiac function – which consider electrophysiology, Ca^{2+} signalling and contraction – are comparatively difficult to develop due to both the relative inflexibility of software implementations, and the expert knowledge required to couple and tune individual models. Nonetheless, an integrated model was developed by Rice et al. (2000), which coupled a model of electrophysiology and Ca^{2+} handling (Jafri et al., 1998) with a model of cardiac contraction (Rice et al., 1999). This model was later integrated with a model of oxidative phosphorylation to investigate cardiac bioenergetics (Cortassa et al., 2006). Local control of Ca^{2+} release was incorporated into this model to provide a more realistic description of Ca^{2+} release (Gauthier et al., 2012). A separate integrated model was developed by Crampin and Smith (2006), who used the model to assess the impact of pH on contraction. Tran et al. (2015) later modified this model to incorporate metabolism to investigate regulatory mechanisms associated with energy usage within cardiac cells. In contrast to the model developed by Cortassa et al. (2006), the essential energy-consuming components were constrained by the laws of thermodynamics to provide more realistic behaviour. The task of constructing integrated models of the cardiac cell has also provided modellers with the opportunity to test software-based approaches that aid model integration. Advances in this area include using a standardised description of the models to highlight possible inconsistencies (Terkildsen et al., 2008), developing software that automatically recognises points of connections between models (Neal et al., 2015) and developing the Cardiac Electrophysiology Web Lab, a web-based tool that allows simulation protocols to be applied to a large population of cardiac models to aid in benchmarking models (Cooper et al., 2016).

1.3 Model integration

1.3.1 Progress

Given that the construction of complex systems biology models requires a hierarchical approach supporting the reuse and integration of independently developed models, models should be developed with reuse in mind (Cooling et al., 2016; Terkildsen et al., 2008; Mallavarapu et al., 2009; Hunter and Borg, 2003). Model integration requires clear interfaces for models to communicate with each other. Therefore, biophysical models are favoured over empirical models for this purpose as physical conservation laws can be used to specify interfaces between models. Furthermore, biophysical models encode assumptions and approximations, allowing models to be simplified and coarse-grained after they have been incorporated into more comprehensive models (Smith et al., 2007). As a result, many of the tools developed by the Physiome Project are built upon biophysical models.

There has been significant progress in using software and standards to aid model integration. The Physiome Project aims to facilitate the reuse of existing models by developing standards for communicating computational models of biology that (a) promote the consistent and unambiguous description of models; (b) ensure relevant information on the context of a model and the data used for parameter fitting are provided; and (c) provide a consistent naming convention for components so they can be compared between models (Nickerson et al., 2016). To facilitate the consistent and unambiguous description of biological models, a number of markup languages are being developed, including the Systems Biology Markup Language (SBML) for biochemical networks (Hucka et al., 2003) and CellML for more general biophysical systems (Lloyd et al., 2004; Nickerson et al., 2016; Cooling et al., 2016). Both SBML and CellML provide means for coupling models together, although the specifications for this are still under development (Waltemath et al., 2016; Cooling et al., 2016). While SBML and CellML are languages for describing models, the Simulation Experiment Description Markup Language (SED-ML) describes simulation protocols for SBML and CellML models, separating the description of the model from the results generated from the model (Waltemath et al., 2011). To further promote model reuse, databases such as the BioModels Database and Physiome model repository store biophysical models in a standardised format to make them more accessible (Nickerson et al., 2016).

In addition to the development of standards, software is also aiding efforts to integrate models. Because biological components that are common between models represent interconnections between modules, model integration requires the identification of such shared components. Therefore, ontologies that provide unambiguous and systematic descriptions of biological objects and processes can be used to recognise points of connection and automatically couple models together. In particular, the software SemGen is being developed to automatically couple biophysical models by taking advantage of ontological annotations (Beard et al., 2012; Neal et al., 2015).

1.3.2 Challenges

While the Physiome Project has been effective at addressing software and collaborative issues surrounding model reuse and integration, there has been a relative lack of attention directed towards equally important modelling and physical issues (Bassingthwaite, 2000). While software can be used to speed up the process of model integration, it does not ensure that models remain consistent with experimental observations and with the laws of physics when coupled together (Hunter, 2016). In most cases, model integration remains a time-consuming and error-prone process that requires domain-specific knowledge (Neal et al., 2015; Macklin et al., 2014). In the following paragraphs, I discuss three limitations in our ability to couple models.

Limitation 1: Unclear interfaces

A common ground is required for models to be coupled (Patel and Nagl, 2010). However, no single formalism is suitable for modelling all biological systems, therefore models in systems biology may make use of numerous formalisms including Boolean networks, Bayesian networks, Petri nets, constraint-based models, differential equations, rule-based models, agent-based models and cellular automata (Gonçalves et al., 2013; Machado et al., 2011). As a result, it is challenging to couple models of different formalisms as there may not be any common ground for the models to communicate (Machado et al., 2011).

Limitation 2: Modularity and physical plausibility

To date, the notion of modularity in systems biology has largely been borrowed from engineering (independent subsystems) and software development (computational functions), where modules are seen as black boxes that expose a pre-defined set of inputs and outputs while hiding the rest of their structure. However, this notion of modularity can become troublesome when applied to the domain of physiology and especially biochemistry (Mallavarapu et al., 2009; Neal et al., 2014). Whereas in engineering, it is sensible for modules to have pre-defined external connections because they are physically separated, in biochemistry, all chemical species are present within the same compartment and interactions could theoretically occur between any subset of these species. Therefore potentially important components may be inaccessible when modules are coupled together. Furthermore, because biological knowledge is constantly developing, modellers generally cannot determine all of the potential connections to a model *a priori*, which limits reuse (Neal et al., 2014). Therefore the black box approach is often too inflexible for biophysical models (Mallavarapu et al., 2009). There has been a recent shift in the field from black box modules to “grey box” modules that allow computational structures to be hidden within modules while preserving the ability to expose components as needed during model integration, although software implementations of grey box modelling are still under development (Neal et al., 2014; Mallavarapu et al., 2009).

Another issue that arises when using the traditional engineering notion of modularity is that it assumes that information flow is unidirectional. However, whereas engineering systems are designed with buffers to ensure information flow is unidirectional, in general, information flow between coupled physical systems is bidirectional, a phenomenon known as retroactivity. Thus not all connections between biological systems involve the unidirectional flow of information, and in fact, these connections may be specifically bidirectional in order to be energy efficient (Gawthrop and Crampin, 2016). As a result of blindly

applying the notion modularity from engineering, retroactivity is sometimes ignored in modelling biological systems. In synthetic biology, a failure to address retroactivity has been known to cause issues in biomolecular circuit design where modules fail once coupled together (Del Vecchio, 2013). While the notion of modularity continues to be debated in systems biology, it is clear that models that are coupled without consideration for retroactivity cannot be physically plausible (Del Vecchio et al., 2008; Del Vecchio, 2013).

Within the space of biophysical models, physical conservation laws can serve as a common ground for coupling (Crampin et al., 2004a), and models should remain consistent with the laws of physics once coupled (Macklin et al., 2014). While mass conservation is often used in coupling biophysical models, energy conservation is often ignored because models are not built with thermodynamics in mind. This can result in inconsistencies that only manifest when models are coupled together. As a result, many attempts at coupling biophysical models result in physically implausible models, especially in the biochemical domain (Gawthrop and Crampin, 2014; Macklin et al., 2014). A particularly relevant example is the creation of perpetual cycles when inconsistent biochemical models are coupled together (van der Schaft et al., 2013). Furthermore, without the ability to take advantage of energy as the natural link between biophysical domains, it is often challenging to couple biochemistry to other physical domains, for example the electrical domain in the case of ion pumps (Smith and Crampin, 2004; Glynn, 2002) or the mechanical domain in cross-bridge cycling (Hill, 1974).

Limitation 3: Parameter inheritance and loss of model provenance

One of the goals of the Physiome project is “to construct descriptive and quantitative models that integrate the available knowledge so as to determine the inadequacies, inconsistencies, and insecurities in that knowledge” (Bassingthwaighe, 2000). Continuous feedback between models and experiment is required to develop our understanding of biology (Mallavarapu et al., 2009; Fink et al., 2011), however this is only possible if model provenance – loosely defined in the literature as the links between models, their assumptions and data – is preserved (Daly et al., 2018). Currently, models often lose their provenance when incorporated into other models. There is often an ambiguous link between models and data and models are rarely reviewed when new experimental data become available (Smith et al., 2007). The blind reuse of models can lead to misuse, for example coupling together models from different animal species or simulating a model under conditions that violate assumptions made in developing the model (Smith et al., 2007; Fink et al., 2011).

Parameterisation is widely known to be a bottleneck in the development of large-scale differential equation models (Mallavarapu et al., 2009; Babtie and Stumpf, 2017). This problem is often exacerbated by retroactivity, as sub-models may no longer be consistent with the experimental measurements once coupled together into a larger model. As a result, parameters within sub-models often need to be fine-tuned to resolve inconsistencies between models and to provide physiological behaviour – an opaque and tedious process that often requires domain-specific knowledge (Terkildsen et al., 2008; Babtie and Stumpf, 2017; Kolczyk and Conradi, 2016). As a result, parameters may become disconnected from measurements in the literature, preventing one from checking for consistency with that data when the models are later updated (Terkildsen et al., 2008).

1.4 The case for an energetic framework

Because biological systems must obey the laws of physics, there has been growing interest in using energy conservation laws for biophysical modelling (Soh and Hatzimanikatis, 2010; Ederer and Gilles, 2007; Liebermeister and Klipp, 2006; Mason and Covert, 2019). Thermodynamics has proven to be useful in constraining the space of potential parameters in biochemical models (Soh and Hatzimanikatis, 2010; Tran et al., 2009) and in restricting sets of feasible fluxes in flux balance analysis (Beard et al., 2004a). However, the above approaches make use of thermodynamic constraints in a fundamentally kinetic framework, which allows for phenomenological behaviour, but also the possibility of thermodynamic inconsistencies if the constraints are not consistently applied to the entire model. Another approach is to employ an *energetic* framework in which all components and processes are described in terms of thermodynamic quantities. Such frameworks ensure that models obey the laws of thermodynamics without the need to add additional constraints on parameters and variables. Examples of energetic frameworks include bond graphs and port-Hamiltonians for general physical systems (Borutzky, 2010; van der Schaft and Jeltsema, 2014) and the “energetic” or “thermokinetic” formulation for biochemical systems (Ederer and Gilles, 2007; Liebermeister and Klipp, 2006; Mason and Covert, 2019).

In the context of model integration, an energetic framework makes mechanistic and biophysical information explicit, helping to resolve apparent conflicts between different models. In particular, the whole-cell modelling community has long stressed the importance of using experimentally measurable parameters (Babtie and Stumpf, 2017), and as a result begun to employ energetic frameworks as a means of reconciling inconsistencies between data (Mason and Covert, 2019). An energetic framework could address current challenges to model integration (listed in § 1.3.2) in the following ways:

Limitation 1: Unclear interfaces: Energy is the common currency of all physical systems, and it can be used as the common quantity allowing models to communicate with each other.

Limitation 2: Modularity and physical plausibility: Considering energy conservation when coupling models can help to identify inconsistencies and preclude unphysical behaviour. Furthermore, in many cases, energetic frameworks have notions of inter-connectivity that can be used to explore the issue of retroactivity (van der Schaft and Jeltsema, 2014; Borutzky, 2010; Willems, 2007).

Limitation 3: Parameter inheritance and loss of model provenance: Because an energetic framework requires the use of physical components and parameters, it is easier to map any changes to a model or its parameters to specific biological entities or assumptions.

1.5 Bond graph modelling

The bond graph is an energetic framework that has emerged as a promising tool for developing multi-scale models of biological systems (Gawthrop and Crampin, 2014). Because bond graphs are an energetic framework, they automatically satisfy thermodynamic consistency and carry the advantages of energetic frameworks as discussed above. However, bond graphs differ from many existing energetic frameworks in two ways. Firstly, they are a generic representation (Gawthrop and Bevan, 2007), making them well-suited for multi-physics problems that frequently occur in biology, such as cardiac contraction. Secondly, in contrast to many existing energetic approaches that are purely mathematical, bond graphs express models in a hierarchical graph structure with connections between discrete components that can be labelled, replaced, connected or moved. As a result, bond graphs better support model reuse and integration (Gawthrop et al., 2015). As discussed later (§ 1.5.6), this provides modellers with additional means for addressing the challenges associated with model integration.

1.5.1 Definition

A bond graph consists of a graphical structure and a set of mathematical equations describing the dynamic behaviour of the system. The structure of a bond graph consists of the following:

- A set of components (or nodes) $V = \{v_1, v_2, \dots, v_n\}$.
- Each component v_i has a finite number of ports $\mathcal{P}_i = \{(v_i, l_1), (v_i, l_2), \dots, (v_i, l_{g(i)})\}$, where l_j are the labels corresponding to the ports and $g(i)$ is a function that describes the number of ports that v_i has (which must be at least one). The set of all ports within the system is defined as $\mathcal{P} = \bigcup_{i=1}^n \mathcal{P}_i$. Note that while the ports of each component are uniquely labelled in general, it is frequently the case that the behaviour of each of the ports is identical. In such cases, the ports are usually left unlabelled.
- A set of bonds (or connections) $B = \{(a, b) : a, b \in \mathcal{P}, a \neq b\}$. The bonds must be such that each port is connected to exactly one bond. Bonds are represented using the harpoon \rightarrow , with the port a being at the tail and b being at the head.

Therefore, the graph structure of a bond graph is similar to a port graph (Andrei and Kirchner, 2009), but differs in that the edges are directed and each port must be connected to exactly one bond.

In addition to the structure, a bond graph also defines a set of mathematical relationships (Borutzky, 2010):

- Each bond $b_i \in B$ carries the two conjugate variables e_i and f_i , where e_i is the effort (or potential-like) variable and f_i is the flow (or velocity-like) variable. The product of effort and flow is power, i.e. $P = ef$. The effort and flow variables are shared between the two components the bond connects, with the direction of

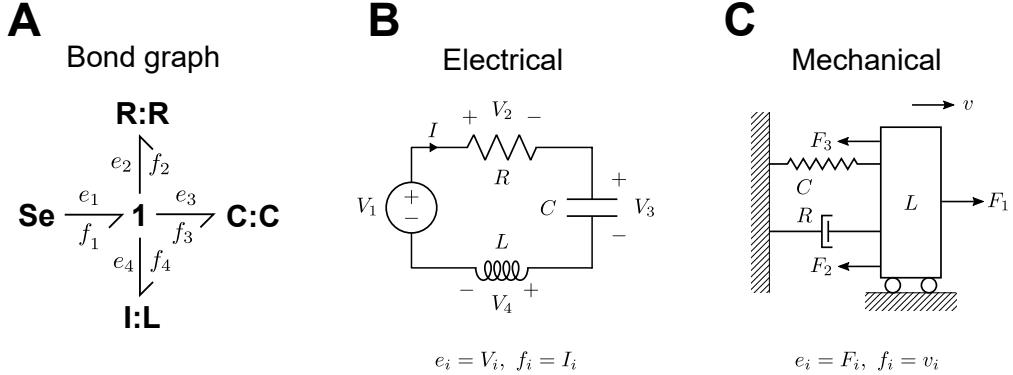


Figure 1.2: A simple bond graph example representing an RLC circuit. (A) Bond graph representation; (B) Electrical circuit representation; (C) Mechanical representation.

the connection specifying the convention of signs; a bond with positive power will release energy at the tail and accumulate or dissipate energy at the head.

- Each component v_i may contain state variables $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,S})$ that record the location of the system. Examples include $q = \int f dt$, the generalised position and $p = \int edt$, the generalised momentum.
- Each component v_i defines a set of constitutive equations that relate the efforts and flows of the connected bonds, i.e. $\Phi_i(\mathbf{e}_i, \mathbf{f}_i, \mathbf{x}_i, \dot{\mathbf{x}}_i) = 0$, where \mathbf{e}_i and \mathbf{f}_i are vectors containing the efforts and flows of the bonds connected to v_i respectively. These constitutive equations may be *atomic* relations that come from fundamental knowledge and assumptions of the system (e.g. Ohm's law) or *composite* relations that are constructed by combining components with atomic relations (Cudmore et al., 2019). The latter provides bond graphs with a notion of modularity, allowing models and equations to be reused.

The constitutive equations of a bond graph can be combined, giving rise to a system of ordinary differential equations, or differential algebraic equations in more complex cases.

1.5.2 Simple examples

In this section, we use some simple examples to illustrate how equations can be derived from bond graphs. In Figure 1.2A we show a bond graph with **R**, **C** and **I** components. Because bond graphs are a generic representation, they can encode a wide range of systems. Therefore, the bond graph in Figure 1.2A could be used to represent an RLC circuit in electrical systems (Figure 1.2B) or a mass-spring-damper system (Figure 1.2C) in mechanics using the analogies that voltages and forces are efforts, and currents and velocities are flows. Furthermore, the series connection in the electrical circuit corresponds to the parallel connection in the mechanical circuit due to the shared flow (current or velocity) between components. This connection is encoded using a **1** junction in the bond graph representation. Therefore, the bond graph structure does not exactly reflect the physical connections within these systems, but how power is transmitted between

components. This abstraction allows the bond graph representation to be both general and compact.

The equations for the dynamics of the system arise from the constitutive equations of the components. The **Se** component constrains the effort of the connected bond to a constant value e^* , thus

$$e_1 = e^* \quad (1.18)$$

The **R**, **C** and **I** components encode the constitutive equations

$$\mathbf{R} : e_2 = Rf_2 \quad (1.19)$$

$$\mathbf{C} : q_3 = Ce_3, \dot{q}_3 = f_3 \quad (1.20)$$

$$\mathbf{I} : p_4 = Lf_4, \dot{p}_4 = e_4 \quad (1.21)$$

where R , C and L are the resistance, capacitance (compliance) and inductance (mass) respectively. Finally, the **1** junction equates the flows of all connected bonds. In order to conserve power, the efforts of those bonds must sum to zero (after accounting for the direction of the bonds), similar to Kirchhoff's voltage law in electrical circuits. Thus

$$f_1 = f_2 = f_3 = f_4 \quad (1.22)$$

$$e_1 - e_2 - e_3 - e_4 = 0 \quad (1.23)$$

Then by substituting Eqs. 1.18–1.21 into Eq. 1.23, we find that

$$e^* - Rf_2 - \frac{1}{C}q_3 - L\dot{f}_4 = 0 \quad (1.24)$$

After substituting Eq. 1.22 and differentiating the above equation, we derive a second-order differential equation for the current through the circuit:

$$L\ddot{f}_1 + R\dot{f}_1 + \frac{1}{C}f_1 = 0 \quad (1.25)$$

The initial conditions of this equation can be expressed in terms of $q_3(0)$ and $p_4(0)$, the initial states of the **C** and **I** components:

$$f_1(0) = f_4(0) = \frac{1}{L}p_4(0) \quad (1.26)$$

$$f'_1(0) = f'_4(0) = \frac{1}{L}e_4(0) = \frac{1}{L}(e_1(0) - e_2(0) - e_3(0)) = \frac{1}{L}\left[e^* - \frac{R}{L}p_4(0) - \frac{1}{C}q_3(0)\right] \quad (1.27)$$

The bond graph representation also provides a framework for developing models in a modular manner. For example, the passive part of the electrical circuit in the above example (containing the **R**, **C** and **I** components with the **1** junction) can be encapsulated in a composite bond graph component that we name **RLC** (Figure 1.3A, left panel). Note that the **Se** component has been replaced with an **SS** port, indicating that the circuit is connected to the external environment, and that the **RLC** module has one port available for connection. The equivalent electrical circuit with an open port is shown in the right panel of Figure 1.3A. Therefore, the circuit in Figure 1.2 can

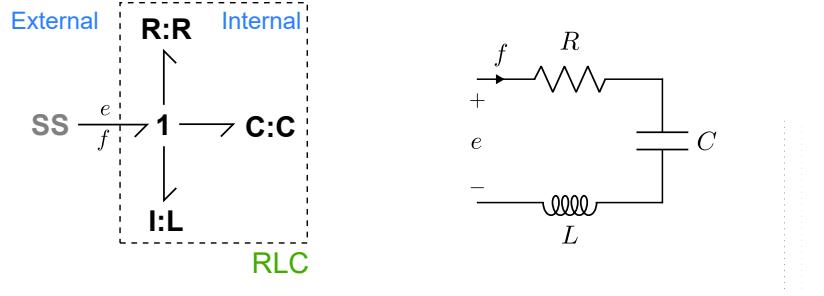
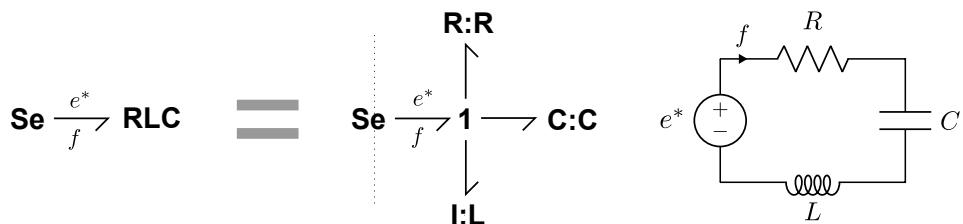
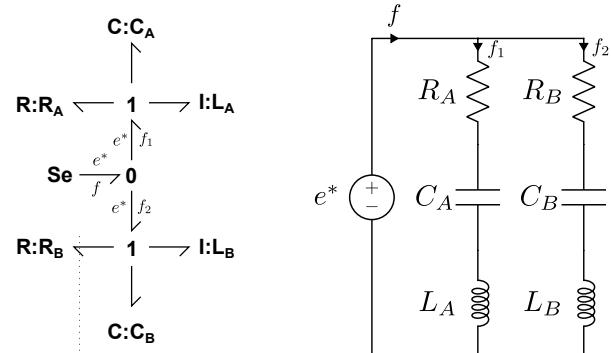
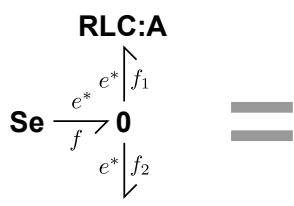
A**Modular representation****B****RLC circuit****C****Coupled circuits**

Figure 1.3: Bond graph modules. (A) The passive part of the electrical circuit in Figure 1.2 can be represented as an open system. The bond graph representation on the left replaces the voltage source with an **SS** port representing an external connection, resulting in a new module that is named **RLC**. The electrical circuit representation with an exposed port is shown on the right. (B) The circuit in Figure 1.2 can be formed by connecting the **RLC** module to an effort (voltage) source. (C) The modularity of bond graphs allows two RLC loops to be coupled in parallel. For (B) and (C), the compact bond graph representation is on the left, the full bond graph representation is in the middle and the equivalent electrical circuit is on the right.

be reproduced by connecting an **Se** component to the **RLC** module (Figure 1.3B). An added advantage of encapsulating parts of a model inside modules is that these modules can be reused. For instance, two **RLC** modules can be connected in parallel to a voltage source (Figure 1.3C). In this case, the parallel connection is encoded using a **0** junction that equates efforts (voltages) and ensures the conservation of flow (current), similar to Kirchhoff's current law in electrical systems:

$$e^* = e_1 = e_2 \quad (1.28)$$

$$f - f_1 - f_2 = 0 \quad (1.29)$$

Then, by using Eq. 1.25 and assuming that the components in each of the **RLC** modules have separate parameters, the current f through the voltage source can be described using the equations

$$f = f_1 + f_2 \quad (1.30)$$

$$L_A \ddot{f}_1 + R_A \dot{f}_1 + \frac{1}{C_A} f_1 = 0 \quad (1.31)$$

$$L_B \ddot{f}_2 + R_B \dot{f}_2 + \frac{1}{C_B} f_2 = 0 \quad (1.32)$$

with initial conditions

$$f_1(0) = \frac{1}{L_A} p_{L_A}(0) \quad (1.33)$$

$$f'_1(0) = \frac{1}{L_A} \left[e^* - \frac{R_A}{L_A} p_{L_A}(0) - \frac{1}{C_A} q_{C_A}(0) \right] \quad (1.34)$$

$$f_2(0) = \frac{1}{L_B} p_{L_B}(0) \quad (1.35)$$

$$f'_2(0) = \frac{1}{L_B} \left[e^* - \frac{R_B}{L_B} p_{L_B}(0) - \frac{1}{C_B} q_{C_B}(0) \right] \quad (1.36)$$

where $p_{L_m}(0)$ is the initial state of the inductor in module m and $q_{C_m}(0)$ is the initial charge of the capacitor in module m .

A summary of commonly used components in bond graph modelling is given in Table 1.1. **TF** components, which were not introduced in the above example, but will be used for describing the stoichiometry of reactions, are transformers that scale effort and flow while maintaining power conservation. For a more detailed description of bond graph theory, the reader is referred to Gawthrop and Smith (1996), Borutzky (2010), and Gawthrop and Bevan (2007).

1.5.3 Properties of bond graphs

As seen in the above examples, bond graphs have the following properties:

- They are a *physics-based* framework in which the connections between components are encoded using mathematical relationships that respect the conservation laws of physics.
- They are *energy-based* because each bond transmits power (the product of effort and flow). This allows the transduction of power to be explicitly tracked and also prevents one from defining models that violate conservation of energy.

Table 1.1: Bond graph components and mathematical relationships. Summarised from Borutzky (2010) and Gawthrop and Crampin (2014). Note that the functions ϕ_C , ϕ_R and ϕ_I may be nonlinear. While three bonds are attached to the **0** and **1** junctions in this figure, they may link an arbitrary number of bonds.

Component	Equations
R	$e = \phi_R(f)$
C	$e = \phi_C(q)$ $f = \dot{q}$
I	$f = \phi_I(p)$ $e = \dot{p}$
$\frac{A^f}{v} \nearrow$ Re $\frac{A^r}{v} \nearrow$	$v = \kappa \left[e^{A^f/(RT)} - e^{A^r/(RT)} \right]$
$\frac{e_1}{f_1} \nearrow$ TF $\frac{e_2}{f_2} \nearrow$	$e_2 = ne_1$ $f_1 = nf_2$
$\frac{e_1}{f_1} \nearrow$ 0 $\frac{e_2}{f_2} \nearrow$	$e_1 = e_2 = e_3$ $f_1 + f_2 + f_3 = 0$
$\frac{e_1}{f_1} \nearrow$ 1 $\frac{e_2}{f_2} \nearrow$	$f_1 = f_2 = f_3$ $e_1 + e_2 + e_3 = 0$

Table 1.2: Summary of effort and flow variables. Summarised from Gawthrop and Bevan (2007); Borutzky (2010).

Bond graph	Electrical	Mechanical	Hydraulic	Chemical
Effort (e)	Voltage V [V]	Force F [N]	Pressure P [Pa]	Chemical potential μ [J/mol]
Flow (f)	Current I [A]	Velocity v [m/s]	Volume flow Q [m ³ /s]	Molar flow rate v [mol/s]
C	Capacitor	Spring	Tank	Chemical species
R	Resistor	Damper	Pipe	Reaction*
I	Inductor	Inertia	Flow inertia	N/A

*Note that unlike electrical resistors or mechanical dampers in which flows are related to a potential difference, i.e. $f = (e_1 - e_2)/R = \Delta e/R$, the flow through a reaction is a function of two potentials $f = \phi(e_1, e_2)$ that cannot be reduced to a function of the potential difference Δe . To account for this, reactions are represented as two-port resistors rather than the typical one-port resistor (see § 1.5.5). However, the analogy still holds as reactions dissipate energy.

- They are a *general-purpose* representation that is not restricted to a single physical domain. Instead, they make use of analogies between domains to abstract the representation of physical connections. A summary of analogies between physical domains is given in Table 1.2. A generic representation provides two advantages. Firstly, they allow intuitions and analyses from other physical domains to be applied to biological systems (Gawthrop and Crampin, 2014). Secondly, cross-domain systems can be modelled, for example, electromechanical systems in engineering or electrogenic transport in electrophysiology.
- They are *modular* as collections of components can be encapsulated within modules with a set of ports exposed for use within an external model.
- They are *declarative* as each component within a bond graph encodes constraints between physical quantities without explicitly defining inputs and outputs. This allows bond graphs to be more flexible than the block diagrams that are commonly used in control theory, as equations can be rearranged depending on context (Gawthrop and Bevan, 2007).

1.5.4 Relationships with other modelling frameworks

Because bond graphs make use of the conjugate variables effort and flow, they bear the most resemblance to other frameworks that make use of conjugate variables. The closest framework to the bond graph is the port-Hamiltonian, as both make use of effort and flow variables and power-conserving interconnections that share variables between components (van der Schaft and Jeltsema, 2014). However, rather than being a graphical representation, port-Hamiltonians are a mathematical and geometric representation. While the two approaches are equivalent for finite-dimensional systems, port-Hamiltonians are able to represent a wider range of systems such as spatially distributed systems. However, the graphical structure of bond graphs provides a more natural framework for labelling components which may prove to be useful in the semantic merging of models.

Another physics-based approach that makes use of conjugate variables is the behavioural approach (Willems, 2007). Similar to bond graphs, the behavioural approach uses physical interconnections that share variables between components. However, the behavioural approach describes connections between *terminals*, in contrast to bond graphs which specify connections between *ports* (which are pairs of terminals in which the flows into the terminals are equal and opposite). As a result, the bond graph approach is more mathematically compact, but sacrifices some flexibility to achieve this; in some cases, it is not clear what appropriate ports for a system are. Nonetheless, this constraint on flexibility does not appear to impact on the ability of bond graphs to represent biological systems. Another distinction between the two approaches is that the behavioural approach does not require interconnections to conserve energy.

System dynamics modelling is another approach that makes use of pairs of variables – stocks and flows (Sterman, 2001). However, in contrast to bond graphs, systems dynamics is neither physics-based nor energy-based. The stock variable, which corresponds most closely to the generalised position q in bond graph modelling, is not related to an effort variable. As a result, while systems dynamics models are suited for modelling complex systems such as population-scale behaviours, they are not appropriate for modelling the constraints within physical systems specifically.

1.5.5 Modelling of biochemical systems

Bond graphs were first applied to biochemical systems by Oster et al. (1973). In this representation, energy is stored within each species. The concentration of each species gives rise to a chemical potential μ [J/mol], which is an effort (Table 1.2). The chemical potentials drive reactions that cause molar flow rates v [mol/s], which are flows (Table 1.2). Since chemical species are energy stores, they are represented by nonlinear **C** components with the constitutive relationship

$$\mu = RT \ln(Kx) \quad (1.37)$$

where $R = 8.314\text{J/mol/K}$ is the universal gas constant, T [K] is the absolute temperature of the system, x [mol] is the molar amount and K [mol^{-1}] is the thermodynamic constant of the chemical species. Reactions are represented by two-port dissipative **Re** components (see Table 1.1) with the constitutive relationship

$$v = \kappa \left[e^{A^f/(RT)} - e^{A^r/(RT)} \right] \quad (1.38)$$

where κ [mol/s] is a reaction-specific constant describing its speed; A^f [J/mol] is the sum of chemical potentials in the reactants, or the forward affinity; and A^r [J/mol] is the sum of chemical potentials in the products, or the reverse affinity. Because reactions will only proceed in the direction of decreasing potential, the affinity of reaction $A = A^f - A^r$ determines the direction of the flux v , which will be positive only if $A > 0$. In the literature, the Gibbs free energy of reaction ΔG is commonly used to quantify the direction of reaction; this is related to the affinity by the relationship $\Delta G = -A$. Combining Eqs. 1.37 and 1.38 results in the mass action relationship familiar to biochemists (Gawthrop and Crampin, 2014).

Bond graph representations of the chemical reaction networks in Eq. 1.6 and Eq. 1.8 are given in Figure 1.4. The bond graph representation for Eq. 1.6 is displayed on the

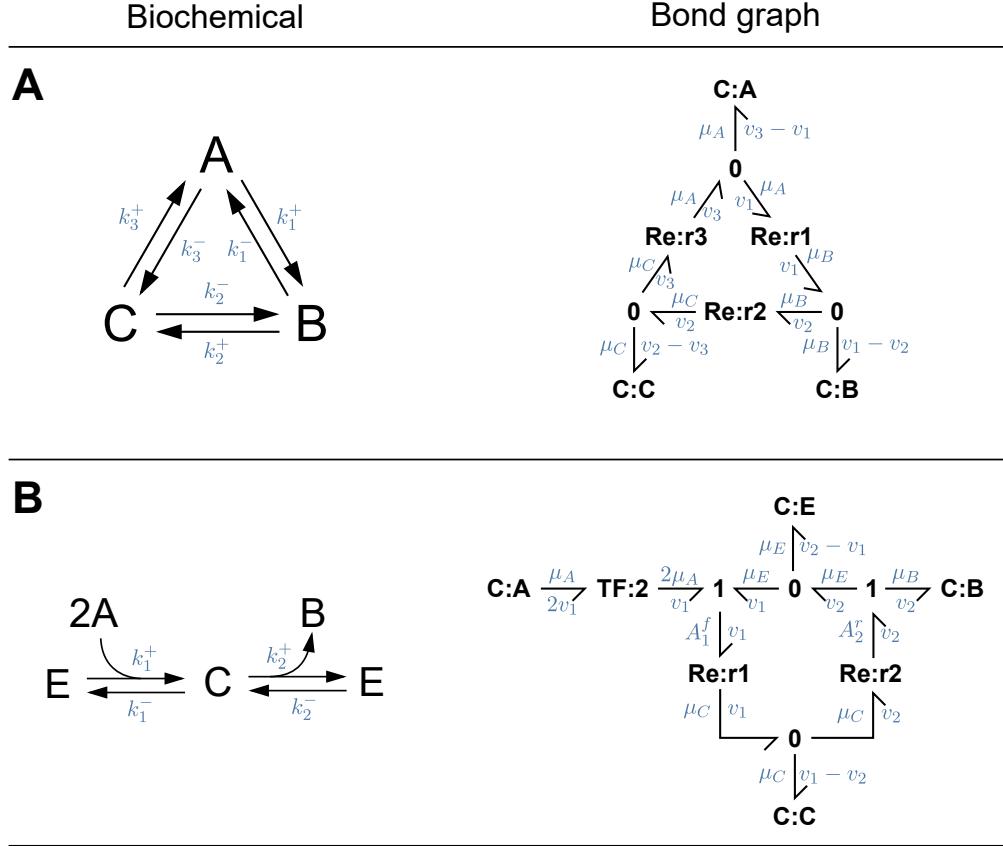


Figure 1.4: Bond graph representations of biochemical systems. Examples are provided for the reaction schemes described in (A) Eq. 1.6 and (B) Eq. 1.8. Biochemical systems are shown on the left panels and bond graph representations are shown on the right panels.

right panel of Figure 1.4A. Note that the standard bond graph notation for components takes on the form *<component type>:<name>*; for example, **C:A** is the **C** component corresponding to the species *A* and **Re:r1** is the **Re** component corresponding to reaction 1. Thus each species has been mapped to a **C** component and each reaction to an **Re** component. The chemical potentials for each species are

$$\mu_A = RT \ln(K_A x_A) \quad (1.39a)$$

$$\mu_B = RT \ln(K_B x_B) \quad (1.39b)$$

$$\mu_C = RT \ln(K_C x_C) \quad (1.39c)$$

and the reaction rates are given by the equations

$$v_1 = \kappa_1 \left(e^{\mu_A/RT} - e^{\mu_B/RT} \right) = \kappa_1 K_A x_A - \kappa_1 K_B x_B \quad (1.40a)$$

$$v_2 = \kappa_2 \left(e^{\mu_B/RT} - e^{\mu_C/RT} \right) = \kappa_2 K_B x_B - \kappa_2 K_C x_C \quad (1.40b)$$

$$v_3 = \kappa_3 \left(e^{\mu_C/RT} - e^{\mu_A/RT} \right) = \kappa_3 K_C x_C - \kappa_3 K_A x_A \quad (1.40c)$$

where the second equalities were derived by substituting Eq. 1.39. The **0** junctions enforce mass conservation, therefore the rates of change of each species are

$$\frac{dx_A}{dt} = v_3 - v_1 \quad (1.41a)$$

$$\frac{dx_B}{dt} = v_1 - v_2 \quad (1.41b)$$

$$\frac{dx_C}{dt} = v_2 - v_3 \quad (1.41c)$$

Therefore Eq. 1.41 gives rise to the same mass-action dynamics as described by the kinetic scheme in Eqs. 1.5,1.7, with

$$k_1^+ = \kappa_1 K_A \quad (1.42a)$$

$$k_1^- = \kappa_1 K_B \quad (1.42b)$$

$$k_2^+ = \kappa_2 K_B \quad (1.42c)$$

$$k_2^- = \kappa_2 K_C \quad (1.42d)$$

$$k_3^+ = \kappa_3 K_C \quad (1.42e)$$

$$k_3^- = \kappa_3 K_A \quad (1.42f)$$

An important distinction, however, is that whereas it is possible to describe kinetic parameters that give rise to thermodynamically infeasible systems, it is impossible do so within the bond graph formalism. Noting that

$$\frac{k_1^+ k_2^+ k_3^+}{k_1^- k_2^- k_3^-} = \frac{\kappa_1 K_A \kappa_2 K_B \kappa_3 K_C}{\kappa_1 K_B \kappa_2 K_C \kappa_3 K_A} = 1 \quad (1.43)$$

all sets of bond graph parameters are by definition thermodynamically feasible as they satisfy the detailed balance constraint in Eq. 1.11. Thus, in practice, whereas modellers using kinetic parameters must adhere to constraints such as Eq. 1.11 to satisfy thermodynamic consistency, these constraints are automatically satisfied if bond graph parameters are used.

The bond graph representation of the reaction scheme in Eq. 1.8 is shown on the right panel of Figure 1.4B. The chemical potentials of this system are

$$\mu_A = RT \ln(K_A x_A) \quad (1.44a)$$

$$\mu_B = RT \ln(K_B x_B) \quad (1.44b)$$

$$\mu_E = RT \ln(K_E x_E) \quad (1.44c)$$

$$\mu_C = RT \ln(K_C x_C) \quad (1.44d)$$

and the reaction rates are

$$v_1 = \kappa_1 \left(e^{A_1^f/RT} - e^{\mu_C/RT} \right) \quad (1.45a)$$

$$v_2 = \kappa_2 \left(e^{\mu_C/RT} - e^{A_2^r/RT} \right) \quad (1.45b)$$

In contrast to the previous example, the presence of multi-substrate reactions requires chemical potentials on the forward and reverse sides of each reaction to be summed. The **1** junctions ensure that this is done in a manner which conserves total potential:

$$A_1^f = \mu_E + 2\mu_A \quad (1.46a)$$

$$A_2^r = \mu_E + \mu_B \quad (1.46b)$$

Note that the double stoichiometry associated with A is described using the **TF** component, which encodes the constraint required for twice its chemical potential to be contributed to A_1^f . By substituting into Eq. 1.45, the reaction rates are therefore

$$v_1 = \kappa_1 K_E K_A^2 x_E x_A^2 - \kappa_1 K_C x_C \quad (1.47a)$$

$$v_2 = \kappa_2 K_C x_C - \kappa_2 K_E K_B x_E x_B \quad (1.47b)$$

Hence the rates of change for each species are described by the equations

$$\frac{dx_A}{dt} = -2v_1 \quad (1.48a)$$

$$\frac{dx_B}{dt} = v_2 \quad (1.48b)$$

$$\frac{dx_E}{dt} = v_2 - v_1 \quad (1.48c)$$

$$\frac{dx_C}{dt} = v_1 - v_2 \quad (1.48d)$$

Note that in addition to the mass conservation constraints enforced by the **0** junctions, the **TF** components encode a further mass conservation constraint for A , where the reaction contributes double stoichiometry. Eq. 1.48 is equivalent to the kinetic representation in Eq. 1.9 with

$$k_1^+ = \kappa_1 K_E K_A^2 \quad (1.49a)$$

$$k_1^- = \kappa_1 K_C \quad (1.49b)$$

$$k_2^+ = \kappa_2 K_C \quad (1.49c)$$

$$k_2^- = \kappa_2 K_E K_B \quad (1.49d)$$

Gawthrop et al. (2015) show that a general biochemical network with forward stoichiometric matrix N^f , reverse stoichiometric matrix N^r and stoichiometric matrix $N = N^r - N^f$ can be represented using the equations

$$\frac{d\mathbf{X}}{dt} = NV(\mathbf{X}) \quad (1.50a)$$

$$V(\mathbf{X}) = \boldsymbol{\kappa} \left(\text{Exp}[N^f \mathbf{Ln}(\mathbf{KX})] - \text{Exp}[N^r \mathbf{Ln}(\mathbf{KX})] \right) \quad (1.50b)$$

where **Exp** and **Ln** are the element-wise exponential and logarithm respectively, $\boldsymbol{\kappa}$ is a diagonal matrix containing reaction rate parameters κ and \mathbf{K} is a diagonal matrix containing species parameters K .

1.5.6 Advantages

In addition to being an energetic framework, bond graphs provide extra benefits for model integration in systems biology because of their graphical and hierarchical nature. Below we discuss how bond graphs could potentially address current limitations in our ability to integrate models:

Limitation 1: Unclear interfaces: Bond graphs are fundamentally made up of components and connections, therefore connections between modules have an explicit and

unambiguous interpretation: effort is shared and flow is transferred. Thus, bond graph components can be assembled into modules that are subsequently connected to form larger models, in contrast to mathematical equations that often require reformulation.

Limitation 2: Modularity and physical plausibility: Because bond graphs contain physical components, inconsistencies between models are revealed once they are coupled. Bond graphs also naturally incorporate retroactivity because the causal relations between effort and flow are bi-directional, ensuring that connections are physically plausible.

Limitation 3: Parameter inheritance and loss of model provenance: Due to their physical nature, each component in a bond graph can be mapped to a specific biological entity such as a species, reaction or cellular structure. Therefore, if a sub-model must be modified when incorporating it into a larger model, this is achieved by either changing the bond graph structure, or the equations of certain components. Because components are transparently related to specific assumptions and approximations, changes to these components are easier to describe and record.

1.5.7 Assumptions and limitations

The theory described in § 1.5.5 assumes that the system is at constant temperature and pressure so that Gibbs free energy determines the chemical potential of each species s via the relationship $\mu_s = \partial G / \partial x_s$. It is also assumed that the species exist within an ideal solution so that the Gibbs free energy may be expressed as a sum of independent contributions from each species (Beard and Qian, 2008):

$$G = \sum_s RT[x_s \ln(K_s x_s) - x_s] \quad (1.51)$$

which gives rise to Eq. 1.37 as seen previously. Being a tool for network-based modelling, bond graphs are well equipped to handle this type of system. However, in general, G is a general nonlinear function of every chemical species within a mixture, i.e. $G = G(x_1, x_2, \dots, x_n)$ (Zielinski et al., 2017). Therefore, in non-ideal biochemical systems, the chemical potential of a species μ_s is a function of all species within the solution as well as the solvent. This presents a problem for the bond graph framework as the system is no longer able to be represented as a network where species are modelled independently. Therefore, bond graphs are not an ideal framework for modelling crowded biochemical systems, and the coupling of such systems would need to be treated in a different manner, although a thermodynamic approach is still feasible.

An additional issue that arises when coupling models together is that even in ideal solutions, the species constant K depends on both the solvent and other species within the solution (Beard and Qian, 2008). This may cause inaccuracies when attempting to couple models developed using data from different cell lines and experimental conditions, and numerical techniques may need to be developed in order to adjust for this. Despite these limitations, we believe that bond graph representations are good first-order approximations of biochemical systems, and models can be switched out for more accurate ones if it is found that crowding effects are significant.

1.5.8 Applications

Since bond graphs were first applied to biochemistry, they have been used to model enzyme kinetics (Oster et al., 1973; Gawthrop and Crampin, 2014) and phosphorylation/dephosphorylation cycles (Gawthrop and Crampin, 2014, 2016). The task of representing existing models as bond graphs can help to correct errors and inconsistencies within models of these systems. For example, in Gawthrop et al. (2015), the process of developing a bond graph model of glycolysis led to the discovery of a thermodynamic inconsistency in an existing model that was subsequently corrected. Because bond graphs are general-purpose, they allow multi-domain processes to be coupled together, and bond graph models have been developed for the action potential (Gawthrop et al., 2017), proton pumps in oxidative phosphorylation (Gawthrop, 2017a), blood vessels (de Bono et al., 2018) and the left ventricle of the heart (Díaz-Zuccarini et al., 2009).

Gawthrop et al. (2015) show that bond graph models of biochemical systems can be coupled together in a modular fashion by exposing connections as external ports. This idea of modularity has been used to couple models of individual enzyme complexes into a model of oxidative phosphorylation (Gawthrop, 2017a). Because retroactivity has implications for the robustness of biological systems, Gawthrop and Crampin (2016) proposed an analysis of retroactivity based on linearisation and transfer functions, and postulated that, as in electrical systems with buffer amplifiers, additional energy consumption is required to isolate one module from the effects of another.

1.6 Conclusions

In this chapter, I reviewed techniques for reusing and integrating mathematical models of biological systems. Biophysical models, and energetic models in particular, are well-suited for developing models that can be later incorporated into multi-scale models because they explicitly encode assumptions and approximations that can be reconciled with each other. Within the space of energetic modelling frameworks, bond graphs are particularly useful for model integration because they provide an additional layer of information that separates biological entities into distinct physical components. Furthermore, they make biological modelling scalable with their hierarchical design. While bond graphs do not remove the duty of the modeller in ensuring that their models are physiologically realistic, they flag physical issues that arise during the modelling process. Thus bond graphs provide a means of integrating biology, physics and experimental data in a consistent manner.

This chapter outlined some physics-related issues that arise when coupling together existing mathematical models of biological systems. Bond graphs are a promising framework for resolving many of these issues, and models of physiological systems have been developed using this framework. However, bond graphs have not been applied to larger-scale systems where independently developed models are coupled together. Therefore the focus of this thesis is to develop methods based on the bond graph methodology to facilitate model reuse and integration in biological systems. In the

following chapters, I make three contributions to the reuse and integration of systems biology models:

- In Chapter 2, I develop methods for modelling membrane transporters using bond graphs. Despite the thermodynamics of membrane transport being a well-known constraint on these systems, many existing kinetic models do not adhere to these constraints. Accordingly, I use some hypothetical models to illustrate these thermodynamic issues, and I apply bond graphs to two existing biophysical models relevant to cardiac physiology: SERCA and the Na^+/K^+ ATPase. Bond graphs reveal physical and thermodynamic inconsistencies within the model of Na^+/K^+ ATPase. Therefore, to correct these inconsistencies, I update the model's equations and refit the parameters to data.
- To generate large-scale bond graph models, separate models must be connected together. Thus in Chapter 3, I use the hierarchical nature of bond graphs to develop a model of cardiac electrophysiology by coupling together separately developed models of ion channels and ion transporters. I show that this approach enables an energetic analysis of the cardiac action potential, and also provides a means of systematically resolving the issues of drift and non-unique steady states that occur in many existing models of cardiac electrophysiology.
- In dynamic models of large-scale systems, model simplification and approximation are important in helping models remain computationally tractable. In Chapter 4, I develop a general method for simplifying cycles of enzyme kinetics and apply this to the development of new bond graph models. Because this method is programmable, it enables model selection, so that simple and thermodynamically consistent models can be rapidly developed with little prior knowledge of these systems.

The methods developed in this thesis are primarily applied to systems within cardiac cells due to the relative maturity of cardiac modelling. However, it is anticipated that these methods could be generalised for a wide range of biological systems. To focus on thermodynamic issues in reasonably simple biophysical models, this thesis deals exclusively with ordinary differential equation (ODE) models.

Chapter 2

A thermodynamic framework for modelling membrane transporters

To couple together existing models of biological systems using bond graphs, existing differential equation models must first be represented as bond graphs. In this chapter, we explore the issue of conversion from kinetic models to bond graphs in the context of membrane transporters. While the laws of thermodynamics are known to constrain the behaviour of these systems, they are often neglected in the development of mathematical models. Accordingly, we argue for the use of bond graphs in describing these transport processes because they naturally satisfy fundamental thermodynamic constraints. We then describe a methodology for converting existing kinetic models into bond graph models and use this approach to represent two existing models of transporters as bond graphs. In cases where the existing model is thermodynamically consistent, a direct conversion into bond graphs is possible, and this is illustrated by converting a model of cardiac SERCA into a bond graph. However, thermodynamically inconsistent models cannot be represented as bond graphs without modification. Despite this, bond graphs provide a useful framework for detecting and resolving physical and thermodynamic inconsistencies. We demonstrate this by updating a thermodynamically inconsistent model of cardiac Na^+/K^+ ATPase to be thermodynamically consistent, a process that requires equations to be corrected and re-fitted to the data. While the methods outlined in this chapter are applied only to membrane transporters, they are more broadly applicable to all dynamic models of biochemical systems that use mass-action kinetic parameters.

This chapter has been published in the *Journal of Theoretical Biology* (Pan et al., 2019). References to Pan et al. (2018a) relate to the material in Chapter 3. The Supplementary Material is included within Appendices 2.A–2.E of this chapter.

2.1 Introduction

The cell membrane is a physical barrier that separates the internal environment of a cell from its external environment, as well as many compartments within the cell. By

transporting chemical species across the membrane, the cell regulates concentration within each compartment, providing the environment necessary for many cellular processes (Keener and Sneyd, 2009). A key contributor to these transport systems are membrane transporters, which are proteins located in cell membranes that maintain cell volume (Glitsch, 2001), set up ionic gradients required for electrical (Glitsch, 2001) and calcium signalling (Periasamy and Kalyanasundaram, 2007; Blaustein and Lederer, 1999), and transport sources of energy into the cell (Mueckler and Thorens, 2013). As with all physical processes, transporters must comply with the principles of thermodynamics (Oster et al., 1973; Hwang, 2004; Beard and Qian, 2008). Because each chemical species within a solution is associated with a chemical potential that increases with concentration, passive transporters can only move substrates from a region of high concentration to a region of low concentration (Keener and Sneyd, 2009; Beard and Qian, 2008). To move a substrate against a concentration gradient, a source of energy must be provided. For example, active transporters use ATP hydrolysis to drive the transport of substrates against the chemical potential associated with the concentration gradient (or the electrochemical gradient in the case of electrogenic ion transport) (Keener and Sneyd, 2009).

Mathematical models of membrane transporters have been developed for the purpose of understanding their transport mechanism, and predicting their behaviour beyond experimental measurements. However, despite the wealth of models available for transporters, thermodynamic consistency has not usually been applied. As a result, many of these models describe physically infeasible systems where for example, species are transported against their potential gradients in the absence of an energy source, therefore generating energy out of nowhere. In the current literature, thermodynamic consistency is commonly violated through the use of equations that describe irreversible transporters, neglect dependence on certain metabolites, or have incorrect equilibrium points (Gawthrop and Crampin, 2014; Smith and Crampin, 2004; Tran et al., 2009). While there are methods for incorporating thermodynamic constraints such as detailed balance (Liebermeister et al., 2010; Beard and Qian, 2008; Smith and Crampin, 2004) and the Nernst potential (Keener and Sneyd, 2009), they tend to be scattered around the literature, and are not universally applied. Furthermore, many transporters are electrogenic, therefore there is an interaction between chemical and electrical potential (Smith and Crampin, 2004; Terkildsen et al., 2007), and the multidomain nature of these transporters can confound efforts to develop thermodynamically consistent models. In some cases, thermodynamic inconsistency may impact on the ability of a model to remain physiological under a wide range of conditions. For example, in the context of heart failure, where ATP is depleted, active transporters such as the Na^+/K^+ ATPase and sarcoplasmic/endoplasmic Ca^{2+} ATPase (SERCA) operate at reduced rates. Whole-cell models of cardiac cells generally do not describe this metabolite dependence because they neglect thermodynamic constraints (Smith and Crampin, 2004), and while thermodynamically consistent models capture these effects (Terkildsen et al., 2007; Tran et al., 2009), they tend to be the exception rather than the rule.

To facilitate the incorporation of thermodynamic constraints into models of membrane transporters, we require a framework that is (a) thermodynamically consistent by design; (b) able to reproduce known thermodynamic constraints; and (c) general enough to model a wide range of transporters. We propose the use of bond graphs, which are a physics-based framework in which energy transfer is explicitly modelled through connections between

physical components, therefore satisfying thermodynamic consistency (Gawthrop and Smith, 1996; Borutzky, 2010). The bond graph representation is also domain-independent, therefore it is general enough to represent a wide range of physical systems. Bond graphs were originally invented by Henry Paynter for use in hydroelectric systems (Paynter, 1961), but they also naturally represent electrical and mechanical systems (Borutzky, 2010). The reader is referred to the texts by Gawthrop and Smith (1996), Borutzky (2010) and Gawthrop and Bevan (2007) for a comprehensive introduction to bond graph theory. More recently, bond graphs have been extended to chemical (Thoma and Bouamama, 2000), biochemical (Oster et al., 1973; Gawthrop and Crampin, 2014) and electrochemical systems (Gawthrop, 2017a), enabling bond graph modelling of membrane transporters such as the sodium-glucose transport protein 1 (SGLT1) (Gawthrop and Crampin, 2017). Models of electrogenic membrane transporters have recently been incorporated into a bond graph model of cardiac electrophysiology (Pan et al., 2018a). A further advantage of modelling transporters as bond graphs is that they are modular, therefore individual models of transporters are easily coupled to other bond graph models to build comprehensive models of biological systems (Gawthrop et al., 2015).

In this chapter, we review the bond graph theory required for modelling membrane transporters and use bond graphs to build simple hypothetical models of transporters (Section 2.2), demonstrating that bond graphs capture important thermodynamic concepts such as detailed balance, free energy of reaction and the Nernst potential. We then develop a bond graph model of cardiac SERCA based on existing work by Tran et al. (2009), and use the bond graph model to assess energy consumption and efficiency (§ 2.3.1). Bond graphs also provide a framework for detecting thermodynamic inconsistencies within existing models. In § 2.3.2 we develop a new model of Na^+/K^+ ATPase based on existing work by Terkildsen et al. (2007), and verify that the model complies with thermodynamic constraints. These examples illustrate that bond graphs are a unifying framework for accounting for thermodynamic constraints in models of membrane transporters. We believe that the bond graph approach will prove to be a powerful tool in the development of thermodynamically consistent models of transporters, and other cellular processes in which energy transduction plays an important role.

2.2 Hypothetical models

2.2.1 Enzyme cycle

In biochemical systems, Gibbs free energy (also called Gibbs energy or free energy) is transmitted between species and reactions (Atkins and De Paula, 2006). Like all physical systems, biochemical systems must comply with the laws of thermodynamics, therefore reactions can only proceed in the direction of decreasing chemical potential. However, chemical potential is ignored by many existing models, often resulting in physically unrealistic behaviour (Gawthrop and Crampin, 2014; Gawthrop et al., 2015). Bond graphs account for chemical potential by decomposing the transfer of power into a product of chemical potential (μ [J/mol]) and molar flow rate (v [mol/s]) (Gawthrop and Crampin, 2014). Power is transmitted between components within a bond graph by connecting them

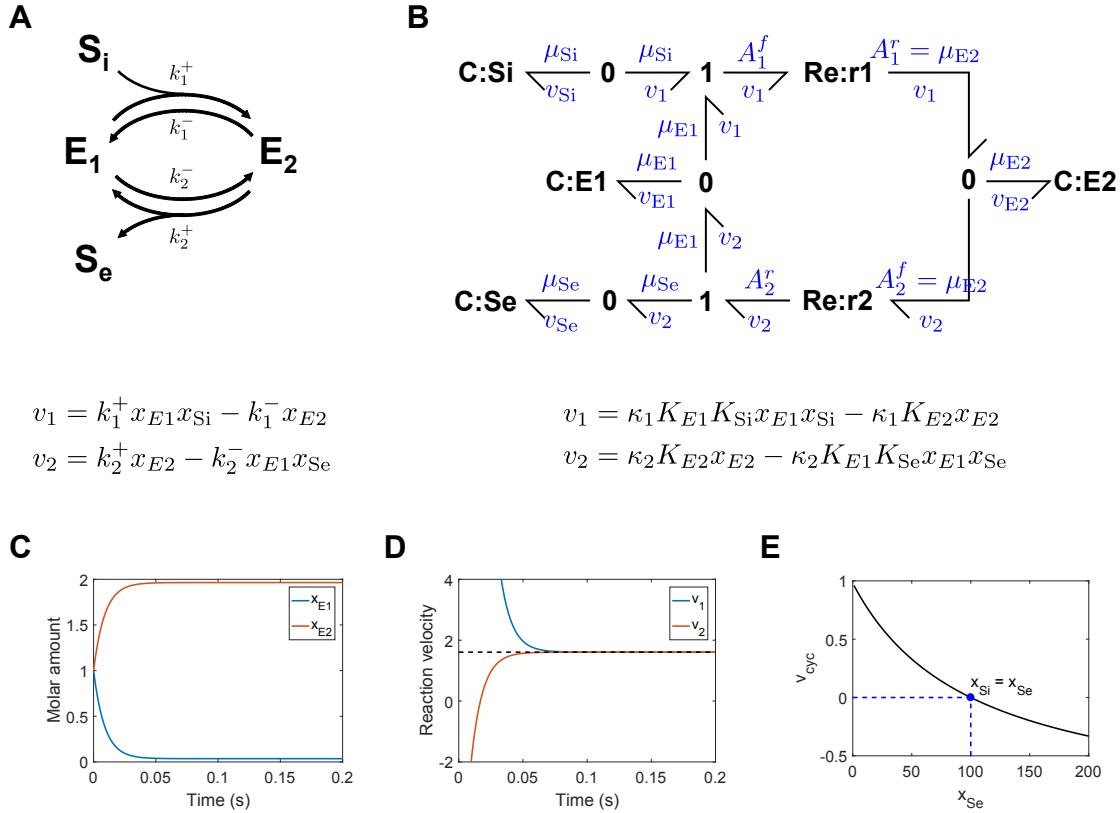


Figure 2.1: A simple enzyme cycle. (A) Kinetic model; (B) Bond graph model. S_i and S_e are the intracellular and extracellular substrate respectively, E_1 is the unbound transporter, and E_2 is the transporter bound to substrate. A simulation was run with unity parameters, the initial conditions $x_{E1} = x_{E2} = 1$ and $x_{Si} = 100$ and $x_{Se} = 10$. The resulting molar amounts of transporter states (C) and reaction velocities (D) were plotted against time. (E) Using $x_{Si} = 100$, the steady-state cycling rates were plotted against x_{Se} , with the blue dot indicating the equilibrium point.

with bonds that carry chemical potential and molar flow rate. Because the power leaving one component is transmitted to another, bond graphs are thermodynamically consistent, that is, the flow of energy is explicitly accounted for. In bond graph terminology, chemical potential is known as the effort variable, and molar flow rate is known as the flow variable (Gawthrop and Crampin, 2014).

In this section, we use hypothetical models (that do not correspond to any real biological system) with arbitrary units to outline the essential aspects of modelling biochemical systems using bond graphs. The first of these is the simple transporter model in Figure 2.1A. This transporter model contains a two-state enzyme cycle, where an uncharged substrate S binds to the enzyme on the intracellular (i) side, and unbinds on the extracellular (e) side. We denote the substrate in a specific compartment as S_x , where $x \in \{i, e\}$. The kinetic model is shown in Figure 2.1A, and the bond graph representation of this reaction cycle is shown in Figure 2.1B. Because a thermodynamic approach implies reversibility for all reactions (Polettini and Esposito, 2014; Gawthrop and Crampin, 2014), the transport of the substrate can proceed in both directions, depending on the direction of the chemical gradient.

Chemical energy is stored within the species of a biochemical system (Atkins and De Paula, 2006). This energy allows the system to do work through chemical reactions. The ability of a species to drive a reaction is given by the chemical potential. A number of factors contribute to the chemical potential, including temperature, the internal energy within the species, interactions between the species and the solvent, interactions between molecules of the same species as well as interactions between different species. However, in dilute biochemical systems, it is generally appropriate to assume that the chemical potential of a species depends only on molar concentration. Under these conditions, the chemical potential μ_s [J/mol] of a species s depends on the molar amount of species x_s [mol] in a logarithmic manner (Keener and Sneyd, 2009; Atkins and De Paula, 2006):

$$\mu_s = RT \ln(K_s x_s), \quad (2.1)$$

where $R = 8.314 \text{ J/K/mol}$ is the ideal gas constant, $T [\text{K}]$ is the absolute temperature and $K_s [\text{mol}^{-1}]$ is the species thermodynamic constant. Therefore, the chemical potentials of the species in the enzyme cycle are

$$\mu_{\text{Si}} = RT \ln(K_{\text{Si}} x_{\text{Si}}) \quad (2.2)$$

$$\mu_{\text{Se}} = RT \ln(K_{\text{Se}} x_{\text{Se}}) \quad (2.3)$$

$$\mu_{\text{E}1} = RT \ln(K_{\text{E}1} x_{\text{E}1}) \quad (2.4)$$

$$\mu_{\text{E}2} = RT \ln(K_{\text{E}2} x_{\text{E}2}). \quad (2.5)$$

Because species store energy, they are represented as **C** components in the bond graph model (Figure 2.1B), as they are analogous to capacitors in electrical circuits.

The chemical energy stored within the species of a biochemical system is used to drive reactions, which convert chemical species into other chemical species, dissipating chemical energy in the process. The rate of a reaction is related to the chemical potentials of its reactants and products using the Marcelin-de Donder equation (Gawthrop and Crampin, 2014; Oster et al., 1973):

$$v_R = \kappa_R (e^{A_R^f/RT} - e^{A_R^r/RT}), \quad (2.6)$$

where for a reaction R , v_R [mol/s] is the rate of reaction, κ_R [mol/s] is the reaction rate constant, A_R^f [J/mol] is the forward affinity (representing the total chemical potential of the reactants) and A_R^r [J/mol] is the reverse affinity (representing the total chemical potential of the products). For the enzyme cycle example in Figure 2.1B, the reaction rates of the upper and lower reactions are given by the equations

$$v_1 = \kappa_1 (e^{A_1^f/RT} - e^{A_1^r/RT}) \quad (2.7)$$

$$v_2 = \kappa_2 (e^{A_2^f/RT} - e^{A_2^r/RT}). \quad (2.8)$$

As seen in Figure 2.1B, reactions are represented as **Re** components in bond graph modelling.

In the enzyme cycle, some species are involved in more than one reaction. Therefore we require constraints to ensure that (a) the same chemical potential is used for the involvement of the same species in different reactions and (b) the contributions of each reaction are summed when calculating the rate of change of each species. These

constraints are captured by the **0** junction in bond graph modelling, which accounts for the first constraint by setting the chemical potentials of all connected bonds to be equal. In Figure 2.1B, there are four **0** junctions corresponding to each species. The second constraint is captured through the conservation equations

$$v_{\text{Si}} = -v_1 \quad (2.9)$$

$$v_{\text{Se}} = v_2 \quad (2.10)$$

$$v_{\text{E}1} = v_2 - v_1 \quad (2.11)$$

$$v_{\text{E}2} = v_1 - v_2. \quad (2.12)$$

0 junctions are analogous to parallel connections in electric circuits, and the Eqs. 2.9–2.12 represent the biochemical version of Kirchhoff's current law. Note that the molar flow rates of the bonds connecting to each **0** junction sum to zero, therefore the **0** junction conserves power (Borutzky, 2010).

In this example, both reactions involve the combination or dissociation of chemical species, and therefore have multiple reactants or products. To account for this, we require constraints so that (a) the rate at which each species is produced/consumed by the reaction is equal to the rate of reaction and (b) the affinities of each reaction are the sums of the chemical potentials of the reactants or products. These constraints are captured by **1** junctions in bond graph modelling, which satisfy the first constraint by fixing the molar flow rates of each connected bond to equal values. In Figure 2.1B, there are two **1** junctions representing the association and dissociation of substrate. The second constraint is accounted for through the conservation laws

$$A_1^f = \mu_{\text{Si}} + \mu_{\text{E}1} \quad (2.13)$$

$$A_2^r = \mu_{\text{Se}} + \mu_{\text{E}1}. \quad (2.14)$$

1 junctions are analogous to series connections in electric circuits, and Eqs. 2.13–2.14 are biochemical versions of Kirchhoff's voltage law. Note that since the chemical potentials of all bonds connected to **1** junctions sum to zero, they are power conserving, like the **0** junction (Borutzky, 2010).

Using the components described above, it is possible to derive a differential equation for the enzyme cycle. By substituting Eqs. 2.13–2.14 into Eqs. 2.7–2.8, we find that the reaction rates are

$$v_1 = \kappa_1(e^{(\mu_{\text{Si}}+\mu_{\text{E}1})/RT} - e^{\mu_{\text{E}2}/RT}) = \kappa_1 K_{\text{Si}} K_{\text{E}1} x_{\text{Si}} x_{\text{E}1} - \kappa_1 K_{\text{E}2} x_{\text{E}2} \quad (2.15)$$

$$v_2 = \kappa_2(e^{\mu_{\text{E}2}/RT} - e^{(\mu_{\text{Se}}+\mu_{\text{E}1})/RT}) = \kappa_2 K_{\text{E}2} x_{\text{E}2} - \kappa_2 K_{\text{Se}} K_{\text{E}1} x_{\text{Se}} x_{\text{E}1}. \quad (2.16)$$

Hence the bond graph components are able to reproduce the mass action equations, with the kinetic parameters

$$k_1^+ = \kappa_1 K_{\text{E}1} K_{\text{Si}} \quad (2.17)$$

$$k_1^- = \kappa_1 K_{\text{E}2} \quad (2.18)$$

$$k_2^+ = \kappa_2 K_{\text{E}2} \quad (2.19)$$

$$k_2^- = \kappa_2 K_{\text{E}1} K_{\text{Se}}. \quad (2.20)$$

By using the conservation laws from Eqs. 2.9–2.12, the rates of change for each species are

$$\begin{aligned}\dot{x}_{Si} &= v_{Si} = -v_1 \\ &= -k_1^+ x_{Si} x_{E1} + k_1^- x_{E2}\end{aligned}\tag{2.21}$$

$$\begin{aligned}\dot{x}_{Se} &= v_{Se} = v_2 \\ &= k_2^+ x_{E2} - k_2^- x_{Se} x_{E1}\end{aligned}\tag{2.22}$$

$$\begin{aligned}\dot{x}_{E1} &= v_{E1} = v_2 - v_1 \\ &= -k_1^+ x_{Si} x_{E1} + k_1^- x_{E2} + k_2^+ x_{E2} - k_2^- x_{Se} x_{E1}\end{aligned}\tag{2.23}$$

$$\begin{aligned}\dot{x}_{E2} &= v_{E2} = v_1 - v_2 \\ &= k_1^+ x_{Si} x_{E1} - k_1^- x_{E2} - k_2^+ x_{E2} + k_2^- x_{Se} x_{E1}.\end{aligned}\tag{2.24}$$

The above equations are applicable to an isolated system. However, since transporters are dissipative systems, i.e. they release energy as heat, it is impossible for these to operate continuously at steady state without an external power supply. Therefore, we model these external power supplies by holding certain species at constant concentrations rather than allowing them to dynamically vary. External connections of this type are treated as constant-potential boundary conditions, which are also known as chemostats in the literature (Polettini and Esposito, 2014; Gawthrop and Crampin, 2016). In this example, we assume that the concentrations of the substrates S_i and S_e are constant, so that Eqs. 2.21 and 2.22 are replaced by

$$\dot{x}_{Si} = 0\tag{2.25}$$

$$\dot{x}_{Se} = 0.\tag{2.26}$$

Because an external flow of species is required to keep the concentration of the species constant, chemostats are not energy conserving, but rather represent the influx or release of energy into the external environment. Thus chemostats turn a biochemical system from an isolated system into an open system that communicates with its external environment, and can therefore be used to couple models together. In bond graph modelling, **C** components are replaced with **Se** (effort source) components when they become chemostats to indicate the transfer of power with an external source. Since species can be represented using **C** or **Se** components interchangeably depending on the purpose of the model (Gawthrop et al., 2017; Gawthrop, 2017a), we represent species as **C** components in all diagrams within this chapter, and mention in text whether certain species are treated as “chemostats” in the analysis.

For the remainder of Section 2.2, we assume that:

1. Apart from the states of the transporter (e.g. E_1 , E_2), the concentrations of all species are constant (i.e. they are modelled as chemostats).
2. The volumes of all compartments are equal, so that the amount of each species directly corresponds to concentration. In real biological systems, compartments will have different volumes, and we show how to incorporate these effects in Appendix B of the Supplementary Material.
3. The parameters K and κ for each species and reaction take on unity values.

Assumptions 2 and 3 are made to simplify the analysis in this section, although we show how to generalise beyond these assumptions in later sections.

Figure 2.1C,D shows simulations of this model when the amount of enzyme is small relative to the amount of substrate, thus we expect the system to achieve a steady state relatively quickly. At steady state, the amounts of each pump state are constant, and the two reaction velocities converge towards the same value, as dictated by the pathway analysis of Gawthrop and Crampin (2017). We are often interested in the cycling (or turnover) rate at steady state, v_{cyc} , which is given by $v_{\text{cyc}} = V/e_0$, where V is the steady-state reaction rate (represented by the dotted black line in Figure 2.1D) and $e_0 = x_{E1} + x_{E2}$ is the total amount of transporter (Atkins and De Paula, 2006).

Figure 2.1E shows the effect of changing x_{Se} on cycling rate, indicating an inverse relationship: as x_{Se} increases, the transporter operates at a lower rate, and eventually operates in the reverse direction as expected. The bond graph model captures the fundamental physical constraint that the equilibrium point between the forward and reverse regimes of operation occurs when $x_{Se} = x_{Si}$. Therefore the simple transporter can only allow transport of substrate down its concentration gradient.

The direction in which each reaction proceeds is determined by the Gibbs free energy of reaction ΔG . Because reactions can only run in the direction of decreasing chemical potential, they will only proceed in the forward direction if ΔG is negative. The Gibbs free energy of reaction relates to the affinities $A = A^f - A^r$ of the **Re** components in bond graph models. Since efforts in the biochemical domain are associated with Gibbs free energy, the Gibbs free energy of reaction is the negative of affinity:

$$\Delta G = -A = A^r - A^f. \quad (2.27)$$

Therefore, the free energies of each of the reactions in this example are

$$\Delta G_1 = \mu_{E2} - \mu_{E1} - \mu_{Si} \quad (2.28)$$

$$\Delta G_2 = \mu_{E1} + \mu_{Se} - \mu_{E2}. \quad (2.29)$$

At steady state, the transporter's direction of operation is determined by the Gibbs free energy of the overall reaction $S_i \rightleftharpoons S_e$, which is also the sum of all reactions in its cycle (Gawthrop and Crampin, 2017):

$$\Delta G = \Delta G_1 + \Delta G_2 = \mu_{Se} - \mu_{Si}. \quad (2.30)$$

By substituting Eq. 2.1 and setting $\Delta G = 0$, the equilibrium of the system can be found:

$$\Delta G = RT \ln(K_{Se}x_{Se}) - RT \ln(K_{Si}x_{Si}) = RT \ln(x_{Se}/x_{Si}) = 0 \quad (2.31)$$

$$\Rightarrow x_{Se} = x_{Si}. \quad (2.32)$$

Therefore, as expected, the point at which the free energy is zero corresponds to the equilibrium of the transporter.

To ensure that the equilibrium occurs at $x_{Se} = x_{Si}$, we must specify the equilibrium between the substrate in each compartment:

$$K_{Si}/K_{Se} = K^{\text{eq}} = 1 \quad (2.33)$$

which is exactly equivalent to the detailed balance constraint used in kinetic models (Keener and Sneyd, 2009; Smith and Crampin, 2004; Tran et al., 2009):

$$\frac{k_1^+ k_2^+}{k_1^- k_2^-} = 1 \quad (2.34)$$

(this can be easily checked by substituting Eq. 2.33 into Eqs. 2.17–2.20). Because bond graphs are thermodynamically consistent, the thermodynamic constraint is simpler and more intuitive when compared to that for the kinetic parameters.

2.2.2 Coupled reactions

In the previous example, we observed that a passive transporter was unable to move a substrate against its concentration gradient. However, many transporters, including active transporters and exchangers, are able to move substrates against their concentration gradients. Such transporters couple the movement of substrate against a concentration gradient to a process that provides sufficient energy to enable transport. In Figure 2.2A, we show a simple mechanism for coupling together these processes. In addition to binding S_i and unbinding S_e , the transporter also binds A and unbinds B , giving rise to the overall reaction



The reaction $A \rightleftharpoons B$ provides energy for the transport of substrate, and may for example represent ATP hydrolysis in active transporters, or the transport of another species down its concentration gradient in exchangers or cotransporters. The bond graph representation of this chemical reaction network is given in Figure 2.2B.

To achieve the correct equilibrium point, we use the following thermodynamic constraint, describing the equilibria between side species:

$$K_{Si}/K_{Se} = 1 \quad (2.36)$$

$$K_A/K_B = 1. \quad (2.37)$$

Note that K_A/K_B can be set to different values depending on the process that $A \rightleftharpoons B$ represents; it has a value of 1 if it represents the transport of an uncharged substrate (as the equilibrium point occurs when the substrate has the same concentration on either side of the membrane), whereas it is determined by the standard Gibbs free energy of reaction if it represents a reaction such as ATP hydrolysis (see Appendix B of the Supplementary Material).

For a kinetic model with the parameters k_1^+ , k_1^- , k_2^+ , k_2^- , k_3^+ , k_3^- , k_4^+ and k_4^- , the corresponding thermodynamic constraint is

$$\frac{k_1^+ k_2^+ k_3^+ k_4^+}{k_1^- k_2^- k_3^- k_4^-} = \frac{K_{Si}}{K_{Se}} \frac{K_A}{K_B} = 1. \quad (2.38)$$

We note here that the thermodynamic constraint between kinetic parameters is more complicated than that in § 2.2.1 because there are more reactions in the cycle. When the number of reactions in biochemical cycles increases, these thermodynamic constraints

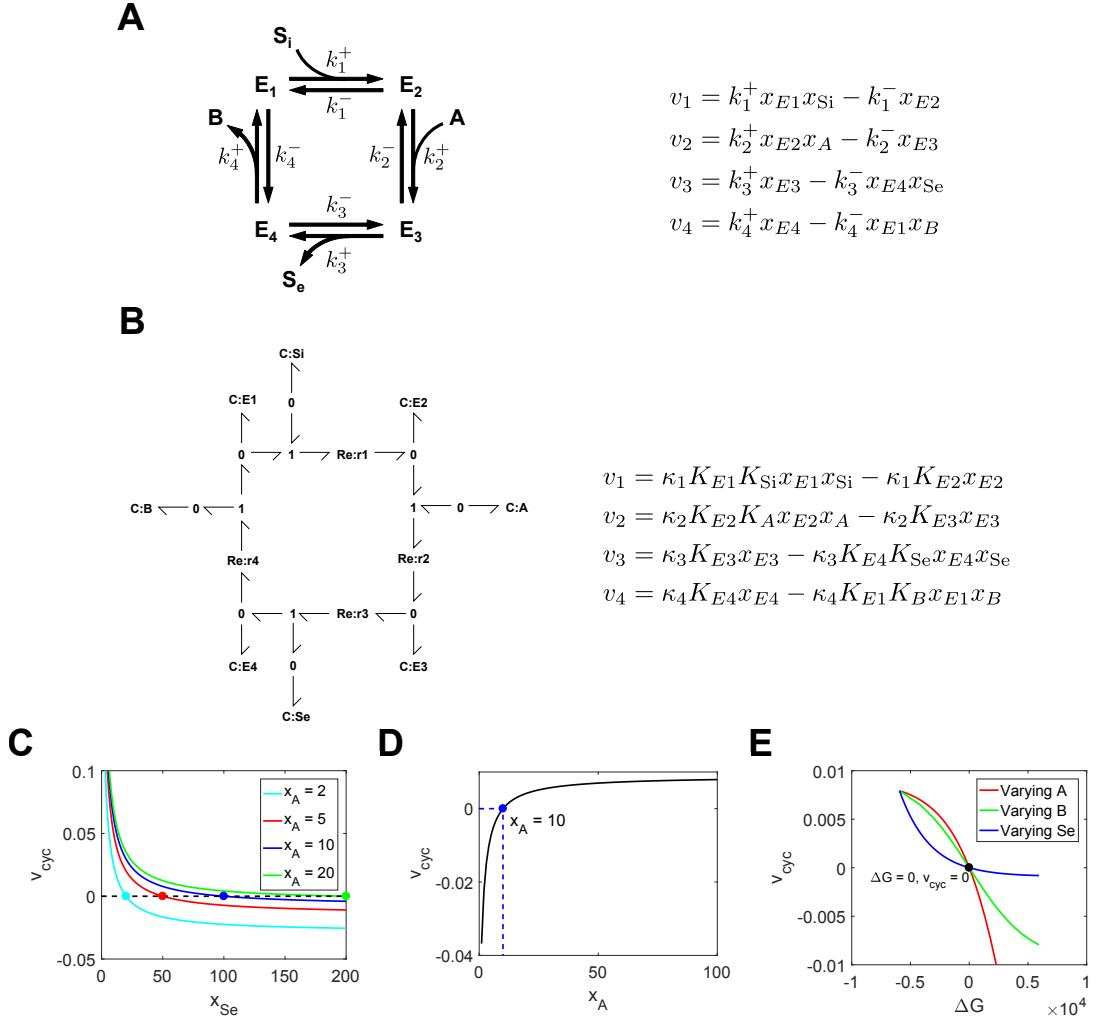


Figure 2.2: Transport of a species coupled to an energy supply. (A) Kinetic model; (B) Bond graph model. S_i and S_e are the intracellular and extracellular substrate respectively, A and B are species that provide power for the transport cycle, and E_1 , E_2 , E_3 and E_4 are the transporter states. Simulations were run using the initial conditions $x_{E1} = x_{E2} = x_{E3} = x_{E4} = 0.5$, and chemostats $x_B = 1$, $x_{S_i} = 10$. (C) The effect of x_{S_e} on steady state cycling rate for different x_A . Equilibrium points, where the cycling rates are zero, are denoted by the dots. (D) Effect of x_A on the cycling rate, with $x_{S_e} = 100$. The equilibrium (blue dot) occurs at $x_A = x_{S_e}/(x_{S_i}x_B) = 10$. (E) The relationship between the Gibbs free energy and cycling rate of the transporter, when A , B and S_e are varied.

become more complex, and harder to derive. By contrast, when bond graph parameters are used, thermodynamic constraints remain simple regardless of cycle complexity, and are more intuitive to formulate.

In Figure 2.2C we simulated the bond graph model to show the steady-state cycling rates under different values of x_A and x_{Se} . We note that for a passive transporter the equilibrium point occurs at $x_{Se} = x_{Si} = 10$, however in this example with $x_A > x_B$, the equilibrium points (represented by dots) occur at $x_{Se} > x_{Si}$. With all thermodynamic parameters set to 1, the equilibrium is given by the equation

$$x_{Se} = x_{Si}x_A/x_B. \quad (2.39)$$

Thus when $x_A > x_B$ (i.e. the coupled process releases energy), x_{Se} is greater than x_{Si} at equilibrium. Therefore there is a region of operation (when $x_{Si} < x_{Se} < x_{Si}x_A/x_B$) where the transporter can transport S against its concentration gradient. As the value of x_A increases, the equilibrium point shifts towards the right (Figure 2.2C), indicating that the transporter is able to move S against a greater concentration gradient. This is because with a greater x_A , the reaction $A \rightleftharpoons B$ provides more power, driving the transport in the forward direction (Figure 2.2D).

The Gibbs free energies of each of the reactions are

$$\Delta G_1 = \mu_{E2} - \mu_{E1} - \mu_{Si} \quad (2.40)$$

$$\Delta G_2 = \mu_{E3} - \mu_{E2} - \mu_A \quad (2.41)$$

$$\Delta G_3 = \mu_{E4} + \mu_{Se} - \mu_{E3} \quad (2.42)$$

$$\Delta G_4 = \mu_{E1} + \mu_B - \mu_{E4} \quad (2.43)$$

and the free energy of the transporter is

$$\Delta G = \Delta G_1 + \Delta G_2 + \Delta G_3 + \Delta G_4 = \mu_{Se} + \mu_B - \mu_{Si} - \mu_A. \quad (2.44)$$

By substituting Eq. 2.1 and setting $\Delta G = 0$, we can recover the equilibrium relationship in Eq. 2.39:

$$\begin{aligned} \Delta G &= RT \ln(K_{Se}x_{Se}) + RT \ln(K_Bx_B) - RT \ln(K_{Si}x_{Si}) - RT \ln(K_Ax_A) \\ &= RT \ln\left(\frac{K_{Se}x_{Se}K_Bx_B}{K_{Si}x_{Si}K_Ax_A}\right) \\ &= RT \ln\left(\frac{x_{Se}x_B}{x_{Si}x_A}\right) = 0 \end{aligned} \quad (2.45)$$

$$\Rightarrow \frac{x_{Se}x_B}{x_{Si}x_A} = 1. \quad (2.46)$$

Therefore, the equilibrium for steady-state operation corresponds to the point where the Gibbs free energy of the transport process is zero. We verify this fundamental physical constraint by plotting Gibbs free energy against cycling rate (Figure 2.2E). We varied the concentrations of A, B and S_e to generate three different curves. Despite differences in the shape of each curve, they each pass through the equilibrium point $\Delta G = 0$, $v_{cyc} = 0$, verifying that the bond graph model correctly captures the relationship between Gibbs free energy and equilibrium. Furthermore, the transporter operates in the forward direction ($v_{cyc} > 0$) with a negative (favourable) Gibbs free energy, and in the reverse direction ($v_{cyc} < 0$) with a positive (unfavourable) free energy.

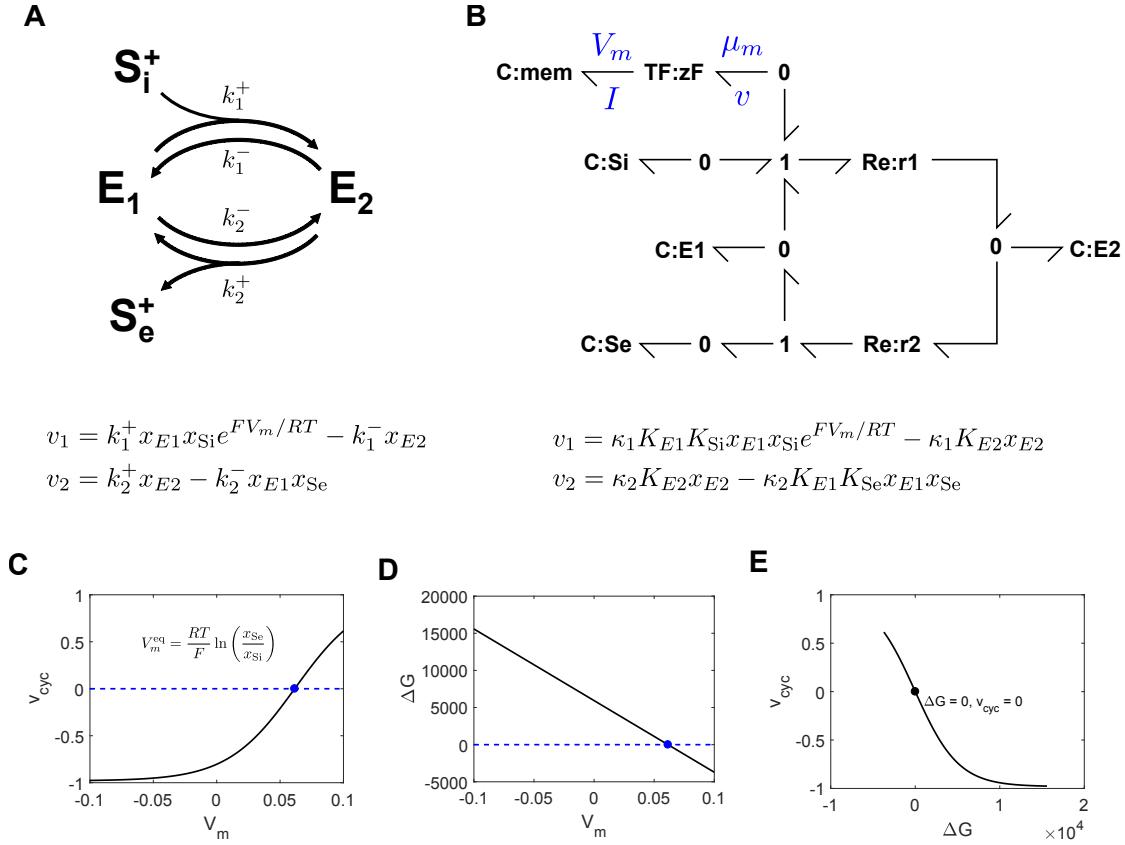


Figure 2.3: Transport of a charged species (A) Kinetic model; (B) Bond graph model. S_i^+ and S_e^+ are the intracellular and extracellular substrate respectively, E_1 is the unbound transporter, and E_2 is the transporter bound to substrate. Simulations were run with the initial conditions $x_{E1} = x_{E2} = 1$, and chemostats $x_{Si} = 10$, $x_{Se} = 100$. (C) Relationship between the membrane potential and cycling rate. The blue dot indicates the equilibrium potential, as predicted by the Nernst equation. (D) The effect of membrane potential on the Gibbs free energy of the transporter. (E) Relationship between the Gibbs free energy and cycling rate under the simulation conditions, with varying membrane potential.

2.2.3 Electrogenic transport

Many membrane transporters, including ion exchangers, cotransporters and active ion transporters, move a charged species across a membrane. For membranes within the cell that are charged (such as the plasma membrane, and inner mitochondrial membrane), the membrane potential influences both the kinetics and thermodynamics of transport. In this section we explore the impact on transport by modifying the enzyme example in § 2.2.1, so that the transported substrate has a single unit of positive charge (i.e. $z = 1$) (Figure 2.3A). For simplicity, we choose to assign the entirety of the electrical dependence to the forward side of the first reaction, which results in an exponential dependence on membrane voltage, arising from thermodynamic constraints (Keener and Sneyd, 2009; Smith and Crampin, 2004).

Because the bond graph representation is domain-independent, physical processes from other physical domains such as electrical or mechanical domains can be modelled by changing the units of the effort and flow variables. In the bond graph model (Figure 2.3B), the electrical dependence is incorporated by adding a **C** component representing membrane capacitance. This **C** component has similar energy-storage properties to the **C** component for each chemical species, however, in place of Eq. 2.1, it has a linear constitutive equation $V_m = q_m/C$, where V_m [V] is the membrane potential and q_m [C] is the charge. We use a membrane capacitance of $C = F^2 = 9309355225$ C, where $F = 96485$ C/mol is Faraday's constant. Since the capacitor is an electrical component, its power must be converted to chemical power to appropriately describe its influence on reaction kinetics. This conversion is given by Faraday's constant F , which relates the membrane potential and current to chemical potential and molar flow rate (Gawthrop et al., 2017):

$$\mu_m = zFV_m \quad (2.47)$$

$$I = zFv. \quad (2.48)$$

These transformations are described by the **TF** component in Figure 2.3B. Since $\mu_m v = V_m I$, the **TF** component is a power-conserving transformation.

By substituting the relevant constitutive equations, the rate of reaction 1 is given by

$$v_1 = \kappa_1(e^{(\mu_{\text{Si}} + \mu_{E1} + \mu_m)/RT} - e^{\mu_{E2}/RT}) = \kappa_1 K_{\text{Si}} K_{E1} x_{\text{Si}} x_{E1} e^{FV_m/RT} - \kappa_1 K_{E2} x_{E2}, \quad (2.49)$$

which corresponds to the exponential dependence of the kinetic reaction scheme. Because adding an electrical component affects the thermodynamics of transport, the equilibrium becomes dependent on membrane potential. As a result, the kinetics of the transporter must depend on membrane voltage to account for changes in equilibrium, although in practice modellers often make assumptions about where the electrical dependence lies due to a lack of experimental data (Smith and Crampin, 2004).

Bond graphs incorporate thermodynamic constraints for electrogenic transport. Because the chemical reaction structure is similar to the system in § 2.2.1, the constraint in Eq. 2.33 holds. However, because a single unit of charge is moved across the membrane in a single cycle, there is an additional constraint for charge:

$$z_1^f - z_1^r + z_2^f - z_2^r = z = 1, \quad (2.50)$$

where z indicates charge, subscripts indicate the reaction number, and superscripts indicate the side of the reaction. In this example, $z_1^f = 1$ and $z_1^r = z_2^f = z_2^r = 0$, therefore Eq. 2.50 is satisfied.

At equilibrium, the free energy of the overall reaction is zero (Gawthrop and Crampin, 2017), and because $K_{\text{Si}} = K_{\text{Se}}$ (Eq. 2.33),

$$\Delta G = \mu_{\text{Se}} - \mu_{\text{Si}} - \mu_m = RT \ln(K_{\text{Se}} x_{\text{Se}}) - RT \ln(K_{\text{Si}} x_{\text{Si}}) - FV_m = 0. \quad (2.51)$$

Therefore the membrane potential at equilibrium is

$$V_m = \frac{RT}{F} \ln \left(\frac{x_{\text{Se}}}{x_{\text{Si}}} \right), \quad (2.52)$$

which is the familiar Nernst potential (Keener and Sneyd, 2009).

In Figure 2.3C, this system was simulated under varying membrane potentials. Because the species is positively charged, changing the membrane potential in the positive direction drives transport of the species out of the cell. The transporter achieved an equilibrium at a membrane potential of approximately 0.062, consistent with the Nernst equation in Eq. 2.52.

In Figure 2.3D we plot the Gibbs free energy (defined in Eq. 2.51) against membrane potential. The membrane potential has a linear contribution to the free energy of the transporter, becoming more negative for positive membrane potentials due to the positively charged substrate. Importantly, the zero of the Gibbs free energy coincides with the zero of the cycling rate in Figure 2.3C, thus the transporter is dissipative, and therefore thermodynamically consistent. Figure 2.3E shows this directly, where the transporter operates in the positive direction only if the Gibbs free energy is negative.

2.3 Thermodynamic models of cardiac cell transporters

2.3.1 SERCA

The sarcoplasmic/endoplasmic Ca^{2+} ATPase (SERCA) is an active ion transporter that pumps calcium from the cytosol into the sarcoplasmic reticulum (SR), an intracellular Ca^{2+} store, restoring the calcium released from the SR with each heart beat. Because SERCA pumps calcium against a concentration gradient ($[\text{Ca}^{2+}]_{\text{sr}} > [\text{Ca}^{2+}]_{\text{i}}$), it couples calcium transport to the hydrolysis of ATP in order to obtain the energy required for transport. The overall reaction of the pump is



Since the SR membrane is uncharged, the pump is driven purely by chemical energy.

SERCA has been included in a wide variety of models of cardiac cell Ca^{2+} cycling, excitation-contraction coupling and electrophysiology. Our analysis here is based on the model by Tran et al. (2009) which has since been incorporated into a number of subsequent cardiac cell models because it describes the dependence of cycling rate on all metabolites (Tran et al., 2015; Williams et al., 2011; Walker et al., 2014). The chemical reaction network and bond graph representation of the Tran et al. (2009) model are shown in Figure 2.4. The structure of the bond graph closely resembles the cyclic nature of the chemical reaction scheme. However, the bond graph representation explicitly shows that the H^+ ions involved in multiple reactions are linked. Note that the bond graph representation uses **TF** components to describe multiple copies of a species in a single reaction; this is discussed further in Appendix A of the Supplementary Material.

Discussion of how bond graph parameters may be determined from an existing kinetic model are presented in Appendix B of the Supplementary Material. The kinetic parameters were thermodynamically constrained, therefore it was possible to find a set of corresponding bond graph parameters by using Eq. 2.74 of the Supplementary Material,

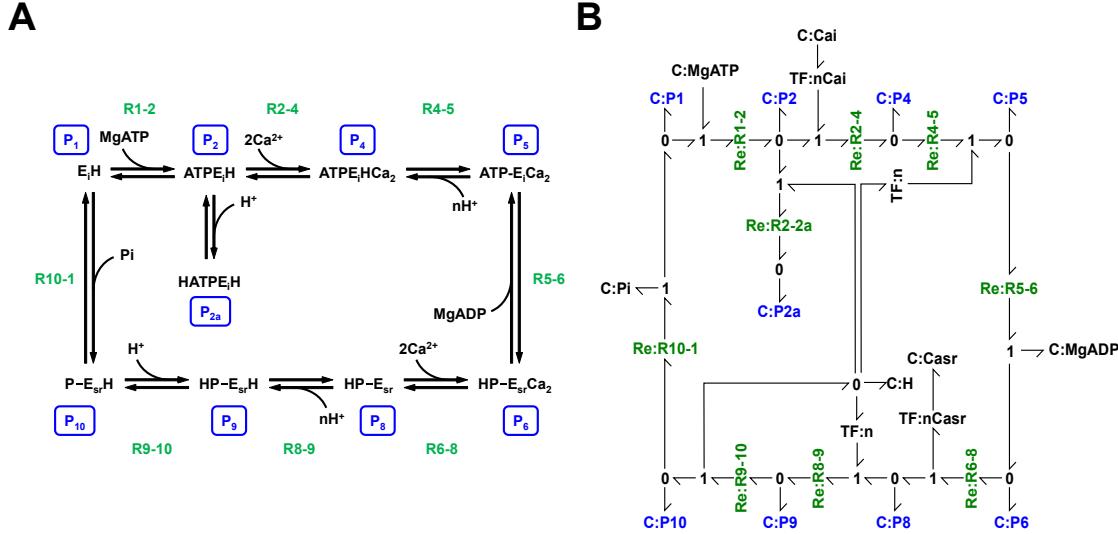


Figure 2.4: Kinetic and bond graph representations of the cardiac SERCA model from Tran et al. (2009). (A) Kinetic model, adapted from Fig. 2 of Tran et al. (2009). The numbers for each pump state are shown in blue boxes, and the names for each reaction are labelled in green. (B) Bond graph representation, with pump states shown in blue, and reactions in green.

with the stoichiometric matrix and resulting bond graph parameters given in Appendix D. A comparison of the kinetic and bond graph models (Figure 2.5A) confirms that the two models match closely, with only minor discrepancies at higher cycling rates due to the assumption of rapid equilibrium in the Tran et al. (2009) model.

SERCA accounts for approximately 10–15% of the energy consumption within cardiomyocytes (Tran et al., 2015; Schramm et al., 1994), and is seen as a major contributor to myocardial energy expenditure. In heart failure, SERCA activity decreases, which causes a higher proportion of Ca^{2+} to be removed via an alternative pathway – the less efficient $\text{Na}^+/\text{Ca}^{2+}$ exchanger (NCX) – increasing energy expenditure (Kawase and Hajjar, 2008). Therefore it is important to introduce the notion of energy and efficiency into models of transporters in cardiac cells to incorporate the metabolite dependencies required for studying changes in flux under heart failure, and to compare the efficiencies of transporters that move the same substrates. While the parameters for kinetic models can be chosen such that they are thermodynamically consistent, energy-related quantities such as Gibbs free energy and power consumption do not arise naturally from these parameters. Because bond graphs explicitly model energy transfer, an advantage of using this framework is that the power consumption and efficiency are easily calculated from the model. At steady state, the Gibbs free energy of the pump can be calculated from the chemical potentials of metabolites (Gawthrop and Crampin, 2017):

$$\Delta G = 2\mu_{\text{Casr}} + \mu_{\text{MgADP}} + \mu_{\text{Pi}} + \mu_{\text{H}} - 2\mu_{\text{Cai}} - \mu_{\text{MgATP}}. \quad (2.54)$$

The relationship between free energy and SR Ca^{2+} is shown in Figure 2.5B. Note that the Gibbs free energy is zero at the equilibrium point of the pump in Figure 2.5A, as expected of a thermodynamically consistent system. It is important to note that when $\Delta G > 0$, SERCA has been experimentally observed to operate in the reverse direction where Ca^{2+} flows from the SR to cytosol, and ATP is synthesised (Makinose and Hasselbach,

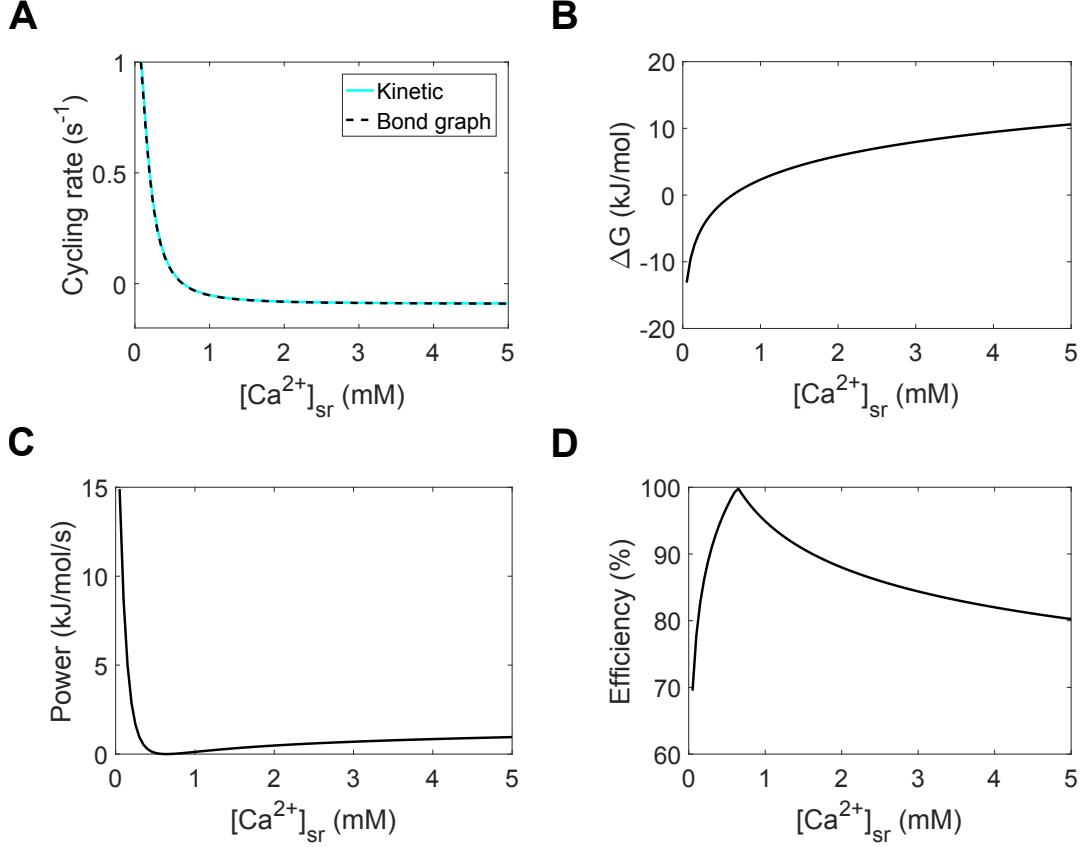


Figure 2.5: Simulation of the SERCA pump. (A) Comparison of cycling rates for kinetic and bond graph models, reproducing part of Fig. 13 in Tran et al. (2009); (B) Gibbs free energy; (C) Power consumption per mol of pump; (D) Pump efficiency. Simulations were run with [Ca²⁺]_i = 150 nM, pH = 4, [MgADP] = 0.0363 mM, [MgATP] = 0.1 mM, [Pi] = 15 mM. Cycling rates were estimated by initialising each pump state to 1/9 fmol, and running the simulation to its steady state.

1971). We observe in Figure 2.5B that a thermodynamic framework captures this reversal mode of operation, as it is a fundamental physical constraint. The product of free energy and reaction rate gives the rate of power dissipation. As seen in Figure 2.5C, the power consumption is positive under all conditions except at equilibrium, where it is zero.

Because the bond graph approach can split energetic contributions from different sources, it is possible to assess the efficiency of the pump. For this, we define the affinity of ATP hydrolysis as

$$A_{\text{hyd}} = \mu_{\text{MgATP}} - \mu_{\text{MgADP}} - \mu_{\text{Pi}} - \mu_{\text{H}}, \quad (2.55)$$

and the affinity of calcium transport as

$$A_{\text{tr}} = 2\mu_{\text{Cai}} - 2\mu_{\text{Casr}}. \quad (2.56)$$

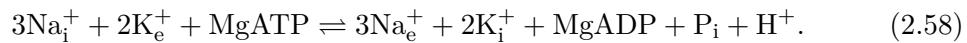
Using the notion of pumping efficiency introduced in Gawthrop and Crampin (2018), we define the efficiency ρ as

$$\rho = \begin{cases} -A_{\text{tr}}/A_{\text{hyd}}, & A_{\text{hyd}} \geq -A_{\text{tr}} \\ -A_{\text{hyd}}/A_{\text{tr}}, & A_{\text{hyd}} < -A_{\text{tr}}. \end{cases} \quad (2.57)$$

Thus if the pump operates in the forward direction, efficiency is the proportion of energy from ATP hydrolysis that is converted into energy for calcium transport, and if the pump operates in the reverse direction, efficiency is the proportion of energy supplied from calcium transport that is used to generate ATP. We therefore expect a cusp point at $A_{\text{hyd}} = -A_{\text{tr}}$. The efficiency of the pump is plotted in Figure 2.5D. This SERCA model operates reasonably efficiently under the simulated conditions, ranging from 70–100%, consistent with previous estimates that the pump is 85–90% efficient (Pinz et al., 2011). We note also a negative relationship between cycling rate and efficiency, with the pump becoming less efficient the further it moves away from equilibrium. Conversely, the pump approaches an efficiency of 100% as it nears equilibrium. However, it should be noted that in reality, SERCA pumps may exhibit “slippage”, where the pump hydrolyses ATP without transporting any Ca^{2+} (de Meis, 2002). Incorporating such behaviour into the model would likely reduce the maximum operating efficiency of the pump (Gawthrop and Crampin, 2017).

2.3.2 Na^+/K^+ ATPase

The Na^+/K^+ ATPase is responsible for maintaining the Na^+ and K^+ gradients that drive ionic currents during the action potential of cardiac cells and many other excitable cells. Na^+ and K^+ are pumped against their electrochemical gradients, therefore their transport requires the supply of energy from ATP hydrolysis. The overall reaction of the pump is



In contrast to SERCA, in which only chemical potentials determine the direction of operation, the Na^+/K^+ ATPase is driven by both chemical and electrical potentials because the plasma membrane is electrically charged. Therefore a thermodynamically consistent model of Na^+/K^+ ATPase must account for both the free energy of ATP hydrolysis as well as the energetic contribution of the membrane potential.

In this section, we outline a new model of cardiac Na^+/K^+ ATPase, based on the earlier model by Terkildsen et al. (2007) (Figure 2.6). While the model of Terkildsen et al. (2007) is a biophysically detailed model that incorporates some thermodynamic principles, the final model is thermodynamically inconsistent. We updated the model, correcting the equilibrium constants used for identical binding sites, the equilibrium constant for ATP hydrolysis and mathematical expressions arising from the rapid equilibrium approximation, as described in Appendix C of the Supplementary Material.

We refitted the new model to the data used to parameterise the original Terkildsen et al. (2007) model (see Appendix C of the Supplementary Material for further detail). Comparisons of model simulations to data are given in Figures 2.7 and 2.8. The updated model provides a good fit to the data, and has a comparable fit compared to the

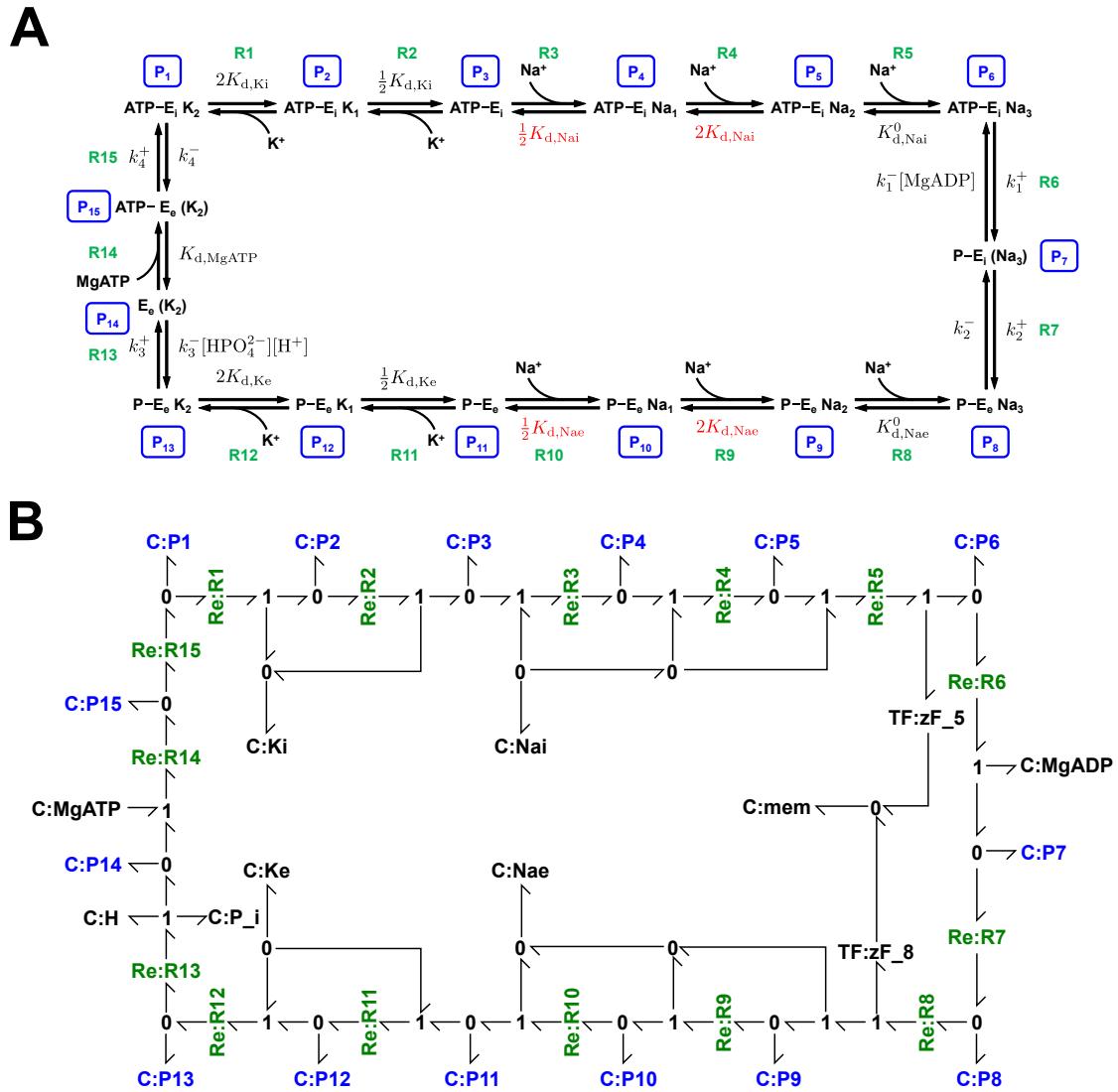


Figure 2.6: The cardiac Na⁺/K⁺ ATPase model. (A) Kinetic model, with numbers for each pump state (blue boxes) and reaction names (green) are labelled, with corrected parameters shown in red. (B) Bond graph model, with pump states coloured in blue, and reactions coloured in green.

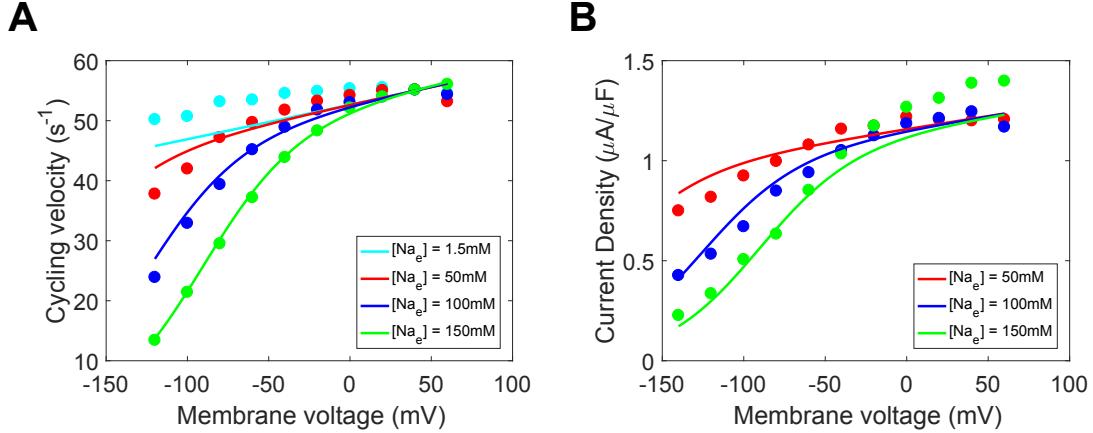


Figure 2.7: Fit of the cardiac Na^+/K^+ ATPase model to current-voltage measurements. (A) Comparison of the model to extracellular sodium and voltage data (Nakao and Gadsby, 1989, Fig. 3), with cycling velocities scaled to a value of 55s^{-1} at $V = 40\text{mV}$. (B) Comparison of the model to whole-cell current measurements (Nakao and Gadsby, 1989, Fig. 2A). $[\text{Na}^+]_i = 50 \text{ mM}$, $[\text{K}^+]_i = 0 \text{ mM}$, $[\text{K}^+]_e = 5.4 \text{ mM}$, $\text{pH} = 7.4$, $[\text{Pi}]_{\text{tot}} = 0 \text{ mM}$, $[\text{MgATP}] = 10 \text{ mM}$, $[\text{MgADP}] = 0 \text{ mM}$, $T = 310 \text{ K}$.

original model. The fit in Figure 2.7A was slightly worse than the original model at lower extracellular Na^+ concentrations, but the model seems to be more consistent with experimental evidence that the saturated pump velocity has little sensitivity to extracellular Na^+ at positive membrane potentials (Nakao and Gadsby, 1989).

The updated model is thermodynamically consistent, therefore it has a bond graph representation (Figure 2.6B). The definition of bond graph parameters and stoichiometric matrices from the fitting process are given in Appendix E of the Supplementary Material. At steady state, the Gibbs free energy of the cycle is

$$\Delta G = 3\mu_{\text{Nae}} + 2\mu_{\text{Ki}} + \mu_{\text{MgADP}} + \mu_{\text{Pi}} + \mu_{\text{H}} - 3\mu_{\text{Nai}} - 2\mu_{\text{Ke}} - \mu_{\text{MgATP}} - \mu_{\text{mem}} \quad (2.59)$$

$$\begin{aligned} &= \Delta G_{\text{MgATP}}^0 + RT \ln \left(\frac{[\text{MgADP}][\text{Pi}][\text{H}^+]}{[\text{MgATP}]} \right) + 3RT \ln \left(\frac{[\text{Na}^+]_e}{[\text{Na}^+]_i} \right) \\ &\quad + 2RT \ln \left(\frac{[\text{K}^+]_i}{[\text{K}^+]_e} \right) - FV_m. \end{aligned} \quad (2.60)$$

To verify the thermodynamic consistency of the bond graph model, we simulated the steady-state operation of the pump at voltages near equilibrium, under physiological (6.95 mM) and ischaemic (1 mM) MgATP concentrations (Figure 2.9A). The equilibrium potentials of the pump, where cycling rates are zero, are indicated by dots. An expression for the equilibrium potential can be derived by solving for $\Delta G = 0$:

$$V_m^{\text{eq}} = \frac{1}{F} \Delta G_{\text{MgATP}}^0 + \frac{RT}{F} \left(\ln \frac{[\text{MgADP}][\text{Pi}][\text{H}^+]}{[\text{MgATP}]} + 3 \ln \frac{[\text{Na}^+]_e}{[\text{Na}^+]_i} + 2 \ln \frac{[\text{K}^+]_i}{[\text{K}^+]_e} \right). \quad (2.61)$$

The predicted equilibrium potentials of -224.5 mV and -276.3 mV for $[\text{MgATP}] = 1 \text{ mM}$ and 6.95 mM respectively are consistent with the equilibria of the steady-state simulations, verifying thermodynamic consistency. We plotted the free energy against

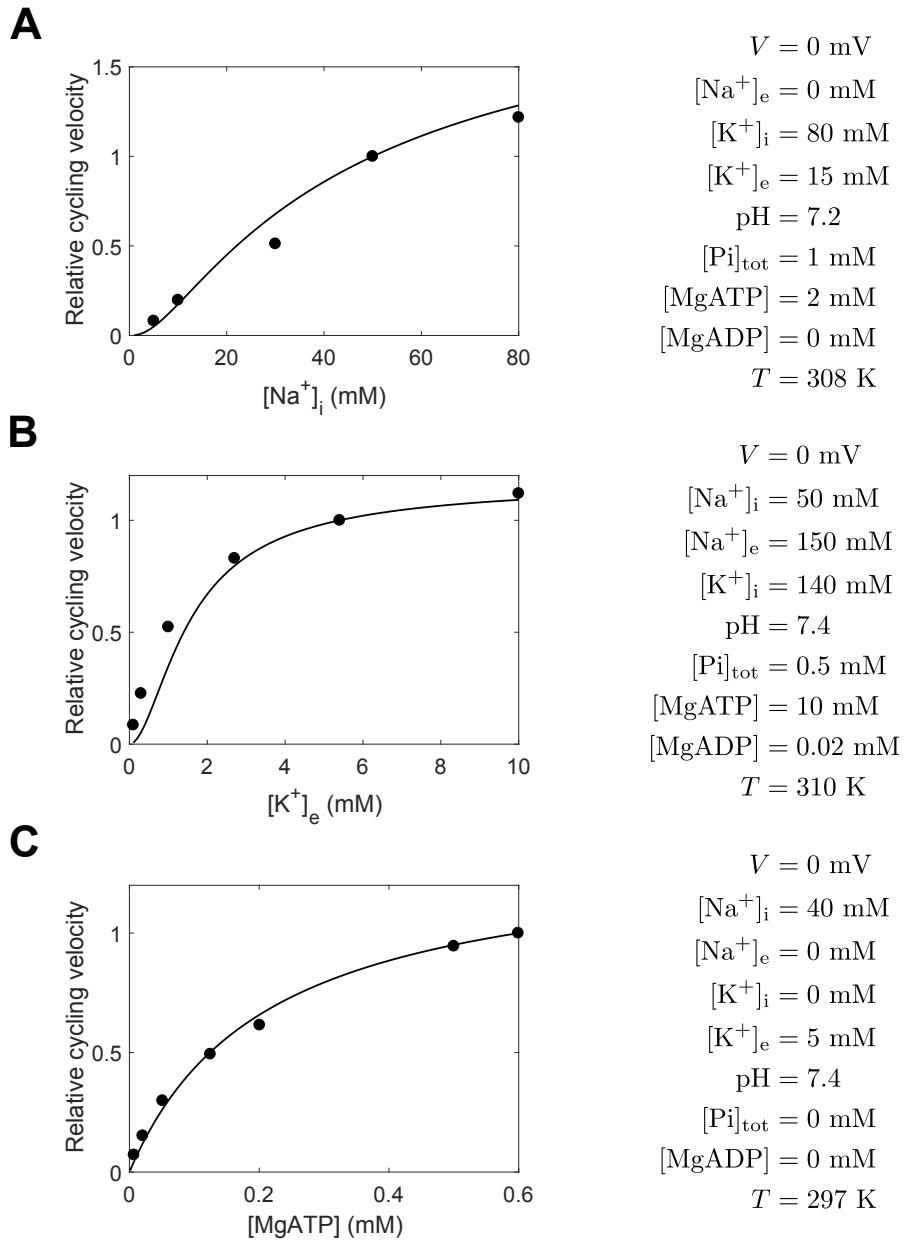


Figure 2.8: Fit of the cardiac Na⁺/K⁺ ATPase model to metabolite dependence data. Simulation conditions are displayed on the right of each figure. (A) Comparison of the model to data with varying intracellular sodium concentrations (Hansen et al., 2002, Fig. 7A), normalised to the cycling velocity at [Na⁺]_i = 50 mM. (B) Comparison of the model to data with varying extracellular potassium (Nakao and Gadsby, 1989, Fig. 11A), normalised to the cycling velocity at [K⁺]_e = 5.4 mM. (C) Comparison of the model to data with varying ATP (Friedrich et al., 1996, Fig. 3B), normalised to the cycling velocity at [MgATP] = 0.6 mM.

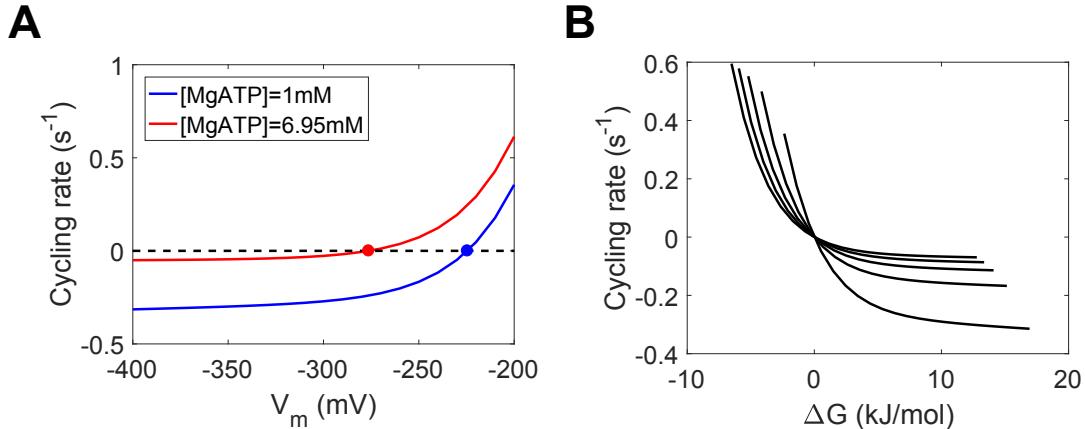


Figure 2.9: Simulation of the Na^+/K^+ ATPase. (A) Cycling rates of the pump near reversal potential; (B) Relationship between Gibbs free energy and cycling rate. The curves represent different concentrations of MgATP, from a concentration of 1 mM on the right, with increments of 1 mM up to a concentration of 5mM on the left. The Gibbs free energy was varied by changing the membrane potential. For (A) and (B), simulations were run using $[\text{Na}^+]_i = 10\text{ mM}$, $[\text{Na}^+]_e = 140\text{ mM}$, $[\text{K}^+]_i = 145\text{ mM}$, $[\text{K}^+]_e = 5.4\text{ mM}$, $\text{pH} = 7.095$, $[\text{Pi}] = 0.3971\text{ mM}$, $[\text{MgATP}] = 6.95\text{ mM}$, $[\text{MgADP}] = 0.035\text{ mM}$. Each pump state was initialised to 1/15 fmol, and steady states were estimated by running each simulation to steady state.

cycling rate for a number of MgATP concentrations, and varying the membrane potential to generate the curves in Figure 2.9B. Despite differences in the curves, they all satisfy the fundamental constraint that $v_{cyc} = 0$ when $\Delta G = 0$, and v_{cyc} is positive only at negative free energies. Thus Figure 2.9 verifies that the updated model is thermodynamically consistent. Like SERCA, the Na^+/K^+ ATPase is reversible, and has been experimentally observed to synthesise ATP under artificially large ionic gradients (Glynn, 2002). A thermodynamic framework is ideal for describing such phenomena.

2.4 Discussion

In this chapter we have discussed the essential thermodynamic principles underlying membrane transporters, and how the bond graph framework captures thermodynamic constraints in these systems. We illustrated using hypothetical models that bond graphs can be used to model simple membrane transport mechanisms, incorporating relevant physical and thermodynamic constraints for both nonelectrogenic and electrogenic transport. The bond graph approach provides a single framework that naturally incorporates known thermodynamic constraints in these systems, including detailed balance (Keener and Sneyd, 2009; Liebermeister et al., 2010) and the Nernst potential (Keener and Sneyd, 2009).

We applied the bond graph approach to two membrane transporters. Using the model of cardiac SERCA by Tran et al. (2009), we demonstrated that thermodynamically consistent models can be represented as bond graphs, and that energetic quantities are

easily calculated from the bond graph model (§ 2.3.1). We then developed a bond graph model of cardiac Na^+/K^+ ATPase based on earlier work by Terkildsen et al. (2007) and Smith and Crampin (2004).

2.4.1 Thermodynamic constraints

Because bond graphs are a general-purpose modelling framework for physical systems, they unify a number of known thermodynamic constraints in the literature under a single framework, allowing thermodynamic constraints to be applied more consistently across models of biological systems. In kinetic models, Wegscheider conditions and detailed balance are used to constrain kinetic parameters in models of membrane transporters (Keener and Sneyd, 2009; Smith and Crampin, 2004; Tran et al., 2009). By contrast, in bond graph models, the thermodynamic parameters automatically account for these constraints (Gawthrop and Crampin, 2014). In electrogenic transport, the Nernst equation (Keener and Sneyd, 2009) is a thermodynamic constraint that captures the electrochemical equilibrium. This constraint is automatically accounted for in bond graph modelling, because the Nernst potential naturally arises from the constitutive equations of the **C** components for charged species and the membrane potential (Gawthrop et al., 2017).

We believe that thermodynamically consistent models are more likely to remain robust when incorporated into whole-cell models and multi-scale models. It is often uncertain whether models remain predictive under conditions beyond the data they were parameterised on (Beattie et al., 2018; Fink et al., 2011). As discussed in Smith and Crampin (2004) and Tran et al. (2009), many existing models of transporters do not incorporate dependence for all metabolites and are therefore unable to accurately predict activity under varying metabolite concentrations. Because bond graphs enforce physical constraints on models, they provide useful constraints that dictate the behaviour of the model outside of the range of experimental data (Soh and Hatzimanikatis, 2010), being particularly useful if the experimental data were captured away from equilibrium. In order to satisfy thermodynamic consistency, bond graphs force the modeller to incorporate dependence for all metabolites in models of membrane transporters. Adherence to conservation laws may also prove important when transporter models are incorporated into multi-scale models. In models of cardiac electrophysiology, mass and charge conservation impact on long-term behaviour, and violating these conservation laws may cause the model to drift (Hund et al., 2001; Livshitz and Rudy, 2009; Pan et al., 2018a). Thus in multi-scale models where many individual submodules are simulated over time-scales far greater than they were originally developed for, accounting for conservation laws such as conservation of energy may be an important factor in maintaining the stability of these models.

2.4.2 Bond graph models of biological systems

In this chapter, we described the representation of existing models into bond graphs, using a model of SERCA (Tran et al., 2009) as an example. These methods can be applied to a wide range existing thermodynamically consistent transporter models. A bond graph representation reveals the underlying energetics of these models (Gawthrop and Crampin,

2017). In addition to membrane transporters, models of ion channels (Gawthrop et al., 2017), redox reactions (Gawthrop, 2017a), metabolic systems (Gawthrop et al., 2015) and signalling pathways (Gawthrop and Crampin, 2016) can also be represented as bond graphs. As the bond graph methodology continues to develop, we anticipate that more biological processes such as cross-bridge cycling, diffusion and gene regulation will be able to be modelled using bond graphs. Using the modularity inherent in bond graphs (Gawthrop, 2017a), these models can be coupled together to assemble whole-cell models that can assess how cells deal with fundamental trade-offs between energy efficiency, speed and robustness.

While the bond graph approach can detect thermodynamic inconsistency, it is unable to detect if its thermodynamic parameters are correct. Therefore, in cases such as the model of Na^+/K^+ ATPase by Terkildsen et al. (2007), parameterising the model with incorrect species constants or equilibrium constants can lead to incorrect equilibrium points even if the model itself is internally consistent. However, we note that the bond graph framework will flag these issues to the modeller when two models with conflicting thermodynamic parameters are coupled together.

2.4.3 Future work

To date, a large proportion of bond graph models of biochemical systems have been derived from existing kinetic models (Gawthrop et al., 2015; Gawthrop, 2017a; Gawthrop and Crampin, 2017). However, as outlined in this chapter, developing new models using the bond graph framework is a more powerful approach as it guarantees thermodynamic consistency. A challenge in this process is estimating bond graph parameters from data, as kinetic information alone is insufficient to uniquely determine bond graph parameters (Gawthrop et al., 2015). Therefore future work will focus on using existing information and optimisation procedures to derive appropriate estimates of bond graph parameters from experimental data when thermodynamic data is unavailable, and designing efficient experimental protocols for identifying bond graph parameters when thermodynamic data is available. In this process, there is a key role for quantifying the uncertainty of biological parameters (Babtie and Stumpf, 2017; Beattie et al., 2018). Given that bond graph parameters are physical quantities, it would be interesting to investigate whether these uncertainties propagate as individual models are coupled together.

We modelled transporters using multiple states with mass action equations, giving rise to steady-state velocities. However, many models of transporters are simplifications of this approach, as they are described by a single equation for the transport rate rather than multiple equations for transitions between enzyme states (Keener and Sneyd, 2009). Smith and Crampin (2004) describe a method for reducing multi-state models of membrane transporters into a single equation for transport rate by using rapid equilibrium and steady-state simplifications, while using detailed balance to ensure thermodynamic consistency. As we move towards increasingly complex multi-scale models, it is important to develop methods of simplifying models to reduce computational cost (Smith et al., 2007). An advantage of using bond graphs in the simplification process is that they impose constraints to ensure that simplified models remain thermodynamically consistent (Gawthrop and Crampin, 2014).

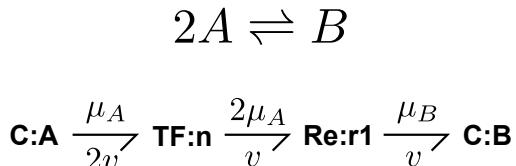


Figure 2.10: Bond graph of the reaction $2A \rightleftharpoons B$.

2.5 Conclusion

We have shown that bond graphs can be used to model membrane transporters while capturing fundamental thermodynamic and physical constraints. We apply this framework to SERCA and Na^+/K^+ ATPase to develop models that are thermodynamically consistent while revealing the underlying energetics of these transporters. Combined with their inherent modularity, we believe that bond graphs are a powerful tool for incorporating models of membrane transporters into other models while maintaining thermodynamic consistency. Future work will focus on extending the methods presented here to develop integrated models of cardiac cells. The function of cardiac cells is dependent on electrical and chemical signalling as well as chemical to mechanical transduction. The multi-domain nature of bond graphs allows them to represent each of these processes under a single framework. Cardiac cells are highly dependent on a source of energy, and their metabolism is altered during heart failure. Because bond graphs explicitly model energy transfer, they provide a principled method of analysing the energy consumption of cardiac cells. More broadly, our approach to modelling transporters can be extended to enzymes and proteins in general, allowing models of cells to be built in a modular fashion.

Data availability

The code associated with this chapter is available from GitHub (https://github.com/uomsystemsbiology/transporter_thermodynamics), and is archived on Zenodo (<https://doi.org/10.5281/zenodo.1422867>) (Pan et al., 2018b). The repositories contain MATLAB code that generates the figures and CellML code containing equations for the models. A virtual reference environment (Hurley et al., 2015) is available from <https://doi.org/10.5281/zenodo.1433958> (Pan et al., 2018c).

2.A TF components in biochemical networks

For some reactions such as dimerisation, two or more of the same species may be involved in one side of the reaction. To represent this, the bond connected to that species is fed through a **TF** component. The **TF** component represents a transformer, which can transmit and convert energy while maintaining energy conservation. The equation for

the transformer is

$$e_2 = ne_1 \quad (2.62)$$

$$f_1 = nf_2, \quad (2.63)$$

where n represents the stoichiometry of the species. Since $e_1 f_1 = e_2 f_2$, the transformer is energy-conserving as the power in is equal to the power out of the transformer.

When applied to the biochemical reaction $2A \rightleftharpoons B$, with the bond graph representation in Figure 2.10, the constitutive equations of the **Re** and **C** components are

$$\mu_A = RT \ln(K_A x_A) \quad (2.64)$$

$$\mu_B = RT \ln(K_B x_B) \quad (2.65)$$

$$v = \kappa(e^{A^f/RT} - e^{A^r/RT}). \quad (2.66)$$

The equation corresponding to the transformer (with $n = 2$) is

$$A_1^f = 2\mu_A \quad (2.67)$$

$$\dot{x}_A = -2v, \quad (2.68)$$

thus the transformer states that the rate of consumption of A is twice the rate of reaction. Because two molecules of A are consumed by the reaction, the forward affinity is twice the chemical potential of A. By combining these equations the law of mass action can be derived:

$$v = \kappa(e^{2\mu_A/RT} - e^{\mu_B/RT}) = \kappa K_A^2 x_A^2 - \kappa K_B x_B \quad (2.69)$$

$$\dot{x}_A = -2v = 2\kappa K_B x_B - 2\kappa K_A^2 x_A^2 \quad (2.70)$$

$$\dot{x}_B = v = \kappa K_A^2 x_A^2 - \kappa K_B x_B. \quad (2.71)$$

2.B Converting existing kinetic models to bond graphs

Because there are a number of existing models of membrane transporters that are thermodynamically consistent, the ability to convert kinetic models of transporters into bond graphs can save one from the effort of building new bond graph models to explain existing data. In § 2.2.1, we outlined how a chemical reaction network can be represented as a bond graph. While the structure of the system can be naturally translated into a bond graph, the parameters of existing models do not directly translate into bond graph parameters K and κ . While bond graph parameters define a unique set of kinetic parameters, identifying bond graph parameters from a set of kinetic parameters is more involved. In this section, we outline how to infer bond graph parameters from kinetic parameters, and discuss the issues that arise during this process.

We first use the enzyme cycle example in § 2.2.1 to illustrate the essential concepts of inferring bond graph parameters. The relationship between the kinetic and bond graph parameters is given by Eqs. 2.17–2.20. Additional constraints are enforced by the detailed balance constraint in Eq. 2.33. Therefore, given a set of kinetic constants $k_1^+, k_2^+, k_1^-, k_2^-$,

an equivalent set of bond graph parameters can be derived by solving the Eqs. 2.17–2.20, 2.33. If we define \mathbf{Ln} as the element-wise logarithm operator, these constraints can be packaged into the matrix equation

$$\mathbf{Ln} \begin{bmatrix} k_1^+ \\ k_2^+ \\ k_1^- \\ k_2^- \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \end{bmatrix} \mathbf{Ln} \begin{bmatrix} \kappa_1 \\ \kappa_2 \\ K_{\text{Si}} \\ K_{\text{Se}} \\ K_{E1} \\ K_{E2} \end{bmatrix} \quad (2.72)$$

which can be solved through linear methods (Anton and Rorres, 2014). However, it should be noted that the rows of the right hand side are linearly dependent, as $R1 + R2 - R3 - R4 - R5 = 0$. Therefore, in order for a solution to exist, the kinetic parameters must satisfy the same dependency between rows:

$$\begin{bmatrix} 1 & 1 & -1 & -1 & -1 \end{bmatrix} \mathbf{Ln} [k_1^+ \ k_2^+ \ k_1^- \ k_2^- \ 1]^T = 0 \quad (2.73)$$

which is easily seen to be the detailed balance constraint in Eq. 2.34. As a result, a set of bond graph parameters exist for this kinetic system only if the kinetic parameters are thermodynamically consistent, reflecting the fact that bond graphs can only represent the subset of kinetic systems that are thermodynamically consistent.

The equations shown above can be generalised to all chemical reaction networks described by mass action. The relationships between kinetic and bond graph parameters are captured in the general equation (Gawthrop et al., 2015)

$$\mathbf{Ln}(\mathbf{k}) = \mathbf{M} \mathbf{Ln}(\mathbf{W} \boldsymbol{\lambda}), \quad (2.74)$$

with partitions defined as

$$\mathbf{k} = \begin{bmatrix} k^+ \\ k^- \\ K^c \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} I_{n_r \times n_r} & N^f T \\ I_{n_r \times n_r} & N^r T \\ 0 & N^c T \end{bmatrix}, \quad \boldsymbol{\lambda} = \begin{bmatrix} \kappa \\ K \end{bmatrix}. \quad (2.75)$$

Here, k^+ and k^- are vectors of the forward and reverse kinetic rate constants respectively, $I_{n \times n}$ is an identity matrix of length n , n_r is the number of reactions, N^f and N^r are the forward and reverse stoichiometric matrices respectively, κ is a vector of the reaction rate constants, K is a vector of the species thermodynamic constants, K^c is a vector of known equilibrium constants between the species defined in the matrix N^c , and \mathbf{W} is a diagonal matrix that accounts for differences in volume between compartments. Therefore the first two rows represent the relationships between kinetic and bond graph constants, and the final row represents additional information that is known about the thermodynamic constants.

For the enzyme cycle, Eq. 2.74 gives rise to Eq. 2.72 through the substitutions

$$k^+ = \begin{bmatrix} k_1^+ \\ k_2^+ \end{bmatrix}, \quad k^- = \begin{bmatrix} k_1^- \\ k_2^- \end{bmatrix}, \quad \kappa = \begin{bmatrix} \kappa_1 \\ \kappa_2 \end{bmatrix}, \quad K = \begin{bmatrix} K_{\text{Si}} \\ K_{\text{Se}} \\ K_{E1} \\ K_{E2} \end{bmatrix}, \quad K^c = 1 \quad (2.76)$$

$$N^f = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad N^r = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad N^c = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{W} = I_{6 \times 6}. \quad (2.77)$$

As seen in the enzyme cycle example, Eq. 2.74 is not always solvable because there may exist sets of kinetic rate constants that do not satisfy detailed balance constraints, and are therefore thermodynamically inconsistent. Because the bond graph approach forces the modeller to use extra discipline in accounting for energy transfer, the exercise of converting kinetic parameters into bond graph parameters can be used to verify the thermodynamic consistency of a kinetic model. We show here that bond graph parameters can be derived from a set of kinetic parameters if they satisfy a thermodynamic constraint, known in the literature as Wegscheider conditions (Liebermeister et al., 2010). By subtracting from the first n_r rows of Eq. 2.74 the next n_r rows, we find that

$$\mathbf{Ln} \begin{bmatrix} K^{\text{eq}} \\ K^c \end{bmatrix} = \begin{bmatrix} -N^T \\ N^{cT} \end{bmatrix} \mathbf{Ln}(K), \quad (2.78)$$

where $N = N^r - N^f$ is the stoichiometric matrix. If we define $/$ as the element wise quotient operator for vectors, $K^{\text{eq}} = k^+/k^-$ is a vector containing the equilibrium constants of the reactions. Thus if Z is a right nullspace matrix (Anton and Rorres, 2014) of $[-N \ N^c]$, multiplying both sides of Eq. 2.78 by Z^T on the left gives

$$Z^T \mathbf{Ln} \begin{bmatrix} K^{\text{eq}} \\ K^c \end{bmatrix} = 0. \quad (2.79)$$

Thus for any biochemical system with mass action equations, Eq. 2.79 can be used to check that the kinetic parameters are thermodynamically consistent. As discussed in Gawthrop et al. (2015), Eq. 2.79 is a form of the Wegscheider condition, which has been associated with thermodynamic consistency (Polettini and Esposito, 2014), therefore only thermodynamically consistent kinetic parameters can be converted into bond graph parameters.

Assuming that a solution exists, one solution to Eq. 2.74 can be found using the equation

$$\boldsymbol{\lambda}_0 = \mathbf{W}^{-1} \mathbf{Exp}(\mathbf{M}^\dagger \mathbf{Ln}(\mathbf{k})), \quad (2.80)$$

where \mathbf{M}^\dagger is the Moore-Penrose pseudoinverse of \mathbf{M} (Hung and Markham, 1975) and \mathbf{Exp} is the element-wise exponential. If a solution exists, the set of bond graph parameters is generally not unique. However, as demonstrated in Gawthrop et al. (2015), the affinity and velocity of each reaction remains the same regardless of the parameters chosen.

To account for differences in volume between compartments, we include a matrix \mathbf{W} in Eq. 2.74. If a species exists within a compartment, the diagonal entry corresponding

to that species is equal to the volume of that compartment. For reaction rate constants and species that do not exist within a volume, the value of the corresponding diagonal element is 1.

In Section 2.3, we use the constraints

$$\frac{W_i K_{\text{Nai}}}{W_e K_{\text{Nae}}} = 1 \quad (2.81)$$

$$\frac{W_i K_{\text{Ki}}}{W_e K_{\text{Ke}}} = 1 \quad (2.82)$$

$$\frac{W_i K_{\text{Cai}}}{W_{\text{sr}} K_{\text{Casr}}} = 1 \quad (2.83)$$

$$\frac{W_i K_{\text{MgATP}}}{W_i K_{\text{MgADP}} W_i K_{\text{Pi}} W_{\text{isr}} K_{\text{H}}} = \exp\left(-\frac{\Delta G_{\text{MgATP}}^0}{RT}\right) M^2 = 9881 \text{ mM}^2, \quad (2.84)$$

where $\Delta G_{\text{MgATP}}^0 = 11.9 \text{ kJ/mol}$ is the standard free energy of MgATP hydrolysis at 310 K (Guynn and Veech, 1973; Tran et al., 2009). We used the volumes $W_i = 38.0 \text{ pL}$, $W_e = 5.182 \text{ pL}$, and $W_{\text{sr}} = 2.28 \text{ pL}$ (Luo and Rudy, 1994a). For the ATP hydrolysis reaction, we used the intracellular volume for MgATP, MgADP and Pi. In the Terkildsen et al. (2007) model, a volume of $W_{\text{isr}} = W_i$ was used for H^+ . The Tran et al. (2009) model describes the countertransport of Ca^{2+} and H^+ , however the steady-state equations were derived assuming that the cytosol and SR shared the same H^+ concentration. Therefore we chose to use the volume $W_{\text{isr}} = W_i + W_{\text{sr}}$.

In examples where charged species are involved (such as in § 2.2.3), the electrical dependence can be ignored for the purpose of identifying bond graph parameters, since the components corresponding to the electrical dependence (the linear **C** component, and **TF** components) have a natural bond graph representation.

The Tran et al. (2009) and Terkildsen et al. (2007) models assumed that some reactions were at rapid equilibrium, therefore we approximated them in the bond graph model by replacing the equilibrium constants with two sufficiently fast kinetic rate constants with the same equilibrium constant.

2.C Modifications to the Terkildsen et al. model of the Na^+/K^+ ATPase

2.C.1 Updates to equations

The Terkildsen *et al.* model is a 15-state model of the Na^+/K^+ ATPase (Figure 2.6). The model was reduced to a 4-state model using rapid equilibrium (Smith and Crampin, 2004), and then further simplified using steady state approximations. We found kinetic and thermodynamic issues in the implementation of this model, and resolved these issues as follows:

1. The equilibrium constants of the chemical reaction network were inconsistent with the number of binding sites. If we assume that the binding/unbinding events are identical, the kinetic rate constants are proportional to the number of binding sites (Keener and Sneyd, 2009). We have corrected the equilibrium constants, shown in red in Figure 2.6A.
2. The original model by Terkildsen et al. (2007) used an incorrect standard free energy of $\Delta G_{\text{MgATP}}^0 = -30.2 \text{ kJ/mol}$ for the hydrolysis of ATP, which resulted in an incorrect equilibrium constant. The authors appeared to adjust $\Delta G_{\text{MgATP}}^0$ to a physiological pH rather than a pH of 0. Because $\Delta G_{\text{MgATP}}^0$ was used in the detailed balance constraint

$$\frac{k_1^+ k_2^+ k_3^+ k_4^+ K_{d,\text{Na}_e}^0 (K_{d,\text{Na}_e})^2 (K_{d,\text{K}_i})^2}{k_1^- k_2^- k_3^- k_4^- K_{d,\text{Na}_i}^0 (K_{d,\text{Na}_i})^2 (K_{d,\text{K}_e})^2 K_{d,\text{MgATP}}} = \exp\left(-\frac{\Delta G_{\text{MgATP}}^0}{RT}\right), \quad (2.85)$$

the model was thermodynamically inconsistent. This error would cause over a 10^7 -fold increase in the equilibrium constant at 310 K, therefore we corrected the value to $\Delta G_{\text{MgATP}}^0 = 11.9 \text{ kJ/mol}$ (Guynn and Veech, 1973; Tran et al., 2009).

3. Terkildsen et al. (2007) applied a rapid equilibrium approximation to reduce the 15-state model into a 4-state model with modified kinetic constants that were functions of metabolite concentrations. However, due to algebraic errors, the equations for some of these modified kinetic rate constants ($\alpha_1^+, \alpha_3^+, \alpha_2^-$ and α_4^-) were incorrect. In our updated model, we have corrected the following modified rate constants to the equations below:

$$\alpha_1^+ = \frac{k_1^+ \tilde{\text{Na}}_{i,1} \tilde{\text{Na}}_{i,2}^2}{\tilde{\text{Na}}_{i,1} \tilde{\text{Na}}_{i,2}^2 + (1 + \tilde{\text{Na}}_{i,2})^2 + (1 + \tilde{K}_i)^2 - 1} \quad (2.86)$$

$$\alpha_3^+ = \frac{k_3^+ \tilde{K}_e^2}{\tilde{\text{Na}}_{e,1} \tilde{\text{Na}}_{e,2}^2 + (1 + \tilde{\text{Na}}_{e,2})^2 + (1 + \tilde{K}_e)^2 - 1} \quad (2.87)$$

$$\alpha_2^- = \frac{k_2^- \tilde{\text{Na}}_{e,1} \tilde{\text{Na}}_{e,2}^2}{\tilde{\text{Na}}_{e,1} \tilde{\text{Na}}_{e,2}^2 + (1 + \tilde{\text{Na}}_{e,2})^2 + (1 + \tilde{K}_e)^2 - 1} \quad (2.88)$$

$$\alpha_4^- = \frac{k_4^- \tilde{K}_i^2}{\tilde{\text{Na}}_{i,1} \tilde{\text{Na}}_{i,2}^2 + (1 + \tilde{\text{Na}}_{i,2})^2 + (1 + \tilde{K}_i)^2 - 1}, \quad (2.89)$$

where

$$\tilde{\text{Na}}_{i,1} = \frac{[\text{Na}^+]_{\text{i}}}{K_{d,\text{Na}_i}^0 e^{\Delta FV_m/RT}} \quad \tilde{\text{Na}}_{i,2} = \frac{[\text{Na}^+]_{\text{i}}}{K_{d,\text{Na}_i}} \quad (2.90)$$

$$\tilde{\text{Na}}_{e,1} = \frac{[\text{Na}^+]_{\text{e}}}{K_{d,\text{Na}_e}^0 e^{(1+\Delta)zFV_m/RT}} \quad \tilde{\text{Na}}_{e,2} = \frac{[\text{Na}^+]_{\text{e}}}{K_{d,\text{Na}_e}} \quad (2.91)$$

$$\tilde{K}_i = \frac{[\text{K}^+]_{\text{i}}}{K_{d,\text{K}_i}} \quad \tilde{K}_e = \frac{[\text{K}^+]_{\text{e}}}{K_{d,\text{K}_e}} \quad (2.92)$$

and Δ is the charge translocated by reaction R5.

α_1^+ was one of the modified rate constants that was updated, described in Terkildsen et al. (2007) by the incorrect equation

$$\alpha_1^+ = \frac{k_1^+ \tilde{\text{Na}}_{i,1} \tilde{\text{Na}}_{i,2}^2}{(1 + \tilde{\text{Na}}_{i,1})(1 + \tilde{\text{Na}}_{i,2})^2 + (1 + \tilde{K}_i)^2 - 1}. \quad (2.93)$$

Since the modified rate constant is k_1^+ scaled by the proportion of state 6 with respect to the total amount of states 1 to 6, a correct expression for the modified rate constant can be derived as follows:

$$\begin{aligned}
 \alpha_1^+ &= k_1^+ \frac{x_6}{x_6 + x_5 + x_4 + x_3 + x_2 + x_1} \\
 &= k_1^+ \frac{1}{1 + x_5/x_6 + x_4/x_6 + x_3/x_6 + x_2/x_6 + x_1/x_6} \\
 &= \frac{k_1^+}{1 + 2\tilde{N}a_{i,1}^{-1} + 2\tilde{N}a_{i,1}^{-1}\tilde{N}a_{i,2}^{-1} + \tilde{N}a_{i,1}^{-1}\tilde{N}a_{i,2}^{-2} + 2\tilde{N}a_{i,1}^{-1}\tilde{N}a_{i,2}^{-2}\tilde{K}_i + \tilde{N}a_{i,1}^{-1}\tilde{N}a_{i,2}^{-2}\tilde{K}_i^2} \\
 &= \frac{k_1^+\tilde{N}a_{i,1}\tilde{N}a_{i,2}^2}{\tilde{N}a_{i,1}\tilde{N}a_{i,2}^2 + (1 + \tilde{N}a_{i,2})^2 + (1 + \tilde{K}_i)^2 - 1}. \tag{2.94}
 \end{aligned}$$

The mathematical expressions for the other modified rate constants can be derived using a similar procedure.

2.C.2 Reparameterising the Terkildsen et al. model

Using the updated equations of the Terkildsen *et al.* model, and setting $\Delta G_{\text{MgATP}}^0 = 11.9$ kJ/mol for the detailed balance constraint, we reparameterised the model to the same data that the Terkildsen *et al.* model was trained on. We minimised an objective function that summarised differences between model predictions and data, using similar methods to those in Terkildsen (2006), with some minor changes:

1. For the extracellular K^+ data of Nakao and Gadsby (1989) (Figure 2.8B), the weighting for extracellular potassium above 5.4 mM was increased from $6\times$ to $15\times$ to obtain a reasonable fit at physiological concentrations.
2. To ensure that the magnitude of cycling velocity matched that of Nakao and Gadsby (1989), we chose to fit the curve for $[Na^+]_e = 150$ mM (Figure 2.7A) without normalisation.
3. Instead of using a local optimiser to minimise the objective function, we used particle swarm optimisation (Kennedy and Eberhart, 1995) followed by a local optimiser to find a global minimum.

2.D Stoichiometry and parameters for SERCA

Table 2.1: Forward stoichiometric matrix for the Tran et al. (2009) SERCA model.

Table 2.2: Reverse stoichiometric matrix for the Tran et al. (2009) SERCA model.

Table 2.3: Parameters for the Tran SERCA model. Adapted from Tran et al. (2009).

Parameter	Value	Modified kinetic constants
k_1^+	$25900 \text{ mM}^{-1} \text{s}^{-1}$	-
k_2^+	2540 s^{-1}	-
k_3^+	20.5 s^{-1}	-
k_1^-	2 s^{-1}	-
k_2^-	$67200 \text{ mM}^{-1} \text{s}^{-1}$	-
k_3^-	$149 \text{ mM}^{-1} \text{s}^{-1}$	-
$K_{d,Cai}^{\text{eq}}$	0.91 mM	$k^+ = 3.1276 \times 10^{10} \text{ mM}^{-2} \text{s}^{-1}$ $k^- = 2.5900 \times 10^{10} \text{ s}^{-1}$
$K_{d,Casr}^{\text{eq}}$	2.24 mM	$k^+ = 1.2996 \times 10^{11} \text{ s}^{-1}$ $k^- = 2.5900 \times 10^{10} \text{ mM}^{-2} \text{s}^{-1}$
$K_{d,H1}^{\text{eq}}$	$1.09 \times 10^{-5} \text{ mM}$	$k^+ = 2.3761 \times 10^{15} \text{ mM}^{-1} \text{s}^{-1}$ $k^- = 2.5900 \times 10^{10} \text{ s}^{-1}$
$K_{d,Hi}^{\text{eq}}$	$3.54 \times 10^{-3} \text{ mM}^2$	$k^+ = 2.5900 \times 10^{10} \text{ s}^{-1}$ $k^- = 7.3164 \times 10^{12} \text{ mM}^{-2} \text{s}^{-1}$
$K_{d,Hsr}^{\text{eq}}$	$1.05 \times 10^{-8} \text{ mM}^2$	$k^+ = 2.4667 \times 10^{18} \text{ mM}^{-2} \text{s}^{-1}$ $k^- = 2.5900 \times 10^{10} \text{ s}^{-1}$
$K_{d,H}^{\text{eq}}$	$7.24 \times 10^{-5} \text{ mM}$	$k^+ = 2.5900 \times 10^{10} \text{ s}^{-1}$ $k^- = 3.5773 \times 10^{14} \text{ mM}^{-1} \text{s}^{-1}$
n	2.0	-

Table 2.4: Parameters for the bond graph model of the SERCA pump. The reaction parameters have unit fmol/s, and the species parameters have unit fmol⁻¹.

Species/Reaction	Parameter	Value
R1 → 2	κ_{12}	0.00053004
R2 → 4	κ_{24}	1567.7476
R2 → 2a	κ_{22a}	8326784.0537
R4 → 5	κ_{45}	1567.7476
R5 → 6	κ_{56}	3063.4006
R6 → 8	κ_{68}	130852.3839
R8 → 9	κ_{89}	11612934.8748
R9 → 10	κ_{910}	11612934.8748
R10 → 1	κ_{101}	0.049926
P ₁	K_1	5263.6085
P ₂	K_2	3803.6518
P _{2a}	K_{2a}	3110.4445
P ₄	K_4	16520516.1239
P ₅	K_5	0.82914
P ₆	K_6	993148.433
P ₈	K_8	37.7379
P ₉	K_9	2230.2717
P ₁₀	K_{10}	410.6048
H ⁺	K_H	1862.5406
Ca _i	K_{Cai}	1.9058
Ca _{sr}	K_{Casr}	31.764
MgATP	K_{MgATP}	244.3021
MgADP	K_{MgADP}	5.8126×10^{-7}
P _i	K_{Pi}	0.014921

2.E Stoichiometry and parameters for Na^+/K^+ ATPase

Table 2.5: Forward stoichiometric matrix for the Terkildsen et al. (2007) model of Na^+/K^+ ATPase.

Table 2.6: Reverse stoichiometric matrix for the Terkildsen et al. (2007) model of Na^+/K^+ ATPase.

	Reactions														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
P2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
P3	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
P4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
P5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
P6	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
P7	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
P8	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
P9	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
P10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
P11	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
P12	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
P13	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
P14	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
P15	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
K_i^+	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
K_e^+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Na_i^+	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Na_e^+	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
MgATP	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MgADP	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
Pi	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
H^+	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0

Table 2.7: Kinetic parameters for the updated cardiac Na^+/K^+ ATPase model.

Parameter	Description	Value
k_1^+	Forward rate constant of reaction R6	1423.2 s^{-1}
k_1^-	Reverse rate constant of reaction R6	225.9048 s^{-1}
k_2^+	Forward rate constant of reaction R7	$11564.8064 \text{ s}^{-1}$
k_2^-	Reverse rate constant of reaction R7	$36355.3201 \text{ s}^{-1}$
k_3^+	Forward rate constant of reaction R13	194.4506 s^{-1}
k_3^-	Reverse rate constant of reaction R13	$281037.2758 \text{ mM}^{-2}\text{s}^{-1}$
k_4^+	Forward rate constant of reaction R15	$30629.8836 \text{ s}^{-1}$
k_4^-	Reverse rate constant of reaction R15	$1.574 \times 10^6 \text{ s}^{-1}$
$K_{d,\text{Nai}}^0$	Voltage-dependent dissociation constant of intracellular Na^+	579.7295 mM
$K_{d,\text{Nae}}^0$	Voltage-dependent dissociation constant of extracellular Na^+	0.034879 mM
$K_{d,\text{Nai}}$	Voltage-independent dissociation constant of intracellular Na^+	5.6399 mM
$K_{d,\text{Nae}}$	Voltage-independent dissociation constant of extracellular Na^+	10616.9377 mM
$K_{d,\text{Ki}}$	Dissociation constant of intracellular K^+	16794.976 mM
$K_{d,\text{Ke}}$	Dissociation constant of extracellular K^+	1.0817 mM
$K_{d,\text{MgATP}}$	Dissociation constant of MgATP	140.3709 mM
Δ	Charge translocated by reaction R5	-0.0550
Pump density	Number of pumps per μm^2	$1360.2624 \mu\text{m}^{-2}$

Table 2.8: Parameters for the bond graph version of the updated cardiac Na^+/K^+ ATPase model.

Component	Description	Parameter	Value
R1	Reaction R1	κ_1	330.5462 fmol/s
R2	Reaction R2	κ_2	132850.9145 fmol/s
R3	Reaction R3	κ_3	200356.0223 fmol/s
R4	Reaction R4	κ_4	2238785.3951 fmol/s
R5	Reaction R5	κ_5	10787.9052 fmol/s
R6	Reaction R6	κ_6	15.3533 fmol/s
R7	Reaction R7	κ_7	2.3822 fmol/s
R8	Reaction R8	κ_8	2.2855 fmol/s
R9	Reaction R9	κ_9	1540.1349 fmol/s
R10	Reaction R10	κ_{10}	259461.6507 fmol/s
R11	Reaction R11	κ_{11}	172042.3334 fmol/s
R12	Reaction R12	κ_{12}	6646440.3909 fmol/s
R13	Reaction R13	κ_{13}	597.4136 fmol/s
R14	Reaction R14	κ_{14}	70.9823 fmol/s
R15	Reaction R15	κ_{15}	0.015489 fmol/s
P ₁	Pump state ATP-E _i K ₂	K_1	101619537.2009 fmol ⁻¹
P ₂	Pump state ATP-E _i K ₁	K_2	63209.8623 fmol ⁻¹
P ₃	Pump state ATP-E _i	K_3	157.2724 fmol ⁻¹
P ₄	Pump state ATP-E _i Na ₁	K_4	14.0748 fmol ⁻¹
P ₅	Pump state ATP-E _i Na ₂	K_5	5.0384 fmol ⁻¹
P ₆	Pump state ATP-E _i Na ₃	K_6	92.6964 fmol ⁻¹
P ₇	Pump state P-E _i (Na ₃)	K_7	4854.5924 fmol ⁻¹
P ₈	Pump state P-E _e Na ₃	K_8	15260.9786 fmol ⁻¹
P ₉	Pump state P-E _e Na ₂	K_9	13787022.8009 fmol ⁻¹
P ₁₀	Pump state P-E _e Na ₁	K_{10}	20459.5509 fmol ⁻¹
P ₁₁	Pump state P-E _e	K_{11}	121.4456 fmol ⁻¹
P ₁₂	Pump state P-E _e K ₁	K_{12}	3.1436 fmol ⁻¹
P ₁₃	Pump state P-E _e K ₂	K_{13}	0.32549 fmol ⁻¹
P ₁₄	Pump state E _e (K ₂)	K_{14}	156.3283 fmol ⁻¹
P ₁₅	Pump state ATP-Ee (K ₂)	K_{15}	1977546.8577 fmol ⁻¹
K _i	Intracellular K _i ⁺	K_{Ki}	0.0012595 fmol ⁻¹
K _e	Extracellular K _e ⁺	K_{Ke}	0.009236 fmol ⁻¹
Nai	Intracellular Na _i ⁺	K_{Nai}	0.00083514 fmol ⁻¹
Nae	Extracellular Na _e ⁺	K_{Nae}	0.0061242 fmol ⁻¹
MgATP	Intracellular MgATP	K_{MgATP}	2.3715 fmol ⁻¹
MgADP	Intracellular MgADP	K_{MgADP}	7.976×10^{-5} fmol ⁻¹
P _i	Free inorganic phosphate	K_{Pi}	0.04565 fmol ⁻¹
H	Intracellular H ⁺	K_H	0.04565 fmol ⁻¹
mem	Membrane capacitance	C_m	153400 fF
zF_5	Charge translocated by R5	z_5	-0.0550
zF_8	Charge translocated by R8	z_8	-0.9450

Chapter 3

Bond graph modelling of the cardiac action potential

In the previous chapter (Chapter 2), we examined the issue of converting existing models into bond graphs. In this chapter, we investigate the issue of coupling separate bond graph models together, assuming that they have either been converted from an existing model or formulated using bond graphs from the beginning. Because bond graphs are inherently modular, separate models can be linked together using bonds to build larger models. We explore the utility of bond graphs in model integration by coupling together separate models of ion channels and ion transporters (including the model of Na^+/K^+ ATPase described in § 2.3.2) to construct a simple model of cardiac electrophysiology. This model reproduces several essential features of the cardiac action potential, and the energetic cost of an action potential is easily calculated because bond graphs explicitly model energy. We then examine how bond graphs could be used to deal with the issues of drift and non-unique steady states observed in many existing models, both of which have been linked to the conservation of charge. We show that charge conservation is generalised by the notion of “conserved moieties” within bond graph models. Therefore, bond graphs provide a rigorous and complete explanation of drift and non-unique steady states, allowing these issues to be resolved in a systematic manner. These results demonstrate that bond graphs help to flag potential inconsistencies between models when they are coupled together.

This chapter has been published in *Proceedings of the Royal Society A* (Pan et al., 2018a). The Supporting Material is included within Appendices 3.A–3.F of this chapter.

3.1 Introduction

Models of the cardiac action potential have been developed to study cardiac diseases such as arrhythmia (Luo and Rudy, 1994a,b; Faber and Rudy, 2000), ischaemia (Terkildsen et al., 2007) and acidosis (Crampin and Smith, 2006). Increasing model complexity has led to concerns over the occurrence of drift and non-unique steady states (Guan et al.,

1997; Fraser and Huang, 2007; Kneller et al., 2002), particularly for extensions of the DiFrancesco and Noble (1985) and Luo-Rudy (Luo and Rudy, 1994a; Faber and Rudy, 2000) models. While solutions to these issues have been proposed using conservation principles (Hund et al., 2001; Livshitz and Rudy, 2009), they have not been universally applied for more recent models, many of which still use nonconservative stimulus currents that predispose them to drift (Aslanidi et al., 2009; Carro et al., 2011; Grandi et al., 2010). More recently, the Food and Drug Administration (FDA) has initiated plans to use cardiac action potential models to assess potential drug side-effects on cardiac instability through the human *ether-à-go-go*-related gene (hERG) K⁺ channel. Thus, with an increasing emphasis on model robustness and accuracy, there is a renewed incentive to resolve the issues of drift and non-unique steady states (Sager et al., 2014; Colatsky et al., 2016).

Drift is the failure of a model to reach a consistent limit cycle when simulated over long periods, and is often caused by a nonconservative stimulus containing current with no charge carrier (Guan et al., 1997; Hund et al., 2001). Hund et al. (2001) derived a charge conservation law, and found that nonconservative stimulus currents violate this conservation law, hence they proposed K⁺ ions as the current charge carrier to resolve this. A related issue in many models where drift has been resolved is that steady state limit cycles under constant pacing depend upon the initial conditions and are therefore non-unique (Fraser and Huang, 2007; Hund et al., 2001; Kneller et al., 2002). Thus, depending on the initial conditions, the same model may lead to different conclusions. Like drift, authors have suggested that charge conservation can constrain initial conditions such that they lead to the same steady state (Hund et al., 2001; Kneller et al., 2002; Livshitz and Rudy, 2009).

While the studies by Hund et al. (2001) and Livshitz and Rudy (2009) suggest measures to eliminate drift and attain a unique steady state by using conservation laws, their analyses are limited in their scope and not a comprehensive solution for all models. Because existing studies (Hund et al., 2001; Livshitz and Rudy, 2009) explore charge conservation only in specific models, and the conservation laws were derived from physical intuition rather than a principled mathematical approach, it is difficult to generalise their findings to other models where charge conservation is routinely neglected. Furthermore, because these studies focus only on conservation of charge, they may miss other conservation laws relevant for long-term behaviour, such as those corresponding to ions, ion channels and buffers. A general approach is, therefore, desirable to deal with the issues of drift and steady states in a more systematic manner and for a broader range of models.

To facilitate a general approach, we propose the use of bond graphs which explicitly model energy transfer across physical systems to ensure compliance with conservation principles. Bond graphs were initially invented to model hydroelectric systems (Paynter, 1961) and they have subsequently been extended to model chemical (Borutzky et al., 1995), biochemical (Oster et al., 1973; Gawthrop and Crampin, 2014) and electrochemical systems (Gawthrop et al., 2017). As with all physical systems, biological processes must obey the fundamental principles of physics and thermodynamics (Omholt and Hunter, 2016), therefore bond graphs are well-suited for constraining models of biological systems to physically plausible solutions (Gawthrop et al., 2015), and also for inferring the energetic cost of biological processes (Gawthrop et al., 2017; Gawthrop, 2017a; Gawthrop

et al., 2015; Gawthrop and Crampin, 2017). Because the bond graph representation emphasises analogies between different physical domains, electrophysiological systems can be analysed as an analogous biochemical system with a stoichiometric matrix that describes the stoichiometry of each reaction within its columns (Beard et al., 2002, 2004a; Gawthrop and Crampin, 2014; van der Schaft et al., 2013). In this context, the “conservation principle” described in earlier studies is an example of the more general principle of a conserved moiety in metabolic and bond graph analysis (Haraldsdóttir and Fleming, 2016; Gawthrop et al., 2015).

In this study, we develop an abridged bond graph model of the cardiac action potential and outline a general approach to study the effects of conserved moieties on drift and steady-state behaviour. Our bond graph model simulates the essential features of the cardiac action potential, and because bond graphs are energy-based this easily provides an estimate of the energetic cost (in Joules) of the cardiac action potential. Our analysis reveals conservation of charge as one of the conserved moieties of our model, along with other conserved moieties corresponding to ions, channels, transporters and buffers. We observed that our model solution was subject to drift when the stimulus current violated any conservation laws corresponding to the conserved moieties, and that changes to the initial conditions led to different steady states if the value of any conserved moiety was changed. To demonstrate that our approach is general, we analyse variants of our bond graph model where different ions have been fixed at a constant concentration (corresponding to “chemostats”). It should be noted that fixing an ion concentration can change the conserved moieties of a system, therefore influencing a model’s susceptibility to drift and non-unique steady states. The bond graph approach is a useful and general method to identify and interpret conservation principles, and it can link conserved moieties to individual steady states. Our approach can be used to automatically derive charge conservation laws that are frequently neglected in existing models of cardiac electrophysiology, and we build upon existing reports (Hund et al., 2001; Livshitz and Rudy, 2009) to propose solutions for drift and non-unique steady states which work for all cardiac action potential models that can be represented using bond graphs.

3.2 Methods

3.2.1 Model components

To study the issues of drift and non-unique steady states, we built an abridged bond graph model of the cardiac action potential, with the minimal number of channels and pumps required to simulate the essential features of a cardiac action potential, and maintain ionic concentrations over long periods of simulation. We note that a series of mathematical equations can only be described using bond graphs if they describe a thermodynamically consistent system, and that no existing model of cardiac electrophysiology is entirely thermodynamically consistent. As seen in Gawthrop et al. (2017), an exact translation of existing ion channel models into the bond graph framework is generally not possible. Therefore, for some transport processes and particularly for ion channels, it was necessary to build new models under the bond graph framework, although their parameters and

structure can be chosen based on equations in existing models. Accordingly, many of the essential components of our model were based primarily on the Luo-Rudy 1994 dynamic model (Luo and Rudy, 1994a), although it is possible to use other models and/or model more sub-cellular processes. We argue in § 3.2.3 that some aspects of the equations for the ion channels described in the Luo and Rudy model are not thermodynamically consistent, and unable to be represented as bond graphs. We describe our approach to modelling ion channels in § 3.2.4. As we use simple and elementary bond graph structures based on physical principles, our bond graph representation of ion channels is a physically constrained approximation of the equations described in Luo and Rudy, although it is still able to simulate the essential features of the cardiac action potential.

Model components are shown in Figure 3.1A, together with the overall bond graph structure (Figure 3.1B). Ion channels and Ca^{2+} buffering components were based upon their representations in Luo and Rudy (1994a). The L-type Ca^+ channel in the Luo-Rudy model is permeable to calcium, sodium and potassium, but we neglected its sodium conductance as this has a relatively small contribution to the action potential. The Na^+/K^+ ATPase model was based on the model by Terkildsen et al. (2007), with modifications suggested by Pan et al. (2017) to allow conversion into a bond graph model. The equation for the $\text{Na}^+/\text{Ca}^{2+}$ exchanger (NCX) current in Luo and Rudy did not have an obvious correspondence to a bond graph structure, thus we modelled this component using a simple bond graph module that was fitted to experimental data (Kimura et al., 1987; Beuckelmann and Wier, 1989). The general approach for modelling ion channels is described later in the methods, with details on the other components, as well as further detail on ion channel modelling given in the Supporting Material.

3.2.2 Bond graph modelling

Here we briefly outline bond graph components as used in electrophysiological modelling. For a more comprehensive introduction, the texts by Gawthrop and Smith (1996) and Borutzky (2010) provide detailed descriptions of bond graph theory, and Gawthrop and Bevan (2007) provide a short tutorial for engineers. Theory for bond graph modelling of biochemical systems can be found in Oster et al. (1973); Gawthrop and Crampin (2014); Gawthrop et al. (2015); Gawthrop (2017b).

Bond graphs consist of components (representing physical objects and processes), bonds (representing the transfer of energy), and junctions (representing network structure). Each bond carries two variables: an effort e and a flow f , such that their product determines the power of the bond (i.e. $p = ef$). Thus bond graphs explicitly account for energy transfer, and are thermodynamically consistent. Because effort and flow are generalised variables, they can represent quantities from a variety of physical systems, including mechanical (e = force [N], f = velocity [m/s]), electrical (e = voltage [V], f = current [A]) and hydraulic systems (e = pressure [Pa], f = volumetric flow rate [m^3/s]) (Borutzky, 2010).

The network structure of a bond graph is specified by 0 and 1 junctions. The 0 (or effort) junctions specify that efforts of all connected bonds are equal, and thus to ensure conservation of energy through this junction, the flows of the bonds must sum to zero. In the electrical and hydraulic domains, 0 junctions represent parallel connections, whereas

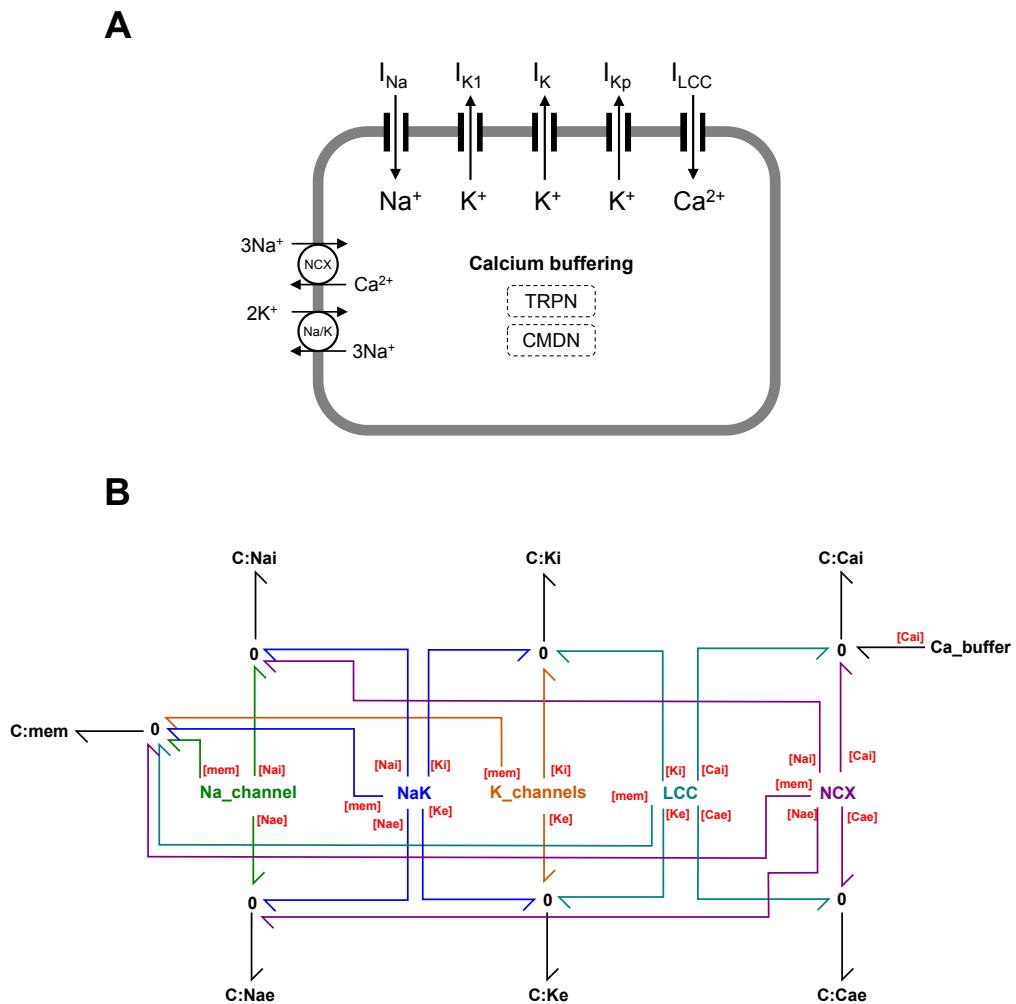


Figure 3.1: Action potential model. (A) Cell schematic; (B) Overall bond graph structure. The bond graph modules $Na_channel$, NaK , $K_channels$, LCC , NCX and Ca_buffer contain more detailed aspects of the bond graph structure which are described further in the Methods and Supporting Material. Coloured bonds link bond graph modules to the appropriate chemical species. Definitions: I_{Na} , sodium current; I_{K1} , time-independent K^+ current; I_K , time-dependent K^+ current; I_{Kp} , plateau K^+ current; I_{LCC} , L-type Ca^{2+} current; NCX , Na^+-Ca^{2+} exchanger; Na/K , Na^+/K^+ ATPase; TRPN, troponin; CMDN, calmodulin.

they represent series connections in the mechanical domain. By a similar principle, 1 (or flow) junctions specify that the flows of all connected bonds are equal, ensuring that their efforts sum to zero. Thus, 1 junctions correspond to series connections in the electrical and hydraulic domains, and parallel connections in the mechanical domain.

To illustrate the use of a bond graph for electric circuit analysis, we consider the electric circuit where two capacitors are connected to a resistor in series (Figure 3.2A). All components are linear, described by the equations:

$$V_A = \frac{q_A}{C_A} \quad (\text{capacitor}) \quad (3.1)$$

$$V_B = \frac{q_B}{C_B} \quad (\text{capacitor}) \quad (3.2)$$

$$I = \frac{V_R}{R} \quad (\text{resistor}) \quad (3.3)$$

The 1 junction enforces Kirchhoff's voltage law, such that:

$$V_R = V_A - V_B \quad (3.4)$$

Combining these equations gives rise to a system of first-order differential equations:

$$\frac{dq_A}{dt} = -I = \frac{V_B - V_A}{R} = \frac{q_B}{RC_B} - \frac{q_A}{RC_A} \quad (3.5)$$

$$\frac{dq_B}{dt} = I = \frac{V_A - V_B}{R} = \frac{q_A}{RC_A} - \frac{q_B}{RC_B} \quad (3.6)$$

It should be noted that in this example, the R and C components are linear. More generally (as required in biochemical and electrophysiological systems), bond graphs allow the definition of components with nonlinear constitutive equations to give rise to nonlinear ordinary differential equations (ODEs) that adhere to thermodynamic and physical constraints (Borutzky, 2010).

More recently, bond graphs have been extended to model biochemical systems (Oster et al., 1973; Gawthrop and Crampin, 2014) where the chemical potential μ [J/mol] is the effort variable, and molar flow rate v [mol/s] is the flow variable. Since temperature and pressure are assumed to be constant in biochemical systems, the measure of thermodynamic potential μ corresponds to Gibbs free energy (Atkins and De Paula, 2006). For bond graph modelling of more general systems where temperature and pressure change, the reader is directed to the text by Thoma and Bouamama (Thoma and Bouamama, 2000).

Because mass action equations are nonlinear in general, nonlinear components are required to model biochemical systems. Each chemical species is represented as a capacitor. However, in contrast to the electrical domain, the constitutive equation for the capacitor representing each species is logarithmic:

$$\mu = RT \ln(Kx) \quad (3.7)$$

where x [mol] is the molar amount of the species, K [mol⁻¹] is a species thermodynamic constant, $R = 8.314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$ is the gas constant and T is the absolute temperature

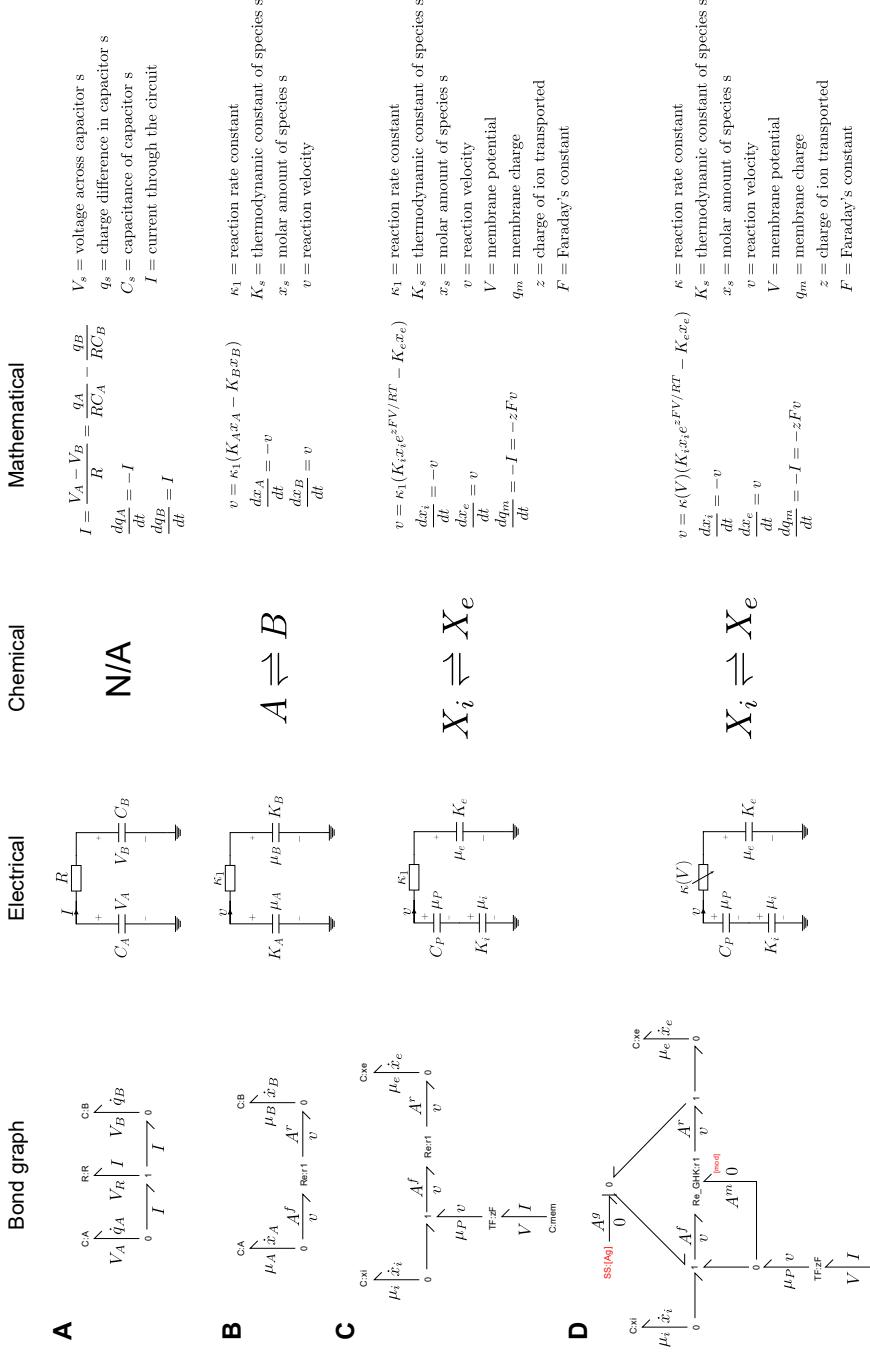


Figure 3.2: Conceptual representations of physical systems. **(A)** A bond graph for the illustrated simple electric circuit with two capacitors and a resistor in series. **(B)** A bond graph analogous to the electric circuit in (A) can also represent the chemical reaction $A \rightleftharpoons B$. **(C)** Bond graphs can also model interaction of components in both the chemical and physical domains, such as the transport of an ion across a membrane. **(D)** Transport of an ion across a membrane through an ion channel involves gating which modulates the rate of reaction. Thus the ion channel is analogous to a potentiometer.

of the system. Reactions are modelled as two-port resistors using the Marcelin-de Donder equation as the constitutive equation:

$$v = \kappa(e^{A^f/RT} - e^{A^r/RT}) \quad (3.8)$$

where κ [mol/s] is a reaction rate constant and A^f [J/mol] and A^r [J/mol] are the forward and reverse affinities, respectively. The two affinities represent the potential energies present in the reactants and products, and the reaction proceeds in the direction of decreasing potential. As illustrated by the example in Figure 3.2B, the reaction $A \rightleftharpoons B$ has a physical analogy to Figure 3.2A, with the same equivalent electric circuit. By using the constitutive equations in Eqs. 3.7 and 3.8, the reaction velocity for the bond graph model follows mass-action kinetics:

$$v = \kappa_1(e^{A^f/RT} - e^{A^r/RT}) = \kappa_1(e^{\mu_a/RT} - e^{\mu_b/RT}) = \kappa_1(K_a x_a - K_b x_b) = k^+ x_a - k^- x_b \quad (3.9)$$

where the forward and reverse rate constants are $k^+ = \kappa_1 K_a$ and $k^- = \kappa_1 K_b$. For more general chemical reaction networks, 1 junctions describe the presence of multiple reactants or products in a single reaction, whereas 0 junctions describe the involvement of a single species in multiple reactions (Gawthrop and Crampin, 2014). For some models, we may wish to keep the amount x of a species constant and this is achieved by defining the species as a “chemostat” (Polettini and Esposito, 2014). Because chemostats can be interpreted as an external flow that balances internal flows, they require energy to be pumped into (or out of) the system (Gawthrop et al., 2015).

The bond graph framework for biochemistry can be extended to electrochemical systems (Gawthrop et al., 2017) as demonstrated in Figure 3.2C, which models the transport of a positively charged species X across a membrane. It should be noted that chemical species are described with C components that have a logarithmic association, whereas the C component corresponding to the (electric) membrane potential has a linear constitutive relationship. A transformer (TF) is used to convert the membrane voltage into an equivalent chemical potential through Faraday’s constant $F = 96485$ C/mol, such that:

$$\mu_P = zFV \quad (3.10)$$

$$I = zFv \quad (3.11)$$

Thus, the reaction velocity is:

$$v = \kappa_1(e^{A^f/RT} - e^{A^r/RT}) = \kappa_1(e^{(\mu_i + \mu_P)/RT} - e^{\mu_e/RT}) = \kappa_1(K_i x_i e^{zFV/RT} - K_e x_e) \quad (3.12)$$

By setting $v = 0$ the familiar Nernst equation for the equilibrium potential can be derived (Gawthrop et al., 2017):

$$V_{\text{eq}} = \frac{RT}{zF} \ln \left(\frac{K_e x_e}{K_i x_i} \right) = \frac{RT}{zF} \ln \left(\frac{[X_e]}{[X_i]} \right) \quad (3.13)$$

where the final equality results from the relation $K_i W_i = K_e W_e$, where W_i and W_e are the intracellular and extracellular volumes, respectively (Gawthrop et al., 2017). Therefore the thermodynamic consistency of the bond graph approach enforces constraints on the

equilibrium of the ion channel. However, it should be noted that where electrical circuit representations of the membrane Nernst potential use voltage sources, the bond graph approach necessarily accounts for possible changes in ionic concentrations, and thus this “voltage source” is split into two capacitors that provide an equivalent voltage difference.

We chose to represent ion channels such that conductance was modulated by membrane voltage, both directly and indirectly through gating processes. A bond graph representation for this relationship is given in Figure 3.2D. As shown, this model has the same electrical representation as Figure 3.2C however it uses a variable resistor. The bond graph representation contains the same states, with C:xi, C:xe, and C:mem (with a transformer) connected through 0 junctions. In this case however, the Re components that describe the constitutive relation have been changed, such that Re_GHK:r1 is connected to an additional effort that modulates its velocity, and the gating affinity A^g is added to both the forward and reverse affinities to describe changes in permeability due to gating. Further detail on modelling ion channels using bond graphs is given in § 3.2.4.

3.2.3 Modelling approach

Because bond graphs constrain the equations of a model to ensure thermodynamic consistency, they only allow the representation of the subset of general mathematical models that are thermodynamically consistent. In such models, dissipative processes such as reactions and ionic currents can only proceed in the direction of decreasing chemical or electrochemical potential, and when there is no potential gradient (i.e. the process is at equilibrium), the process must stop. However, many existing models of cardiac electrophysiology are not constrained to ensure this behaviour, and in some cases, they describe physically infeasible systems that create energy out of nowhere (Gawthrop and Crampin, 2014). Therefore many existing models do not have a direct bond graph representation (Gawthrop et al., 2015). An advantage of the bond graph approach is that the discipline required to convert an existing model into a bond graph helps to highlight thermodynamic issues and inconsistencies in existing models that would have otherwise been missed or ignored (see Gawthrop et al. (2015) for an example of this process applied to a model of glycolysis). For the currents in this study, the Luo-Rudy equations for the time-dependent K⁺ and L-type Ca²⁺ channels are not thermodynamically consistent, and therefore cannot be directly converted into a bond graph model. As shown in Eq. 3.13, a consequence of the thermodynamic consistency of the bond graph framework is that there is a constraint on the equilibrium point for each ion channel that is determined by the Nernst equation. Since the equilibrium points of the time-dependent K⁺ and L-type Ca²⁺ channels do not correspond to the Nernst potential (see A.1.3 and A.1.5 of the Supporting Material), they are not thermodynamically consistent and therefore their equations cannot be translated into a dissipative bond graph component. These issues were addressed by finding a thermodynamically consistent approximation that satisfies the Nernst equation. A second issue is that it is difficult to simultaneously model open-channel currents and channel gating using the elementary components described in § 3.2.2 and Figure 3.2, thus we approximated these relations using the components described in that section rather than using more complicated components. Therefore rather than attempting to reproduce the Luo-Rudy equations exactly, we built a bond graph structure as implied by the equations in the Luo and Rudy model, and chose

parameters of our bond graph model to fit aspects of the Luo-Rudy model as closely as possible, specifically the current-voltage (I-V) curves and gating parameters. For all other components conversion into a bond graph model was more straightforward, and we used the methods of Gawthrop et al. (2015). Further information on the bond graph model, and parameter identification is given in the Supporting Material.

3.2.4 Ion channel modelling

3.2.4.1 Bond graph structure

In this section, we discuss decisions made in developing models of ion channels. The bond graph structure for the K_p channel is shown in Figure 3.3. The other channels have similar structures that follow from the discussion in this section.

3.2.4.2 Current-voltage relations

While thermodynamic properties can be used to determine how membrane voltage and ionic concentrations relate at equilibrium, they do not specify behaviour away from equilibrium. For this purpose, the current-voltage (I-V) relationship defines how the membrane voltage relates to the current through a specific channel. In many electrophysiology models (Luo and Rudy, 1994a), currents through ion channels are described through the use of linear I-V relationships based on empirical fits to data rather than fundamental physical principles. Using bond graphs, it is difficult to incorporate the effects of gating using a linear I-V equation. While linear I-V equations can be thermodynamically consistent provided they satisfy the Nernst equation in Eq. 3.13, and such equations can be represented by using a linear R component, the modulation of its conductance would require the use of signal bonds that do not necessarily represent physical processes (Borutzky, 2010). Therefore we use the Goldman-Hodgkin-Katz (GHK) equation to model ion channels, as it enables relatively simple incorporation of ion channel gating as a physics-based biochemical module (Gawthrop et al., 2017). The GHK equation defines a non-linear relationship between current I and membrane voltage V :

$$I = P \frac{z^2 F^2}{RT} V \left(\frac{c_i - c_e e^{-zFV/RT}}{1 - e^{-zFV/RT}} \right) \quad (3.14)$$

where c_i and c_e are the ion's intracellular and extracellular concentrations respectively (Keener and Sneyd, 2009). A bond graph formulation of this relationship requires three potentials: the forward and reverse affinities A^f and A^r that correspond to the intracellular and extracellular ion respectively, and a modulation affinity A^m to provide the nonlinear scaling of conductance required to describe the GHK equation. Thus, in a bond graph, the GHK equation for current can be described by a modulated Re component (Re_GHK) with a single modulator (see Figure 3.2D, left panel and Figure 3.3A), using the constitutive equation from Gawthrop et al. (2017):

$$v = \begin{cases} \kappa \frac{\frac{A^m}{RT}}{\exp(\frac{A^m}{RT}) - 1} \left[\exp\left(\frac{A^f}{RT}\right) - \exp\left(\frac{A^r}{RT}\right) \right], & A^m \neq 0 \\ \kappa \left[\exp\left(\frac{A^f}{RT}\right) - \exp\left(\frac{A^r}{RT}\right) \right], & A^m = 0 \end{cases} \quad (3.15)$$

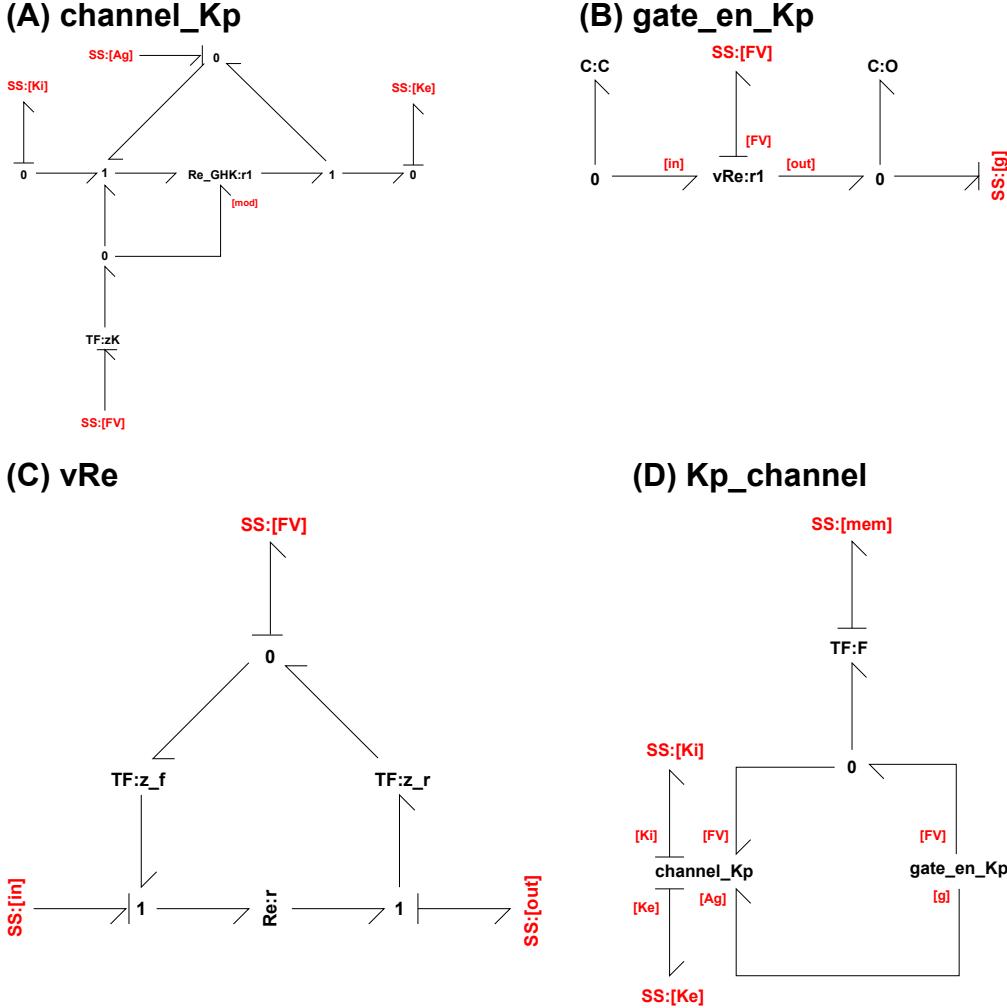


Figure 3.3: The bond graph model of a plateau K^+ channel. (A) The **channel_Kp** module describes the current through the ion channel. (B) The **gate_en_Kp** contains the states required for gating. (C) The **vRe** module contains a voltage-dependent reaction used to describe channel state transitions. (D) The channel current and gating modules are combined into an ion channel model (**Kp_channel**).

The flux for $A^m = 0$ is defined to be $\lim_{A^m \rightarrow 0} v$ to avoid a zero denominator. Then, as discussed in Gawthrop et al. (2017), setting

$$A^f = \mu_i + zFV \quad (3.16)$$

$$A^r = \mu_e \quad (3.17)$$

$$A^m = zFV \quad (3.18)$$

gives rise to the GHK equation. This can be seen by substituting Eqs. 3.16–3.18 and Eq. 3.7 into Eq. 3.15, noting that $K_i W_i = K_e W_e$. Therefore a bond graph implementation of the GHK equation is achieved by connecting the respective ports of the Re_GHK component such that the correct affinities for A^f , A^r and A^m are obtained (Figure 3.2D).

Since many ion channels in the Luo-Rudy model are described using a linear I-V relationship, the use of GHK equations requires approximation. A comparison of the

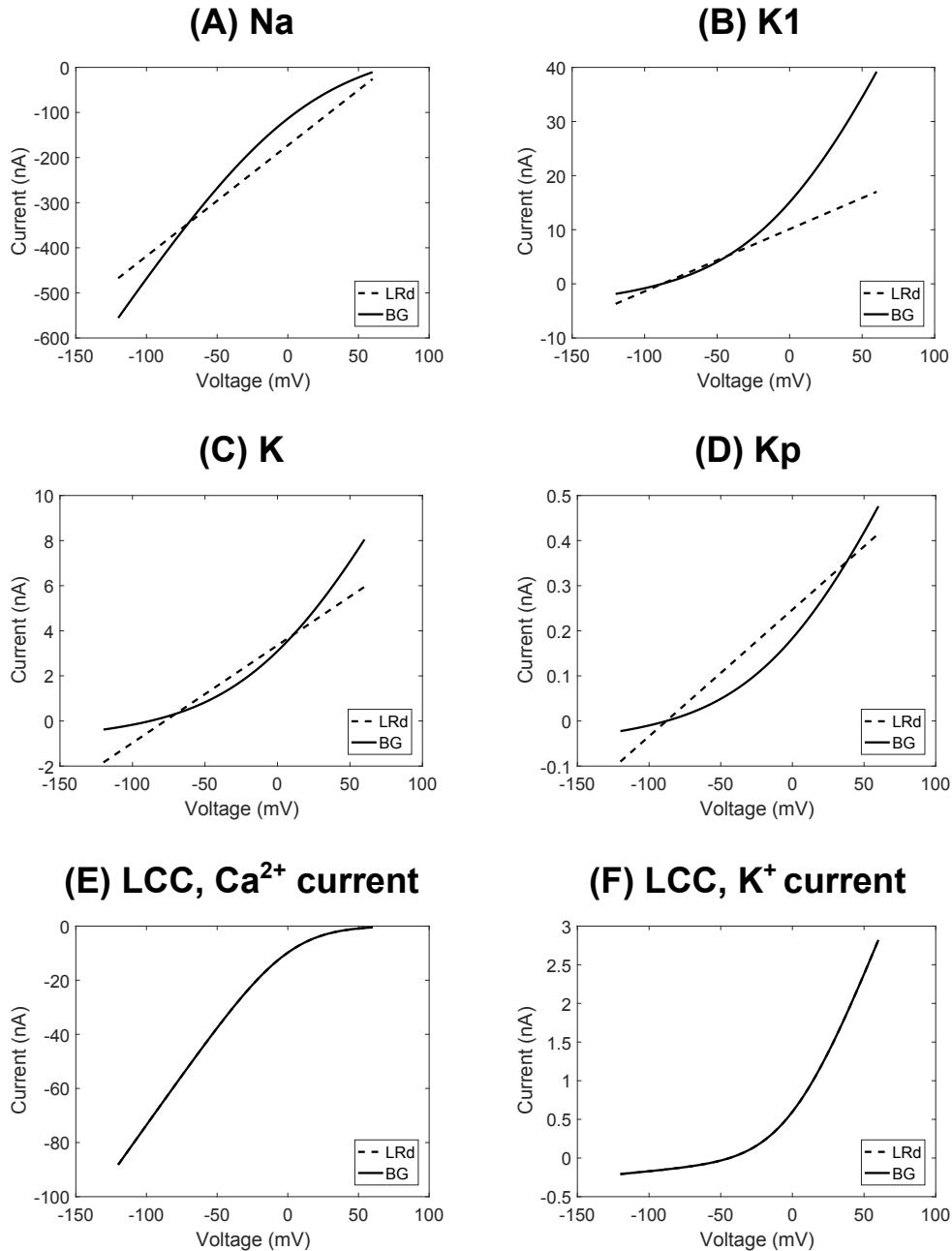


Figure 3.4: Comparison of I-V curves between the Luo-Rudy (LRd) and bond graph (BG) models. (A) I_{Na} ; (B) $I_{\text{K}1}$; (C) I_{K} ; (D) I_{Kp} ; (E) $I_{\text{Ca,L}}$; (F) $I_{\text{K,L}}$. The standard concentrations in Luo and Rudy (Luo and Rudy, 1994a) ($[\text{Na}_i^+] = 10 \text{ mM}$, $[\text{Na}_e^+] = 140 \text{ mM}$, $[\text{K}_i^+] = 145 \text{ mM}$, $[\text{K}_e^+] = 5.4 \text{ mM}$, $[\text{Ca}_i^+] = 0.12 \mu\text{M}$, $[\text{Ca}_e^+] = 1.8 \text{ mM}$) were used to match I-V curves.

resulting I-V curves is given in Figure 3.4. Details of the fitting process, as well as the fitted bond graph parameters are given in the Supporting Material. The Na^+ channel I-V curves appeared to match reasonably well (Figure 3.4A), with some discrepancies at positive membrane potentials. For K^+ channels (Figure 3.4B–D), we attempted to optimise the fit across voltages that correspond to their physiological function, so that their currents would be most similar to those of Luo and Rudy when the ion channels are open. Accordingly, for $I_{\text{K}1}$ ($-90 \text{ mV} \leq V < -30 \text{ mV}$), I_{K} ($-20 \text{ mV} \leq V \leq 30 \text{ mV}$) and I_{Kp} ($V > 0 \text{ mV}$) the I-V curves matched reasonably well in these regions. Discrepancies occurred outside these ranges of voltages, but appeared to only cause minor differences to the currents. In their implementation of I_{K} , Luo and Rudy (1994a) use a thermodynamically inconsistent I-V equation where the current is nonzero at the Nernst potential for K^+ . Despite this, bond graph parameters could still be chosen to give a reasonable fit to this I-V equation (Figure 3.4C). Because the Luo-Rudy model based their L-type Ca^{2+} I-V curves on the GHK equation, there was a far closer match between the bond graph and Luo-Rudy models for these currents, (Figure 3.4E,F) and the K^+ curve was matched exactly (Figure 3.4F).

3.2.4.3 Modulation

While the I-V curves describe currents through open ion channels, a formulation for gating is required to describe the number of open ion channels at any given time. In the Hodgkin-Huxley framework, gating is modelled as differential equations that give the proportion of open gates at any given time. We incorporated the effects of gating through a gating affinity A^g , which is added to both the forward and reverse affinities of a reaction (Figure 3.3A) to modulate its rate without changing the equilibrium (Gawthrop et al., 2017).

3.2.4.4 State models

Ion channel models must account for gating and bond graphs require the use of physical components to achieve this. We model gating as transitions between channel states, known in the literature as Markov models (Rudy and Silva, 2006; Fink and Noble, 2009). To illustrate, we use the example of a typical Na^+ channel in which the current I is described by the equation

$$I = m^3 h \bar{I} \quad (3.19)$$

where \bar{I} is the current when all channels are open. This can be described using the reaction scheme in Figure 3.5, where S_{31} represents the open channel. Because individual channel states are modelled, the current depends only on the amount of S_{31} and not any of the other closed states. Thus, incorporation into the gating framework described above is intuitive; each state represents a structural conformation of the ion channel and the number of channels in each state are explicitly tracked, facilitating a simple approach to account for the energetics of gating under varying ion channel densities.

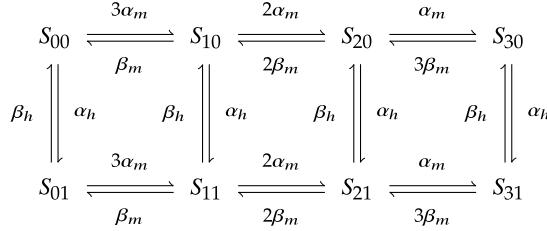


Figure 3.5: Channel states of a Na^+ channel.

3.2.4.5 Voltage dependence of state transitions

The transition rates between open and closed states are voltage-dependent for ion channels. Hodgkin-Huxley models describe state transitions using ODEs of the form

$$\frac{dg}{dt} = \alpha(V)(1 - g) - \beta(V)g \quad (3.20)$$

where g is a gating variable such as m , h or n . In typical Hodgkin-Huxley models, these rates $\alpha(V)$ and $\beta(V)$ are modelled using mathematical expressions on the basis of empirical fits to data. However, the gate transitions of the ion channels described in this study are voltage-dependent, thus they must result from the movement of charge through an electric field (Hodgkin and Huxley, 1952). Accordingly, due to the physics-based nature of bond graphs, the open and closed states of channels must be explicitly modelled as physical components with a restricted set of constitutive equations. We note that the equations used to fit to data are not necessarily thermodynamically consistent. This is particularly common for fits to experimental measurements that do not capture the equilibria of the processes examined. Therefore, because common expressions for $\alpha(V)$ and $\beta(V)$ do not obey the laws of thermodynamics, bond graphs are unable to perfectly replicate existing mathematical expressions used for ion channel transition rates. We chose to model state transitions by using the module described in Figure 3.3B, which results in an exponential voltage-dependence for both the forward and reverse reactions. In the case of the plateau K^+ channel, the rate of transition from the closed state (C) to the open state (O) is:

$$v = \alpha_0 \exp\left(\frac{z_f FV}{RT}\right) x_A - \beta_0 \exp\left(\frac{z_r FV}{RT}\right) x_B \quad (3.21)$$

where

$$\alpha_0 = \kappa K_C \quad (3.22)$$

$$\beta_0 = \kappa K_O \quad (3.23)$$

The parameters α_0 , z_f , β_0 and z_r are fitted against mathematical equations in the original model, and then used to determine the bond graph parameters. It is acceptable to fit the kinetic parameters α_0 and β_0 to reduce computation time since the equilibrium points of state transitions are not specified. The incorporation of voltage-dependence for the f -gate was more involved, and discussed in further detail in the Supporting Material.

To assess the quality of fit we compare steady-state open probabilities $g_{ss} = \alpha(V)/(\alpha(V) + \beta(V))$ and time constants $\tau = 1/(\alpha(V) + \beta(V))$ (Figure 3.6). The curves for g_{ss} and τ were

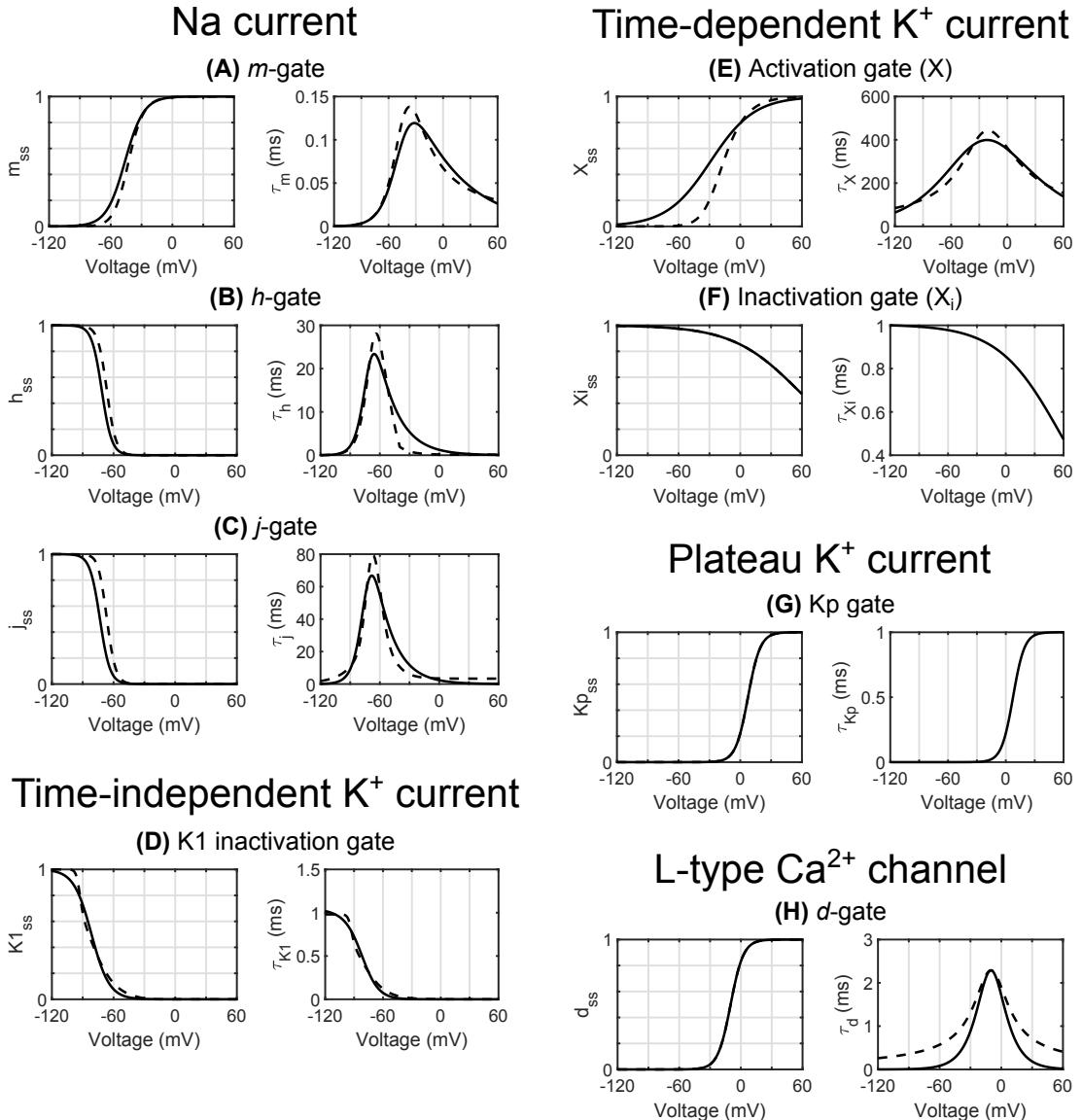


Figure 3.6: Fits for bond graph (BG) parameters against corresponding gating equations from the Luo-Rudy (LRd) model. Steady-state open probabilities are shown on left panels, and time constants are shown on the right. The bond graph equations are plotted with solid lines, and the Luo and Rudy equations in dashed lines. Gates include (A) m , sodium activation; (B) h , sodium inactivation; (C) j , slow sodium inactivation; (D) K_1 , time-independent K^+ activation; (E) X , time-dependent K^+ activation; (F) X_i , time-dependent K^+ inactivation; (G) K_p , plateau K^+ activation; (H) d , L-type Ca^{2+} channel activation. Note that the X_i and K_p gates were originally formulated as steady-state equations, thus time constants are shown only for matched bond graph parameters.

generally in agreement however there were some exceptions. In particular, time constants for the Na^+ channel gates have lower peaks in the bond graph model when compared to the Luo-Rudy model (Figure 3.6A–C), but this did not appear to significantly affect Na^+ channel function as the peaks were all decreased by a similar proportion, facilitating coordination between opening and closing. Similarly, the time constant τ_d (Figure 3.6H) was lower in the bond graph model for some voltages, but given that discrepancies occur at time constants much smaller than the time course of a cardiac action potential we expect that the effects would be negligible. Finally, for the time-dependent K^+ current X_{ss} is substantially higher at negative voltages so that the bond graph model can provide a better match at positive voltages (Figure 3.6E). The effects of this difference are partially offset by the lower GHK current at negative voltages which are still above the Nernst potential of K^+ (Figure 3.4C).

3.2.5 Finding conserved moieties

Within a biochemical model, conserved moieties are chemical structures that are neither created, removed nor broken down. A common example in energy-dependent metabolic networks is the adenosine moiety found in AMP, ADP and ATP (Haraldsdóttir and Fleming, 2016; Gawthrop et al., 2015). Mass balance specifies that the total amount of each conserved moiety remains constant, and if information on the molecular structure of each species of a reaction network is available, these conservation laws can be derived by counting the number of moieties across all species (Haraldsdóttir and Fleming, 2016). In practice, many models do not contain this structural information and this approach cannot be used, however the conservation laws still hold. Here we outline a method to find conserved moieties using stoichiometric information rather than chemical structures.

Models of cardiac electrophysiology can be represented by the differential equation

$$\dot{X} = N\mathbf{V}(X) \quad (3.24)$$

where $X(t)$ is a vector of each state (such as species, or charge difference across a membrane), N is the stoichiometric matrix (Beard et al., 2002, 2004a; Gawthrop and Crampin, 2014; van der Schaft et al., 2013), and $\mathbf{V}(X)$ is a vector of fluxes (such as reaction velocities, or ion channel currents) (Gawthrop and Crampin, 2014; Liebermeister et al., 2010; Beard et al., 2002), which is in general a nonlinear function of X . If the model contains chemostats, the entries of X , and rows of N corresponding to the chemostats are deleted prior to performing the above analysis (Polettini and Esposito, 2014). Using results from biochemical systems (Gawthrop and Crampin, 2014), if g is a row vector in the left nullspace of N , i.e. $gN = 0$, then

$$g\dot{X} = gN\mathbf{V}(X) = 0 \quad (3.25)$$

Therefore the linear combination gX is constant for the duration of the simulation. We call the linear combination of species represented by gX a conserved moiety. The space of all conserved moieties can be described by a left nullspace matrix G , whereby linear combinations of the rows of G give all possible conserved moieties gX (Palsson, 2006; Klipp et al., 2009). Because the left nullspace of N is a vector subspace and the rows of G form a basis for this subspace, G accounts for all conservation laws. A further advantage of using the left nullspace matrix is that it does not require information on chemical structures.

The left nullspace matrix for any given system is generally not unique, however there are many well-established techniques for calculating nullspace matrices (Anton and Rorres, 2014), specialised approaches for finding meaningful conserved moieties in biochemical networks (Schuster and Hilgetag, 1995; Schuster and Höfer, 1991; Haraldsdóttir and Fleming, 2016) as well as methods for finding conserved moieties from the junction structure of a bond graph (Gawthrop, 2017b). In this study, we chose conserved moieties with clear physical interpretations (Schuster and Höfer, 1991), but our conclusions hold regardless of our choice of the left nullspace matrix.

3.2.6 Stimulus currents

The cardiac action potential model was stimulated using a constant current stimulus that contained enough charge to raise the membrane potential by 30 mV over 0.1 ms. The nonconservative stimulus currents consist of non-specific charge that is not carried by any specific ion. As recommended by Kneller et al. (2002), conservative stimulus currents contained K⁺ ions as the charge carrier.

3.3 Results

3.3.1 Simulation of a single action potential

To verify that our bond graph model reproduced the features of a typical action potential we simulated the model over a single beat (Figure 3.7A–C). The membrane potential (Figure 3.7A, with stimulation indicated by the arrow) resembled a typical cardiac action potential, with a distinct peak and plateau phase. The contributions of ion channel currents reproduce some common features of cardiac action potentials (Figure 3.7B). Once the action potential is initiated by a stimulus current, the sodium current I_{Na} briefly activates to give rise to a voltage spike. Following this, the plateau phase occurs where depolarising L-type Ca²⁺ currents oppose the repolarising K⁺ currents I_K and I_{K_p} . Towards the end of the action potential, I_{K1} activates to restore the resting potential (Noble and Rudy, 2001). Our model also simulates the reversal of NCX current across the action potential, and the consistent outward current of the Na⁺/K⁺ ATPase to maintain ionic gradients (Figure 3.7C). As a consequence of incorporating the voltage-dependence of gating transitions in a physical framework, transitions between channel states are associated with a gating current resulting from the movement of such charged residues. Our model reveals that the total gating current across all channels I_{gate} has minimal contribution to total current (Figure 3.7C).

Figure 3.7E shows the power consumption of the membrane model over three cardiac cycles which was integrated to estimate the energetic cost of the cardiac action potential (Figure 3.7F). Note that energy continues to be consumed even during the resting state due the presence of currents associated with ion transporters. Thus while energy is predominantly consumed during the action potential, there is a positive gradient between action potentials (Figure 3.7F). By setting the energy consumption at the start of the

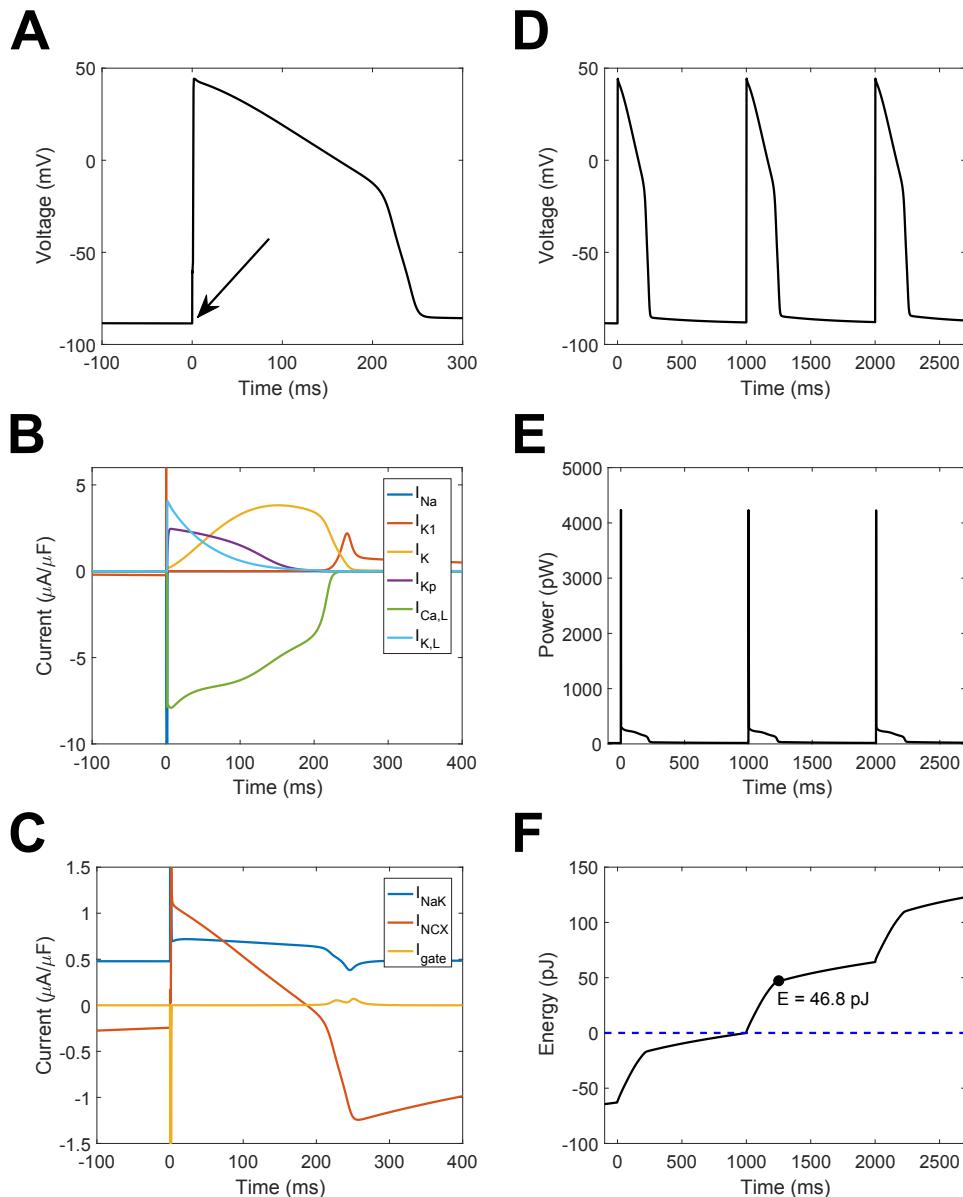


Figure 3.7: A simulation of the cardiac action potential using a bond graph model. (A) Membrane voltage, following stimulation with a conservative stimulus current (arrow); (B) Ion channel currents; (C) Transporter and gating currents; (D) Membrane voltage over three cycles, for comparison with (E) and (F); (E) Power consumption; (F) Energy dissipated, with the variable E representing the energy consumption over the duration of the action potential. The model was run initially for 300 ms to allow the membrane potential and channel gates to stabilise. The intracellular ion concentrations were dynamic variables with initial concentrations $[\text{Na}_i^+]$ = 10 mM, $[\text{K}_i^+]$ = 145 mM and $[\text{Ca}_i^+]$ = 0.12 μM . Constant concentrations were $[\text{Na}_e^+]$ = 140 mM, $[\text{K}_e^+]$ = 5.4 mM, $[\text{Ca}_e^+]$ = 1.8 mM, $[\text{MgATP}]$ = 6.95 mM, $[\text{MgADP}]$ = 0.035 mM, $[\text{Pi}]$ = 0.3971 mM and pH = 7.095. T = 310 K.

second action potential to zero (Figure 3.7F, dotted blue line), we calculated the energetic cost over the duration of the action potential to be 46.8 pJ. Since the capacitive area of membrane for this model is $1.534 \times 10^{-4} \text{ cm}^2$, the energy consumed per unit membrane area is 305 nJ/cm². When compared to Gawthrop et al.'s (2017) estimate of 173 nJ/cm² for the energetic cost of an action potential in the giant axon of a squid, the cardiac action potential uses 76% more energy. The main reason for this difference is that in contrast to a neuron, the cardiac action potential contains a plateau phase with opposing currents. Despite the relatively slow rate of change in voltage, the Ca²⁺ and K⁺ currents remain relatively high, therefore a large amount of energy is dissipated during the plateau phase.

3.3.2 Chemostats influence the conserved moieties of cardiac action potential models

Because the earliest models of the cardiac action potential did not include active transporters, they used constant intracellular concentrations to maintain ionic gradients across multiple cardiac cycles (DiFrancesco and Noble, 1985; Luo and Rudy, 1991). Later models incorporated ion transporters, allowing them to represent physiological conditions with dynamic intracellular ion concentrations, and constant extracellular ion concentrations to model washout from the circulatory system (Luo and Rudy, 1994a; Faber and Rudy, 2000). Under ischaemic conditions, washout is greatly inhibited, thus models of ischaemia use dynamic extracellular ion concentrations (Terkildsen et al., 2007). We investigated the issue of drift in three classes of model: those with (A) dynamic ion concentrations on both sides of the membrane, representing models of myocytes under ischaemic conditions; (B) dynamic intracellular ion concentrations but constant extracellular ion concentrations, representing models of myocytes under physiological conditions; and (C) constant ion concentrations, representing models without transporters.

We used our bond graph model to represent these classes of models, selecting ions to fix at constant concentrations that resulted in three variants representative of the classes listed above. Thus variant A represents models of cardiomyocytes under ischaemic conditions (Terkildsen et al., 2007), variant B represents models of cardiomyocytes under physiological conditions (Luo and Rudy, 1994a; Faber and Rudy, 2000), and variant C represents models without ion transporters (DiFrancesco and Noble, 1985; Luo and Rudy, 1991). Conserved moieties of each variant were found using the left nullspace matrix of the stoichiometric matrix (Table 3.1), and these include for example, the total amount of K₁ channel (moiety 1). Because the channel is neither synthesised nor degraded in our model, the total amount of channel, i.e. the sum of its closed (C_{K_1}) and open (O_{K_1}) states, remains constant over the course of a simulation.

Similarly, moiety 10 for variant (A) represents the total amount of K⁺ ions, which includes intracellular K⁺, extracellular K⁺ and the K⁺ ions bound to Na⁺/K⁺ ATPase. The total amount of K⁺ is constant when ion concentrations are dynamic. However, because fixing the concentration of K⁺ requires an additional external flux, the conservation law is broken in variants (B) and (C). Because the membrane capacitance is included in the stoichiometry of the system, our method automatically identifies a charge conservation law (moiety 13 for variant (A), and moiety 10 for variant (B)).

Table 3.1: Conserved moieties associated with chemostat selection. Across some biochemical subgroups (“Moiety”), there are constraints (“Conserved quantity”) on a corresponding sum of species representing the total of the moiety. The conserved quantities remain constant over the course of a simulation. Q represents contributions of other species to charge imbalance across the membrane. The symbol Σ represents charge contributions from Markov states of channels and transporters. The definition of Σ , and all species can be found in the Supporting Material and code.

Moiety	Conserved quantity
<i>Conserved moieties common to all variants (A,B,C)</i>	
1 K1 channel	$C_{K1} + O_{K1}$
2 K channel	$S_{00,K} + S_{10,K} + S_{20,K} + S_{01,K} + S_{11,K} + S_{21,K}$
3 Kp channel	$C_{Kp} + O_{Kp}$
4 Na channel	$S_{000,Na} + S_{100,Na} + S_{200,Na} + S_{300,Na} + S_{010,Na} + S_{110,Na} + S_{210,Na} + S_{310,Na} + S_{001,Na} + S_{101,Na} + S_{201,Na} + S_{301,Na} + S_{011,Na} + S_{111,Na} + S_{211,Na} + S_{311,Na}$
5 LCC	$S_{000,LCC} + S_{010,LCC} + S_{020,LCC} + S_{100,LCC} + S_{110,LCC} + S_{120,LCC} + S_{001,LCC} + S_{011,LCC} + S_{021,LCC} + S_{101,LCC} + S_{111,LCC} + S_{121,LCC}$
6 Na^+/K^+ ATPase	$P1_{NaK} + P2_{NaK} + P3_{NaK} + P4_{NaK} + P5_{NaK} + P6_{NaK} + P7_{NaK} + P8_{NaK} + P9_{NaK} + P10_{NaK} + P11_{NaK} + P12_{NaK} + P13_{NaK} + P14_{NaK} + P15_{NaK}$
7 NCX	$P1_{NCX} + P2_{NCX} + P3_{NCX} + P4_{NCX} + P5_{NCX} + P6_{NCX}$
8 Troponin	$TRPN + TRPNCa$
9 Calmodulin	$CMDN + CMDNCa$
<i>Dynamic ion concentrations (A)</i>	
Chemostats	$MgADP, MgATP, Pi, H^+$
10 K^+ ion	$K_i^+ + K_e^+ + 2P1_{NaK} + P2_{NaK} + P12_{NaK} + 2P13_{NaK} + 2P14_{NaK} + 2P15_{NaK}$
11 Na^+ ion	$Na_i^+ + Na_e^+ + P4_{NaK} + 2P5_{NaK} + 3P6_{NaK} + 3P7_{NaK} + 3P8_{NaK} + 2P9_{NaK} + P10_{NaK} + 3P1_{NCX} + 3P6_{NCX}$
12 Ca^{2+} ion	$Ca_i^{2+} + Ca_e^{2+} + 2S_{001,LCC} + 2S_{011,LCC} + 2S_{021,LCC} + 2S_{101,LCC} + 2S_{111,LCC} + 2S_{121,LCC} + P3_{NCX} + P4_{NCX} + TRPNCa + CMDNCa$
13 Charge	$Q - K_i^+ - Na_i^+ - 2Ca_i^{2+} + 2TRPN + 2CMDN + \Sigma$
<i>Dynamic intracellular ion concentrations (B)</i>	
Chemostats	$MgADP, MgATP, Pi, H^+, K_e^+, Na_e^+, Ca_e^{2+}$
10 Charge	$Q - K_i^+ - Na_i^+ - 2Ca_i^{2+} + 2TRPN + 2CMDN + \Sigma$
<i>Constant ion concentrations (C)</i>	
Chemostats	$MgADP, MgATP, Pi, H^+, K_i^+, K_e^+, Na_i^+, Na_e^+, Ca_i^{2+}, Ca_e^{2+}$

While it is reassuring that our approach reveals the obvious conserved moieties described above, it also reveals the nontrivial charge conserved moiety that is missed by many existing cardiac electrophysiology models. The overall amount of intracellular charge can be described as a sum of contributions from intracellular K^+ , Na^+ , Ca^{2+} (and its buffers) and Markov states from ion channels and transporters (Σ), similar to forms found in previous studies (Hund et al., 2001; Varghese and Sell, 1997). It should be noted, however, that when all ion concentrations were held constant charge conservation was broken, as indicated by the absence of a conserved charge moiety in the bottom partition of Table 3.1. In general, holding the concentration of a species constant disrupts conserved quantities within a system (Polettini and Esposito, 2014), therefore the number of conserved moieties progressively decreases as more ion concentrations are modelled as chemostats. We discuss the consequences of this in later sections.

3.3.3 Nonconservative stimulus currents cause drift in models with a charge conservation law

An important feature of cardiac electrophysiology models is that they must be simulated for extended periods to examine physiologically relevant changes in behaviour, thus we tested how the type of stimulus current affected each variant of the cardiac action potential model by pacing at 1 Hz for 30 minutes. As illustrated (Figure 3.8A,B), a nonconservative stimulus resulted in drift when the model had dynamic ion concentrations either for all compartments, or only within the intracellular compartment. The drift was particularly pronounced when all ion concentrations were dynamic (Figure 3.8A), as extracellular concentrations changed faster than intracellular concentrations. In contrast, the model was resistant to drift from a nonconservative stimulus when all ion concentrations were held constant (Figure 3.8C).

These results suggested that drift arose due to violations of the conserved charge moiety. Charge is a conserved moiety (Table 3.1) in model variants where drift occurred with a nonconservative stimulus. In this situation nonconservative stimulus currents cause drift because every stimulus causes a stepwise increase in the value of the conserved charge moiety (Figure 3.8A,B bottom panels). However, because conservation laws are broken as more species are represented as chemostats (Polettini and Esposito, 2014), charge is no longer a conserved moiety when all ion concentrations are constant (Table 3.1). Thus an observation which may not be obvious to intuition is that under these conditions charge is no longer constant between stimuli, and therefore free to return to its original value after each stimulus (Figure 3.8C, bottom panel), allowing such models to achieve a steady-state limit cycle.

Using the observations from the bottom row of Figure 3.8, it is possible to develop a systematic and automated check for drift. Let v_s be a row matrix representing the stoichiometry of the stimulus current (with chemostats removed), N^{cd} be the stoichiometric matrix after removing rows corresponding to chemostats, and G be the left nullspace matrix of N^{cd} . As seen in this section, a stimulus current will cause drift if it results in any change to the conserved moieties GX . Therefore to avoid altering any of the conserved moieties, the stimulus current must have zero contribution to them, i.e. $Gv_s = 0$ (or equivalently, v_s needs to lie in the image of N^{cd}). Thus the model drifts if $Gv_s \neq 0$.

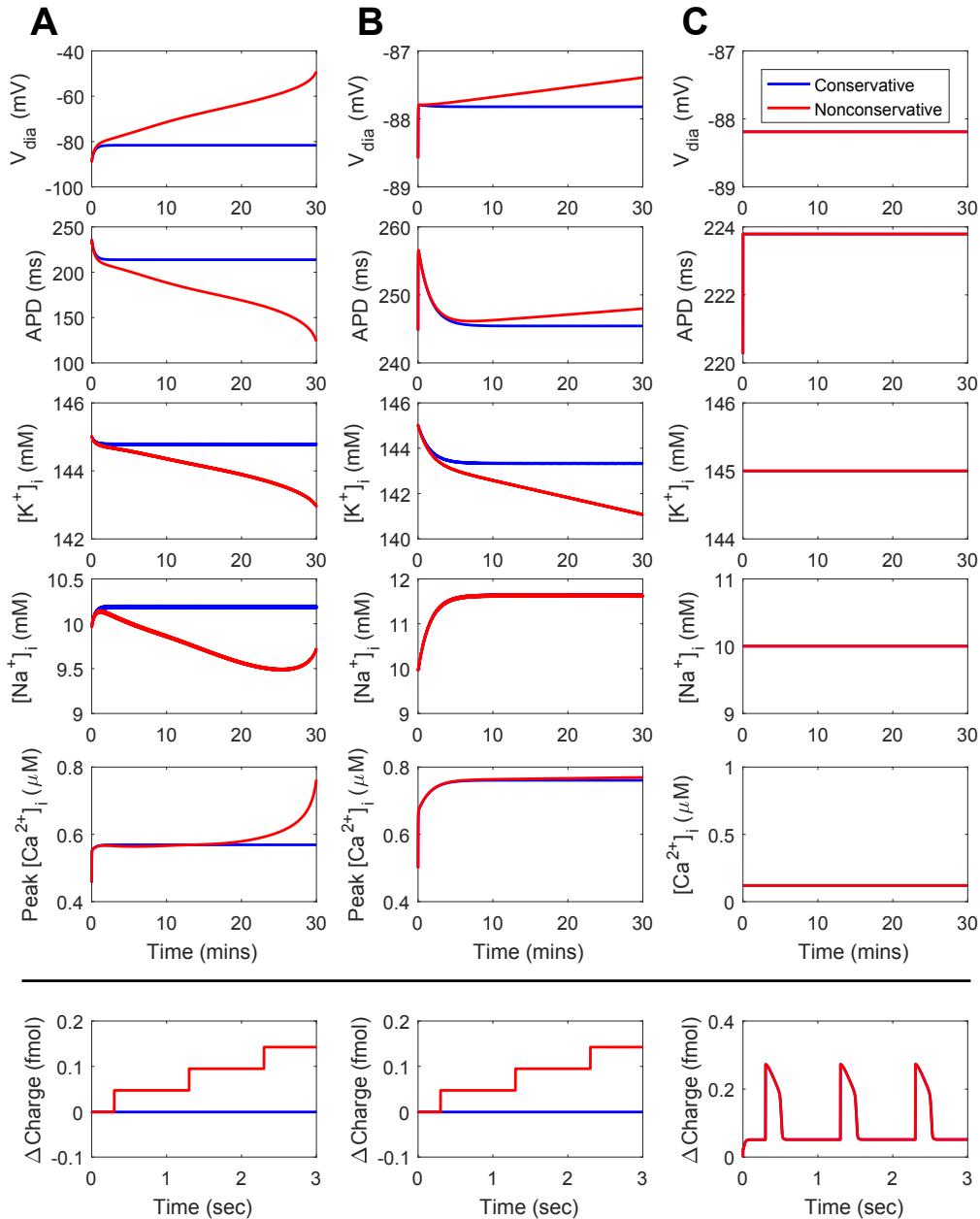


Figure 3.8: Effect of stimulus type and variable ion concentrations on drift in the bond graph model. (A) Dynamic ion concentrations; (B) Dynamic intracellular ion concentrations; (C) Constant ion concentrations. Results are shown for stimuli that conserve overall charge (blue) and those that do not conserve charge (red). Nonconservative stimulus currents contain non-specific charge, and conservative stimulus currents contain charge and potassium ions. Charge values are given as differences from the initial value of -5882.2 fmol. $T = 310$ K. Definitions: V_{dia} , diastolic membrane potential; APD, action potential duration at 90% repolarisation.

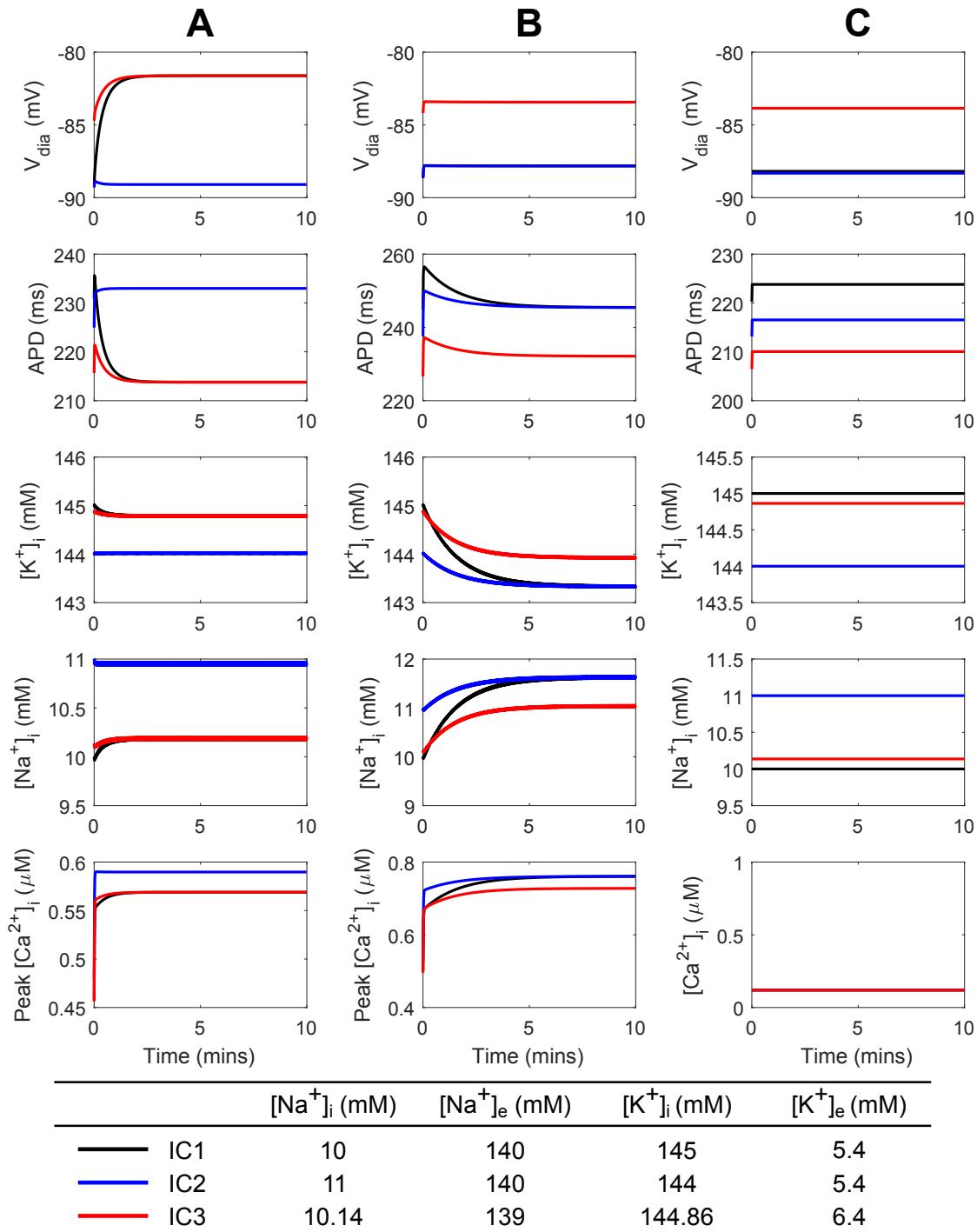


Figure 3.9: Effect of initial conditions on steady-state behaviour. (A) Dynamic ion concentrations; (B) Dynamic intracellular ion concentrations; (C) Constant ion concentrations. The models were paced at 1Hz for 30 minutes using a conservative stimulus current. $[MgATP] = 6.95$ mM, $[MgADP] = 0.035$ mM, $[P_i] = 0.3971$ mM, pH = 7.095, $T = 310$ K. Definitions: V_{dia} , diastolic membrane potential; APD, action potential duration at 90% repolarisation.

3.3.4 Initial conditions influence steady states through conserved moieties and chemostats

Next, for different sets of conserved moieties (as determined by constrained/dynamic ionic concentrations) we tested how the steady-state behaviour of the cardiac action potential was altered under three different initial conditions (Figure 3.9). The first set of initial conditions (IC1) are common values for comparison (Figure 3.9; *at bottom*). IC2 is the same as IC1 but with 1mM intracellular K⁺ exchanged for 1mM of intracellular Na⁺, such that charge is conserved but K⁺ and Na⁺ are not conserved. Similarly, IC3 is IC1, but with some K⁺ extruded and an equal amount of Na⁺ moved into the cell such that charge, Na⁺, and K⁺ are all conserved. When all ion concentrations are dynamic IC1 and IC3 lead to the same steady state, but IC2 results in a different steady state (Figure 3.9A). If only intracellular ion concentrations are dynamic, however, IC1 and IC2 result in identical steady states, but IC3 leads to a different steady state (Figure 3.9B). Finally, keeping all ion concentrations constant leads to different steady states for all initial conditions (Figure 3.9C).

These results demonstrate that the summed amount for each conserved moiety and/or chemostat value determines the steady-state behaviour of cardiac action potential models. To investigate this further, we calculated the values for conserved moieties and chemostats that resulted from each initial condition (Table 3.2; *differences from IC1 indicated in bold*). For two sets of initial conditions to achieve identical steady states, all conserved moieties and chemostats must have the same value. Thus under dynamic ion concentrations (Figure 3.9A), IC3 results in the same steady state as IC1 because all conserved moieties have been preserved (Table 3.2), whereas, IC2 causes a different steady state because the K⁺ and Na⁺ conserved moieties take on different values. Similarly, when only intracellular ion concentrations are dynamic, IC2 preserves the value of all conserved moieties and chemostats, but IC3 changes the values of the chemostats corresponding to extracellular Na⁺ and K⁺ concentrations (Table 3.2), hence the different steady state. When all ion concentrations were held constant, changes in the chemostat values (Table 3.2) were associated with different steady states for all three initial conditions (Figure 3.9C).

3.4 Discussion

In this study, we developed a bond graph model of the cardiac action potential with the aim of resolving the issues of drift and non-unique steady states. Analysis using conserved moieties enabled the discovery of all conservation laws within the model, including the charge conservation law neglected by many existing studies. In addition to the conservation of charge law from previous studies (Varghese and Sell, 1997; Hund et al., 2001; Endresen et al., 2000), we found conservation laws corresponding to ions, states of Markov models of channels and transporters, and buffers, demonstrating the comprehensiveness of our approach. Because the bond graph approach requires species and processes to be resolved in biophysical detail, calculation of conservation laws is more straightforward than working purely with the mathematical equations of a model. Two key advantages of our approach over existing analyses are that it reveals all conservation

Table 3.2: The values of conserved moieties and chemostats under different initial conditions. All values are in fmol. Chemostats are indicated with (cs). Values different from IC1 are shown in bold.

Moiety/Chemostat		Value		
		IC1	IC2	IC3
<i>Dynamic ion concentrations (A)</i>				
K ⁺	5538.1	5500.1	5538.1	
Na ⁺	1105.6	1143.6	1105.6	
Charge	-5882.2	-5882.2	-5882.2	
<i>Dynamic intracellular ion concentrations (B)</i>				
K _e ⁺ (cs)	27.98	27.98	38.35	
Na _e ⁺ (cs)	5510	5510	715.12	
Charge	-5882.2	-5882.2	-5882.2	
<i>Constant ion concentrations (C)</i>				
K _e ⁺ (cs)	27.98	27.98	38.35	
Na _e ⁺ (cs)	725.48	725.48	715.12	
K _i ⁺ (cs)	5510	5472	5500	
Na _i ⁺ (cs)	380	418	390.36	

laws in a comprehensive and systematic manner, and that it is general for all models of the cardiac action potential that can be represented as bond graphs. As a result, our approach can be scaled to more complex cardiac models without manual examination of the equations of a model to identify each individual conservation law, and the nontrivial charge conservation law will be accounted for when it appears as a conserved moiety. When simulated over long periods with a nonconservative stimulus our bond graph model displayed solution drift, but it became resistant to drift when ion concentrations were held constant, demonstrating that changes in the value of a conserved charge moiety drive model drift. We also found that two sets of initial conditions can lead to different steady states if the values of their corresponding conserved moieties and chemostats are different, suggesting a strong link between conserved moieties and the steady-state limit cycles of cardiac action potential models. To demonstrate that our approach is general, we tested how the selection of chemostats (i.e. fixed concentrations) influenced drift and steady states by using variants of our model that were representative of existing models in the literature. Our approach highlights the subtle but relevant observation that holding ion concentrations constant changes the conserved moieties of the model, which in turn has an effect on the susceptibility of a model to drift and non-unique steady states. Because chemostats represent connections between a system and its external environment, they are essential to coupling together biological processes (Gawthrop et al., 2015). The coupling of biological processes generally causes changes in the conservation laws of a system which may be difficult to capture through observation. Our approach using bond graphs provides a systematic method of dealing with changes to conservation laws as a result of coupling models together.

3.4.1 Drift

When paced with a nonconservative stimulus, variants of the model with a charge conservation law (A and B) underwent drift (Figure 3.8A,B) consistent with previous studies on the Luo-Rudy model (Hund et al., 2001; Livshitz and Rudy, 2009). Our observations also explain why models with constant ion concentrations (similar to variant C) are more likely to be resistant to drift (Hund et al., 2001). By observing changes in the charge conserved moiety, the bond graph approach attributes drift to regular perturbations in charge that cannot be restored due to the presence of a conservation law. Whereas previous analyses relied solely on intuition to derive a conservation law corresponding to charge (Hund et al., 2001; Livshitz and Rudy, 2009), we note that our approach automatically derives this conservation law, and can also detect other conservation laws that may be relevant for drift.

We note that in order to avoid drift, all conserved moieties (and not only the conserved moiety corresponding to charge) must be preserved by the stimulus current. However, in the examples explored in this study, we found that the stimulus currents preserved the value of all conserved moieties apart from charge, and therefore were not plotted. While intuition may suggest that modelling a stimulus current with K^+ would cause an accumulation of intracellular K^+ , the K^+ is passively transported back through K^+ channels. Furthermore, the total amount of K^+ (the relevant conserved moiety) remains constant because the loss of K^+ from the extracellular side is offset by accumulation on the intracellular side. In contrast, for charge conservation in variants A and B, the model contains no mechanism to reverse the change in charge difference caused by the nonconservative current, thus drift results.

As demonstrated, the bond graph method requires construction of a stoichiometric matrix, providing a simple approach to check whether a stimulus current will cause drift. While it is common practice to use K^+ as the charge carrier for stimulus currents, it is likely that multiple species contribute to the current (Hund et al., 2001; Kneller et al., 2002). Thus the automated approach suggested here is well-suited for checking whether more complex stimulus currents satisfy conservation of charge, as well as other conservation laws within the model. It should be noted however, that while a model satisfying $Gv_s = 0$ will not drift due to violating conservation laws, drift may still occur due to an imbalance of currents throughout the action potential, for instance, in the absence of Na^+/K^+ ATPase, the ionic gradients would gradually disappear in a model with dynamic ion concentrations.

Finally, we believe that this analysis provides a link between the issues of drift and steady states. Our models show that drift due to a nonconservative stimulus current can be attributed to changes in the value of the charge conserved moiety with every stimulus, and accordingly the steady state of the model changes. Model drift then occurs as the solution continually chases a moving steady state.

3.4.2 Effects of initial conditions on steady states

We also found that initial conditions of cardiac action potential models change their steady states through the values of chemostats and conserved moieties (Figure 3.9, Table 3.2).

Accordingly, the same perturbation to initial conditions can have different effects on the steady state depending on which species are held constant. Therefore in addition to ensuring that the concentration of ions are physiological, care should be taken to correctly initialise each state of buffers and Markov models of ion channels and ion transporters, as they may contain a significant fraction of total ion abundance. For example, Ca^{2+} buffers and SERCA can sequester a significant amount of Ca^{2+} and they should be initialised with the correct amount of bound Ca^{2+} when multi-state models are used (Higgins et al., 2006). We note that the difficulty of manually deriving conservation laws increases exponentially as models of cardiac electrophysiology become more complex, and we believe that our approach extends on existing analyses (Hund et al., 2001; Livshitz and Rudy, 2009) to provide a general method for assessing steady-state behaviour by comparing the values of conserved moieties and chemostats that result from each initial condition.

In the field of biochemical network analysis, there is a well-established dependence of quiescent steady states on conserved moieties, and numerous mathematical techniques for assessing the uniqueness and stability of these steady states have been developed (Gross et al., 2016; Feliu and Wiuf, 2012). However, the influence of conserved moieties on limit cycles in an oscillating system that is regularly stimulated has yet to be investigated. Our results hint at similarities between these two fields, and while we only tested the uniqueness of steady states using relatively small perturbations to the initial conditions, it is possible that a set of conserved moieties may have multiple steady states, and greater perturbations may lead to other limit cycles.

3.4.3 The “differential” and “algebraic” methods

The discovery of conservation principles in cardiac electrophysiology has lead to a debate over whether to use the differential or algebraic methods of simulation (Hund et al., 2001; Fraser and Huang, 2007; Livshitz and Rudy, 2009; Varghese and Sell, 1997; Endresen et al., 2000). The differential method is the calculation of membrane voltage by integrating total current, and the algebraic method is the calculation of membrane voltage using an algebraic relationship derived from charge conservation. We chose the differential method over the algebraic method since it better supports model reuse and modularity – in particular it is easier to modify the equations to select different species as chemostats, and to combine equations when two models are coupled. We note, however, that the algebraic method may reduce computational complexity (Gawthrop and Crampin, 2014; Hund et al., 2001). In bond graph modelling, the algebraic method can be implemented by using conserved moieties to turn the system of ODEs into an index-0 differential algebraic equation (DAE) (Eq. 3.48 of (Gawthrop and Crampin, 2014)). This method generalises existing algebraic methods to reduce the system of differential equations by using all conserved moieties and not just the conserved charge moiety. While we did not use the algebraic approach, we emphasise that the choice of method relates to numerical approaches for model simulation rather than the underlying physics of the system (Hund et al., 2001). Therefore, the differential and algebraic methods are equivalent in conservative systems provided that the initial conditions and values of conserved moieties are consistent.

3.4.4 Integration into whole-cell models

Our bond graph model of the cardiac action potential is the first step towards a fully-integrated whole-cell bond graph model of a cardiomyocyte that couples electrophysiology, signalling, metabolism and mechanics. Since energy drives all biological processes, and energy supply can be limited under certain pathophysiological states, it is of interest to examine how cardiomyocytes allocate their energy, and also to estimate the efficiency of processes that are essential for cardiomyocyte function (Neubauer, 2007; Tran et al., 2015). Modelling studies for the energetic regulation of a cardiac cell exist across the literature (Tran et al., 2015), but while some components used in these models are thermodynamically consistent (Tran et al., 2009; Smith and Crampin, 2004), existing whole-cell models are neither energy-based nor thermodynamically consistent throughout the entire model. Furthermore, because existing experimental and modelling studies use ATP consumption as a proxy for energy consumption, they can only estimate the energy consumption of major energy sinks: the Na^+/K^+ ATPase, SERCA, and cross-bridge cycling (Schramm et al., 1994; Tran et al., 2015). A bond graph approach may thus provide more detailed insights into how a cardiac cell uses energy downstream of ATP hydrolysis processes, and help to identify energy-consuming processes. Because the bond graph approach is energy-based, it not only provides the necessary constraints to develop a thermodynamically consistent model, but also allows us to directly assess energy consumption of the model (in Joules). We found that when normalised against membrane area the cardiac action potential consumes approximately 76% more energy than an action potential in the axon of a giant squid. To the authors' knowledge, this is the first account of energy consumed by electrochemical processes during the cardiac action potential.

3.4.5 Limitations

A limitation of our approach is that the physical constraints imposed by the bond graph approach prevent the direct translation of existing models of cardiac electrophysiology into bond graphs. This points at thermodynamic issues that exist within existing models of cardiac electrophysiology; while many models contain components that are thermodynamically consistent, the authors are not aware of any models that are entirely thermodynamically consistent. Nonetheless, the inability to perfectly replicate properties of existing models as bond graphs impedes the creation of bond graph models of cardiac electrophysiology, as it requires new models to be built from the bottom up with existing equations only used as guides for parameter fitting. As a result of this limitation, our bond graph model represents only a subset of the ion transport processes within a cardiomyocyte, and only approximates the behaviour described in existing models. While our model is able to reproduce many essential features of the action potential, it is unlikely to be physiologically realistic under conditions different to those used for our fitting process. Thus in future studies, it would be interesting to explore the use of more complex bond graph components to generate more physiologically realistic models. We chose to use a restricted set of components and constitutive equations to keep the model simple, although the bond graph framework is flexible enough to account for a wider range of equations provided that they are thermodynamically consistent. While many ion channels (including those for I_{Na} , I_{K1} and I_{Kp} in the Luo and Rudy model) are modelled using

linear I-V relationships, the choice of I-V relationship is generally chosen on the basis of providing an empirical fit to data rather than as a result of physical principles. Thus a more ideal approach would be to use nonlinear bond graph components derived from physical principles, and fit their parameters directly to experimental data. Additionally, empirical equations for gating transition parameters are generally not expressed in a thermodynamic framework, and are therefore impossible to replicate exactly with a bond graph model. We used a simple gating mechanism to reduce equation complexity while maintaining thermodynamic consistency, however the quality of fit could be improved by using more complicated gating mechanisms, or other thermodynamically consistent constitutive equations.

Because of physical restrictions imposed by the bond graph framework, we were forced to model ion channels and transporters using Markov states to faithfully represent their underlying physics. However, this produced a model that had numerous states compared to the number of biological processes. While it is reassuring to find that our method of identifying conserved moieties remained robust despite this complexity, simulation of the model was computationally expensive. For the purpose of integrating this action potential model into a larger whole-cell model, it would be useful to have simple model components that reduce computational cost. While current methods for reducing biochemical models in the bond graph framework are not advanced enough to apply to the biological components in this study, we note that bond graphs provide a useful foundation for applying model simplification while ensuring that thermodynamic consistency is maintained (Gawthrop and Crampin, 2014).

We also decided to limit the transport processes included in our model to those considered essential for producing a cardiac action potential, while maintaining a limit cycle using dynamic ion concentrations. Our bond graph model omitted many ionic currents due to their small amplitudes, however these channels may have greater contributions under conditions which vary from those tested here. Thus an obvious extension of this work would be the integration of other electrogenic processes within the cardiac membrane. It would be interesting to investigate whether coupling other models requires further tuning of parameters (Babtie and Stumpf, 2017), and whether the presence of physical bond graph parameters changes this process. A related limitation is that our model does not account for ion concentrations such as H^+ and Cl^- , as well as pH buffers (Crampin and Smith, 2006). While including these ions and their transporters would lead to a more accurate model, the omission of these ions did not cause inconsistencies in the conservation laws described in this study. These ions were assumed to be membrane impermeable in our model, and thus their constant contributions to the membrane potential were accounted for in the initial value of the charge conserved moiety. Similar to how calcium and its buffers were accounted for in our current approach (Table 3.1), our bond graph approach is sufficiently general to account for other ions and their buffers.

When formulating the structure and parameters for a bond graph model of the cardiac action potential (or most other biological processes), it is possible to either fit against existing mathematical models or the underlying experimental measurements. For all processes in this study excluding the NCX, we developed our bond graph model to reproduce the behaviour of an existing model, in an attempt to re-use existing knowledge about these processes. This approach poses constraints on the bond graph structure used, especially for gating structure. Therefore it would be interesting to develop an

approach that assesses bond graph structures as well as bond graph parameters, based on their fits to data (Babtie and Stumpf, 2017). Such an approach may provide a better fit to the data, and uncover insights into the physical mechanisms of ion channels.

3.5 Conclusion

In this study we have developed a bond graph model of the cardiac action potential and used this to explore the issues of drift and non-unique steady states. We demonstrate that the analysis of conserved moieties generalises the concept of charge conservation used in earlier studies, and found that changes in conserved moieties can explain drift as well as changes in steady state behaviour. Importantly, holding ion concentrations constant can have significant consequences on both drift and steady states as they change the conserved moieties in the model. Our approach to resolving drift and non-unique steady states is sufficiently general that it can be applied to any bond graph model of the cardiac action potential. We hope that the bond graph approach outlined here will prove useful for the development of future cardiac electrophysiology models, and eventually whole-cell models of the cardiomyocyte.

Data access: The code associated with this study is available from GitHub (https://github.com/uomsystemsbiology/bond_graph_cardiac_AP), and archived on Zenodo (<https://doi.org/10.5281/zenodo.1172205>) (Pan et al., 2018d). The repository contains MATLAB (The MathWorks, Natick, MA) code that generates the figures, CellML code containing parameters, initial conditions and equations of the model, and full details of the bond graph structure.

3.A Fitting ion channel parameters

3.A.1 I-V equations

A variety of methods were used to fit permeability constants P for the GHK equations used for the bond graph model. For some channels, P could be determined algebraically (such as the Na^+ and L-type Ca^{2+} channels). For others, optimisation was required to reduce error between the fitted I-V curve $I_{\text{GHK}}(V)$ (see Eq. 3.14 of the main text) and Luo-Rudy I-V curve $I_{\text{LR}}(V)$. In these cases, fitting was weighted towards $-90 \text{ mV} \leq V \leq -30 \text{ mV}$ for I_{K_1} , $-20 \text{ mV} \leq V < 30 \text{ mV}$ for I_K , and $0 \text{ mV} \leq V \leq 60 \text{ mV}$ for I_{K_p} . These regions were chosen based on when those channels activated. Where applicable, the optimisation problem was carried out by using particle swarm optimisation (Kennedy and Eberhart, 1995) followed by a local nonlinear optimiser. The permeabilities from fitting I-V curves are summarised in Table 3.3.

Table 3.3: Permeabilities of the GHK equations used for the bond graph model.

Permeability	Value (pL/s)
P_{Na}	9.0602
$P_{\text{K}1}$	1.1200
P_{K}	0.2299
P_{Kp}	0.0136
P_{CaL}	28.2471
P_{KL}	0.0222

3.A.1.1 Sodium current

The permeability was chosen to match the linear equation at the negative of the Nernst potential (Gawthrop et al., 2017):

$$P_{\text{Na}} = \frac{2\bar{G}_{\text{Na}}(1 - \exp[FE_{\text{Na}}/(RT)])}{[\text{Na}_i^+] - [\text{Na}_e^+]\exp[FE_{\text{Na}}/(RT)]}\frac{RT}{F^2} \quad (3.26)$$

where

$$E_{\text{Na}} = \frac{RT}{F} \ln \left(\frac{[\text{Na}_e^+]}{[\text{Na}_i^+]} \right) \quad (3.27)$$

$$\bar{G}_{\text{Na}} = 2.45 \text{ } \mu\text{A/mV} \quad (3.28)$$

3.A.1.2 Time-independent K⁺ current

$$P_{\text{K}1} = \arg \min_P \left\{ \sum_{V=-90}^{-30} [I_{\text{K}1,\text{LR}}(V) - I_{\text{K}1,\text{GHK}}(V, P)]^2 \right\} \quad (3.29)$$

$$I_{\text{K}1,\text{LR}}(V) = \bar{G}_{\text{K}1}(V - E_{\text{K}}) \quad (3.30)$$

$$\bar{G}_{\text{K}1} = 1.1505 \times 10^{-4} \text{ } \mu\text{A/mV} \quad (3.31)$$

$$E_{\text{K}} = \frac{RT}{F} \ln \left(\frac{[\text{K}_e^+]}{[\text{K}_i^+]} \right) \quad (3.32)$$

3.A.1.3 Time-dependent K⁺ current

$$P_{\text{K}} = \arg \min_P \left\{ \sum_{V=-20}^{29} [I_{\text{K},\text{LR}}(V) - I_{\text{K},\text{GHK}}(V, P)]^2 \right\} \quad (3.33)$$

$$I_{\text{K},\text{LR}}(V) = \bar{G}_{\text{K}}(V - E_{\text{K},\text{LR}}) \quad (3.34)$$

$$\bar{G}_{\text{K}} = 4.3259 \times 10^{-5} \text{ } \mu\text{A/mV} \quad (3.35)$$

$$E_{\text{K},\text{LR}} = \frac{RT}{F} \ln \left(\frac{[\text{K}_e^+] + P_{\text{Na},\text{K}}[\text{Na}_e^+]}{[\text{K}_i^+] + P_{\text{Na},\text{K}}[\text{Na}_i^+]} \right) \quad (3.36)$$

$$P_{\text{Na},\text{K}} = 0.01833 \text{ [dimensionless]} \quad (3.37)$$

Note that the Luo and Rudy I-V relationship for the time-dependent K⁺ current (Eq. 3.34) is thermodynamically inconsistent because $E_{K,LR}$ is not the Nernst potential defined in Eq. 3.32. As a result, nonzero currents can result at the Nernst potential, where the current should be zero from thermodynamic constraints. Despite this thermodynamic inconsistency, it is possible to define $I_{K,GHK}(V)$, a thermodynamically consistent approximation to the I-V curve in Luo and Rudy.

3.A.1.4 Plateau K⁺ current

$$P_{K_p} = \arg \min_P \left\{ \sum_{V=0}^{60} [I_{K_p,LR}(V) - I_{K_p,GHK}(V, P)]^2 \right\} \quad (3.38)$$

$$I_{K_p,LR}(V) = \bar{G}_{K_p}(V - E_K) \quad (3.39)$$

$$\bar{G}_{K_p} = 2.8072 \times 10^{-6} \text{ } \mu\text{A}/\text{mV} \quad (3.40)$$

E_K same as for the time-independent K⁺ current.

3.A.1.5 L-type Ca²⁺ channel

For the L-type Ca²⁺ channel, Luo and Rudy (Luo and Rudy, 1994a) use the I-V equation

$$I_{Ca,s} = P_{sL} \frac{z_s^2 F^2 V}{RT} \frac{\gamma_{si}[S_i] \exp(z_s FV/RT) - \gamma_{so}[S_e]}{\exp(z_s FV/RT) - 1} \quad (3.41)$$

which resembles the GHK equation, but allows thermodynamic laws to be broken through the use of different partitioning factors γ_{si} and γ_{so} , where $s \in \{\text{Ca}^{2+}, \text{K}^+\}$. In the case of the Ca²⁺ component of the current, this was resolved by setting both factors to the value of γ_{Cae} , with little effect on the I-V curve. Thus the permeabilities of the GHK equations are calculated as follows:

$$P_{CaL} = P_{CaL,LR} \gamma_{Cae} \quad (3.42)$$

$$P_{KL} = P_{KL,LR} \gamma_{Ke} = P_{KL,LR} \gamma_{Ki} \quad (3.43)$$

where

$$P_{CaL,LR} = 8.2836 \times 10^{-8} \text{ cm}^3/\text{s} \quad (3.44)$$

$$P_{KL,LR} = 2.9606 \times 10^{-11} \text{ cm}^3/\text{s} \quad (3.45)$$

$$\gamma_{Cae} = 0.341 \quad (3.46)$$

$$\gamma_{Ke} = \gamma_{Ki} = 0.75 \quad (3.47)$$

3.A.2 Gating transition parameters

The parameters derived for gate transition are summarised in Table 3.4, with further detail described below.

Table 3.4: Gate transition parameters.

Gate	α_0 (s ⁻¹)	z_f	β_0 (s ⁻¹)	z_r
m	12516.4361	0.4954	79.9996	-2.4284
h	0.00033539	-4.1892	799.9028	1.2995
j	0.00013079	-4.0381	422.7582	1.4281
K1	1127.3395	0.0336	13544806.3586	3.1153
X	2.2317	0.5192	0.5750	-0.7317
Xi	995.8931	0	172.6026	0.8322
Kp	999.8464	0	3497.4018	-4.4669
d	486.7619	2.1404	98.0239	-2.1404
f		See Appendix 3.B.3		

3.A.2.1 m, h, j, K1 and X-gates

A vector quantity $\mathbf{p} = (\alpha_0, z_f, \beta_0, z_r)$ was optimised based on the quality of fits to the transition parameters, steady-state open probability and time constant in the range -120 mV $\leq V \leq 60$ mV:

$$\mathbf{p}_g = \arg \min \left\{ \sum_{V=-120}^{60} a(V) \left(a_\alpha [\alpha_{g,\text{LR}}(V) - \alpha_g(V, \mathbf{p})]^2 + a_\beta [\beta_{g,\text{LR}}(V) - \beta_g(V, \mathbf{p})]^2 + a_{\text{gss}} [g_{\text{ss},\text{LR}}(V) - g_{\text{ss}}(V, \mathbf{p})]^2 + a_\tau [\tau_{g,\text{LR}}(V) - \tau_g(V, \mathbf{p})]^2 \right) \right\} \quad (3.48)$$

where g is replaced with m , h , j , K1 or X depending on the gate. $a(V) = 1$ and $a_\alpha = a_\beta = a_{\text{gss}} = a_\tau = 1$ for the m , h and j gates. For the K1 gate, $a(V) = 1$, $a_\alpha = a_\beta = 0$, $a_\tau = 1$ and $a_{\text{gss}} = 1000$. For the X gate,

$$a_{\text{gss}} = 100, \quad a(V) = \begin{cases} 1, & V < 0 \text{ mV} \\ 25, & V \geq 0 \text{ mV} \end{cases} \quad (3.49)$$

The parameters α_g and β_g have unit ms⁻¹ and τ_g has unit ms. Optimisation was carried out using particle swarm optimisation followed by a local optimiser.

3.A.2.2 Xi-gate

To give a perfect fit for X_{ss} ,

$$\alpha_0 = K_{\text{Xi}} \quad (3.50)$$

$$\beta_0 = K_{\text{Xi}} e^{56.26/32.1} \quad (3.51)$$

$$z_f = 0 \quad (3.52)$$

$$z_r = \frac{RT}{F} \frac{1000 \text{mV/V}}{32.1 \text{mV}} = 0.8322 \quad (3.53)$$

To achieve a time constant of less than 1ms in the range -120 mV $\leq V \leq 60$ mV, we chose

$$K_{\text{Xi}} = 0.9959 \quad (3.54)$$

3.A.2.3 K_p-gate

To give a perfect fit for K_{pss},

$$\alpha_0 = K_{Kp} \quad (3.55)$$

$$\beta_0 = K_{Kp} e^{7.488/5.98} \quad (3.56)$$

$$z_f = 0 \quad (3.57)$$

$$z_r = \frac{RT}{F} \frac{1000\text{mV/V}}{5.98\text{mV}} = -4.4669 \quad (3.58)$$

To achieve a time constant of less than 1ms in the range $-120 \text{ mV} \leq V \leq 60 \text{ mV}$, we chose

$$K_{Kp} = 0.9998 \quad (3.59)$$

3.A.2.4 d-gate

To give a perfect fit for d_{ss},

$$\alpha_0 = K_d e^{10/12.48} \quad (3.60)$$

$$\beta_0 = K_d e^{-10/12.48} \quad (3.61)$$

$$z_f = \frac{RT}{F} \frac{1000\text{mV/V}}{12.48\text{mV}} = 2.1404 \quad (3.62)$$

$$z_r = -\frac{RT}{F} \frac{1000\text{mV/V}}{12.48\text{mV}} = -2.1404 \quad (3.63)$$

K_d was chosen to match the peak time constant because that is where changes would be most likely to make a difference given that the time constant is small:

$$K_d = 0.2184 \quad (3.64)$$

3.B Channel-specific modelling issues

3.B.1 K⁺ regulation of K⁺ currents

For the K and K1 channels, Luo and Rudy (1994a) describe a dependence of the permeability on the square root of extracellular K⁺ concentration:

$$G = G_{ch} \sqrt{[K^+]_e} \quad (3.65)$$

where G_{ch} is a constant value that differs between the two channels. This was incorporated by assigning an additional extracellular K⁺ stoichiometry of 0.5 to both sides of the ion transport reaction, which results in an ion channel current of

$$v = \kappa(V) \left(K_{Ke}^{0.5} x_{Ke}^{0.5} K_{Ki} x_{Ki} e^{zFV/RT} - K_{Ke}^{1.5} x_{Ke}^{1.5} \right) \quad (3.66)$$

$$= \kappa(V) \sqrt{K_{Ke} x_{Ke}} \left(K_{Ki} x_{Ki} e^{zFV/RT} - K_{Ke} x_{Ke} \right) \quad (3.67)$$

3.B.2 Ca²⁺ inactivation of L-type Ca²⁺ current

Luo and Rudy (1994a) describe a mechanism whereby the L-type Ca²⁺ channel is inactivated by intracellular Ca²⁺, using the function:

$$f_{\text{Ca}} = \frac{1}{1 + ([\text{Ca}_i^{2+}]/K_{m,\text{Ca}})^2} \quad (3.68)$$

This mechanism was incorporated into the bond graph framework through the reaction:



with a dissociation constant equal to $K_{m,\text{Ca}}^2$. It can be shown that at equilibrium:

$$\frac{x_A}{x_A + x_I} = \frac{1}{1 + ([\text{Ca}_i^{2+}]/K_{m,\text{Ca}})^2} = f_{\text{Ca}} \quad (3.70)$$

Therefore Ca²⁺ inactivation was incorporated by applying the reaction in Eq. 3.69 to each of the states that result from independent d and f gating, using kinetic constants that were sufficiently high to approximate rapid equilibrium.

3.B.3 f-gate of the L-type Ca²⁺ channel

Luo and Rudy use the equations from Rasmusson et al. (1990) for their L-type Ca²⁺ channel f -gate, resulting in U-shaped functions for both the steady-state open probability f_{ss} and time constant τ_f . Using the exponential dependence in Eq. 3.21 of the main text, f_{ss} must have a monotonic and sigmoidal shape, and τ_f must either be bell-shaped or monotonic. As neither the f_{ss} nor τ_f could be made U-shaped with the current formulation, we used an alternative mechanism to describe the f -gate. We observed that the f -gate activated at both negative and positive voltages, and that the minima of f_{ss} , and τ_f of the Rasmusson equations appeared to coincide. We modelled the gate using

the reaction network $O_1 \xrightleftharpoons[\alpha_1]{\beta_1} C \xrightleftharpoons[\beta_2]{\alpha_2} O_2 \xrightleftharpoons[k_3^+]{k_3^-} O_1$ with the final reaction assumed to be at

quasi-equilibrium. The rationale behind using this three-state model was that: (a) there were two open states, one that activated at negative voltages and one that activated at positive voltages, and; (b) the inactivation parameters could be chosen such that the gate inactivated faster than it activated. The initial spike in membrane potential during an action potential implies that the open probability is unable to change, thus we used a reaction in rapid equilibrium to convert between the two open states; without this, the gate would need to pass the closed states to move between the open states.

Similar to the transition parameters in other gates an exponential dependence on voltage was assumed. Since the mechanism involves a biochemical cycle, a detailed balance constraint was used to determine parameters for the third reaction between the two open states:

$$\frac{k_3^+(V)}{k_3^-(V)} = \frac{\beta_1(V)\alpha_2(V)}{\alpha_1(V)\beta_2(V)} \quad (3.71)$$

The following information was used to parameterise the f -gate:

1. The difference between the steady-state open probabilities in the Luo-Rudy model (f_{ss}) and bond graph model ($f_{ss,BG}$) over the range $-90 \text{ mV} \leq V \leq 50 \text{ mV}$. The open probability of the bond graph formulation was calculated by rapid equilibrium arguments (Smith and Crampin, 2004):

$$f_{ss,BG} = \frac{\alpha_1/\beta_1 + \alpha_2/\beta_2}{1 + \alpha_1/\beta_1 + \alpha_2/\beta_2} \quad (3.72)$$

Differences were taken between the natural logarithms of each of the open probabilities prior to calculating differences to better match lower values.

2. Simulations of the f -gate were run with the voltage held constant. The open probabilities over time were compared to solutions obtained from the Luo-Rudy formulation of the f -gate. The conditions for the simulations are summarised in Table 3.5. For computational efficiency, the third reaction was neglected for the bond graph simulations. All simulations involve either activation/inactivation processes involving one of the open states. It was assumed that very little of the other open state would become open.

Table 3.5: Summary of conditions used to simulate f-gate for fitting parameters. o_1 , c and o_2 represent the proportion of the three states representing the inactivation process.

#	Voltage (mV)	Initial conditions	Description
1	-80	$o_1 = 0, c = 1, o_2 = 0$	Activation at -80 mV
2	-40	$o_1 = 1, c = 0, o_2 = 0$	Inactivation at -40 mV
3	-40	$o_1 = 0, c = 1, o_2 = 0$	Activation at -40 mV
4	0	$o_1 = 1, c = 0, o_2 = 0$	Inactivation at 0 mV from O_1
5	0	$o_1 = 0, c = 0, o_2 = 1$	Inactivation at 0 mV from O_2
6	40	$o_1 = 0, c = 0, o_2 = 1$	Inactivation at 40 mV

The transition rates for the f -gate are

$$\alpha_1(V) = \alpha_{0,1} \exp\left(\frac{z_{f,1}FV}{RT}\right) \quad (3.73)$$

$$\beta_1(V) = \beta_{0,1} \exp\left(\frac{z_{r,1}FV}{RT}\right) \quad (3.74)$$

$$\alpha_2(V) = \alpha_{0,2} \exp\left(\frac{z_{f,2}FV}{RT}\right) \quad (3.75)$$

$$\beta_2(V) = \beta_{0,2} \exp\left(\frac{z_{r,2}FV}{RT}\right) \quad (3.76)$$

$$k_3^+(V) = r_3 K_{3,0} \exp\left(\frac{z_{f,3}FV}{RT}\right) \quad (3.77)$$

$$k_3^- = r_3 \quad (3.78)$$

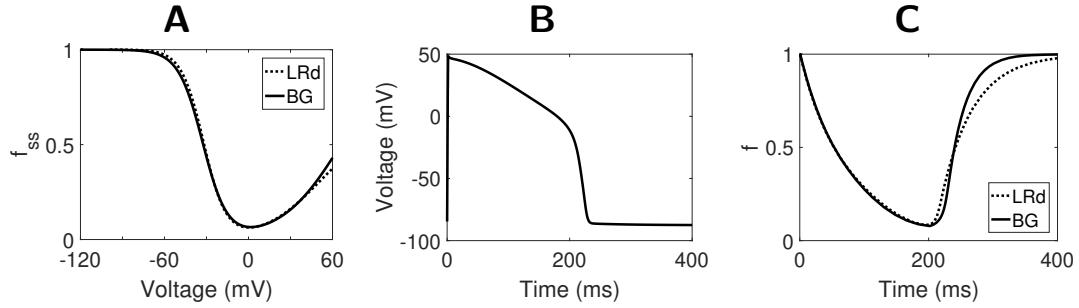


Figure 3.10: Fitting the *f*-gate of the L-type Ca^{2+} channel. (A) The steady-state open probability of the *f*-gate, calculated by adding the proportion of the two open states. (B) The action potential waveform used to compare the behaviour of the Luo and Rudy (LRd) and bond graph (BG) formulations of the *f*-gate. This was obtained by simulating the Luo-Rudy model with the ion channels used in this study, and holding the ion concentrations constant. (C) The response of the *f*-gates to the voltage signal in B.

with the constants

$$\alpha_{0,1} = 0.8140 \text{ s}^{-1}, \quad z_{f,1} = -1.1669 \quad (3.79)$$

$$\beta_{0,1} = 36.1898 \text{ s}^{-1}, \quad z_{r,1} = 1.6709 \quad (3.80)$$

$$\alpha_{0,2} = 1.6369 \text{ s}^{-1}, \quad z_{f,2} = 0.7312 \quad (3.81)$$

$$\beta_{0,2} = 35.5248 \text{ s}^{-1}, \quad z_{r,2} = -0.5150 \quad (3.82)$$

$$r_3 = 10000 \text{ s}^{-1}, \quad K_{3,0} = 2.0485 \quad (3.83)$$

$$z_{f,3} = z_{r,1} + z_{f,2} - z_{f,1} - z_{r,2} = 4.0839 \quad (3.84)$$

The three-state scheme in the bond graph framework produced a similar curve for f_{ss} to the *f*-gate of the Luo-Rudy model (Figure 3.10A). Since there is no direct time constant for our three-state model we compared the dynamic behaviour of the *f*-gates by simulating to an action potential-like voltage waveform (Figure 3.10B). During the depolarised phase of the action potential where the *f*-gate steadily inactivates, the bond graph model provides a very good fit to the Luo and Rudy model (Figure 3.10C). In the resting phase the bond graph model reactivates faster, but still provides a reasonable fit.

3.C Ion transporters

3.C.1 Na^+/K^+ ATPase

We used the 15-state bond graph model described in Pan et al. (2017), with a pump density of $4625 \mu\text{m}^{-2}$ ($0.1178 \text{ fmol per cell}$).

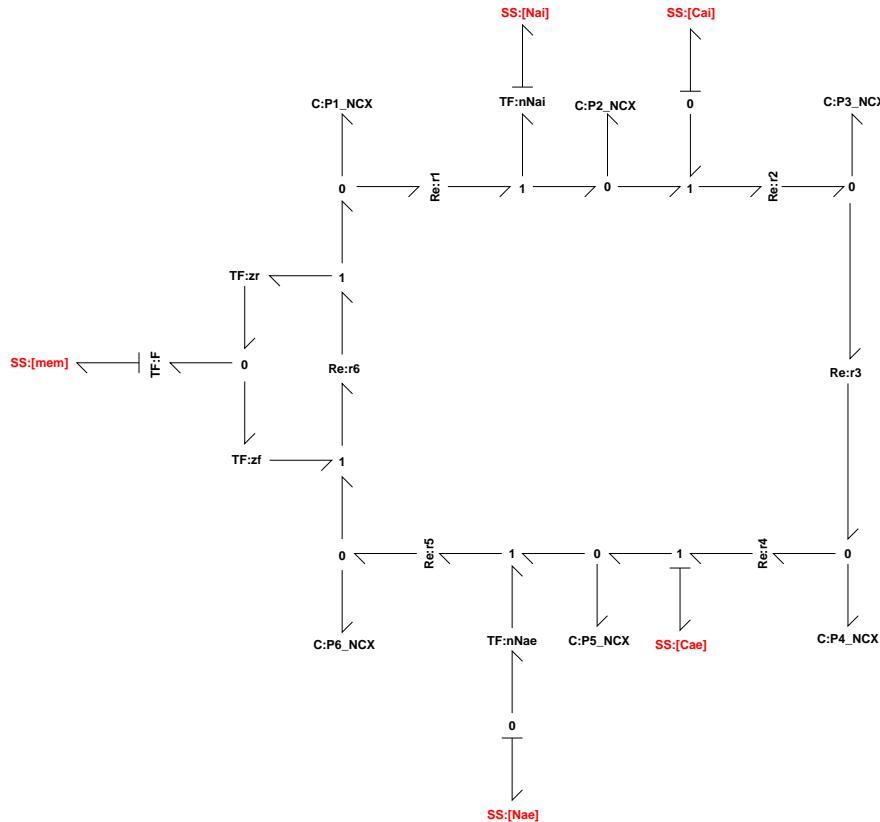


Figure 3.11: The bond graph model of NCX.

3.C.2 Na⁺-Ca²⁺ exchanger

The NCX was modelled using the bond graph shown in Figure 3.11. The reaction scheme was based on the ping-pong mechanism proposed in Giladi et al. (2016), with reactions r1, r2, r4 and r5 modelled by fast rate constants to approximate rapid equilibrium. We assigned voltage dependence to translocation of Na⁺, based on experimental findings from Hilgemann et al. (1992).

Using similar methods to Luo and Rudy (1994a), the NCX model was fitted to the following data, assuming steady-state operation:

1. Dependence of cycling rate on extracellular Na⁺ and voltage, from Kimura et al. (1987).
2. Dependence of cycling rate on extracellular Ca²⁺, from Kimura et al. (1987). Data obtained at $V < -50$ mV and $[Ca^{2+}]_e = 1$ mM were excluded from the fitting process.
3. To incorporate behaviour for another intracellular Ca²⁺ concentration, data from Beuckelmann and Wier (1989) were used. Data obtained at $V < -120$ mV were excluded from the fitting process.

Parameters of the model were identified using particle swarm optimisation followed by a local optimiser, and a comparison between the model and data is shown in Fig-

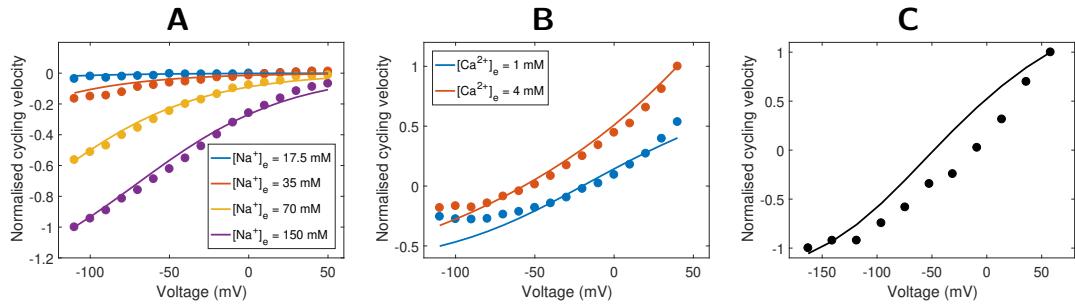


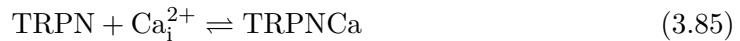
Figure 3.12: Fit of NCX model to data. (A) Comparison of model to Fig. 8B of Kimura et al. (1987). Fluxes were normalised to the value at $[\text{Na}^+]_{\text{e}} = 140 \text{ mM}$ and $V = -110 \text{ mV}$. $[\text{Na}^+]_{\text{i}} = 0 \text{ mM}$, $[\text{Ca}^{2+}]_{\text{e}} = 1 \text{ mM}$, $[\text{Ca}^{2+}]_{\text{i}} = 430 \text{ nM}$. (B) Comparison of model to Fig. 9A of Kimura et al. (1987). Fluxes were normalised to the value at $[\text{Ca}^{2+}]_{\text{e}} = 4 \text{ mM}$ and $V = 40 \text{ mV}$. $[\text{Na}^+]_{\text{e}} = 140 \text{ mM}$, $[\text{Na}^+]_{\text{i}} = 10 \text{ mM}$, $[\text{Ca}^{2+}]_{\text{i}} = 172 \text{ nM}$. (C) Comparison of model to Fig. 6B of Beuckelmann and Wier (1989). Fluxes were normalised to the value at $V = 60 \text{ mV}$. $[\text{Na}^+]_{\text{e}} = 135 \text{ mM}$, $[\text{Na}^+]_{\text{i}} = 15 \text{ mM}$, $[\text{Ca}^{2+}]_{\text{e}} = 2 \text{ mM}$, $[\text{Ca}^{2+}]_{\text{i}} = 450 \text{ nM}$.

ure 3.12. The model closely matched the data describing extracellular Na^+ dependence (Figure 3.12A). Reasonable fits were obtained for the other data, although there was some discrepancy at negative voltages in Figure 3.12B. There was some difference between the model and data from Beuckelmann and Wier (1989) (Figure 3.12C), although this appears to have resulted from differences in the equilibrium point.

The cycling velocity was normalised to 700 s^{-1} at the normalisation point of Figure 3.12A to approximately match experimental currents at a membrane capacitance of 200 pF and 300 sites per μm^{-2} . To ensure that the exchanger current had a similar magnitude to that of Luo and Rudy (1994a), we used a site density of $170 \mu\text{m}^{-2}$ ($0.0043 \text{ fmol per cell}$) in our cardiac action potential model.

3.D Ca^{2+} buffering

The model of Ca^{2+} buffering was based on the equations described in Luo and Rudy (1994a). These equations represent the reactions



with the dissociation constants $K_{d,\text{TRPN}} = 0.5 \mu\text{M}$ and $K_{d,\text{CMDN}} = 2.38 \mu\text{M}$. The total concentrations of each buffer were $70 \mu\text{M}$ for troponin and $50 \mu\text{M}$ for calmodulin. The reactions were modelled using sufficiently fast kinetic constants to approximate rapid equilibrium, and the amount of Ca^{2+} bound to each buffer was initialised to the value at equilibrium for the initial intracellular Ca^{2+} concentration of $0.12 \mu\text{M}$.

3.E Bond graph parameters

Bond graph parameters were found by using an extension of the method presented in Gawthrop et al. (2015). The kinetic parameters and bond graph parameters can be related through the matrix equation

$$\mathbf{Ln}(\mathbf{k}) = \mathbf{M}\mathbf{Ln}(\mathbf{W}\boldsymbol{\lambda}) \quad (3.87)$$

where

$$\mathbf{k} = \begin{bmatrix} k^+ \\ k^- \end{bmatrix}, \quad \mathbf{M} = \left[\begin{array}{c|c} I_{n_r \times n_r} & N^f{}^T \\ \hline I_{n_r \times n_r} & N^r{}^T \end{array} \right], \quad \boldsymbol{\lambda} = \begin{bmatrix} \kappa \\ K \end{bmatrix} \quad (3.88)$$

k^+ is a column vector consisting of the forward kinetic constants, k^- is a column vector consisting of the reverse kinetic constants, N^f and N^r are the forward and reverse stoichiometric matrices respectively, κ is a column vector of bond graph reaction rate constants, and K is column vector of thermodynamic constants. To account for the volumes of each compartment, \mathbf{W} is a diagonal matrix where the i -th diagonal element is the volume corresponding to i -th bond graph component (either a reaction or species). Depending on compartment, the elements corresponding to each ion were set to either the intracellular volume of $W_i = 38$ pL or the extracellular volume of $W_e = 5.182$ pL. All other diagonal entries were set to 1. Assuming that detailed balance constraints are satisfied, a solution to Eq. 3.87 is

$$\boldsymbol{\lambda}_0 = \mathbf{W}^{-1} \mathbf{Exp}(\mathbf{M}^\dagger \mathbf{Ln}(\mathbf{k})) \quad (3.89)$$

where \mathbf{M}^\dagger is the pseudo-inverse of \mathbf{M} . All parameters were identified using $T = 310$ K.

For reactions involved in ion transport that use the GHK equation, both the forward and reverse rate constants were set to P/x_{ch} , where P is the permeability constant found by fitting to Eq. 3.14 of the main text, and x_{ch} is the total number of channels. The values of x_{ch} used for each channel are given in Table 3.6. Since the bond graph parameters of the NCX model were fitted to kinetic data, the bond graph parameters were converted back to kinetic parameters (Gawthrop et al., 2015) to parameterise the action potential model.

Table 3.6: Amounts of each ion channel per cell. A geometric area of 0.767×10^{-4} cm² was used to convert between channel density and channels per cell (x_{ch}).

*Quantity cited from reference.

Ion channel	Channel density (μm ⁻²)	Channels per cell	Reference
Na	16*	122720	Reuter (1984)
K1	1.8*	4261	Sakmann and Trube (1984)
K	0.7*	5369	Shibasaki (1987)
Kp	0.095	725*	Yue et al. (1996)
LCC	6.5	50000*	Hinch et al. (2004)

3.F Charge conserved moiety

In Table 3.1 of the main text, Σ is defined as

$$\begin{aligned}
 \Sigma = & + 3.0818C_{K1} - 1.6697S_{00,K} - 0.4188S_{10,K} + 0.8322S_{20,K} - 2.5019S_{01,K} \\
 & - 1.2509S_{11,K} - 4.4669C_{Kp} + 2.1835S_{000,Na} + 5.1073S_{100,Na} + 8.0311S_{200,Na} \\
 & + 10.9549S_{300,Na} - 3.3052S_{010,Na} - 0.3814S_{110,Na} + 2.5424S_{210,Na} + 5.4662S_{310,Na} \\
 & - 3.2827S_{001,Na} - 0.3589S_{101,Na} + 2.5649S_{201,Na} + 5.4887S_{301,Na} - 8.7714S_{011,Na} \\
 & - 5.8476S_{111,Na} - 2.9238S_{211,Na} - 1.5253S_{000,LCC} - 4.5742S_{010,LCC} - 0.2808S_{020,LCC} \\
 & + 2.7555S_{100,LCC} - 0.2933S_{110,LCC} + 4S_{120,LCC} - 5.5253S_{001,LCC} - 8.5742S_{011,LCC} \\
 & - 4.2808S_{021,LCC} - 1.2445S_{101,LCC} - 4.2933S_{111,LCC} + P2_{NaK} + P3_{NaK} + P4_{NaK} \\
 & - 0.9450P6_{NaK} - 0.9450P7_{NaK} - 0.9450P8_{NaK} - P1_{NCX} + 2P2_{NCX}
 \end{aligned} \tag{3.90}$$

Chapter 4

Generalised model reduction of enzyme kinetics

The preceding chapters demonstrated how bond graphs can be used to represent biomolecular systems and to couple together models into more comprehensive models. However, large-scale dynamic models are often difficult to analyse mathematically and computationally expensive (sometimes infeasible) to simulate. Model reduction, in which simplifying assumptions are used to reduce the dimensionality of the equations, is often used to improve computational efficiency while retaining the essential features of a model (Snowden et al., 2017). However, in systems biology, many current approaches to model reduction do not preserve essential physical and thermodynamic information. In this chapter, we outline a general method for simplifying the equations of enzyme kinetics models with the intent of applying it to whole-cell modelling. Bond graphs form the underlying framework for this method, ensuring that links to fundamental physical and mechanistic information are retained within the simplified models. Because the method is general and programmable, reduced equations can be generated for arbitrarily large enzyme cycles. We demonstrate the utility of this generalised model reduction approach by comparing numerous multi-state models for individual enzymes, an issue that is often overlooked in the development of these models.

4.1 Introduction

Whole-cell models, which are comprehensive mechanistic models that predict the dynamics of the entire biochemical networks within a cell, hold great promise for biology and medicine because they could be used to generate interesting experimental hypotheses and to screen drugs *in silico* (Babtie and Stumpf, 2017). To date, the most advanced and only complete whole-cell model has been achieved in the *Mycoplasma genitalium* bacterium (Karr et al., 2012). However, the more complex human whole-cell model remains out of reach for two reasons. Firstly, we lack both experimental data and models for a substantial number of biochemical networks (Szigeti et al., 2018). Secondly, the enzymes that catalyse most reactions within the cell are themselves highly diverse and

complex. Despite the extensive research conducted on enzyme kinetics (Segel, 1975), there is no one-size-fits-all model that is appropriate for all enzymes (Liebermeister et al., 2010). The vast number and complexity of these processes implies a need to automate model development to facilitate whole-cell modelling.

As demonstrated in the first whole-cell model (Karr et al., 2012), the only computationally tractable way to construct a whole-cell model at present is through the independent development of smaller sub-models which are subsequently coupled together. However, the task of integrating sub-models is seen as a major roadblock to whole-cell modelling (Szigeti et al., 2018). The interfaces between models are often obscure, and even when models can be coupled together, the parameters within the original models often need to be fine-tuned. Computational biologists have acknowledged that an adherence to the laws of physics and thermodynamics is required to generate whole-cell models (Smith and Crampin, 2004; Smith et al., 2007). However, many existing models break the laws of thermodynamics, through the use of either irreversible reactions or ill-defined parameters that violate detailed balance (Gawthrop and Crampin, 2014; Beard and Qian, 2008; Smith and Crampin, 2004; Ederer and Gilles, 2007). In an attempt to resolve this issue, Mason and Covert (2019) reformulated kinetic rate laws in terms of independent energetic quantities, and found that this formulation was useful in mapping experimental measurements to parameter determinacy and in reconciling apparent inconsistencies within the training data.

Another roadblock to whole-cell modelling is the lack of tools for model development (Szigeti et al., 2018). The development of a mathematical model involves identifying both an appropriate model structure and model parameters. While techniques for parameter estimation are widely employed, the task of model selection is often overlooked in systems biology (Babtie and Stumpf, 2017; Toni et al., 2009). The lack of consideration for structural uncertainty impedes our ability to revise our knowledge of biological systems as new experimental data become available. We believe that model selection is underused because the development of separate models is time-consuming. Therefore, a framework that automatically generates equations for several candidate models would make model selection more accessible.

In this study, we present a strategy for reducing the time required to develop models of enzymes for whole-cell modelling. At the core of our strategy is a general method for simplifying models of enzymes. The use of simplified rate laws rather than differential equations makes parameter estimation computationally efficient, and because our method is general for a wide range of enzyme cycles, it can be used to test the validity of numerous potential mechanisms against experimental data. We represent enzyme kinetics using bond graphs, a framework that is both energetic and modular (in sense that separate models can be sensibly coupled together into larger models, even if the individual models behave differently once coupled) (Gawthrop and Crampin, 2014). Our approach enables the development of simple, biochemically plausible and thermodynamically consistent models that explain the available experimental data.

In § 4.2.1.1, we review the essential concepts of reducing models of enzymes using a simple reversible Michaelis-Menten model. We then motivate the use of an energetic framework for representing enzyme kinetics (§ 4.2.1.2–4.2.1.3), and outline a generalised methodology for reducing models of enzyme kinetics under this energetic framework

(§ 4.2.2). In § 4.3.1–4.3.3, we verify that our approach can be used to generate thermodynamically consistent rate laws for a range of enzyme cycles. In § 4.3.4, we show how rate laws generated by our approach can be used to perform parameter estimation. Finally, we show how results can be compared between multiple potential mechanisms to perform model selection, and apply this approach to two examples: a hypothetical enzyme with a bi-bi mechanism in which substrates bind individually in an ordered sequence before the products are produced and sequentially released (§ 4.3.5–4.3.6); and a physiological example in the cardiac Na⁺/K⁺ ATPase (§ 4.3.7).

4.2 Methods

4.2.1 Reversible Michaelis-Menten enzyme

4.2.1.1 Kinetic formulation

We first review the essential principles of model reduction that we use in our approach. A simple Michaelis-Menten model (Figure 4.1) is used to illustrate how these principles can be applied to enzyme kinetics. Because all enzymes are thermodynamically reversible, we differ from the typical formulation in using a reversible second reaction (Figure 4.1A). Using the law of mass action, the rates of the two reactions are

$$v_1 = k_1^+ x_E x_S - k_1^- x_C \quad (4.1a)$$

$$v_2 = k_2^+ x_C - k_2^- x_E x_P \quad (4.1b)$$

and accumulation of free enzyme (E) and complex (C) are given by the ordinary differential equations (ODEs)

$$\dot{x}_E = -v_1 + v_2 = -k_1^+ x_E x_S + k_1^- x_C + k_2^+ x_C - k_2^- x_E x_P \quad (4.2a)$$

$$\dot{x}_C = v_1 - v_2 = k_1^+ x_E x_S - k_1^- x_C - k_2^+ x_C + k_2^- x_E x_P \quad (4.2b)$$

where k_1^+ , k_1^- , k_2^+ and k_2^- are the kinetic constants, and x_s is the molar amount of $s \in \{E, C, S, P\}$. Here we assume that all of the species exist within the same volume so that their molar amounts are proportional to their concentration.

In the context of whole-cell models, enzymes are typically smaller in concentration than their substrates, and substrates are generally maintained at relatively constant concentrations (Smith and Crampin, 2004). We therefore assume that the dynamics of the metabolites are slow relative to the dynamics of the enzyme, so that the enzyme is operating under a quasi-steady-state. Under this assumption – the quasi-steady-state approximation – we assume that (Keener and Sneyd, 2009):

Assumption A (*Quasi-steady-state assumption*) *The enzyme is at a steady-state for any given concentration of substrate (S) and product (P).*

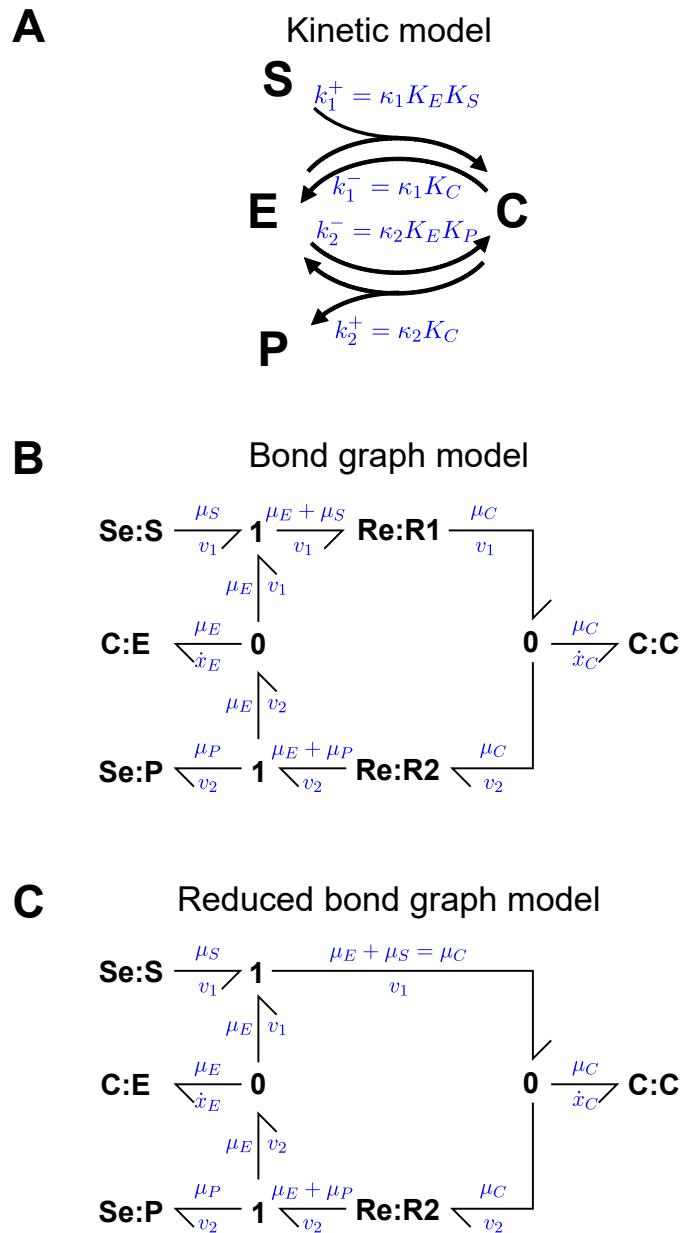


Figure 4.1: A reversible Michaelis-Menten model. **(A)** Kinetic representation, with the mapping between kinetic and bond graph parameters shown in blue; **(B)** Bond graph representation of the full model; **(C)** Bond graph representation of the model simplified using rapid equilibrium.

Under these conditions, the left hand side of Eq. 4.2 is zero:

$$0 = -k_1^+ x_E x_S + k_1^- x_C + k_2^+ x_C - k_2^- x_E x_P \quad (4.3a)$$

$$0 = k_1^+ x_E x_S - k_1^- x_C - k_2^+ x_C + k_2^- x_E x_P \quad (4.3b)$$

By solving the equations simultaneously, the amounts of enzyme and complex are

$$x_E = \frac{e_0(k_1^- + k_2^+)}{k_1^+ x_s + k_2^- x_P + k_1^- + k_2^+} \quad (4.4a)$$

$$x_C = \frac{e_0(k_1^+ x_S + k_2^- x_P)}{k_1^+ x_s + k_2^- x_P + k_1^- + k_2^+} \quad (4.4b)$$

where

$$e_0 = x_C + x_E \quad (4.5)$$

is the total amount of enzyme. Therefore, by substituting into Eq. 4.1a, the enzymatic reaction rate is

$$v = v_1 = \frac{e_0(k_1^+ k_2^+ x_S - k_1^- k_2^- x_P)}{k_1^+ x_S + k_2^- x_P + k_1^- + k_2^+} \quad (4.6)$$

A key advantage of the quasi-steady-state approximation is that all references to the individual enzyme states E and C have been discarded in the rate law, and the enzyme only affects the rate law through the total amount of enzyme e_0 (Gunawardena, 2014).

Rate laws can be further simplified by assuming that some steps occur significantly faster than others so that they are effectively at equilibrium. This is commonly applied to the Michaelis-Menten model through the rapid equilibrium approximation:

Assumption B (*Rapid equilibrium assumption*) *The first reaction is in rapid equilibrium.*

In this approximation, the forward and reverse fluxes of the first reaction are equal:

$$k_1^+ x_E x_S = k_1^- x_C \quad (4.7)$$

so the total amount of enzyme can be expressed as

$$e_0 = x_C + x_E = \left(1 + \frac{k_1^-}{k_1^+ x_S}\right) x_C \quad (4.8)$$

By substituting this into Eq. 4.1b, it can be shown that the reaction rate is

$$\begin{aligned} v = v_2 &= k_2^+ x_C - k_2^- x_P (e_0 - x_C) \\ &= \frac{e_0(k_1^+ k_2^+ x_S - k_1^- k_2^- x_P)}{k_1^+ x_S + k_1^-} \end{aligned} \quad (4.9)$$

We note that the rapid equilibrium approximation is distinct from the quasi-steady-state approximation; neither approximation implies the other, and they can be applied

separately or in conjunction with each other. In this example, the two approximations were applied to the full model separately. Both of these approximations have their strengths and weaknesses. The quasi-steady-state approximation is able to reduce enzyme kinetics to a single rate law, but does not reduce the number of parameters. By contrast, in general, the rapid equilibrium approximation reduces the number of parameters but does not reduce the kinetics of enzymes to a single rate law as individual enzyme states may still accumulate at different rates. Therefore, in later sections, we use rapid equilibrium in conjunction with the quasi-steady-state approximation to make use of the strengths of both.

4.2.1.2 Energetic formulation

While the kinetic formulation can be used to apply model reduction to enzyme kinetics, a weakness of this formulation is that kinetic parameters are not independent, as they contain both species and reaction-related properties. As a result, these parameters cannot be perturbed individually and modellers frequently need to apply additional detailed balance constraints to satisfy thermodynamic consistency (Ederer and Gilles, 2007). Here we describe an alternative formulation, the energetic formulation, which separates thermodynamic quantities from kinetic quantities (Oster et al., 1971; Gauthrop and Crampin, 2014). Accordingly, species and reactions are separate components with independent parameters. Each species s is associated with a chemical potential μ_s [J/mol], described by the equation

$$\mu_s = RT \ln(K_s x_s) \quad (4.10)$$

where $R = 8.314 \text{ J/mol/K}$ is the universal gas constant, $T [\text{K}]$ is the absolute temperature, $K_s [\text{mol}^{-1}]$ is the species thermodynamic constant, and $x_s [\text{mol}]$ is the molar amount of species. This chemical potential describes the propensity of the chemical species to drive reactions. Reactions are dissipative processes that operate in the direction of decreasing chemical potential. In the energetic formalism, a reaction r is described by the Marcelin-de Donder equation

$$v_r = \kappa_r \left(e^{A_r^f/RT} - e^{A_r^r/RT} \right) \quad (4.11)$$

where $v_r [\text{mol/s}]$ is the rate of reaction (which can be positive or negative), $\kappa_r [\text{mol/s}]$ is the reaction rate constant, $A_r^f [\text{J/mol}]$ is the forward affinity (the sum of chemical potentials of the reactants), and $A_r^r [\text{J/mol}]$ is the reverse affinity (the sum of chemical potentials of the products).

In the Michaelis-Menten model, the chemical potentials (labelled in Figure 4.1B) are

$$\mu_S = RT \ln(K_S x_S) \quad (4.12a)$$

$$\mu_P = RT \ln(K_P x_P) \quad (4.12b)$$

$$\mu_E = RT \ln(K_E x_E) \quad (4.12c)$$

$$\mu_C = RT \ln(K_C x_C) \quad (4.12d)$$

and the reaction affinities are

$$A_1^f = \mu_E + \mu_S \quad (4.13a)$$

$$A_1^r = \mu_C \quad (4.13b)$$

$$A_2^f = \mu_C \quad (4.13c)$$

$$A_2^r = \mu_E + \mu_P \quad (4.13d)$$

Using Eqs. 4.11–4.13, it can be shown that the reaction rates of the full system are

$$v_1 = \kappa_1 \left(e^{(\mu_E + \mu_S)/RT} - e^{\mu_C/RT} \right) = \kappa_1 K_E K_S x_E x_S - \kappa_1 K_C x_C \quad (4.14a)$$

$$v_2 = \kappa_2 \left(e^{\mu_C/RT} - e^{(\mu_E + \mu_P)/RT} \right) = \kappa_2 K_C x_C - \kappa_2 K_E K_P x_E x_P \quad (4.14b)$$

and the differential equations for the rates of change for the enzyme states are

$$\dot{x}_E = v_2 - v_1 = -\kappa_1 K_E K_S x_E x_S + \kappa_1 K_C x_C + \kappa_2 K_C x_C - \kappa_2 K_E K_P x_E x_P \quad (4.15a)$$

$$\dot{x}_C = v_1 - v_2 = \kappa_1 K_E K_S x_E x_S - \kappa_1 K_C x_C - \kappa_2 K_C x_C + \kappa_2 K_E K_P x_E x_P \quad (4.15b)$$

Therefore the energetic formulation gives rise to the mass action behaviour in Eq. 4.2 with the kinetic constants

$$k_1^+ = \kappa_1 K_E K_S \quad (4.16a)$$

$$k_1^- = \kappa_1 K_C \quad (4.16b)$$

$$k_2^+ = \kappa_2 K_C \quad (4.16c)$$

$$k_2^- = \kappa_2 K_E K_P \quad (4.16d)$$

It should be noted that while the energetic parameters K_s and κ_r give rise to the same kinetic behaviour as the kinetic constants, they also determine the chemical potentials associated with each species and reaction. Because energetic parameters are thermodynamically independent, they are guaranteed to be thermodynamically consistent and can be perturbed or updated independently (Liebermeister and Klipp, 2006; Gawthrop et al., 2015; Mason and Covert, 2019; Ederer and Gilles, 2007).

The analysis for the quasi-steady-state approximation follows that for the kinetic formulation with the substitutions in Eq. 4.16, therefore the rate of the catalysed reaction is

$$v = \frac{e_0 \kappa_1 \kappa_2 K_E K_C (K_S x_S - K_P x_P)}{\kappa_1 K_E K_S x_S + \kappa_2 K_E K_P x_P + \kappa_1 K_C + \kappa_2 K_C} \quad (4.17)$$

With some algebraic manipulation, Eq. 4.17 is equivalent to Eq. 5.16 of Gawthrop and Crampin (2014), which is an equation for the reaction rate derived from the same system.

The analysis for the rapid equilibrium approximation also follows a similar pattern to the kinetic framework. However, it is useful to note that the rapid equilibrium constraint in Eq. 4.7 arises due to an equilibration of affinities on both sides of reaction 1:

$$\mu_E + \mu_S = \mu_C \quad (4.18)$$

Therefore, by using Eq. 4.12, the constraint can be alternatively formulated as

$$\frac{x_C}{x_E x_S} = \frac{K_E K_S}{K_C} \quad (4.19)$$

This constraint can be used to show that the reaction rate is

$$v = \frac{\kappa_{R2} K_C K_E e_0 (K_S x_S - K_P x_P)}{K_C + K_E K_S x_S} \quad (4.20)$$

An advantage of formulating rapid equilibrium in terms of energetic parameters is that the approximation can be linked back to the parameter for the corresponding reaction. We note that Eq. 4.20 is Eq. 4.17 when we take $\kappa_1 \rightarrow \infty$. Because the steady-state cycling velocity is proportional to e_0 (Hill, 1989; Gunawardena, 2014), it is useful to define its normalised form $v_{\text{cyc}} = v/e_0$ as the cycling rate for use in later sections.

4.2.1.3 Bond graph representation

We use bond graphs to represent the energetic formulation because they are an effective and generalisable framework for encoding thermodynamic information. There are two advantages to using bond graphs for modelling enzymes. Firstly, they necessarily contain mechanistic information in their structure (see Chapter 2; and Gawthrop et al. (2015)). Secondly, bond graphs are inherently modular, allowing them to construct larger-scale systems such as glycolysis, oxidative phosphorylation and electrophysiology from smaller models (see Gawthrop et al. (2015); Gawthrop (2017a); Gawthrop et al. (2017); and Chapter 3). A bond graph representation of the reversible Michaelis-Menten model is shown in Figure 4.1B. In this representation, components are linked by bonds (\rightarrow) that represent the transfer of power. For biochemical systems, the power carried by each bond is decomposed into a chemical potential μ and a molar flow rate v . Because bond graphs explicitly model energy transfer, thermodynamic consistency is automatically satisfied along with many other conservation laws (Gawthrop and Bevan, 2007; Gawthrop and Smith, 1996; Borutzky, 2010). The quantities carried by each bond are constrained by the components they are connected to. These constraints dictate the behaviour of the system. The constraints for the Michaelis-Menten model (Eqs. 4.10–4.15) are captured using **C**, **Se**, **Re**, **0** and **1** components.

Dynamic chemical species are represented using **C** components, which are energy storage components that are named as such due to their analogy with electrical capacitors. There are two **C** components in the bond graph representation corresponding to the dynamic species E and C. **C** components encode the constraints in Eqs. 4.12c–4.12d. The two **C** components are dynamic as the molar amounts of species are allowed to vary based on the molar flow rates of their connected bonds.

The metabolites S and P are represented using **Se** components. **Se** components relate chemical potential and molar amount in the same way as **C** components (Eqs. 4.12a–4.12b). However, in contrast to the enzyme states, concentrations of metabolites are assumed to be approximately constant on the timescale of enzyme kinetics. **Se** components approximate this behaviour by fixing the concentrations of these metabolites to constant values ($\dot{x}_S = \dot{x}_P = 0$). These species with fixed concentrations are known as chemostats. Because external flows are required to keep the concentrations of chemostats constant, chemostats represent connections to external energy sources or sinks. When incorporating enzyme kinetics models into broader whole-cell models with variable metabolite concentrations, external connections can be formalised in terms of **SS** components (Gawthrop et al., 2015).

Reactions are represented using **Re** components. There are two **Re** components in Figure 4.1B corresponding to the two reactions of the Michaelis-Menten model, and these encode Eq. 4.11. Because reactions dissipate energy, they are analogous to electrical resistors.

Conservation laws are represented using **0** and **1** junctions. The **0** junction is used when a species is involved in multiple reactions. This junction encodes the conservation of mass law seen in the first equality of Eq. 4.15, and also ensures that the chemical potential contributed by the species to each reaction is equal (see chemical potentials connected to the **0** junction in Figure 4.1B). Therefore the **0** junction is a generalisation of Kirchhoff's current law. Similarly, **1** junctions are used when more than one species is present on either side of a reaction. This junction encodes the conservation of potential law in Eqs. 4.13a,4.13d while ensuring that the molar flow rates of the connected bonds are equal (see Figure 4.1B). Thus the **1** junction is a generalisation of Kirchhoff's voltage law. Via the analysis in § 4.2.1.2, the constraints encoded by the bond graph components are sufficient to derive the dynamics of the Michaelis-Menten model.

Because the bond graph representation is energetic, it can naturally represent the rapid equilibrium approximation. As seen in Eq. 4.18, rapid equilibrium is the equilibration of the forward and reverse affinities. Bond graphs can represent this constraint through the removal of the corresponding **Re** component (Figure 4.1C) (Gawthrop and Crampin, 2014).

For a more comprehensive description of bond graph theory, the reader is referred to the textbooks by Gawthrop and Smith (1996) and Borutzky (2010). The application of bond graphs to biochemical systems is detailed in Gawthrop and Crampin (2014), Gawthrop et al. (2015) and Chapter 2.

4.2.2 Model reduction

To extend our analysis to arbitrarily large enzyme cycles, we use timescale separation and stoichiometric matrices (Klipp et al., 2009) to derive explicit equations for the cycling rate. Because these equations are derived explicitly, they are easily generated using software and scalable to complex enzyme models. We briefly summarise the method here, with the full method described in Appendix 4.A.

As seen in § 4.2.1.1, the derivation of a rate law involves three constraints corresponding to the rapid equilibrium approximation (Eq. 4.19), quasi-steady-state approximation (Eq. 4.3), and the conservation law arising from conservation of enzyme (Eq. 4.5). The *rapid equilibrium* constraint arises from the equilibrium point of the fast timescale

$$\mathbf{D}_{\text{re}}X^{\text{cd}} = 0 \quad (4.21)$$

while the *quasi-steady-state* constraint arises from the steady state of the slow timescale

$$\mathbf{D}_{\text{ss}}X^{\text{cd}} = 0 \quad (4.22)$$

X^{cd} is a vector containing the amounts of each enzyme state, and \mathbf{D}_{re} and \mathbf{D}_{ss} are matrices that depend on the energetic parameters, metabolite concentrations and the

stoichiometry of reactions within the enzyme cycle. Eqs. 4.21 and 4.22 are derived in Appendix 4.A, and their full forms are given in Eqs. 4.42 and 4.51 respectively. Note that for computational efficiency, the rapid equilibrium constraint is derived prior to deriving the quasi-steady-state constraint to reduce the dimensionality of the equations. The steady state of the system is dependent on the amount of enzyme present in the initial condition of the system:

$$GX^{\text{cd}} = GX^{\text{cd}}(0) \quad (4.23)$$

For the enzyme cycles examined in this study, G is a row vector of ones, and $e_0 = GX^{\text{cd}}(0)$ is the total amount of enzyme. Because \mathbf{D}_{re} , \mathbf{D}_{ss} and G are independent of X^{cd} , linear methods can be used to simultaneously solve Eqs. 4.21–4.23 for X^{cd} , which can then be used to derive an analytical expression for the cycling rate.

4.2.3 Parameter estimation

Energetic parameters were estimated by minimising a cost function. The cost function was minimised using differential evolution (Price et al., 2005) as the global optimiser, the results of which were used as initial guesses for a local optimiser using a quasi-Newton method (Nocedal and Wright, 2006, Ch. 6).

Energetic parameters describe the independent contribution of each species to equilibrium constants. As a result, there are families of energetic parameters that give rise to the same kinetic behaviour (see Chapter 2 and Gawthrop et al. (2015)) and therefore families of parameter values rather than specific parameter values arise from fitting kinetic measurements (see Chapter 2 and Mason and Covert (2019)). By analysing the nullspaces of the stoichiometric matrix (Gawthrop et al., 2015; Klipp et al., 2009), we developed a fitting procedure that expresses the resulting parameters as mathematical expressions with free parameters that can be assigned arbitrarily (Appendix 4.B). The energetic parameters associated with species and reactions are positive, therefore they were log transformed during optimisation for computational efficiency (Hass et al., 2019).

A squared residuals cost function was used for the four-state model (§ 4.3.4). For the Na^+/K^+ ATPase (§ 4.3.7), the weighted squared residuals cost function described in Appendix 2.C was used.

4.2.4 Model selection

The performance of numerous alternative models can be compared by performing parameter estimation on each and recording the cost function. To perform model selection, we added a penalty for model complexity. We chose a penalty that scaled linearly with model complexity so that the adjusted cost would be

$$J_{\text{AIC}} = J_{\text{misfit}} + c_p n_p \quad (4.24)$$

where J_{misfit} is the cost associated with differences between model predictions and experimental measurements, c_p is the cost per parameter and n_p is the number of fitted parameters. This cost function has a similar form to the Akaike information criterion

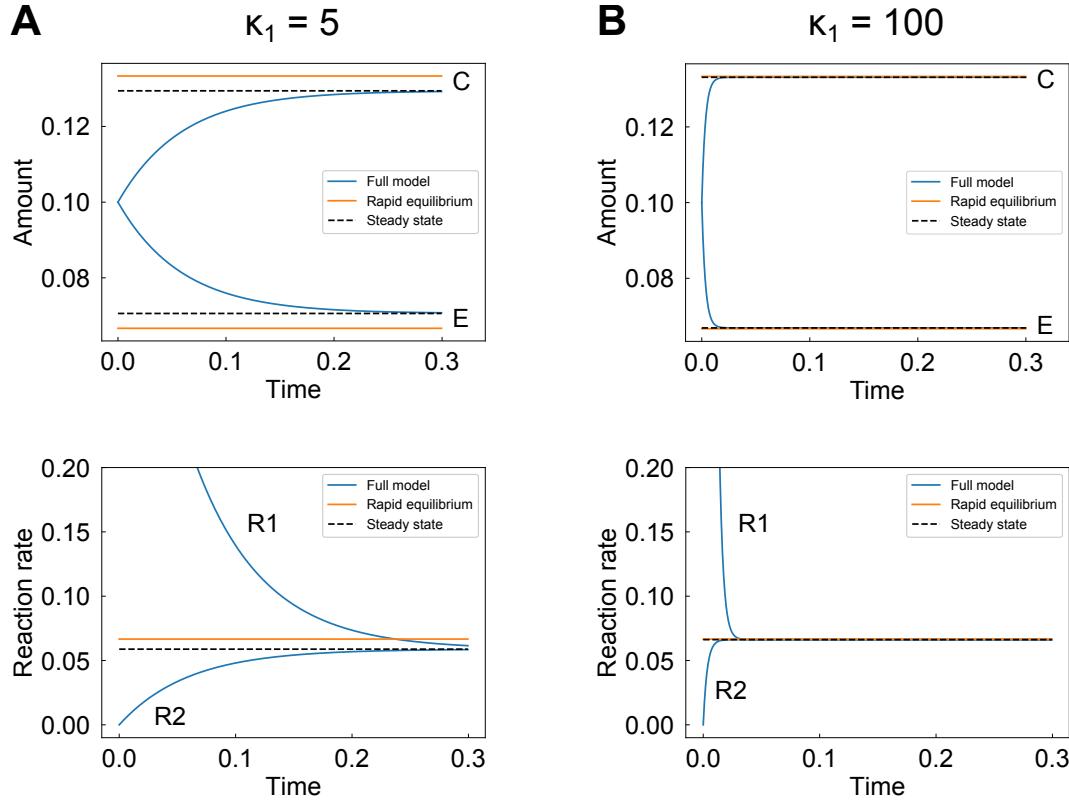


Figure 4.2: Simulations of the Michaelis-Menten model. Simulations were performed with (A) $\kappa_1 = 5$ and (B) $\kappa_1 = 100$. The amounts of the enzyme states E and C are shown on the top panel, and the reaction rates are shown on the bottom panel. Because the reaction rates converge to the same value at steady state, there is only one reaction rate for the rapid equilibrium and quasi-steady-state approximations.

(AIC) (Akaike, 1974). However, while the AIC uses statistical reasoning to choose an appropriate value of c_p , we use a more utility-driven approach to optimise for model complexity by setting c_p to the cost associated with a 1% misfit error.

4.3 Results

4.3.1 Simplifications of the reversible Michaelis-Menten model are thermodynamically consistent

To verify that the rapid equilibrium and quasi-steady-state approximations accurately represent limiting behaviours of the Michaelis-Menten model in the energetic formulation, we simulate the full model to steady state, and compare the results to values predicted by the reduced equations from the rapid equilibrium and quasi-steady-state approximations (Figure 4.2). We set the energetic parameters to $\kappa_2 = 1$, $K_S = 2$, $K_E = K_C = K_P = 1$. The metabolite amounts were fixed at $x_S = x_P = 1$, and the enzymes states were

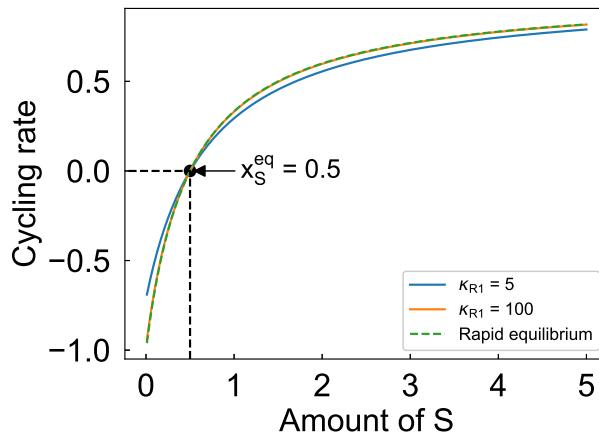


Figure 4.3: Variation of the reversible Michaelis-Menten model with the amount of substrate. The dotted black lines indicate the equilibrium of the enzyme ($v = 0$), which corresponds to the thermodynamic equilibrium of $x_S = K_P x_P / K_S = 0.5$.

initialised to $x_C(0) = x_E(0) = 0.1$. The full model is the only model with dynamic behaviour. Because the total amount of enzyme is much smaller than that of the metabolites, the enzyme achieves a steady state relatively quickly. At steady state, the amounts of each enzyme state achieve a constant value, and the reaction velocities converge to the same value. When we set the rate of the first reaction to $\kappa_1 = 5$, quantities within the full model approach those specified by the quasi-steady-state approximation as time progresses, as expected (Figure 4.2A). However, there is a noticeable difference between the quasi-steady-state approximation and the rapid equilibrium approximation because the rate of the first reaction is not substantially faster than the second reaction. When κ_1 is increased to 100, the system achieves a steady state far quicker (Figure 4.2B). As with $\kappa_1 = 5$, the steady state of the full model matches that of the quasi-steady-state approximation. However, in contrast to Figure 4.2A, the rapid equilibrium and quasi-steady-state approximations are similar, and therefore the full model also approaches the rapid equilibrium approximation. Therefore, the quasi-steady-state approximation accurately captures the steady-state behaviour of the system while the rapid equilibrium approximation is accurate for higher values of κ_1 .

Using the equations for the rapid equilibrium and quasi-steady-state approximations, we plot how the cycling rate v_{cyc} varies with the amount of substrate. Curves for the full model (with $\kappa_1 = 5$ and 100) and rapid equilibrium approximation are plotted in Figure 4.3. Following from Figure 4.2, the cycling rates at rapid equilibrium match those for $\kappa = 100$ more closely than the cycling rates for $\kappa_1 = 5$. The curves resemble typical Michaelis-Menten curves in that the cycling rate plateaus for high amounts of substrate. However, because an energetic approach requires reversibility, there is a reverse mode of operation when the chemical potential of the substrate is lower than that of the product. The dotted black lines indicate this equilibrium, which occurs at $x_S = 0.5$. This corresponds to the thermodynamic equilibrium of $x_S = K_P x_P / K_S = 0.5$. Thus, regardless of whether a full model or a reduced form of the model is used, the energetic

formulation accounts for the equilibrium point, which explains observations such as product inhibition and enzyme reversibility (Frieden and Walter, 1963).

4.3.2 The principles of model reduction are generalisable to complex enzyme cycles

To demonstrate that our method is generalisable, we apply it to the more complex four-state enzyme cycle shown in Figure 4.4A, which represents an ordered bi-bi mechanism (Segel, 1975). Reactions 1 and 4 (enclosed in dotted blue boxes) are assumed to be in rapid equilibrium. The cycling rate of this enzyme is given by

$$v_{\text{cyc}} = K_{E1}K_{E2}K_{E3}K_{E4}e_0\kappa_{R2}\kappa_{R3}(K_Ax_AK_Bx_B - K_Cx_CK_Dx_D)/\Sigma(x_A, x_B, x_C, x_D) \quad (4.25)$$

where

$$\begin{aligned} \Sigma(x_A, x_B, x_C, x_D) = & K_{E1}K_{E2}K_{E3}K_Dx_D\kappa_{R2} + K_{E1}K_{E2}K_{E3}K_Dx_D\kappa_{R3} \\ & + K_{E1}K_{E2}K_{E4}K_Ax_AK_Bx_B\kappa_{R2} + K_{E1}K_{E2}K_{E4}K_Cx_CK_Dx_D\kappa_{R3} \\ & + K_{E1}K_{E3}K_{E4}K_Ax_A\kappa_{R2} + K_{E1}K_{E3}K_{E4}K_Ax_A\kappa_{R3} \\ & + K_{E2}K_{E3}K_{E4}\kappa_{R2} + K_{E2}K_{E3}K_{E4}\kappa_{R3} \end{aligned} \quad (4.26)$$

This expression is far more complicated than the expression for the Michaelis-Menten example, underlining the importance of using software to generate these equations, as manually deriving these equations is slow and error-prone.

4.3.3 The reduced rate laws generated by our approach are thermodynamically consistent

To verify that the cycling rate (Eqs. 4.25–4.26) of the four-state model exhibits the expected limiting behaviour as the rates of reactions 1 and 4 increase, we plot the cycling rate for the full model and the model approximated by rapid equilibrium (Figure 4.5A). We use the parameters $\kappa_2 = \kappa_3 = 1$, $K_{E1} = 4$, $K_{E2} = 2$, $K_{E3} = 1$, $K_{E4} = 2$, $K_A = 2$, $K_B = 1$, $K_C = 1$ and $K_D = 2$ and set the rates of reactions 1 and 4 to the same value κ that we vary. As κ increases, the magnitude of the flux increases in both the forward and reverse modes of operation, and the curve approaches that of the rapid equilibrium approximation, as expected.

Because the model is specified within an energetic formalism, the simplified rate law is thermodynamically consistent. The free energy of the overall reaction $A + B \rightleftharpoons C + D$ is

$$\Delta G = \mu_C + \mu_D - \mu_A - \mu_B = RT \ln \left(\frac{K_C K_D x_C x_D}{K_A K_B x_A x_B} \right) \quad (4.27)$$

At equilibrium, $\Delta G = 0$, hence the metabolites must satisfy the constraint

$$K_A K_B x_A x_B = K_C K_D x_C x_D \quad (4.28)$$

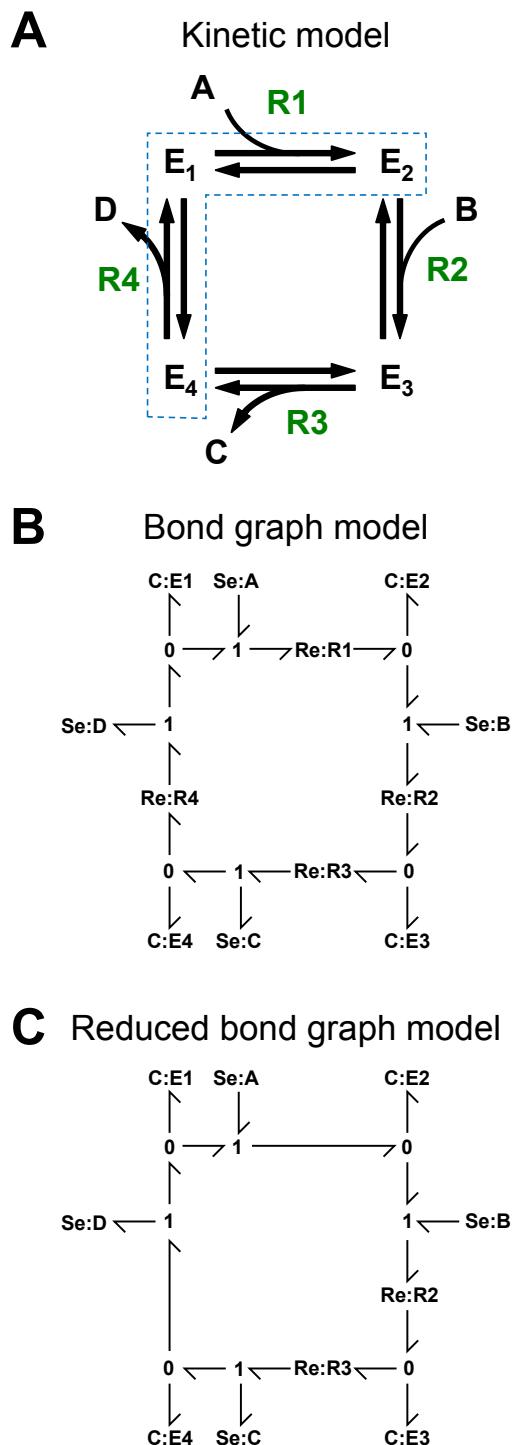


Figure 4.4: A four-state enzyme cycle. (A) Kinetic representation, with the reaction names shown in green. Reactions assumed to be in rapid equilibrium are enclosed in the dotted blue box. (B) Bond graph representation of the full model; (C) Bond graph representation of the model with the rapid equilibrium approximation.

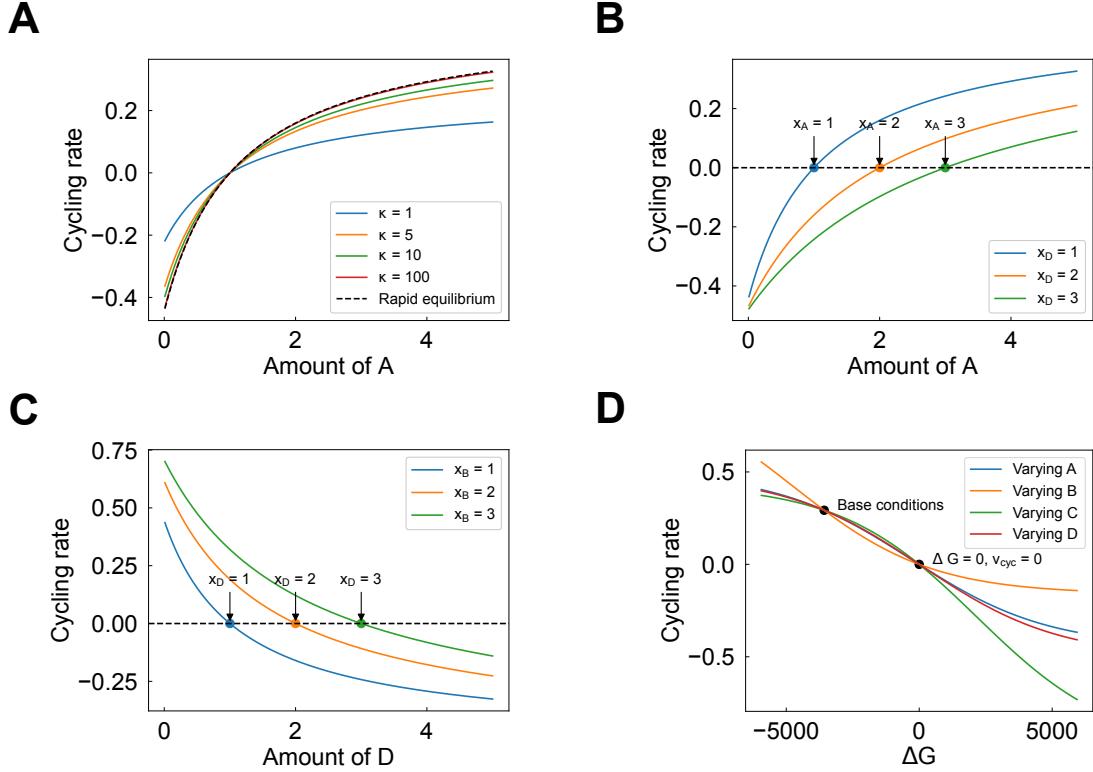


Figure 4.5: The cycling rate of the four-state enzyme model. (A) Comparison of cycling rates of the full model with different values of κ , when compared to the model with the rapid equilibrium approximation. Using the reduced form of the model at rapid equilibrium, the cycling rate is plotted against (B) the amount of A, and (C) the amount of D. Equilibrium points occur at $v_{cyc} = 0$, which are indicated by dots annotated by the amount at that point. For (A)–(C), the metabolites A, B, C and D are fixed at an amount of 1 unless otherwise indicated in the figure axes and legend. (D) The relationship between Gibbs free energy and cycling rate when different metabolites are varied to change the free energy. Metabolite amounts were varied from the base conditions $(x_A, x_B, x_C, x_D) = (2, 1, 1, 1)$.

This is illustrated in Figure 4.5A, where regardless of whether the full or reduced model is used, the cycling rate is zero at the equilibrium point

$$x_A = \frac{K_C K_D x_C x_D}{K_A K_B x_B} = 1 \quad (4.29)$$

which corresponds to the thermodynamic equilibrium described by Eq. 4.28.

We plot how the cycling rate varies with the amount of A (Figure 4.5B). Since A is a substrate for the overall reaction, increasing its amount drives the enzyme cycle forward. Equilibrium points are indicated by the dots. Increasing the amount of D reduces the cycling rate, and moves the equilibrium point to the right. In all three cases, the equilibrium point satisfies the thermodynamic equilibrium described in Eq. 4.28.

Similarly, we plot how the cycling rate varies with the amount of D (Figure 4.5C). As D is a product, increasing D drives the enzyme in the reverse direction. Three curves are

Table 4.1: Energetic parameters estimated using synthetic data for the four-state enzyme model. The variables t_0 , t_1 , t_2 and t_3 are free parameters that can be set to any positive value. During optimisation, these free variables were set to 1.0. The constants for the estimated values were equal to 1 up to 7 significant figures.

Parameter	True value	Estimated value
κ_{R1}	1	$1.0 \cdot t_1/(t_0 t_2)$
κ_{R2}	1	$1.0/(t_0 t_2)$
κ_{R3}	1	$1.0/(t_0 t_2)$
κ_{R4}	1	$1.0/t_0$
K_{E1}	1	$1.0 \cdot t_0/t_3$
K_{E2}	1	$1.0 \cdot t_0 t_2/t_1$
K_{E3}	1	$1.0 \cdot t_0 t_2$
K_{E4}	1	t_0
K_A	1	$1.0 \cdot t_2 t_3/t_1$
K_B	1	t_1
K_C	1	t_2
K_D	1	t_3

generated for different amounts of B, and increasing B shifts the equilibrium to the right, as expected. As with Figure 4.5B, the simplified expression for the cycling rate obeys the thermodynamic equilibrium described by Eq. 4.28.

To further check the thermodynamic consistency of the system, we plotted the Gibbs free energy of the overall reaction against cycling rate (Figure 4.5D). Since ΔG can be changed by varying any of the metabolites, four curves were generated by varying each individually from a non-equilibrium regime of operation indicated by the black dot in the upper left quadrant. While varying each of the metabolites resulted in different kinetic behaviour, all curves pass through the equilibrium point $\Delta G = 0$, $v_{\text{cyc}} = 0$, showing that the system is thermodynamically consistent under a variety of conditions.

4.3.4 Reduced rate laws and steady-state measurements can be used to parameterise energetic models of enzymes

Because parameter estimation is a key process in model development, we explore how this process applies to energetic parameters using the four-state model described in § 4.3.2. We assume that the true model has all parameters set to unity. To simulate an enzyme assay, we generate synthetic data for the concentrations $(x_A, x_B, x_C, x_D) \in \{1, 2, 3, 4, 5\}^4$, resulting in 625 steady-state measurements. As seen in Figure 4.6B, the four-state model provides a perfect fit to synthetic data generated from itself. The true and estimated parameters are compared in Table 4.1. Because the mapping between kinetic and energetic parameters is non-unique, the possible energetic parameters lie on a set described by four free parameters. Up to the selection of these free parameters, the energetic parameters were estimated almost perfectly (Table 4.1).

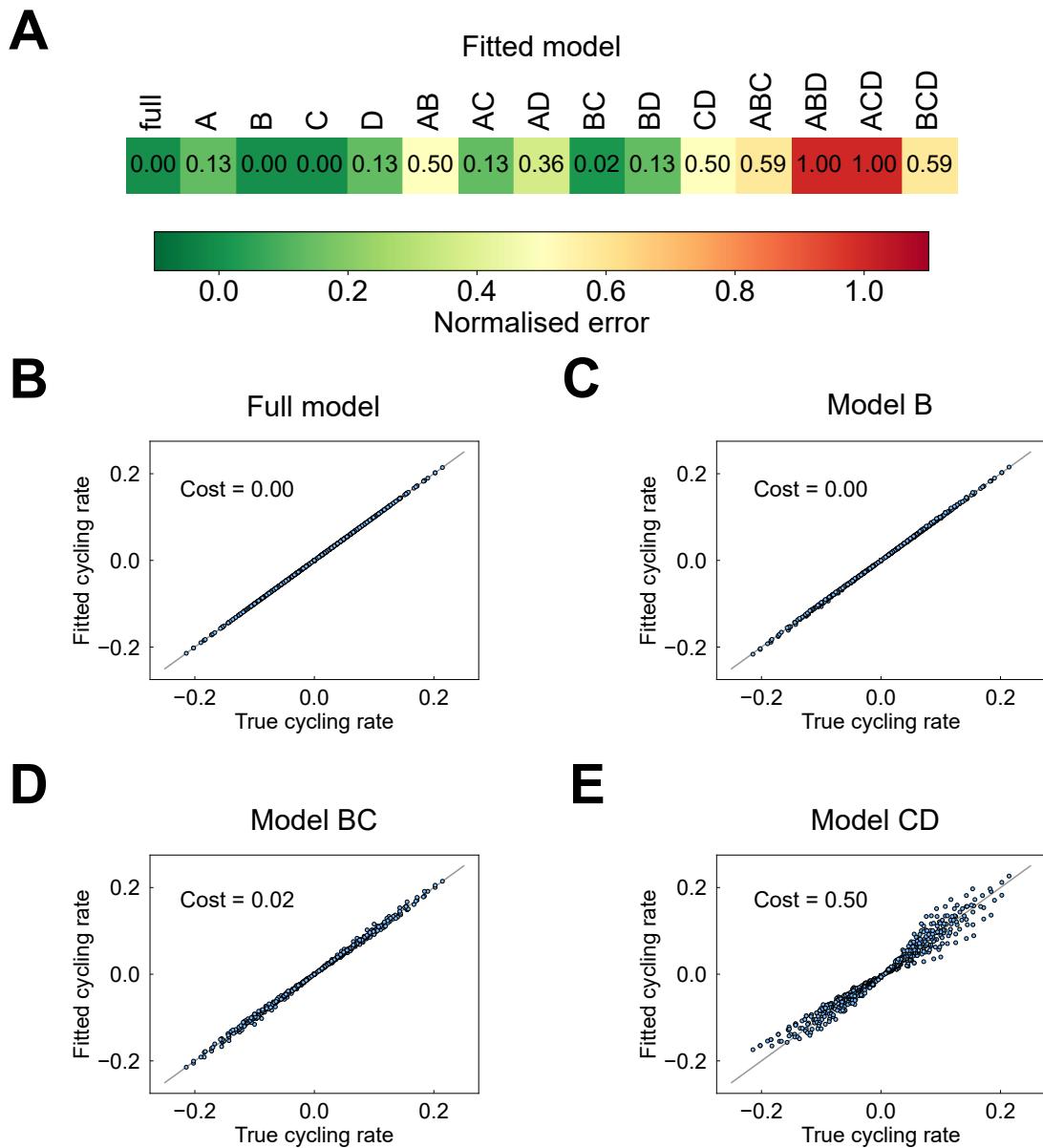


Figure 4.6: Fit of reduced versions of the four-state model to synthetic data generated from the full model. (A) A table of the costs associated with fitting each of the reduced models. The reduced models are named via the metabolites involved in the fast reactions, for example, AC refers to the model with reactions 1 and 3 at rapid equilibrium. Comparisons between the synthetic data and fitted models are plotted for (B) the full model; (C) Model B; (D) Model BC; (E) Model CD. The dots indicate model predictions, and the grey line indicates equality between the synthetic data and fitted model. Costs are normalised to the maximum value of 0.3042.

4.3.5 Models simplified using rapid equilibrium may provide good fits to steady-state measurements

To investigate the impact of rapid equilibrium approximations on the quality of fit, we fitted all possible approximations of the four-state model to synthetic data generated from the full model (Figure 4.6A). The reduced models are named according to the binding steps at rapid equilibrium, i.e. Model AB represents the model with reactions 1 and 2 at rapid equilibrium. There is a gradual increase in cost towards the right side of Figure 4.6A, indicating that simplifying models weakens their quality of fit as one would expect. More specifically, models that are simplifications of other models provide a worse fit, i.e. Model AB fits the data worse than Model A. Despite this, there are several approximations that fit the synthetic data well. Model B provides a near-perfect fit (Figure 4.6C), and Model BC, which contains two reactions in rapid equilibrium, fits the data reasonably well (Figure 4.6D). This was surprising given that the original model did not have any reactions operating on faster timescales, suggesting that simplified models may be sufficient in describing steady-state measurements.

4.3.6 Automatic derivation of rate laws allows multiple mechanisms to be tested against data

Because our method automatically generates cycling rates from enzymatic mechanisms, it is possible to test several alternative models against data. We test the feasibility of this approach by generating synthetic data using the four-state model in Figure 4.4 with the parameters

$$(\kappa_1, \kappa_2, \kappa_3, \kappa_4, K_A, K_B, K_C, K_D, K_{E1}, K_{E2}, K_{E3}, K_{E4}) = (1, 2, 3, 4, 1, 2, 3, 4, 1, 1, 1, 1) \quad (4.30)$$

The synthetic data generated from this “true” model was used to parameterise all possible mechanisms with 1–4 states (100 models in total). The fits are plotted in Figure 4.7, where each model is represented as a point and the models are grouped according to their complexity (increasing from left to right). With 81 measurements (Figure 4.7A), both the minimum and mean costs improve with complexity, although the effect plateaus after three states. The correct mechanism (red dot; four states) is not unique in providing a good fit to data, and many other mechanisms fit the data similarly well. This suggests that three-state mechanisms are sufficient for explaining this particular set of data. When the number of measurements is increased to 625 (Figure 4.7B), there is a similar trend of improved fits with complexity. However, in comparison to the case with 81 measurements, the correct mechanism appears to be slightly more accurately resolved. While the correct mechanism is still not unique in providing a perfect fit, there are fewer models that provide similar levels of fit. These results suggest that testing the fits of multiple models could be used to reject those that are inconsistent with the data. With more measurements, a greater proportion of models can be rejected, although it is unclear whether the correct mechanism can be uniquely isolated using steady-state measurements alone. Nonetheless, a comparison of the costs associated with fitting could potentially be used in a model selection scheme, as discussed in the next section.

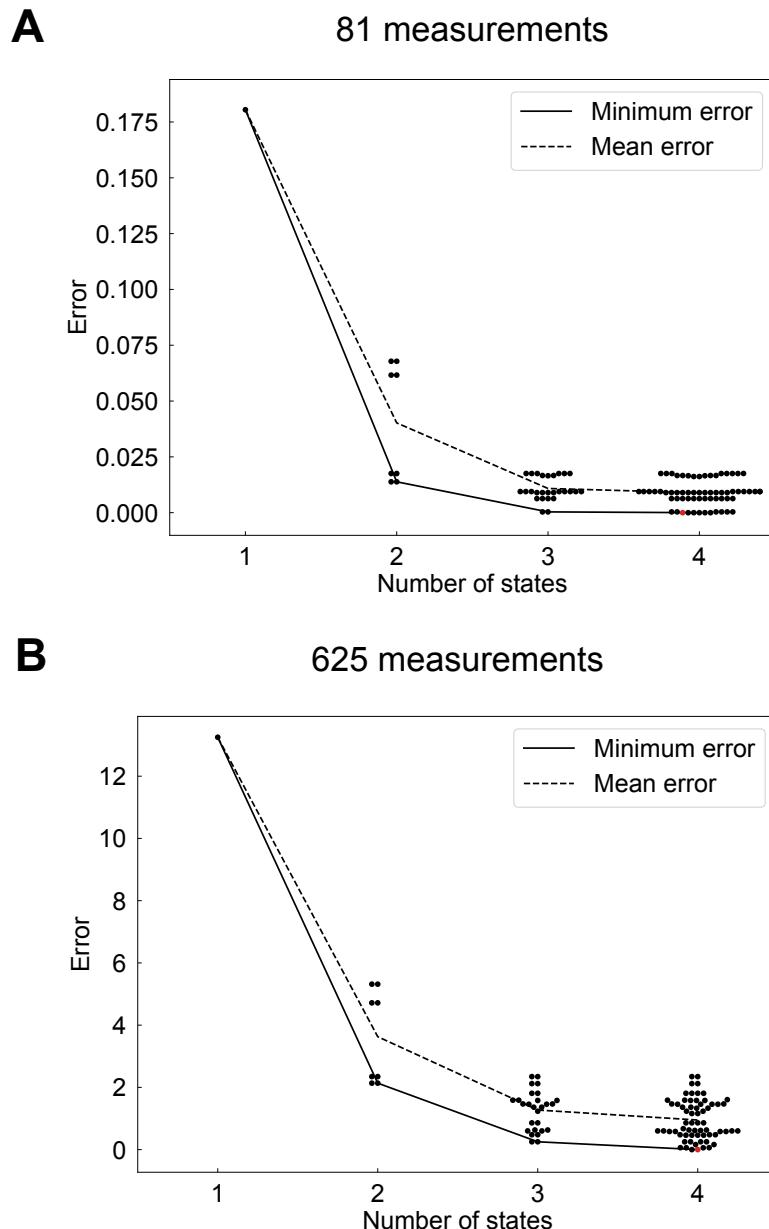


Figure 4.7: Fits of all possible enzymatic mechanisms to synthetic data generated from the four-state model. (A) Models fitted to 81 data points with concentrations varied individually from an initial set of values; (B) Models fitted to 625 data points. Dots represent potential models and are grouped according to the number of enzyme states. The model with the correct mechanism is coloured in red. The solid black line represents the minimum error for each number of states, and the dashed line represents the mean error of all models with a certain number of states.

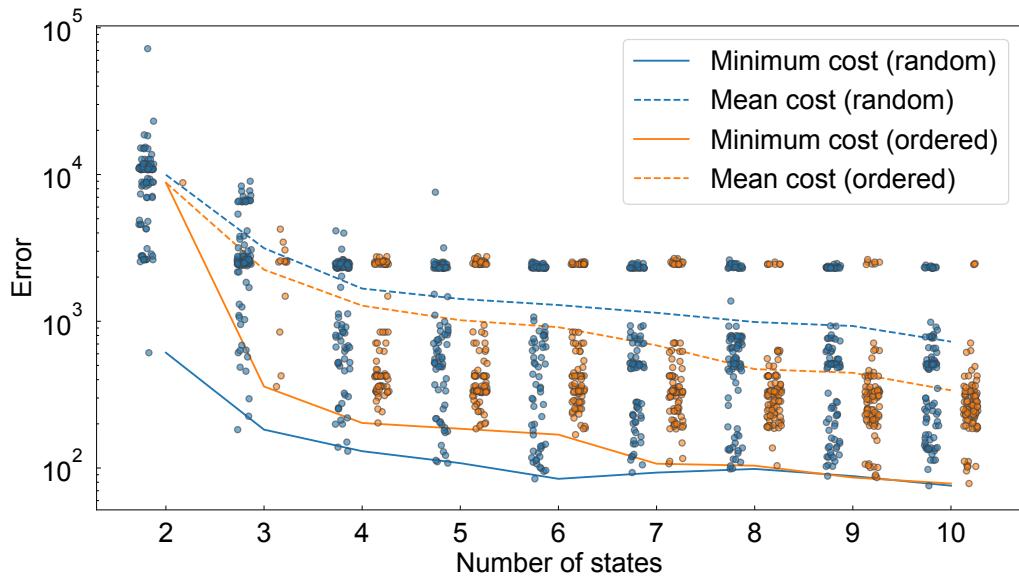
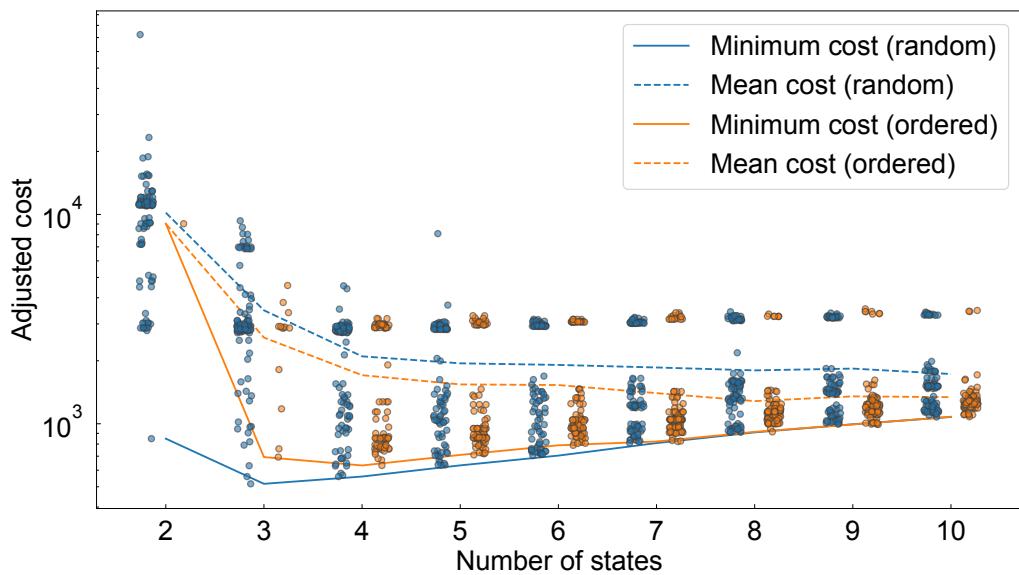
A**B**

Figure 4.8: Fits of possible models of the Na^+/K^+ ATPase to data. For each number of states, up to 100 models were sampled. Models that were sampled randomly with unordered binding steps are represented as blue dots, and models with binding steps ordered using the Post-Albers scheme are represented as orange dots. **(A)** Misfit cost; **(B)** Costs after adjusting for model complexity.

4.3.7 Automatic derivation of rate laws enables model selection for the cardiac Na⁺/K⁺ ATPase

The exact mechanism of an enzyme is often unknown prior to the process of model development. The results of the previous section suggest that there is a degree of mechanistic uncertainty associated with fitting models to steady-state measurements that are typically gathered in enzyme assays (Schnell et al., 2006). A practical choice for dealing with this uncertainty is to select the simplest model that fits the data well, as simple models are less prone to over-fitting (Crampin et al., 2004b). We applied this approach to a physiological example by performing both model selection and parameter estimation to experimental data gathered from the cardiac Na⁺/K⁺ ATPase.

The Na⁺/K⁺ ATPase is an active ion transporter that pumps Na⁺ out of the cell in exchange for the transport of K⁺ into the cell. As both of these processes are against the typical electrochemical gradients within cells, a source of energy is required to drive this process: ATP hydrolysis. Therefore the reaction catalysed by this enzyme is



Because ions are transported across a charged membrane, there is interplay between chemical and electrical energy in this process. An advantage of the bond graph representation is that it is general enough to represent both chemical and electrical processes, making it possible to model electrogenic enzymes such as the Na⁺/K⁺ ATPase. Further detail on the incorporation of the membrane potential and thermodynamic constraints are described in Appendix 4.C.

The Na⁺/K⁺ ATPase is believed to act via the Post-Albers mechanism (Apell, 1989) that is used in many existing models (Smith and Crampin, 2004; Terkildsen et al., 2007; Pan et al., 2019). Alternative models of the Na⁺/K⁺ ATPase can be generated by assigning individual reactants and products to various binding and unbinding steps in different orders. To examine the effects of incorporating known mechanistic information, we employed two model generation strategies: the “random” (or “unordered”) and “ordered” strategies. In the random strategy, binding steps are assigned in a random order. In the ordered strategy, we only sample from the subset of models in which the binding steps are ordered according to the Post-Albers cycle, and adhere to further structural constraints to prevent reactions from containing binding steps on both the intracellular and extracellular compartments (see Appendix 4.C for further detail). Thus the random sampling strategy may represent a situation where we are performing model development in the absence of prior mechanistic information, whereas the ordered sampling strategy is a representation of a situation where model development is informed by known mechanistic information. Because the number of possible mechanisms of the Na⁺/K⁺ ATPase are too numerous for an exhaustive search to be computationally feasible, we limited our analysis to 100 models for each number of states, generated using the random and ordered sampling strategies.

We plot the misfit cost in Figure 4.8A. For both sampling strategies, increasing model complexity generally improves their fits to data, however the benefits decrease with the number of states. For the unordered (blue) models, there is little improvement in the quality of fit after six states. A similar pattern emerges in the ordered (orange) models, although there is a more consistent improvement in fitting with the number of states.

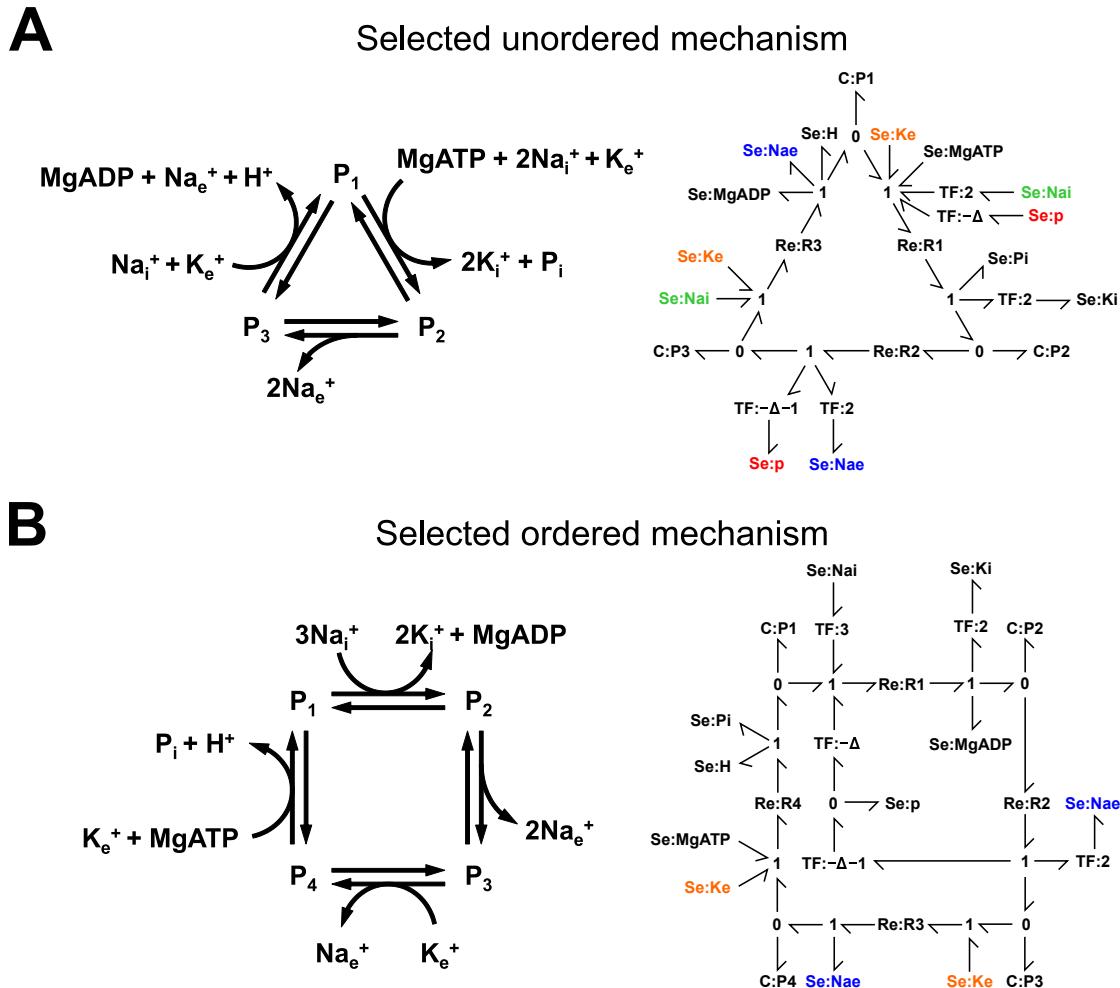


Figure 4.9: Selected mechanisms from fitting potential models to data for the Na^+/K^+ ATPase. (A) Selected unordered model; (B) Selected ordered model. Biochemical networks are shown on the left, and bond graph representations are shown on the right. Due to space constraints, a bond graph shorthand was used, where **Se** components with the same name (emphasised with colour) are connections to the same **Se** component through a common **0** junction (P.J. Gawthrop, personal communication, 2018).

If the minimum cost is used as a measure of the quality of fit, randomly sampled models outperform ordered models (Figure 4.8A, solid lines) with 7 or less states, while the two approaches achieve a similar performance with more states. While unordered models outperform ordered models in terms of minimum cost, ordered models perform better in terms of mean cost (Figure 4.8A, dotted lines). A large contributor to this discrepancy is the population of models that fit the data poorly; these can be seen in the band of models in Figure 4.8A with a cost between 2000 and 3000.

We used Eq. 4.24 to penalise model complexity ($n_p = 2n + 1$, where n is the number of states), and plot the adjusted costs in Figure 4.8B. There is now a clear minimum for the best performing model against the number of states. The optimal randomly sampled model has three states (Figure 4.9A), while the optimal ordered model has four

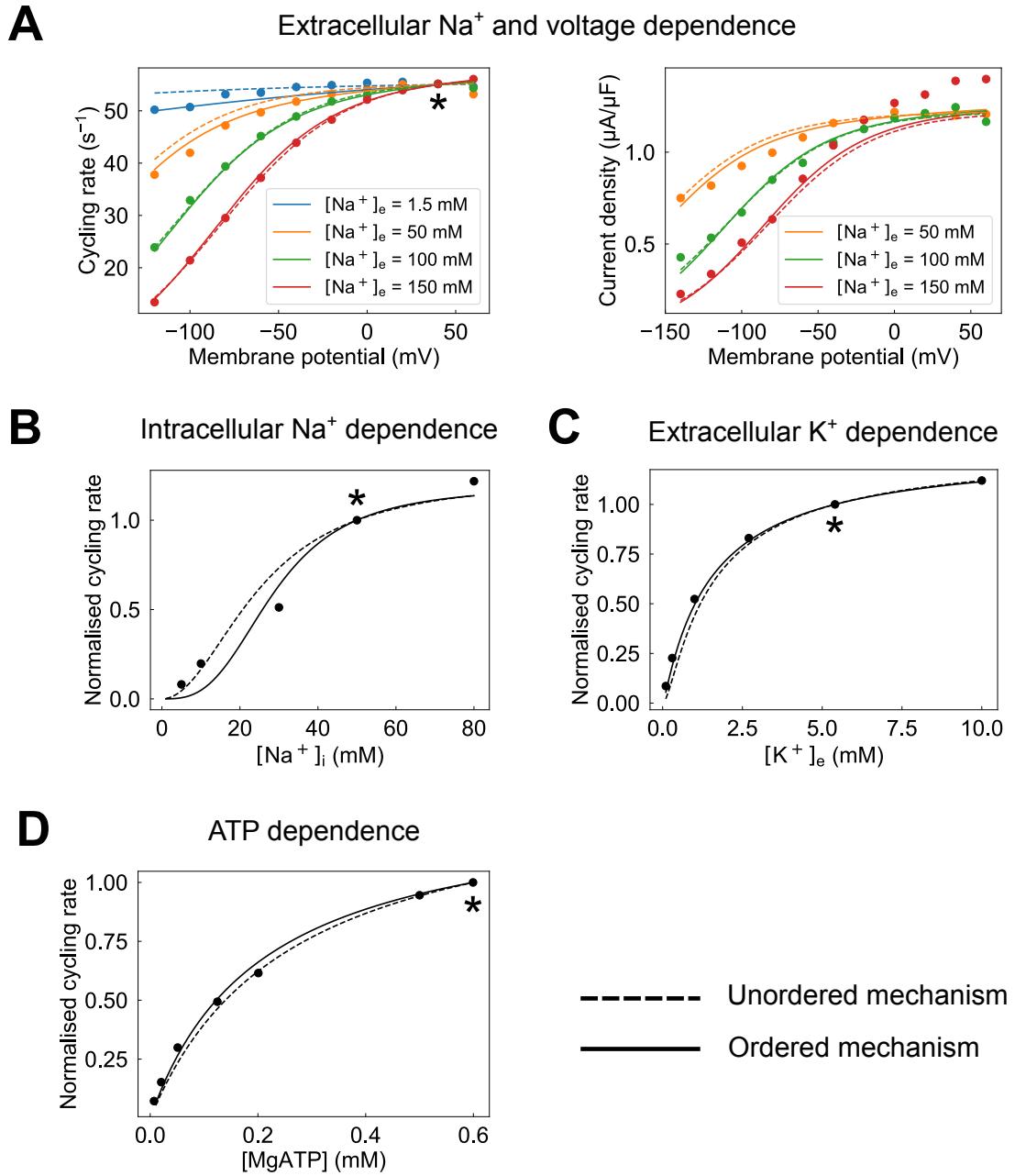


Figure 4.10: Fits of selected Na^+/K^+ ATPase mechanisms to data. Model simulations were compared to (A) extracellular Na^+ data (Nakao and Gadsby, 1989, Figs. 3 and 2A) containing cycling rate (left) and current density (right) measurements; (B) intracellular Na^+ data (Hansen et al., 2002, Fig. 7A); (C) extracellular K^+ data (Nakao and Gadsby, 1989, Fig. 11A); (D) ATP dependence data (Friedrich et al., 1996, Fig. 3B). Simulations were performed under the conditions specified in Table 4.2 and $T = 310$ K. Where present, the asterisks (*) indicate normalisation points.

Table 4.2: Simulation conditions for Figure 4.10. Voltages (V) have unit mV, concentrations have unit mM and the pH is dimensionless. Empty entries indicate that the variable was experimentally varied.

Figure	V	$[Na^+]_i$	$[Na^+]_e$	$[K^+]_i$	$[K^+]_e$	$[MgATP]$	$[MgADP]$	$[Pi]_{tot}$	pH
4.10A	–	50	–	0	5.4	10	0	0	7.4
4.10B	0	–	0	80	15	2	0	1	7.2
4.10C	0	50	150	140	–	10	0.02	0.5	7.4
4.10D	0	40	0	0	5	–	0	0	7.4

states (Figure 4.9B). We compare results from simulations of these models to the data in Figure 4.10, and confirm that both models provide good fits despite being simpler than many existing models of this system (Terkildsen et al., 2007; Pan et al., 2019). Given the similarity of these results, both the random and ordered strategies appear to be viable for performing model selection, depending on the mechanistic information that is available.

4.4 Discussion

In this study we outlined an approach to developing models of enzymes for whole-cell modelling. The two key features of our approach were the automatic generation of rate laws and the use of an energetic framework. We developed a methodology for automatically deriving simplified rate laws for a wide range of enzyme cycles with arbitrary complexity, allowing modellers to test the fits of multiple models against experimental measurements while avoiding the time-consuming process of manually deriving rate laws. In contrast to kinetic parameters, energetic parameters provide a natural basis for incorporating thermodynamic constraints and additional data that may further refine the parameters (see Chapter 2 and Gawthrop et al. (2015); Mason and Covert (2019)). Therefore, our method facilitates both the development of enzyme models and their integration into whole-cell models. We encoded information within the energetic formulation using bond graphs, a framework that explicitly models energy and therefore preserves links to fundamental thermodynamic and mechanistic information when models are simplified. The approach was applied to the Na^+/K^+ ATPase, generating two simple, thermodynamically consistent models that fitted the training data well while making minimal assumptions about the mechanism of the enzyme. Our approach could form part of a workflow to automatically build and update components within a whole-cell model.

While there are numerous methods for simplifying biochemical systems (Okino and Mavrovouniotis, 1998; Snowden et al., 2017), the link between the underlying biochemical mechanism and mathematical equations is often non-existent or lost in the process of simplification (Bassingthwaighe et al., 2009; Ermentrout, 2001) and few approaches account for thermodynamic constraints (Smith and Crampin, 2004; Liebermeister and Klipp, 2006). We therefore based our method on existing approaches that use the King-Altman-Hill method coupled with detailed balance (Smith and Crampin, 2004; Segel, 1975; Hill, 1989; Beard and Qian, 2008). We improved on these approaches in two

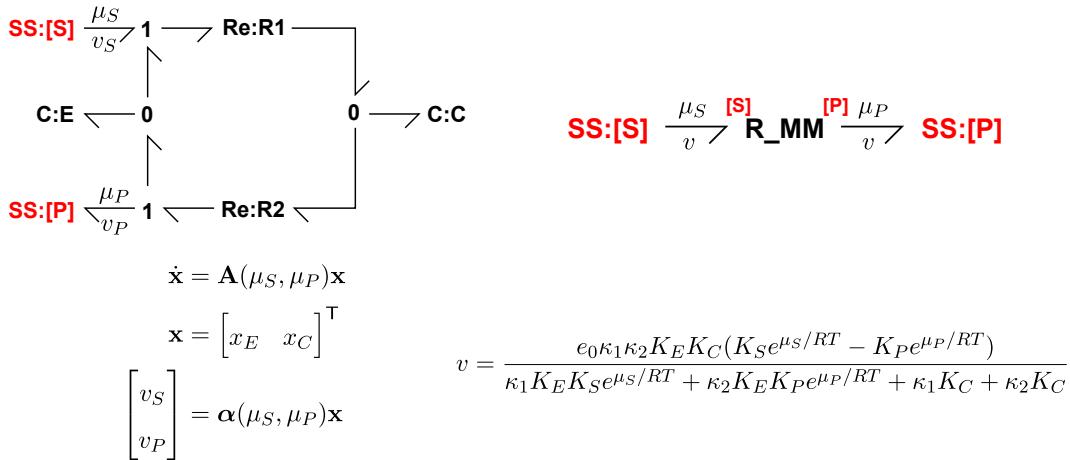
ways: by using energetic parameters to automatically satisfy thermodynamic constraints, and by deriving an explicit form for the reduced model so that the method is easily programmable. Since bond graphs are energetic in nature, the efficiency of enzymes can be studied regardless of whether the full or reduced models are used (Chapter 2; Gawthrop and Crampin 2018). While the analysis of enzyme kinetics reduces to linear dynamics under our assumptions, our approach to model reduction can be easily generalised to other biochemical systems with nonlinear dynamics.

Our model selection approach inferred that three and four-state models were sufficient to explain steady-state measurements collected from the Na^+/K^+ ATPase, depending on whether the sampled models were restricted to the consensus Post-Albers mechanism. Both selected models were simpler than many of the other models that were used to explain the same data (Terkildsen et al., 2007; Pan et al., 2019). However, given the proof-of-concept nature of our approach, we caution against replacing existing models of the Na^+/K^+ ATPase with these simplified models as more potential mechanisms need to be sampled and further validation is required before either of the selected models can be seen as suitable replacements for existing models. Despite this, these results demonstrate the utility of our approach in model development as it is flexible enough to be adapted depending on the mechanistic information that is available – the random strategy could be used in cases where there is no prior information on the mechanism, whereas the ordered sampling strategy may be more appropriate when mechanistic information is available. Furthermore, the ability to compare models in a systematic manner may enable the refinement of existing models in the future.

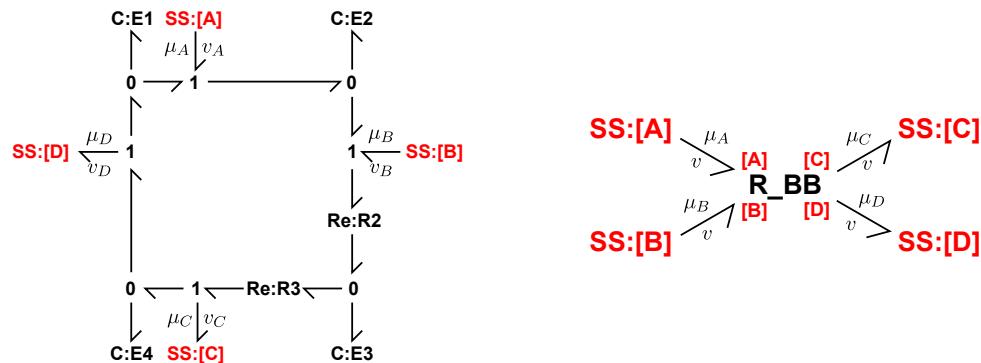
We believe that bond graphs are ideal for whole-cell modelling because their modularity allows full models of enzymes to be interchanged with reduced representations. Figure 4.11 shows examples of how a dynamic model can be replaced by a reduced model described by a rate law. For a reversible Michaelis-Menten enzyme (Figure 4.11A) and the four-state model in § 4.3.2 (Figure 4.11B), the full models are shown on the left, and the reduced models are shown on the right. Whereas the full models represent the dynamics of the system, the reduced models contain no dynamics. The reduced components are dissipative like typical **Re** components, but they have different nonlinear constitutive equations. Because the parameters of the constitutive equation are independent of the metabolite parameters and are only related to the metabolites via their chemical potential, the simplified model can be expressed in a modular fashion. Bond graphs are modular in nature, and external connections are represented as red **SS** components (see Figure 4.11). These **SS** components can be replaced by any metabolite, therefore bond graphs naturally enable operations that make biochemical modelling scalable, such as switching between full and reduced representations of an enzyme and reusing the same mechanisms for different enzyme-catalysed reactions. Liebermeister et al. (2010) argue for developing a relatively small library of thermodynamically consistent rate laws that can approximately represent the kinetics of most enzymes. Simplified rate laws such as those shown in Figure 4.11 could be used to form such a library containing relatively simple mechanisms such as the Michaelis-Menten, ping-pong and bi-bi mechanisms. These can be used as first choices to fit new data from enzymes, with further refinements possible when experimental data become comprehensive enough to require more complex mechanisms. Thus bond graphs appear to be an ideal framework for scaling up from models of enzymes to whole-cell models.

A

Michaelis-Menten model

**B**

Ordered sequential bi-bi mechanism



$$\dot{x} = \mathbf{A}(\mu_A, \mu_B, \mu_C, \mu_D)x$$

$$x = \begin{bmatrix} x_{E1} & x_{E2} & x_{E3} & x_{E4} \end{bmatrix}^\top$$

$$\begin{bmatrix} v_A \\ v_B \\ v_C \\ v_D \end{bmatrix} = \boldsymbol{\alpha}(\mu_A, \mu_B, \mu_C, \mu_D)x$$

$$v = \frac{e_0 K_{E1} K_{E2} K_{E3} K_{E4} e_0 \kappa_{R2} \kappa_{R3} (e^{(\mu_A+\mu_B)/RT} - e^{(\mu_C+\mu_D)/RT})}{\Sigma(\mu_A, \mu_B, \mu_C, \mu_D)}$$

Figure 4.11: Reduction of bond graph models of enzymes. (A) The Michaelis-Menten model described in Figure 4.1; (B) the four-state enzyme model described in Figure 4.4. Full models are shown on the left panels, and reduced models are shown on the right panels. Equations for each model are displayed under the corresponding bond graph. The matrices \mathbf{A} and $\boldsymbol{\alpha}$ can be derived from Eq. 4.50 and Eq. 4.60 in Appendix 4.A respectively. $\Sigma(\mu_A, \mu_B, \mu_C, \mu_D)$ can be derived by substituting chemical potentials into Eq. 4.26. The red **SS** components represent general external connections that can be connected to the metabolites of choice for different enzymes with the same mechanism.

Bond graphs are modular in the computational sense, in that two well-behaving models can be sensibly coupled together. However, there is another notion of modularity, termed behavioural modularity, that describes systems that retain essential behavioural properties when coupled together (Gawthrop and Crampin, 2016). Biological systems generally do not exhibit behavioural modularity because they are affected by retroactivity, where modules communicate in a bi-directional manner when coupled together. Retroactivity is seen as an obstacle to designing cellular systems (Del Vecchio, 2013), however it has been postulated that retroactivity arises from the need to minimise energy expenditure, therefore understanding how modularity relates to retroactivity is essential for cellular design (Gawthrop and Crampin, 2016).

We found that simplified versions of a hypothetical four-state model were sufficient in providing good fits to synthetic data generated from the model (Figures 4.6–4.7). Given that the synthetic data was both noiseless and comprehensive, these results would appear to suggest that simple models are sufficient for describing steady-state experimental data. This idea is supported by the relative simplicity of the selected mechanisms for the Na^+/K^+ ATPase, which contained three and four states (Figure 4.9). Thus, our results suggest that in the absence of transient measurements or further mechanistic information, simplified enzyme models with at most four states strike the right balance between flexibility and simplicity.

Our approach facilitates the use of model selection, which could form the basis of a method for identifying the correct mechanism of an enzyme in the future. We have described a model selection procedure that can rank models based on their consistency with the data, allowing the space of plausible models to be limited. For example, the models of Na^+/K^+ ATPase within the upper band seen in Figure 4.8A (and possibly many more) could be rejected by this approach. Our results suggest that the steady-state measurements that are common in enzyme assays are generally insufficient to perfectly resolve the mechanism of an enzyme, even in the presence of noiseless and comprehensive data. Therefore, additional information from transient measurements, time series measurements and structural data would likely be required to confidently infer the correct mechanism (Schnell et al., 2006; Mescam et al., 2011; Reytor González et al., 2017). Because our simplified models are explicitly linked to a full biophysical mechanism with a system of differential equations, it is relatively easy to extend our approach to model identification to account for time-dependent behaviour. Furthermore, by comparing the predictions of mechanisms against each other, our analysis has the potential to inform subsequent experiments to maximise the mechanistic information that could be gained. Here it may be useful to use existing mathematical analyses for mechanism distinguishability (Schnell et al., 2006) as well as Bayesian computation to quantify mechanistic uncertainty (Toni et al., 2009). We anticipate that a method for mechanistic inference could be used to inform rational drug designs that inhibit specific conformations of enzymes (Atkins and De Paula, 2006).

We believe that our strategy could be improved in two ways. Firstly, while we used optimisation for parameter estimation due to its computational efficiency, it does not naturally account for measurement noise or parameter uncertainty. In cases where the confidence in a model needs to be quantified, Bayesian estimation may be preferable over optimisation, although it comes at the expense of increased computational complexity (Babtie and Stumpf, 2017). We note that the rate laws generated using our approach could

be easily integrated into an approximate Bayesian computation framework (Toni et al., 2009). Because Bayesian estimation outputs probability distributions over the parameter and model space, it provides a framework for updating the uncertainty associated with models and parameters as more experimental data are collected (Liepe et al., 2013). Such a framework may be especially useful for favouring specific mechanisms consistent with knowledge in the literature while still allowing other mechanisms to be inferred in the presence of conflicting experimental evidence.

Secondly, we assumed that the concentrations of enzymes were small so that they would be in quasi-steady-state. However, recent studies have cast doubt over the validity of this assumption (Schnell and Maini, 2003). Surovtsova et al. (2012) used computational singular perturbation to assess the validity of rapid equilibrium and quasi-steady-state approximations. While their approach did not incorporate thermodynamic information, integrating such a method into a physical and modular framework may allow simulation software to automatically switch between full and reduced models as needed.

4.5 Conclusion

We have developed a generalised approach for model reduction based on bond graphs, an energetic framework. Our approach is applicable to a wide range of enzyme cycles, allowing numerous models to be simultaneously tested against experimental measurements. Because our analysis is based on a physical and energy-based framework, the reduced models are thermodynamically consistent and retain links to the underlying biochemical mechanism. As a result, it is relatively straightforward to incorporate thermodynamic information into the development of enzyme models and to reconcile measurements from different experiments. Therefore, our approach contributes to the automation of model development for enzymes while providing an explicit physical framework for coupling models together.

Software

Bond graph models were produced using the BondGraphTools (<https://github.com/BondGraphTools/BondGraphTools>) package. Symbolic manipulations were performed in SymPy (Meurer et al., 2017). Optimisation was performed using the BlackBoxOptim (<https://github.com/robertfeldt/BlackBoxOptim.jl>) and Optim (Mogensen and Riseth, 2018) packages in Julia.

4.6 Future work

In this section we outline plans to improve our method for model development. We demonstrated our approach on the Na^+/K^+ ATPase, however, more physiological examples are required to demonstrate that the approach can be generalised. Therefore,

we plan to apply this method to experimental data from cardiac transporters such as the $\text{Na}^+ \text{-Ca}^{2+}$ exchanger (Blaustein and Lederer, 1999) and SERCA (Periasamy and Kalyanasundaram, 2007) as well as commonly-modelled enzymes within the glycolysis pathway (Lambeth and Kushmerick, 2002). It would be interesting to observe the extent to which mechanisms inferred from our method agree with literature, and whether integrating data from enzyme assays provides new mechanistic insights.

As mentioned in the discussion (Section 4.4), another extension for this work is to integrate it into a Bayesian framework (specifically approximate Bayesian computation), which estimates probability distributions for the space of possible models and parameters (Liepe et al., 2014). An advantage of Bayesian inference is that an ensemble of models can be used to make predictions, allowing uncertainties in predictions to be quantified (Babtie and Stumpf, 2017) while also providing avenues for directing future experiments to optimally extract mechanistic and parametric information. In Bayesian inference, existing mechanistic information can be described using prior distributions. In contrast to our approach where mechanisms inconsistent with what we believe are excluded from analysis, Bayesian inference allows one to favour models consistent with the existing body of knowledge. Therefore, mechanistic information can be incorporated while also allowing for the possibility of rejecting consensus mechanisms in the presence of conflicting observations. Furthermore, the posteriors obtained from Bayesian inference can be used as priors for subsequent model development, providing a natural method of updating models when new data become available.

In order to achieve effective model selection, we plan to develop more efficient strategies for sampling the space of plausible models. We found that it was infeasible to test every possible mechanism for complex catalysed reactions. However, if only a small population of randomly sampled models are tested, it is likely that many plausible models would be missed. Therefore it would be desirable to develop methods for restricting the search space based on the performance of models that have already been tested. There has been some development into search algorithms in large model spaces for Bayesian model selection (Berger and Molina, 2005).

4.A Model reduction

4.A.1 Summary

The model reduction approach involves finding an explicit mathematical expression for the steady-state cycling rate of an enzyme cycle after it has been reduced using rapid equilibrium. An outline of the procedure is as follows:

- Step 1: Separate the system into fast and slow subsystems (Appendix 4.A.3)
- Step 2: Analyse the system on the fast timescale to find the rapid equilibrium constraints (Appendix 4.A.4)
- Step 3: Find an expression for the rates of change for the slow variables on the slow timescale (Eq. 4.44)

- Step 4: Using the mapping between enzyme states and slow variables from Step 3, derive a linear differential equation for the enzyme states (Eq. 4.50)
- Step 5: Derive the steady-state and conservation constraints from the differential equation on the slow subsystem (Appendix 4.A.5)
- Step 6: Derive expressions for the enzyme states at the steady state of the slow subsystem (Eq. 4.53)
- Step 7: Calculate the steady-state reaction rates using the steady-state enzyme states (Eq. 4.58)
- Step 8: Find the cycling rate of the enzyme (Eq. 4.61)

4.A.2 Notation

The stoichiometry of a network of biochemical reactions is denoted by the forward and reverse matrices N^f and N^r , which have rows corresponding to species and columns corresponding to reactions. The stoichiometry of the i th species on the forward side of the j th reaction is recorded in the (i, j) entry in N^f , and a similar convention follows for N^r . The stoichiometric matrix is given by $N = N^r - N^f$. In the upcoming analysis, it is useful to break these stoichiometric matrices into partitions. We denote such partitions as $N_{\alpha, \beta}$ where $\alpha \in \{F, S\}$ and $\beta \in \{\text{cd}, \text{cs}\}$. The symbol α indicates that the matrix is limited to reactions operating on the fast (F) or slow (S) timescale, and β indicates that the matrix is limited to species corresponding to metabolites (chemostatic, or cs) or enzyme states (chemodynamic, or cd). If a symbol is not present, all reactions or species are included. A similar convention applies to N^f and N^r .

We define κ and K as vectors containing the reaction and species energetic parameters respectively. The bold symbols $\boldsymbol{\kappa}$ and \mathbf{K} denote diagonal matrices containing these parameters. Using a similar notation to stoichiometric matrices, κ_α , K_β , $\boldsymbol{\kappa}_\alpha$, \mathbf{K}^β and X^β represent vectors and matrices restricted to a subset of species and reactions.

$I_{n \times n}$ is the $n \times n$ identity matrix, and $0_{m \times n}$ is the $m \times n$ zero matrix. The functions **Exp** and **Ln** are the element-wise exponential and logarithm respectively, and \circ is the element-wise multiplication operator.

4.A.3 Timescale separation

In models of enzyme kinetics, we often assume that some reactions operate on timescales far quicker than the rest. We can depict a vector bond graph of the full biochemical network in Figure 4.12. The components are represented in italics, denoting the fact that they contain multiple potentials and flows. In the context of a whole-cell model, we see the reactions at rapid equilibrium as occurring on a “fast” timescale, the evolution of enzyme kinetics occurring on a “slow” timescale and the evolution of metabolites occurring on an “ultraslow” timescale (Gunawardena, 2014).

Since the full model contains both dynamic species and chemostats, they have been placed in separate vector components. The transformers described by matrices T^{cd} and T^{cs} are used to merge the chemical potentials into a single vector. To merge the n_{cd}

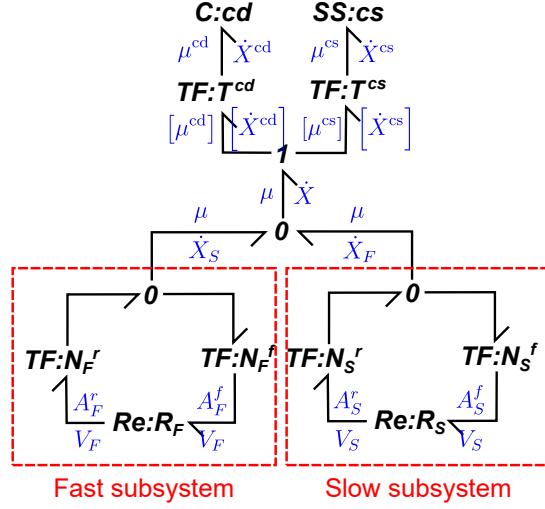


Figure 4.12: Generalised model of full enzyme kinetics.

dynamic species and n_{cs} chemostats into a vector of length $n_s = n_{\text{cd}} + n_{\text{cs}}$, we define T^{cd} , a $n_s \times n_{\text{cd}}$ logical matrix with 1's where the species match. The $n_s \times n_{\text{cs}}$ matrix T^{cs} is defined similarly. As a result, the effort coming out of the 1 junction in Figure 4.12 can be written as

$$\mu = T^{\text{cd}}\mu^{\text{cd}} + T^{\text{cs}}\mu^{\text{cs}} \quad (4.32)$$

As shown in Gawthrop and Crampin (2016), the equations of the full model can be described by

$$\dot{X}^{\text{cd}} = N_{\text{cd}}V \quad (4.33)$$

where X^{cd} is a vector containing the amounts of the enzyme states and V is a vector of the reaction rates. We then split V into V_S and V_F , which are the components corresponding to the fast and slow reactions respectively. Then

$$\dot{X}^{\text{cd}} = N_{F,\text{cd}}V_F + N_{S,\text{cd}}V_S \quad (4.34)$$

$$\begin{aligned} &= N_{F,\text{cd}}\kappa_F \left(\text{Exp} \left[N_F^{f,T} \text{Ln}(\mathbf{K}X) \right] - \text{Exp} \left[N_F^{r,T} \text{Ln}(\mathbf{K}X) \right] \right) \\ &\quad + N_{S,\text{cd}}\kappa_S \left(\text{Exp} \left[N_S^{f,T} \text{Ln}(\mathbf{K}X) \right] - \text{Exp} \left[N_S^{r,T} \text{Ln}(\mathbf{K}X) \right] \right) \end{aligned} \quad (4.35)$$

We also split the chemical potentials into those for the chemodynamic species, and the chemostats, so that

$$\begin{aligned} \dot{X}^{\text{cd}} &= N_F^{\text{cd}}\kappa_F \left(B_F^f \text{Exp} \left[N_F^{f,\text{cd},T} \text{Ln}(\mathbf{K}^{\text{cd}}X^{\text{cd}}) \right] - B_F^r \text{Exp} \left[N_F^{r,\text{cd},T} \text{Ln}(\mathbf{K}^{\text{cd}}X^{\text{cd}}) \right] \right) \\ &\quad + N_S^{\text{cd}}\kappa_S \left(B_S^f \text{Exp} \left[N_S^{f,\text{cd},T} \text{Ln}(\mathbf{K}^{\text{cd}}X^{\text{cd}}) \right] - B_S^r \text{Exp} \left[N_S^{r,\text{cd},T} \text{Ln}(\mathbf{K}^{\text{cd}}X^{\text{cd}}) \right] \right) \end{aligned} \quad (4.36)$$

where \mathbf{K}^{cd} is a diagonal matrix with entries corresponding to the chemodynamic species. For $a \in \{S, F\}$ and $d \in \{f, r\}$, we define

$$B_a^d = \text{diag} \left(\text{Exp} \left[N_a^{d,\text{cs},T} \text{Ln}(\mathbf{K}^{\text{cs}}X^{\text{cs}}) \right] \right) \quad (4.37)$$

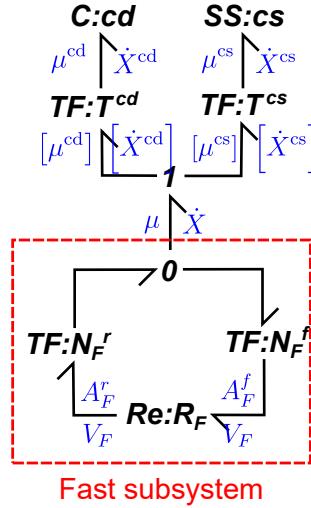


Figure 4.13: Fast timescale of enzyme kinetic models.

In models of enzymes, all reactions are transitions between enzyme states, therefore N_{cd}^f and N_{cd}^r contain a single 1 in each column, and zeros elsewhere. As a result, the model is a linear system with the equations

$$\dot{X}^{cd} = \left[N_F^{cd} \kappa_F \left(B_F^f N_F^{f,cdT} - B_F^r N_F^{r,cdT} \right) + N_S^{cd} \kappa_S \left(B_S^f N_S^{f,cdT} - B_S^r N_S^{r,cdT} \right) \right] \mathbf{K}^{cd} X^{cd} \quad (4.38)$$

In the next sections, we separate the fast and slow timescales, and show how this separation of timescales can be used to reduce the system of equations.

4.A.4 Fast timescale

The fast timescale contains the reactions that proceed at faster kinetics. We write $\kappa_F = (1/\varepsilon)\kappa_F^*$, and observe the behaviour of the system as $\varepsilon \rightarrow 0$. On the fast timescale $T = t/\varepsilon$, Eq. 4.38 becomes

$$\frac{dX^{cd}}{dT} = N_F^{cd} \left[\kappa_F^* \left(B_F^f N_F^{f,cdT} - B_F^r N_F^{r,cdT} \right) + \varepsilon \kappa_S \left(B_S^f N_S^{f,cdT} - B_S^r N_S^{r,cdT} \right) \right] \mathbf{K}^{cd} X^{cd} \quad (4.39)$$

Then in the limit as $\varepsilon \rightarrow 0$, the rates of the slow reactions vanish, thus

$$\frac{dX^{cd}}{dT} = N_F^{cd} V_F^{cd} \quad (4.40)$$

$$= N_F^{cd} \kappa_F^* \left(B_F^f N_F^{f,cdT} - B_F^r N_F^{r,cdT} \right) \mathbf{K}^{cd} X^{cd} \quad (4.41)$$

This equation corresponds to the reduced bond graph in Figure 4.13, where the slow subsystem has been removed.

In timescale separation, the steady state of the fast timescale is coupled to the slow dynamics. While steady states in biochemical networks are not necessarily in equilibrium,

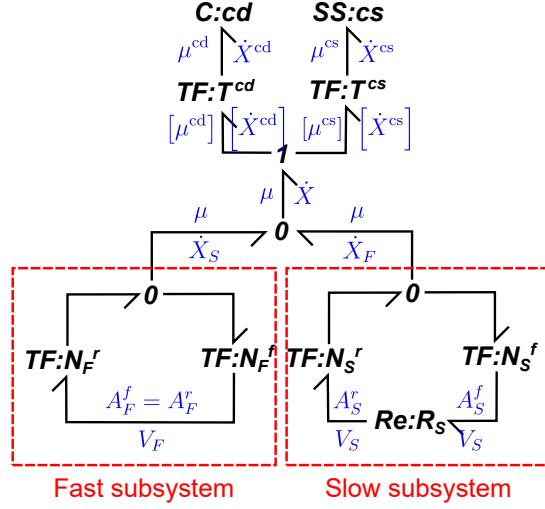


Figure 4.14: Slow timescale of enzyme kinetic models.

we make the additional assumption that the steady state is in equilibrium on the fast timescale (otherwise fluxes would be drawn from the chemostats at arbitrarily high rates). This leads to the *rapid equilibrium* constraint

$$\left(B_F^f N_F^{f,cdT} - B_F^r N_F^{r,cdT} \right) \mathbf{K}^{cd} X^{cd} = 0 \quad (4.42)$$

which will be used on the slow timescale.

4.A.5 Slow timescale

If we take $\varepsilon \rightarrow 0$, then on the slow timescale, the fast reactions will have run to equilibrium (i.e. Eq. 4.42 holds). A bond graph representation of this reduced system is shown in Figure 4.14. To model rapid equilibrium, the **Re** components corresponding to the fast reactions have been removed, so that the forward and reverse affinities are equal (Gawthrop and Crampin, 2014).

Because many of the states in X^{cd} are fast variables in the limit $\varepsilon \rightarrow 0$, we formulate the differential equations of the slow timescale in terms of the slow variables Z . As seen in Schauer and Heinrich (1983), the slow variables can be defined as $Z = G_F X^{cd}$, where G_F is the left nullspace matrix of N_F^{cd} . Then the rates of change of the slow variables are

$$\dot{Z} = G_F \left[N_F^{cd} \kappa_F \left(B_F^f N_F^{f,cdT} - B_F^r N_F^{r,cdT} \right) + N_S^{cd} \kappa_S \left(B_S^f N_S^{f,cdT} - B_S^r N_S^{r,cdT} \right) \right] \mathbf{K}^{cd} X^{cd} \quad (4.43)$$

$$= G_F N_S^{cd} \kappa_S \left(B_S^f N_S^{f,cdT} - B_S^r N_S^{r,cdT} \right) \mathbf{K}^{cd} X^{cd} \quad (4.44)$$

because $G_F N_F^{cd} = 0$. Because the equality holds regardless of the value of ε , the equation also holds in the limit as $\varepsilon \rightarrow 0$.

In order to express Eq. 4.44 as a proper ODE, we convert between X^{cd} and Z . On the slow timescale, this is possible because the rapid equilibrium constraint (Eq. 4.42)

and $Z = G_F X^{\text{cd}}$ imply

$$\begin{bmatrix} 0_{n_F,1} \\ Z \end{bmatrix} = CX^{\text{cd}} \quad (4.45)$$

where

$$C = \begin{bmatrix} \left(B_F^f N_F^{f,\text{cd}T} - B_F^r N_F^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} \\ G_F \end{bmatrix} \quad (4.46)$$

If the fast subsystem contains no reaction cycles, C is square and invertible because its rows and columns can be rearranged to form a block diagonal matrix with invertible subblocks. Then

$$X^{\text{cd}} = C^{-1} \begin{bmatrix} 0_{n_F,1} \\ Z \end{bmatrix} \quad (4.47)$$

$$= E_Z Z \quad (4.48)$$

where n_F is the number of fast reactions and E_Z is a matrix of the final $n_{\text{cd}} - n_F$ rows of C^{-1} .

Using Eq. 4.48, we can write a differential equation for the slow variables:

$$\dot{Z} = G_F N_S^{\text{cd}} \kappa_S \left(B_S^f N_S^{f,\text{cd}T} - B_S^r N_S^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} E_Z Z \quad (4.49)$$

We can also write an ODE for all of the chemodynamic species:

$$\dot{X}^{\text{cd}} = E_Z G_F N_S^{\text{cd}} \kappa_S \left(B_S^f N_S^{f,\text{cd}T} - B_S^r N_S^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} X^{\text{cd}} \quad (4.50)$$

At steady state, the system follows the *quasi-steady-state* constraint

$$\dot{Z} = G_F N_S^{\text{cd}} \kappa_S \left(B_S^f N_S^{f,\text{cd}T} - B_S^r N_S^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} X^{\text{cd}} = 0 \quad (4.51)$$

The final constraint dictating the steady state relates to conserved moieties within the enzyme cycle. These can be summarised using G , the left nullspace matrix of N^{cd} so that $G \dot{X}^{\text{cd}} = G N^{\text{cd}} X^{\text{cd}} = 0$. Then the steady state must satisfy

$$G X^{\text{cd}} = G X^{\text{cd}}(0) \quad (4.52)$$

Eq. 4.52 can be seen as *conservation relations* determined by the initial conditions of the system. If there is only a single enzyme, G is a row vector of ones.

4.A.6 Reduced equations

We use the following three constraints to derive expressions for X^{cd} at steady state:

1. The rapid equilibrium constraint in Eq. 4.42

2. The quasi-steady-state constraint in Eq. 4.51
3. The conservation constraints in Eq. 4.52

Since these constraints are all linear functions of X^{cd} , we can combine them into a matrix equation

$$MX^{\text{cd}} = b \quad (4.53)$$

where

$$M = \begin{bmatrix} \left(B_F^f N_F^{f,\text{cd}T} - B_F^r N_F^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} \\ G_F N_S^{\text{cd}} \boldsymbol{\kappa}_S \left(B_S^f N_S^{f,\text{cd}T} - B_S^r N_S^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} \\ G \end{bmatrix} \quad (4.54)$$

and

$$b = \begin{bmatrix} 0_{n_F \times 1} \\ 0_{n \times 1} \\ GX^{\text{cd}}(0) \end{bmatrix} \quad (4.55)$$

We have defined n_F as the number of fast reactions, and n as the number of slow variables. Eq. 4.53 can be solved to find an expression for X^{cd} .

Using the expression for X^{cd} , the expressions for the slow reactions V_S can be found using the equation

$$V_S = \boldsymbol{\kappa}_S \left(B_S^f N_S^{f,\text{cd}T} - B_S^r N_S^{r,\text{cd}T} \right) \mathbf{K}^{\text{cd}} X^{\text{cd}} \quad (4.56)$$

At steady state, $N^{\text{cd}}V = 0$. By splitting this equation into fast and slow components,

$$N_S^{\text{cd}} V_S + N_F^{\text{cd}} V_F = 0 \quad (4.57)$$

therefore the rates of the fast reactions can be found by solving the linear equation

$$N_F^{\text{cd}} V_F = -N_S^{\text{cd}} V_S \quad (4.58)$$

Thus the rates of all reactions at steady state can be found by joining V_S and V_F :

$$V = T_S V_S + T_F V_F \quad (4.59)$$

where T_S and T_F are matrices that ensure the fast and slow reactions are mapped to the correct entries. The molar flow rates supplied by the chemostats are given by

$$V_{\text{cs}} = N^{\text{cs}} V \quad (4.60)$$

V_{cs} can be used to determine the fluxes to and from each metabolite when coupling into a whole-cell model.

From Gawthrop and Crampin (2017), the steady-state cycling rates can be reduced by pathway analysis. By looking at the right nullspace matrix C_p of N^{cd} , we find that the steady-state reaction rates can be reduced using the equation

$$V = C_p V_p \quad (4.61)$$

where V_p is a reduced vector with length equal to the number of cycles, and contains the information required to reconstruct V . While enzymatic mechanisms in general may have many cycles (Hill, 1989; Gawthrop and Crampin, 2017), the examples we deal with in this study only have a single cycle, so V_p contains only a single entry that we refer to as the steady-state flux within the main text. V_p can be normalised by e_0 to obtain the cycling rate.

4.B Fitting energetic parameters to kinetic data

Energetic parameters are related to kinetic parameters by the equation

$$\mathbf{Ln}(\mathbf{k}) = \mathbf{M} \mathbf{Ln}(\boldsymbol{\lambda}) \quad (4.62)$$

with

$$\mathbf{k} = \begin{bmatrix} k^+ \\ k^- \end{bmatrix}, \quad \mathbf{M} = \left[\begin{array}{c|c} I_{n_r \times n_r} & N^{fT} \\ \hline I_{n_r \times n_r} & N^{rT} \end{array} \right], \quad \boldsymbol{\lambda} = \begin{bmatrix} \kappa \\ K \end{bmatrix} \quad (4.63)$$

where k^+ is the vector of forward kinetic rate constants and k^- is the vector of reverse rate constants (see Chapter 2 and Gawthrop et al. 2015). The mapping between kinetic and energetic parameters is not one-to-one. As a consequence of this, more than one set of energetic parameters map to the same set of kinetic parameters (Gawthrop et al., 2015). Because steady-state measurements depend only on the kinetics of the system and not the thermodynamic quantities specifically, there will be families of energetic parameters with the same kinetic properties, and therefore explain steady-state data equally well. Accordingly, we designed our parameter estimation method to restrict the space of sampled parameters to those with unique kinetic behaviour and output results that indicate the extent to which energetic parameters are determined.

If there are energetic parameters that are known prior to parameter estimation, Eq. 4.62 can be expressed as

$$\mathbf{Ln}(\mathbf{k}) = \mathbf{M}_k \mathbf{Ln}(\boldsymbol{\lambda}_k) + \mathbf{M}_u \mathbf{Ln}(\boldsymbol{\lambda}_u) \quad (4.64)$$

where \mathbf{M}_k and \mathbf{M}_u are the columns of \mathbf{M} corresponding to the known and unknown parameters respectively. Similarly, $\boldsymbol{\lambda}_k$ and $\boldsymbol{\lambda}_u$ are the rows of $\boldsymbol{\lambda}$ corresponding to the known and unknown parameters respectively. The unknown parameters can then be found by solving the equation

$$l = \mathbf{M}_u \mathbf{Ln}(\boldsymbol{\lambda}_u) \quad (4.65)$$

where $l = \mathbf{Ln}(\mathbf{k}) - \mathbf{M}_k \mathbf{Ln}(\boldsymbol{\lambda}_k)$. In many cases, even after previously known energetic parameters are set, the mapping between kinetic and remaining energetic parameters is

still non-unique. To overcome this issue, we choose energetic parameters to set prior to parameter estimation. Energetic parameters associated with linearly dependent rows of \mathbf{M}_u are arbitrarily set until a one-to-one mapping results. The resulting set of energetic parameters can then be generated from this individual point.

For any two vectors $\boldsymbol{\lambda}'_u$ and $\boldsymbol{\lambda}''_u$ that satisfy Eq. 4.65,

$$l = \mathbf{M}_u \mathbf{Ln}(\boldsymbol{\lambda}'_u) = \mathbf{M}_u \mathbf{Ln}(\boldsymbol{\lambda}''_u) \quad (4.66)$$

and therefore

$$\mathbf{M}_u [\mathbf{Ln}(\boldsymbol{\lambda}''_u) - \mathbf{Ln}(\boldsymbol{\lambda}'_u)] = 0 \quad (4.67)$$

Once a set of unknown parameters $\boldsymbol{\lambda}'_u$ arises from fitting, all other possible parameters follow the form

$$\boldsymbol{\lambda}''_u = \mathbf{Exp}(\mathbf{r}) \circ \boldsymbol{\lambda}'_u \quad (4.68)$$

where $\mathbf{r} \in \ker(\mathbf{M}_u)$. If R is a right nullspace matrix of \mathbf{M}_u , then the family of possible energetic parameters is described by the set

$$\mathcal{S} = \left\{ \boldsymbol{\lambda}''_u = \mathbf{Exp}(\mathbf{t}R) \circ \boldsymbol{\lambda}'_u : \mathbf{t} \in \mathbb{R}^{1 \times n_{\text{free}}} \right\} \quad (4.69)$$

In the above equation, $\mathbf{t} = [t_1 \ t_2 \ \dots \ t_{n_{\text{free}}}]$ is a row vector containing the free parameters, with a length equal to the dimension of $\ker(\mathbf{M}_u)$.

4.C Fitting experimental data for Na^+/K^+ ATPase

4.C.1 Incorporation of membrane potential

Unlike many enzymes, the Na^+/K^+ ATPase transports ions across a charged membrane. Therefore, the cycling rate must be dependent on membrane potential in order to be thermodynamically consistent. Bond graphs are highly desirable in this context because they can represent quantities from different domains under a single framework. Under the bond graph representation, chemical potential (μ) and membrane potential (V_m) are analogous quantities. We make use of this analogy to incorporate voltage dependence. While $\mathbf{K}^{\text{cs}} X^{\text{cs}}$ in Eq. 4.37 contains entries of the form $K_s x_s$, these can also be expressed in terms of the chemical potential:

$$\gamma = K_s x_s = e^{\mu_s / RT} \quad (4.70)$$

We call γ the thermokinetic potential (Ederer and Gilles, 2007). To incorporate voltage dependence, the chemical potential μ was substituted for the analogous potential FV_m , where $F = 96485$ C/mol:

$$\gamma = e^{FV_m / RT} \quad (4.71)$$

Table 4.3: The binding steps used to generate candidate models for the Na^+/K^+ ATPase. Due to the cyclic nature of enzyme cycles, if n is equal to the number of states, P_1 appears in the products rather than P_{n+1} . The species p represents the translocation of a single unit of charge from the intracellular to the extracellular side of the membrane, and Δ is a partitioning constant.

Step #	Reaction
1	$P_n \rightleftharpoons P_{n+1} + K_i^+$
2	$P_n \rightleftharpoons P_{n+1} + K_i^+$
3	$P_n + \text{Na}_i^+ \rightleftharpoons P_{n+1}$
4	$P_n + \text{Na}_i^+ \rightleftharpoons P_{n+1}$
5	$P_n + \text{Na}_i^+ - \Delta p \rightleftharpoons P_{n+1}$
6	$P_n \rightleftharpoons P_{n+1} + \text{MgADP}$
7	$P_n \rightleftharpoons P_{n+1} + \text{Na}_e^+ - (1 + \Delta)p$
8	$P_n \rightleftharpoons P_{n+1} + \text{Na}_e^+$
9	$P_n \rightleftharpoons P_{n+1} + \text{Na}_e^+$
10	$P_n + K_e^+ \rightleftharpoons P_{n+1}$
11	$P_n + K_e^+ \rightleftharpoons P_{n+1}$
12	$P_n \rightleftharpoons P_{n+1} + \text{Pi} + \text{H}^+$
13	$P_n + \text{MgATP} \rightleftharpoons P_{n+1}$

4.C.2 Model generation

Na^+/K^+ ATPase mechanisms were generated using the 13 binding and unbinding steps listed in Table 4.3. An unordered mechanism with n states and n reactions was generated by randomly assigning each step to a single reaction. Ordered steps were generated by starting with a cycle with binding steps in the order specified in Table 4.3, and grouping neighbouring steps together into the same reaction to form the required number of reactions. To avoid intracellular and extracellular species from binding and unbinding in the same reaction, we excluded models that grouped steps 13 and 1; or 6 and 7.

4.C.3 Parameter estimation

It is possible to directly incorporate known relationships between energetic parameters into models of enzymes. Since Na^+ and K^+ have a known equilibrium of equal concentration on the intracellular and extracellular sides of the membrane, we set the energetic constants of these species to $K_{\text{Nai}} = K_{\text{Nae}} = K_{\text{Ki}} = K_{\text{Ke}} = 1\text{mM}^{-1}$. To account for the known standard free energy of MgATP hydrolysis of $\Delta_{\text{MgATP}}^0 = 11.9\text{ kJ/mol}$ at 310K (Tran et al., 2009), we set $K_{\text{MgADP}} = K_{\text{MgADP}} = K_{\text{MgADP}} = 1\text{ mM}^{-1}$ and $K_{\text{MgATP}} = 9881\text{ mM}^{-1}$. Since these species are chemostats, their parameters have been expressed in terms of concentration so that they generalise to different compartmental volumes.

Because the concentrations of Pi in Table 4.2 refer to total Pi rather than free Pi, total Pi was converted to free Pi using the equation

$$[\text{Pi}_{\text{free}}] = \frac{[\text{Pi}_{\text{tot}}]}{1 + [\text{K}_i^+]/K_{d,\text{KPi}} + [\text{H}^+]/K_{d,\text{HPi}} + [\text{Na}_i^+]/K_{d,\text{NaPi}}} \quad (4.72)$$

where the dissociation constants are $K_{d,\text{KPi}} = 292$ mM, $K_{d,\text{HPi}} = 10^{-3.77}$ mM and $K_{d,\text{NaPi}} = 224$ mM (Smith and Crampin, 2004). While temperature is known to affect enzyme kinetics, this influence has only been characterised phenomenologically from experimental measurements (Barry, 1914). Due to the difficulty in predicting the precise effect on energetic parameters, we assumed a constant temperature of 310 K across all experimental measurements.

Chapter 5

Conclusions

5.1 Summary of research

The primary issue motivating this thesis relates to how existing models can be reused and integrated to construct large-scale models of biological systems. In Chapter 1, I listed challenges that often arise when separate models are coupled together, including the lack of a common ground for models to communicate, the inability to maintain physical and thermodynamic consistency, and the lack of transparency in updating model parameters. Therefore, it is highly desirable to represent models using the bond graph, an energy-based modelling framework that respects the laws of physics and thermodynamics. The aim of my thesis was therefore to develop methods based on the bond graph framework to facilitate model reuse and integration in systems biology.

In Chapter 2, I described how models of both electroneutral and electrogenic membrane transporters could be represented using bond graphs, and used hypothetical examples to demonstrate that bond graph models inherently satisfy the thermodynamic constraints employed in some (but not all) existing models (Section 2.2). I then formulated a method for converting existing models with kinetic parameters to bond graph models (Appendix 2.B). The main issue associated with this conversion was the transformation from kinetic parameters to bond graph parameters, and I showed that this transformation is possible if and only if the kinetic parameters are thermodynamically consistent. Using this approach I created bond graph models based on two existing ODE-based transporter models: cardiac SERCA and the cardiac Na^+/K^+ ATPase (Section 2.3). In this physiological context, bond graphs proved useful in incorporating known thermodynamic relationships between species and in detecting inconsistencies within existing models. As a result of representing these systems using bond graphs, I found physical and thermodynamic inconsistencies within the model of Na^+/K^+ ATPase that were resolved in an updated model.

In Chapter 3, I used the hierarchical nature of bond graphs to couple independently developed models of ion transporters, ion channels and buffering processes into a simplified model of cardiac electrophysiology. Because a bond graph formulation required thermodynamically consistent descriptions of ion channels that incorporate both gating

and ionic currents, I developed new models of ion channels based on existing models in the literature (§ 3.2.4). In some cases, thermodynamic inconsistencies within existing models needed to be fixed. The resulting model of cardiac electrophysiology generated many essential features of the action potential (§ 3.3.1). Furthermore, because bond graph models explicitly represent energy, the energy consumption of these processes is easily calculated, an analysis that is currently missing in the literature. I then showed that a bond graph formulation of cardiac electrophysiology allowed the issues of drift and non-unique steady states to be explored in a more general context than is currently available in the literature. Whereas the issue of drift was previously explored only through a conservation of charge equation that was derived manually for specific models, the notion of conserved moieties in bond graph modelling allowed this principle to be generalised to arbitrary models. Changes to the amounts of conserved moieties explained the occurrence of drift, providing a means of systematically detecting drift without running any simulations (§ 3.3.3). Finally, I showed that it was possible to predict whether different initial conditions would result in different steady-state behaviours by comparing the values of conserved quantities within the initial conditions (§ 3.3.4).

In Chapter 4, I derived explicit equations for simplified reaction fluxes in arbitrary enzyme cycles. Bond graphs provided an elegant framework for performing model reduction while maintaining thermodynamic consistency (§4.3.1,4.3.3). Because this approach to model reduction was general and programmable, it could be used within a model selection routine, allowing modellers to compare performance across a large population of candidate models rather than testing individual models separately. I first applied this methodology to synthetic data generated from a hypothetical enzyme with a bi-bi mechanism and showed that steady-state data could be used to significantly reduce the space of plausible models (§ 4.3.6). I then applied the model selection approach to data from the Na^+/K^+ ATPase (§ 4.3.7) to generate two simple, biochemically plausible and thermodynamically consistent models that fit experimental data with an accuracy that was comparable to far more complex models (see Chapter 2 and Terkildsen et al. (2007)).

5.2 Implications and further work

5.2.1 Incorporating thermodynamics into models of biological systems

In this thesis, I developed methods for converting existing models into bond graph models. These methods were applied to membrane transporters (Chapter 2) and ion channels (Chapter 3). A strength of the bond graph representation is that thermodynamic quantities within existing models become explicit, allowing these values to be checked against existing experimental measurements. I found that in many cases, the existing models were not thermodynamically consistent. In some cases, these inconsistencies could be resolved with minor changes to the original model. However, in other cases where existing models contained phenomenological components, correcting thermodynamic inconsistencies required deriving entirely new equations and either fitting the parameters to the original experimental data or synthetic data generated from the original model.

Resolving these issues more broadly will likely require further experimental and theoretical investigation to better understand the physics underlying these systems.

A direction for future work would be to extend the methods described within Chapters 2 and 3 to other physical domains that occur within biology. While this thesis focussed on the coupling between chemical and electrical components, biological systems may also be coupled across chemical and mechanical domains, for example during cross-bridge cycling (Huxley, 1957) and mechanosensation (Gillespie and Walker, 2001). Cross-bridge cycling is particularly challenging to represent within an energetic and biophysical framework due to the interaction between different actin-myosin cross-bridges, and the process is often represented as a partial differential equation (PDE) or a system of ODEs with phenomenological components (Huxley, 1957; Rice et al., 2008; Tran et al., 2010).

As we integrate models across greater spatial and temporal scales, there will be an increasing need to account for processes of a stochastic and spatially varying nature. While bond graphs do not naturally account for stochastic or spatially distributed systems, these systems are nonetheless composed of processes that obey the laws of thermodynamics. Therefore, it may be possible to use bond graphs to interface with models of these systems provided they are expressed in an alternative energetic framework. Expressing such processes under an energetic framework is the subject of further work; statistical thermodynamics may be suitable for stochastic processes (Keizer, 1987), while port-Hamiltonians may be suitable for spatially varying processes (van der Schaft and Jeltsema, 2014).

5.2.2 Iterative refinement of biological knowledge

In Chapter 4, I found a substantial degree of mechanistic uncertainty associated with models developed using steady-state measurements gathered from enzyme assays. Many existing models carry errors in both the equations and parameters, but these models often avoid further scrutiny once they are published and they are rarely updated in light of new experimental observations. This has been noted as a missed opportunity in advancing our knowledge of biological systems (Smith et al., 2007). I explored a methodology that could be used to compare alternative models under a thermodynamically consistent framework (Chapter 4). This methodology can help to reveal the structural uncertainty associated with a model, but further work is required to adapt this methodology for the continual refinement of models.

Key to the future goal of continual model revision is the development of a database of models and parameters that updates constantly. Extending on existing databases such as the Physiome Repository and the BioModels database that archive models (Nickerson et al., 2016), a continually updating database would include the following information:

1. The model, together with assumptions and approximations encoded within the model.
2. The experimental data used to parameterise the model as well as experimental conditions under which the data were collected. Details of the parameterisation should also be recorded so that models do not become disconnected from data as they are reused.

3. Probability distributions for individual parameters and model structures.
4. A history of changes made to the model.
5. A description of how models may be related. In particular, models created by coupling existing models should contain a transparent record of how the sub-models were coupled together as well as any modifications that were made to the individual sub-models.

The probability distributions described in this database (item 4) could be incorporated into a probabilistic framework such as Bayesian inference to weigh existing knowledge against new evidence. This provides a basis for systematically updating models and parameters in the presence of new data. Energetic frameworks such as bond graphs are highly likely to be useful in constructing this sort of database as they ensure that parameters are independent of one another, allow inconsistencies between data to be reconciled with each other, and force modellers to be explicit in their assumptions.

5.2.3 Improving workflows for model development

As we work towards large-scale dynamic models, it is imperative that we make workflows for model development more time efficient by automating aspects of model development. In this thesis, I have shown that bond graphs are useful here because they automatically satisfy thermodynamic consistency and easily incorporate existing thermodynamic information (Chapter 2). Furthermore, I have outlined a strategy that could potentially be used to develop new models of enzymes without requiring prior mechanistic knowledge (Chapter 4), which has the potential to avoid time-consuming model development cycles involving the manual manipulation of equations. To make workflows such as these more efficient, future work is required to efficiently extract data from existing publications and databases for use in model development.

Currently, many attempts at coupling models together either fail or take longer than expected because of the tedious process of tuning parameters within the coupled model. While bond graphs ensure that models remain physically plausible once coupled together, they do not guarantee that the integrated model will remain physiologically realistic. Therefore the database described in § 5.2.2 could aid in model integration by detecting when models are no longer consistent with experimental observations once incorporated into a larger model. Accordingly, this database could be useful in benchmarking models to select the most suitable sub-models for incorporation into a larger model, while also providing the opportunity to update models once integrated. The ability to easily swap out one model for another is essential for assessing model accuracy and computational efficiency. We believe that bond graphs are useful in this context because they are inherently modular, and thus naturally allow models to be swapped (discussed in Section 4.4). The ability to swap sub-models may also prove useful in simulating multi-scale models. In order to achieve a good balance between accuracy and computational efficiency, it may be necessary to switch between a fully detailed sub-model and a simplified version during a simulation. Therefore there is future work to be done in developing numerical techniques that allow model solvers to decide whether to use detailed or simplified versions of a model, depending on their relative accuracy.

Bibliography

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19, 716–723.
- Andrei, O., Kirchner, H., 2009. A port graph calculus for autonomic computing and invariant verification. *Electronic Notes in Theoretical Computer Science* 253, 17–38.
- Anton, H., Rorres, C., 2014. *Elementary Linear Algebra: Applications Version*. John Wiley & Sons Inc, Hoboken, NJ.
- Apell, H.J., 1989. Electrogenic properties of the Na,K pump. *The Journal of Membrane Biology* 110, 103–114.
- Aslanidi, O.V., Boyett, M.R., Dobrzynski, H., Li, J., Zhang, H., 2009. Mechanisms of transition from normal to reentrant electrical activity in a model of rabbit atrial tissue: Interaction of tissue heterogeneity and anisotropy. *Biophysical Journal* 96, 798–817.
- Atkins, P.W., De Paula, J., 2006. *Physical Chemistry for the Life Sciences*. Oxford University Press, Oxford, UK.
- Babtie, A.C., Stumpf, M.P.H., 2017. How to deal with parameters for whole-cell modelling. *Journal of The Royal Society Interface* 14, 20170237.
- Bailey, J.E., 1998. Mathematical modeling and analysis in biochemical engineering: Past accomplishments and future opportunities. *Biotechnology Progress* 14, 8–20.
- Baker, R.E., Peña, J.M., Jayamohan, J., Jérusalem, A., 2018. Mechanistic models versus machine learning, a fight worth fighting for the biological community? *Biology Letters* 14, 20170660.
- Barry, F., 1914. The influence of temperature on chemical reaction in general. *American Journal of Botany* 1, 203–225.
- Bassingthwaighte, J., Hunter, P., Noble, D., 2009. The Cardiac Physiome: perspectives for the future. *Experimental Physiology* 94, 597–605.
- Bassingthwaighte, J.B., 2000. Strategies for the Physiome Project. *Annals of Biomedical Engineering* 28, 1043–1058.
- Batushansky, A., Toubiana, D., Fait, A., 2016. Correlation-based network generation, visualization, and analysis as a powerful tool in biological studies: A case study in cancer cell metabolism. *BioMed Research International* 2016.
- Beard, D.A., Babson, E., Curtis, E., Qian, H., 2004a. Thermodynamic constraints for biochemical networks. *Journal of Theoretical Biology* 228, 327–333.
- Beard, D.A., Bassingthwaighte, J.B., Greene, A.S., 2005. Computational modeling of physiological systems. *Physiological Genomics* 23, 1–3.
- Beard, D.A., Liang, S.d., Qian, H., 2002. Energy balance for analysis of complex metabolic networks. *Biophysical Journal* 83, 79–86.
- Beard, D.A., Neal, M.L., Tabesh-Saleki, N., Thompson, C.T., Bassingthwaighte, J.B., Shimoyama, M., Carlson, B.E., 2012. Multiscale modeling and data integration in the Virtual Physiological Rat Project. *Annals of Biomedical Engineering* 40, 2365–2378.
- Beard, D.A., Qian, H., 2008. *Chemical Biophysics: Quantitative Analysis of Cellular Systems*. Cambridge University Press, Cambridge.
- Beard, D.A., Qian, H., Bassingthwaighte, J.B., 2004b. Stoichiometric foundation of large-scale biochemical system analysis, in: Ciobanu, G., Rozenberg, G. (Eds.), *Modelling in Molecular Biology*. Springer, Berlin, Heidelberg, pp. 1–19.

- Beattie, K.A., Hill, A.P., Bardenet, R., Cui, Y., Vandenberg, J.I., Gavaghan, D.J., Boer, T.P., Mirams, G.R., 2018. Sinusoidal voltage protocols for rapid characterisation of ion channel kinetics. *The Journal of Physiology* 596, 1813–1828.
- Beeler, G.W., Reuter, H., 1977. Reconstruction of the action potential of ventricular myocardial fibres. *The Journal of Physiology* 268, 177–210.
- Berger, J.O., Molina, G., 2005. Posterior model probabilities via path-based pairwise priors. *Statistica Neerlandica* 59, 3–15.
- Beuckelmann, D.J., Wier, W.G., 1989. Sodium-calcium exchange in guinea-pig cardiac cells: exchange current and changes in intracellular Ca^{2+} . *The Journal of Physiology* 414, 499–520.
- Bezanilla, F., Armstrong, C.M., 1977. Inactivation of the sodium channel. I. Sodium current experiments. *The Journal of General Physiology* 70, 549–566.
- Birtwistle, M.R., Mager, D.E., Gallo, J.M., 2013. Mechanistic vs. empirical network models of drug action. *CPT: Pharmacometrics & Systems Pharmacology* 2, 72.
- Blaustein, M.P., Lederer, W.J., 1999. Sodium/calcium exchange: Its physiological implications. *Physiological Reviews* 79, 763–854.
- de Bono, B., Safaei, S., Grenon, P., Hunter, P., 2018. Meeting the multiscale challenge: representing physiology processes over ApiNATOMY circuits using bond graphs. *Interface Focus* 8, 20170026.
- Borutzky, W., 2010. Bond Graph Methodology. Springer, London.
- Borutzky, W., Dauphin-Tanguy, G., Thoma, J.U., 1995. Advances in bond graph modelling: theory, software, applications. *Mathematics and Computers in Simulation* 39, 465–475.
- Carro, J., Rodríguez, J.F., Laguna, P., Pueyo, E., 2011. A human ventricular cell model for investigation of cardiac arrhythmias under hyperkalaemic conditions. *Phil. Trans. R. Soc. A* 369, 4205–4232.
- Colatsky, T., Fermini, B., Gintant, G., Pierson, J.B., Sager, P., Sekino, Y., Strauss, D.G., Stockbridge, N., 2016. The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — Update on progress. *Journal of Pharmacological and Toxicological Methods* 81, 15–20.
- Cooling, M.T., Nickerson, D.P., Nielsen, P.M.F., Hunter, P.J., 2016. Modular modelling with Physiome standards. *The Journal of Physiology* 594, 6817–6831.
- Cooper, J., Scharm, M., Mirams, G.R., 2016. The Cardiac Electrophysiology Web Lab. *Biophysical Journal* 110, 292–300.
- Cornish-Bowden, A., Cárdenas, M.L., Letelier, J.C., Soto-Andrade, J., Abarzúa, F.G., 2004. Understanding the parts in terms of the whole. *Biology of the Cell* 96, 713–717.
- Cortassa, S., Aon, M.A., O'Rourke, B., Jacques, R., Tseng, H.J., Marbán, E., Winslow, R.L., 2006. A computational model integrating electrophysiology, contraction, and mitochondrial bioenergetics in the ventricular myocyte. *Biophysical Journal* 91, 1564–1589.
- Crampin, E.J., Halstead, M., Hunter, P., Nielsen, P., Noble, D., Smith, N., Tawhai, M., 2004a. Computational physiology and the physiome project. *Experimental Physiology* 89, 1–26.
- Crampin, E.J., Schnell, S., McSharry, P.E., 2004b. Mathematical and computational techniques to deduce complex biochemical reaction mechanisms. *Progress in Biophysics and Molecular Biology* 86, 77–112.
- Crampin, E.J., Smith, N.P., 2006. A dynamic model of excitation-contraction coupling during acidosis in cardiac ventricular myocytes. *Biophysical Journal* 90, 3074–3090.
- Cudmore, P., Gawthrop, P.J., Pan, M., Crampin, E.J., 2019. Computer-aided modelling of complex physical systems with BondGraphTools. arXiv:1906.10799 .
- Daly, A.C., Clerx, M., Beattie, K.A., Cooper, J., Gavaghan, D.J., Mirams, G.R., 2018. Reproducible model development in the cardiac electrophysiology Web Lab. *Progress in Biophysics and Molecular Biology* 139, 3–14.
- Díaz-Zuccarini, V., Rafirou, D., LeFevre, J., Hose, D.R., Lawford, P.V., 2009. Systemic modelling and computational physiology: The application of Bond Graph boundary conditions for 3D cardiovascular models. *Simulation Modelling Practice and Theory* 17, 125–136.
- Del Vecchio, D., 2013. A control theoretic framework for modular analysis and design of biomolecular networks. *Annual Reviews in Control* 37, 333–345.
- Del Vecchio, D., Ninfa, A.J., Sontag, E.D., 2008. Modular cell biology: retroactivity and insulation. *Molecular Systems Biology* 4, 161.

- DiFrancesco, D., Noble, D., 1985. A model of cardiac electrical activity incorporating ionic pumps and concentration changes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 307, 353–398.
- Ederer, M., Gilles, E.D., 2007. Thermodynamically feasible kinetic models of reaction networks. *Biophysical Journal* 92, 1846–1857.
- Endresen, L.P., Hall, K., Høye, J.S., Myrheim, J., 2000. A theory for the membrane potential of living cells. *European Biophysics Journal* 29, 90–103.
- Ermentrout, B., 2001. Simplifying and reducing complex models, in: Bower, J.M., Bolouri, H. (Eds.), *Computational Modeling of Genetic and Biochemical Networks*. MIT Press, Massachusetts, pp. 307–323.
- Faber, G.M., Rudy, Y., 2000. Action potential and contractility changes in $[Na^+]$ _i overloaded cardiac myocytes: A simulation study. *Biophysical Journal* 78, 2392–2404.
- Fearnley, C.J., Roderick, H.L., Bootman, M.D., 2011. Calcium signaling in cardiac myocytes. *Cold Spring Harbor Perspectives in Biology* 3, a004242.
- Feliu, E., Wiuf, C., 2012. Variable elimination in chemical reaction networks with mass-action kinetics. *SIAM Journal on Applied Mathematics* 72, 959–981.
- Fink, M., Niederer, S.A., Cherry, E.M., Fenton, F.H., Koivumäki, J.T., Seemann, G., Thul, R., Zhang, H., Sachse, F.B., Beard, D., Crampin, E.J., Smith, N.P., 2011. Cardiac cell modelling: Observations from the heart of the cardiac physiome project. *Progress in Biophysics and Molecular Biology* 104, 2–21.
- Fink, M., Noble, D., 2009. Markov models for ion channels: versatility versus identifiability and speed. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 367, 2161–2179.
- Finkelstein, A., Hetherington, J., Li, L., Margoninski, O., Saffrey, P., Seymour, R., Warner, A., 2004. Computational challenges of systems biology. *Computer* 37, 26–33.
- Forger, D., Gonze, D., Virshup, D., Welsh, D.K., 2007. Beyond intuitive modeling: Combining biophysical models with innovative experiments to move the circadian clock field forward. *Journal of Biological Rhythms* 22, 200–210.
- Fraser, J.A., Huang, C.L.H., 2007. Quantitative techniques for steady-state calculation and dynamic integrated modelling of membrane potential and intracellular ion concentrations. *Progress in Biophysics and Molecular Biology* 94, 336–372.
- Frieden, E., Walter, C., 1963. Prevalence and significance of the product inhibition of enzymes. *Nature* 198, 834–837.
- Friedrich, T., Bamberg, E., Nagel, G., 1996. Na^+,K^+ -ATPase pump currents in giant excised patches activated by an ATP concentration jump. *Biophysical Journal* 71, 2486–2500.
- Gauthier, L.D., Greenstein, J.L., Winslow, R.L., 2012. Toward an integrative computational model of the guinea pig cardiac myocyte. *Frontiers in Physiology* 3.
- Gawthrop, P., Bevan, G., 2007. Bond-graph modeling. *IEEE Control Systems* 27, 24–45.
- Gawthrop, P., Smith, L., 1996. Metamodelling: For bond graphs and dynamic systems. Prentice Hall International Series in Systems and Control Engineering, Prentice Hall, London, New York.
- Gawthrop, P.J., 2017a. Bond graph modeling of chemiosmotic biomolecular energy transduction. *IEEE Transactions on NanoBioscience* 16, 177–188.
- Gawthrop, P.J., 2017b. Bond-graph modelling and causal analysis of biomolecular systems, in: Bond Graphs for Modelling, Control and Fault Diagnosis of Engineering Systems. Springer, Cham, pp. 587–623.
- Gawthrop, P.J., Crampin, E.J., 2014. Energy-based analysis of biochemical cycles using bond graphs. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 470, 20140459.
- Gawthrop, P.J., Crampin, E.J., 2016. Modular bond-graph modelling and analysis of biomolecular systems. *IET Systems Biology* 10, 187–201.
- Gawthrop, P.J., Crampin, E.J., 2017. Energy-based analysis of biomolecular pathways. *Proc. R. Soc. A* 473, 20160825.

- Gawthrop, P.J., Crampin, E.J., 2018. Biomolecular system energetics, in: Proceedings of the 13th International Conference on Bond Graph Modeling (ICBGM'18). Society for Computer Simulation, Bordeaux. [arXiv:1803.09231](https://arxiv.org/abs/1803.09231).
- Gawthrop, P.J., Cursons, J., Crampin, E.J., 2015. Hierarchical bond graph modelling of biochemical networks. *Proc. R. Soc. A* 471, 20150642.
- Gawthrop, P.J., Siekmann, I., Kameneva, T., Saha, S., Ibbotson, M.R., Crampin, E.J., 2017. Bond graph modelling of chemoelectrical energy transduction. *IET Systems Biology* 11, 127–138.
- Giladi, M., Shor, R., Lisnyansky, M., Khananshvili, D., 2016. Structure-functional basis of ion transport in sodium–calcium exchanger (NCX) proteins. *International Journal of Molecular Sciences* 17.
- Gillespie, P.G., Walker, R.G., 2001. Molecular basis of mechanosensory transduction. *Nature* 413, 194–202.
- Glitsch, H.G., 2001. Electrophysiology of the sodium-potassium-ATPase in cardiac cells. *Physiological Reviews* 81, 1791–1826.
- Glynn, I.M., 2002. A hundred years of sodium pumping. *Annual Review of Physiology* 64, 1–18.
- Gonçalves, E., Bucher, J., Ryll, A., Niklas, J., Mauch, K., Klamt, S., Rocha, M., Saez-Rodriguez, J., 2013. Bridging the layers: towards integration of signal transduction, regulation and metabolism into mathematical models. *Molecular BioSystems* 9, 1576.
- Grandi, E., Pasqualini, F.S., Bers, D.M., 2010. A novel computational model of the human ventricular action potential and Ca transient. *Journal of Molecular and Cellular Cardiology* 48, 112–121.
- Greenstein, J.L., Hinch, R., Winslow, R.L., 2006. Mechanisms of excitation-contraction coupling in an integrative model of the cardiac ventricular myocyte. *Biophysical Journal* 90, 77–91.
- Gross, E., Harrington, H.A., Rosen, Z., Sturmfels, B., 2016. Algebraic systems biology: A case study for the Wnt pathway. *Bulletin of Mathematical Biology* 78, 21–51.
- Guan, S., Lu, Q., Huang, K., 1997. A discussion about the DiFrancesco–Noble model. *Journal of Theoretical Biology* 189, 27–32.
- Gunawardena, J., 2014. Time-scale separation – Michaelis and Menten’s old idea, still bearing fruit. *FEBS Journal* 281, 473–488.
- Guynn, R.W., Veech, R.L., 1973. The equilibrium constants of the adenosine triphosphate hydrolysis and the adenosine triphosphate-citrate lyase reactions. *Journal of Biological Chemistry* 248, 6966–6972.
- Hansen, P.S., Buhagiar, K.A., Kong, B.Y., Clarke, R.J., Gray, D.F., Rasmussen, H.H., 2002. Dependence of Na^+ - K^+ pump current-voltage relationship on intracellular Na^+ , K^+ , and Cs^+ in rabbit cardiac myocytes. *American Journal of Physiology - Cell Physiology* 283, C1511–C1521.
- Haraldsdóttir, H.S., Fleming, R.M.T., 2016. Identification of conserved moieties in metabolic networks by graph theoretical analysis of atom transition networks. *PLoS Computational Biology* 12.
- Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W., 1999. From molecular to modular cell biology. *Nature* 402, C47–C52.
- Hasenstaub, A., Otte, S., Callaway, E., Sejnowski, T.J., 2010. Metabolic cost as a unifying principle governing neuronal biophysics. *Proceedings of the National Academy of Sciences* 107, 12329–12334.
- Hass, H., Loos, C., Alvarez, E.R., Timmer, J., Hasenauer, J., Kreutz, C., 2019. Benchmark problems for dynamic modeling of intracellular processes. *Bioinformatics* (in press).
- Helland, I.S., 1990. Partial least squares regression and statistical models. *Scandinavian Journal of Statistics* 17, 97–114.
- Higgins, E.R., Cannell, M.B., Sneyd, J., 2006. A buffering SERCA pump in models of calcium dynamics. *Biophysical Journal* 91, 151–163.
- Hilgemann, D.W., Matsuoka, S., Nagel, G.A., Collins, A., 1992. Steady-state and dynamic properties of cardiac sodium-calcium exchange. Sodium-dependent inactivation. *The Journal of General Physiology* 100, 905–932.
- Hill, A.V., 1938. The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London B: Biological Sciences* 126, 136–195.

- Hill, T.L., 1974. Theoretical formalism for the sliding filament model of contraction of striated muscle Part I. *Progress in Biophysics and Molecular Biology* 28, 267–340.
- Hill, T.L., 1989. Free Energy Transduction and Biochemical Cycle Kinetics. Springer, New York.
- Hinch, R., Greenstein, J.L., Tanskanen, A.J., Xu, L., Winslow, R.L., 2004. A simplified local control model of calcium-induced calcium release in cardiac ventricular myocytes. *Biophysical Journal* 87, 3723–3736.
- Hodgkin, A.L., Huxley, A.F., 1952. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology* 117, 500–544.
- Hucka, M., Finney, A., Sauro, H.M., Bolouri, H., Doyle, J.C., Kitano, H., Arkin, A.P., Bornstein, B.J., Bray, D., Cornish-Bowden, A., Cuellar, A.A., Dronov, S., Gilles, E.D., Ginkel, M., Gor, V., Goryanin, I.I., Hedley, W.J., Hodgman, T.C., Hofmeyr, J.H., Hunter, P.J., Juty, N.S., Kasberger, J.L., Kremling, A., Kummer, U., Novère, N.L., Loew, L.M., Lucio, D., Mendes, P., Minch, E., Mjolsness, E.D., Nakayama, Y., Nelson, M.R., Nielsen, P.F., Sakurada, T., Schaff, J.C., Shapiro, B.E., Shimizu, T.S., Spence, H.D., Stelling, J., Takahashi, K., Tomita, M., Wagner, J., Wang, J., 2003. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524–531.
- Hund, T.J., Kucera, J.P., Otani, N.F., Rudy, Y., 2001. Ionic charge conservation and long-term steady state in the Luo–Rudy dynamic cell model. *Biophysical Journal* 81, 3324–3331.
- Hund, T.J., Rudy, Y., 2004. Rate dependence and regulation of action potential and calcium transient in a canine cardiac ventricular cell model. *Circulation* 110, 3168–3174.
- Hung, C.h., Markham, T.L., 1975. The Moore-Penrose inverse of a partitioned matrix. *Linear Algebra and its Applications* 11, 73–86.
- Hunter, P., 2016. The Virtual Physiological Human: The Physiome Project aims to develop reproducible, multiscale models for clinical practice. *IEEE Pulse* 7, 36–42.
- Hunter, P.J., Borg, T.K., 2003. Integration from proteins to organs: the Physiome Project. *Nature Reviews Molecular Cell Biology* 4, 237–243.
- Hurley, D.G., Budden, D.M., Crampin, E.J., 2015. Virtual Reference Environments: a simple way to make research reproducible. *Briefings in Bioinformatics* 16, 901–903.
- Huxley, A.F., 1957. Muscle structure and theories of contraction. *Progress in Biophysics and Biophysical Chemistry* 7, 255–318.
- Hwang, S.T., 2004. Nonequilibrium thermodynamics of membrane transport. *AIChE Journal* 50, 862–870.
- Jafri, M.S., Rice, J.J., Winslow, R.L., 1998. Cardiac Ca^{2+} dynamics: The roles of ryanodine receptor adaptation and sarcoplasmic reticulum load. *Biophysical Journal* 74, 1149–1168.
- Karr, J.R., Sanghvi, J.C., Macklin, D.N., Gutschow, M.V., Jacobs, J.M., Bolival Jr., B., Assad-Garcia, N., Glass, J.I., Covert, M.W., 2012. A whole-cell computational model predicts phenotype from genotype. *Cell* 150, 389–401.
- Kawase, Y., Hajjar, R.J., 2008. The cardiac sarcoplasmic/endoplasmic reticulum calcium ATPase: a potent target for cardiovascular diseases. *Nature Clinical Practice Cardiovascular Medicine* 5, 554–565.
- Keener, J., Sneyd, J., 2009. Mathematical Physiology. Springer, New York.
- Keizer, J., 1987. Statistical Thermodynamics of Nonequilibrium Processes. Springer, New York.
- Kennedy, J., Eberhart, R., 1995. Particle swarm optimization, in: IEEE International Conference on Neural Networks, 1995. Proceedings, pp. 1942–1948 vol.4.
- Kimura, J., Miyamae, S., Noma, A., 1987. Identification of sodium-calcium exchange current in single ventricular cells of guinea-pig. *The Journal of Physiology* 384, 199–222.
- Klipp, E., Liebermeister, W., Wierling, C., Kowald, A., Lehrach, H., Herwig, R., 2009. Systems Biology: A Textbook. Wiley-VCH, Weinheim.
- Kneller, J., Ramirez, R.J., Chartier, D., Courtemanche, M., Nattel, S., 2002. Time-dependent transients in an ionically based mathematical model of the canine atrial action potential. *American Journal of Physiology - Heart and Circulatory Physiology* 282, H1437–H1451.
- Kohl, P., Crampin, E.J., Quinn, T.A., Noble, D., 2010. Systems biology: An approach. *Clinical Pharmacology & Therapeutics* 88, 25–33.

- Kolczyk, K., Conradi, C., 2016. Challenges in horizontal model integration. *BMC Systems Biology* 10.
- Lambeth, M.J., Kushmerick, M.J., 2002. A computational model for glycogenolysis in skeletal muscle. *Annals of Biomedical Engineering* 30, 808–827.
- Liebermeister, W., Klipp, E., 2006. Bringing metabolic networks to life: convenience rate law and thermodynamic constraints. *Theoretical Biology and Medical Modelling* 3, 41.
- Liebermeister, W., Uhlendorf, J., Klipp, E., 2010. Modular rate laws for enzymatic reactions: thermodynamics, elasticities and implementation. *Bioinformatics* 26, 1528–1534.
- Liepe, J., Filippi, S., Komorowski, M., Stumpf, M.P.H., 2013. Maximizing the information content of experiments in systems biology. *PLOS Computational Biology* 9, e1002888.
- Liepe, J., Kirk, P., Filippi, S., Toni, T., Barnes, C.P., Stumpf, M.P.H., 2014. A framework for parameter estimation and model selection from experimental data in systems biology using approximate Bayesian computation. *Nature Protocols* 9, 439–456.
- Livshitz, L., Rudy, Y., 2009. Uniqueness and stability of action potential models during rest, pacing, and conduction using problem-solving environment. *Biophysical Journal* 97, 1265–1276.
- Lloyd, C.M., Halstead, M.D.B., Nielsen, P.F., 2004. CellML: its future, present and past. *Progress in Biophysics and Molecular Biology* 85, 433–450.
- Luo, C.H., Rudy, Y., 1991. A model of the ventricular cardiac action potential. Depolarization, repolarization, and their interaction. *Circulation Research* 68, 1501–1526.
- Luo, C.H., Rudy, Y., 1994a. A dynamic model of the cardiac ventricular action potential. I. Simulations of ionic currents and concentration changes. *Circulation Research* 74, 1071–1096.
- Luo, C.H., Rudy, Y., 1994b. A dynamic model of the cardiac ventricular action potential. II. Afterdepolarizations, triggered activity, and potentiation. *Circulation Research* 74, 1097–1113.
- Machado, D., Costa, R.S., Rocha, M., Ferreira, E.C., Tidor, B., Rocha, I., 2011. Modeling formalisms in systems biology. *AMB express* 1, 1–14.
- Macklin, D.N., Ruggero, N.A., Covert, M.W., 2014. The future of whole-cell modeling. *Current Opinion in Biotechnology* 28, 111–115.
- Makinose, M., Hasselbach, W., 1971. ATP synthesis by the reverse of the sarcoplasmic calcium pump. *FEBS Letters* 12, 271–272.
- Mallavarapu, A., Thomson, M., Ullian, B., Gunawardena, J., 2009. Programming with models: modularity and abstraction provide powerful capabilities for systems biology. *Journal of The Royal Society Interface* 6, 257–270.
- Mason, J.C., Covert, M.W., 2019. An energetic reformulation of kinetic rate laws enables scalable parameter estimation for biochemical networks. *Journal of Theoretical Biology* 461, 145–156.
- McAllister, R.E., Noble, D., Tsien, R.W., 1975. Reconstruction of the electrical activity of cardiac Purkinje fibres. *The Journal of Physiology* 251, 1.
- de Meis, L., 2002. Ca^{2+} -ATPases (SERCA): Energy transduction and heat production in transport ATPases. *The Journal of Membrane Biology* 188, 1–9.
- Mescam, M., Vinnakota, K.C., Beard, D.A., 2011. Identification of the catalytic mechanism and estimation of kinetic parameters for fumarase. *Journal of Biological Chemistry* 296, 21100–21109.
- Meurer, A., Smith, C.P., Paprocki, M., Čertík, O., Kirpichev, S.B., Rocklin, M., Kumar, A., Ivanov, S., Moore, J.K., Singh, S., Rathnayake, T., Vig, S., Granger, B.E., Muller, R.P., Bonazzi, F., Gupta, H., Vats, S., Johansson, F., Pedregosa, F., Curry, M.J., Terrel, A.R., Roučka, Š., Saboo, A., Fernando, I., Kulal, S., Cimrman, R., Scopatz, A., 2017. SymPy: symbolic computing in Python. *PeerJ Computer Science* 3, e103.
- Meyer, P.E., Kontos, K., Lafitte, F., Bontempi, G., 2007. Information-theoretic inference of large transcriptional regulatory networks. *EURASIP Journal on Bioinformatics and Systems Biology* 2007, 1–9.
- Mogensen, P.K., Riseth, A.N., 2018. Optim: A mathematical optimization package for Julia. *The Journal of Open Source Software* 3, 615.
- Mueckler, M., Thorens, B., 2013. The SLC2 (GLUT) family of membrane transporters. *Molecular Aspects of Medicine* 34, 121–138.

- Nakao, M., Gadsby, D.C., 1989. [Na] and [K] dependence of the Na/K pump current-voltage relationship in guinea pig ventricular myocytes. *The Journal of General Physiology* 94, 539–565.
- Nakayama, H., Bodí, I., Maillet, M., DeSantiago, J., Domeier, T.L., Mikoshiba, K., Lorenz, J.N., Blatter, L.A., Bers, D.M., Molkentin, J.D., 2010. The IP₃ receptor regulates cardiac hypertrophy in response to select stimuli. *Circulation Research* 107, 659–666.
- Neal, M.L., Carlson, B.E., Thompson, C.T., James, R.C., Kim, K.G., Tran, K., Crampin, E.J., Cook, D.L., Gennari, J.H., 2015. Semantics-based composition of integrated cardiomyocyte models motivated by real-world use cases. *PLOS ONE* 10, e0145621.
- Neal, M.L., Cooling, M.T., Smith, L.P., Thompson, C.T., Sauro, H.M., Carlson, B.E., Cook, D.L., Gennari, J.H., 2014. A reappraisal of how to build modular, reusable models of biological systems. *PLOS Comput Biol* 10, e1003849.
- Neubauer, S., 2007. The failing heart — An engine out of fuel. *New England Journal of Medicine* 356, 1140–1151.
- Nickerson, D., Atalag, K., de Bono, B., Geiger, J., Goble, C., Hollmann, S., Lonien, J., Müller, W., Regierer, B., Stanford, N.J., Golebiewski, M., Hunter, P., 2016. The Human Physiome: how standards, software and innovative service infrastructures are providing the building blocks to make it achievable. *Interface Focus* 6, 20150103.
- Niederer, S.A., de Oliveira, B.L., Curtis, M.J., 2017. The opportunities and challenges for biophysical modelling of beneficial and adverse drug actions on the heart. *Current Opinion in Systems Biology* 4, 29–34.
- Noble, D., 1962. A modification of the Hodgkin-Huxley equations applicable to Purkinje fibre action and pace-maker potentials. *The Journal of Physiology* 160, 317–352.
- Noble, D., 2002. Modeling the heart—from genes to cells to the whole organ. *Science* 295, 1678–1682.
- Noble, D., Rudy, Y., 2001. Models of cardiac ventricular action potentials: iterative interaction between experiment and simulation. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 359, 1127–1142.
- Nocedal, J., Wright, S.J., 2006. Numerical Optimization. Springer Series in Operations Research. 2nd ed., Springer, New York.
- O'Hara, T., Virág, L., Varró, A., Rudy, Y., 2011. Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLOS Computational Biology* 7, e1002061.
- Okino, M.S., Mavrovouniotis, M.L., 1998. Simplification of mathematical models of chemical reaction systems. *Chemical reviews* 98, 391–408.
- Omholt, S.W., Hunter, P.J., 2016. The Human Physiome: a necessary key for the creative destruction of medicine. *Interface Focus* 6, 20160003.
- Oster, G., Perelson, A., Katchalsky, A., 1971. Network thermodynamics. *Nature* 234, 393–399.
- Oster, G.F., Perelson, A.S., Katchalsky, A., 1973. Network thermodynamics: dynamic modelling of biophysical systems. *Quarterly Reviews of Biophysics* 6, 1–134.
- Palsson, B., 2006. Systems Biology: Properties of Reconstructed Networks. Cambridge University Press, Cambridge.
- Pan, M., Gawthrop, P.J., Cursons, J., Tran, K., Crampin, E.J., 2017. The cardiac Na⁺/K⁺ ATPase: An updated, thermodynamically consistent model. arXiv:1711.00989 [q-bio].
- Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Crampin, E.J., 2018a. Bond graph modelling of the cardiac action potential: implications for drift and non-unique steady states. *Proc. R. Soc. A* 474, 20180106.
- Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Crampin, E.J., 2018b. Supporting code for “A thermodynamic framework for modelling membrane transporters”. Zenodo. URL: <https://zenodo.org/record/1422867>, doi:10.5281/zenodo.1422867.
- Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Crampin, E.J., 2018c. Virtual reference environment for “A thermodynamic framework for modelling membrane transporters”. URL: <https://zenodo.org/record/1433958>, doi:10.5281/zenodo.1433958.
- Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Crampin, E.J., 2019. A thermodynamic framework for modelling membrane transporters. *Journal of Theoretical Biology* 481, 10–23.

- Pan, M., Gawthrop, P.J., Tran, K., Cursons, J., Edmund J. Crampin, 2018d. Supporting code for “Bond graph modelling of the cardiac action potential: Implications for drift and non-unique steady states”. Zenodo. URL: <https://zenodo.org/record/1172205>, doi:10.5281/zenodo.1172205.
- Pandit, S.V., Clark, R.B., Giles, W.R., Demir, S.S., 2001. A mathematical model of action potential heterogeneity in adult rat left ventricular myocytes. *Biophysical Journal* 81, 3029–3051.
- Patel, M.I., Nagl, S., 2010. The Role of Model Integration in Complex Systems Modelling: An Example from Cancer Biology. *Understanding complex systems*, Springer, Berlin.
- Paynter, H.M., 1961. *Analysis and design of engineering systems*. MIT press.
- Periasamy, M., Kalyanasundaram, A., 2007. SERCA pump isoforms: Their role in calcium transport and disease. *Muscle & Nerve* 35, 430–442.
- Pinz, I., Tian, R., Belke, D., Swanson, E., Dillmann, W., Ingwall, J.S., 2011. Compromised myocardial energetics in hypertrophied mouse hearts diminish the beneficial effect of overexpressing SERCA2a. *Journal of Biological Chemistry* 286, 10163–10168.
- Polettini, M., Esposito, M., 2014. Irreversible thermodynamics of open chemical networks. I. Emergent cycles and broken conservation laws. *The Journal of Chemical Physics* 141, 024117.
- Price, K.V., Storn, R.M., Lampinen, J.A., 2005. *Differential Evolution: A Practical Approach to Global Optimization*. *Natural Computing Series*, Springer, Berlin.
- Rasmusson, R.L., Clark, J.W., Giles, W.R., Robinson, K., Clark, R.B., Shibata, E.F., Campbell, D.L., 1990. A mathematical model of electrophysiological activity in a bullfrog atrial cell. *American Journal of Physiology - Heart and Circulatory Physiology* 259, H370–H389.
- Reuter, H., 1984. Ion channels in cardiac cell membranes. *Annual review of physiology* 46, 473–484.
- Reytor González, M.L., Cornell-Kennon, S., Schaefer, E., Kuzmič, P., 2017. An algebraic model to determine substrate kinetic parameters by global nonlinear fit of progress curves. *Analytical Biochemistry* 518, 16–24.
- Rice, J.J., Jafri, M.S., Winslow, R.L., 2000. Modeling short-term interval-force relations in cardiac muscle. *Am. J. Physiol. Heart Circ. Physiol.* 278, H913–H931.
- Rice, J.J., Wang, F., Bers, D.M., de Tombe, P.P., 2008. Approximate model of cooperative activation and crossbridge cycling in cardiac muscle using ordinary differential equations. *Biophysical Journal* 95, 2368–2390.
- Rice, J.J., Winslow, R.L., Hunter, W.C., 1999. Comparison of putative cooperative mechanisms in cardiac muscle: length dependence and dynamic responses. *American Journal of Physiology-Heart and Circulatory Physiology* 276, H1734–H1754.
- Rudy, Y., Silva, J.R., 2006. Computational biology in the study of cardiac ion channels and cell electrophysiology. *Quarterly Reviews of Biophysics* 39, 57–116.
- Sager, P.T., Gintant, G., Turner, J.R., Pettit, S., Stockbridge, N., 2014. Rechanneling the cardiac proarrhythmia safety paradigm: A meeting report from the Cardiac Safety Research Consortium. *American Heart Journal* 167, 292–300.
- Sakmann, B., Trube, G., 1984. Conductance properties of single inwardly rectifying potassium channels in ventricular cells from guinea-pig heart. *The Journal of Physiology* 347, 641–657.
- van der Schaft, A., Jeltsema, D., 2014. Port-Hamiltonian Systems Theory: An Introductory Overview. *Foundations and Trends in Systems and Control* 1, 173–378.
- van der Schaft, A., Rao, S., Jayawardhana, B., 2013. On the mathematical structure of balanced chemical reaction networks governed by mass action kinetics. *SIAM Journal on Applied Mathematics* 73, 953–973.
- Schauer, M., Heinrich, R., 1983. Quasi-steady-state approximation in the mathematical modeling of biochemical reaction networks. *Mathematical Biosciences* 65, 155–170.
- Schnell, S., Chappell, M.J., Evans, N.D., Roussel, M.R., 2006. The mechanism distinguishability problem in biochemical kinetics: The single-enzyme, single-substrate reaction as a case study. *Comptes Rendus Biologies* 329, 51–61.
- Schnell, S., Maini, P.K., 2003. A century of enzyme kinetics. Should we believe in the K_m and v_{max} estimates? *Comments in Theoretical Biology* 8, 169–187.

- Schramm, M., Klieber, H.G., Daut, J., 1994. The energy expenditure of actomyosin-ATPase, Ca^{2+} -ATPase and Na^+,K^+ -ATPase in guinea-pig cardiac ventricular muscle. *The Journal of Physiology* 481, 647–662.
- Schuster, S., Höfer, T., 1991. Determining all extreme semi-positive conservation relations in chemical reaction systems: a test criterion for conservativity. *Journal of the Chemical Society, Faraday Transactions* 87, 2561–2566.
- Schuster, S., Hilgetag, C., 1995. What information about the conserved-moiety structure of chemical reaction systems can be derived from their stoichiometry? *The Journal of Physical Chemistry* 99, 8017–8023.
- Segel, I.H., 1975. *Enzyme Kinetics: Behavior and Analysis of Rapid Equilibrium and Steady-State Enzyme Systems*. Wiley, New York.
- Shaw, R.M., Rudy, Y., 1997. Electrophysiologic effects of acute myocardial ischemia: a theoretical study of altered cell excitability and action potential duration. *Cardiovascular Research* 35, 256–272.
- Shibasaki, T., 1987. Conductance and kinetics of delayed rectifier potassium channels in nodal cells of the rabbit heart. *The Journal of Physiology* 387, 227.
- Smith, N.P., Crampin, E.J., 2004. Development of models of active ion transport for whole-cell modelling: cardiac sodium–potassium pump as a case study. *Progress in Biophysics and Molecular Biology* 85, 387–405.
- Smith, N.P., Crampin, E.J., Niederer, S.A., Bassingthwaite, J.B., Beard, D.A., 2007. Computational biology of cardiac myocytes: proposed standards for the physiome. *Journal of Experimental Biology* 210, 1576–1583.
- Snowden, T.J., van der Graaf, P.H., Tindall, M.J., 2017. Methods of model reduction for large-scale biological systems: A survey of current methods and trends. *Bulletin of Mathematical Biology* 79, 1449–1486.
- Soh, K.C., Hatzimanikatis, V., 2010. Network thermodynamics in the post-genomic era. *Current Opinion in Microbiology* 13, 350–357.
- Sterman, J.D., 2001. System dynamics modeling: Tools for learning in a complex world. *California Management Review* 43, 8–25.
- Surovtsova, I., Simus, N., Hübner, K., Sahle, S., Kummer, U., 2012. Simplification of biochemical models: a general approach based on the analysis of the impact of individual species and reactions on the systems dynamics. *BMC Systems Biology* 6, 14.
- Szigeti, B., Roth, Y.D., Sekar, J.A.P., Goldberg, A.P., Pochiraju, S.C., Karr, J.R., 2018. A blueprint for human whole-cell modeling. *Current Opinion in Systems Biology* 7, 8–15.
- Terkildsen, J., 2006. Modelling Extracellular Potassium Accumulation in Cardiac Ischaemia. Master's thesis. The University of Auckland.
- Terkildsen, J.R., Crampin, E.J., Smith, N.P., 2007. The balance between inactivation and activation of the Na^+,K^+ pump underlies the triphasic accumulation of extracellular K^+ during myocardial ischemia. *American Journal of Physiology - Heart and Circulatory Physiology* 293, H3036–H3045.
- Terkildsen, J.R., Niederer, S., Crampin, E.J., Hunter, P., Smith, N.P., 2008. Using Physiome standards to couple cellular functions for rat cardiac excitation–contraction. *Experimental Physiology* 93, 919–929.
- Thoma, J., Bouamama, B.O., 2000. *Modelling and Simulation in Thermal and Chemical Engineering*. Springer, Berlin.
- Toni, T., Welch, D., Strelkowa, N., Ipsen, A., Stumpf, M.P.H., 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface* 6, 187–202.
- Tracqui, P., 2009. Biophysical models of tumour growth. *Reports on Progress in Physics* 72, 056701.
- Tran, K., Loiselle, D.S., Crampin, E.J., 2015. Regulation of cardiac cellular bioenergetics: mechanisms and consequences. *Physiological Reports* 3, e12464.
- Tran, K., Smith, N.P., Loiselle, D.S., Crampin, E.J., 2009. A thermodynamic model of the cardiac sarcoplasmic/endoplasmic Ca^{2+} (SERCA) pump. *Biophysical Journal* 96, 2029–2042.

- Tran, K., Smith, N.P., Loiselle, D.S., Crampin, E.J., 2010. A metabolite-sensitive, thermodynamically constrained model of cardiac cross-bridge cycling: Implications for force development during ischemia. *Biophysical Journal* 98, 267–276.
- ten Tusscher, K.H.W.J., Noble, D., Noble, P.J., Panfilov, A.V., 2004. A model for human ventricular tissue. *American Journal of Physiology - Heart and Circulatory Physiology* 286, H1573–H1589.
- Varghese, A., Sell, G.R., 1997. A conservation principle and its effect on the formulation of Na–Ca exchanger current in cardiac cells. *Journal of Theoretical Biology* 189, 33–40.
- Villaverde, A.F., Banga, J.R., 2014. Reverse engineering and identification in systems biology: strategies, perspectives and challenges. *Journal of The Royal Society Interface* 11, 20130505.
- Voit, E.O., 2017. The best models of metabolism. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 9, e1391.
- Walker, M.A., Williams, G.S.B., Kohl, T., Lehnart, S.E., Jafri, M.S., Greenstein, J.L., Lederer, W.J., Winslow, R.L., 2014. Superresolution modeling of calcium release in the heart. *Biophysical Journal* 107, 3018–3029.
- Waltemath, D., Adams, R., Bergmann, F.T., Hucka, M., Kolpakov, F., Miller, A.K., Moraru, I.I., Nickerson, D., Sahle, S., Snoep, J.L., Le Novère, N., 2011. Reproducible computational biology experiments with SED-ML - The Simulation Experiment Description Markup Language. *BMC Systems Biology* 5, 198.
- Waltemath, D., Karr, J., Bergmann, F., Chelliah, V., Hucka, M., Krantz, M., Liebermeister, W., Mendes, P., Myers, C., Pir, P., Alaybeyoglu, B., Aranganathan, N., Baghalian, K., Bittig, A., Burke, P., Cantarelli, M., Chew, Y., Costa, R., Cursons, J., Czauderna, T., Goldberg, A., Gomez, H., Hahn, J., Hameri, T., Gardiol, D., Kazakiewicz, D., Kiselev, I., Knight-Schrijver, V., Knupfer, C., Konig, M., Lee, D., Lloret-Villas, A., Mandrik, N., Medley, J., Moreau, B., Naderi-Meshkin, H., Palaniappan, S., Priego-Espinosa, D., Scharm, M., Sharma, M., Smallbone, K., Stanford, N., Song, J.H., Theile, T., Tokic, M., Tomar, N., Toure, V., Uhlendorf, J., Varusai, T., Watanabe, L., Wendland, F., Wolfien, M., Yurkovich, J., Zhu, Y., Zardilis, A., Zhukova, A., Schreiber, F., 2016. Toward community standards and software for whole-cell modeling. *IEEE Transactions on Biomedical Engineering* 63, 2007–2014.
- Weckwerth, W., 2003. Metabolomics in systems biology. *Annual Review of Plant Biology* 54, 669–689.
- Willems, J.C., 2007. The behavioral approach to open and interconnected systems. *IEEE Control Systems Magazine* 27, 46–99.
- Williams, G.S.B., Chikando, A.C., Tuan, H.T.M., Sobie, E.A., Lederer, W.J., Jafri, M.S., 2011. Dynamics of calcium sparks and calcium leak in the heart. *Biophysical Journal* 101, 1287–1296.
- Yue, L., Feng, J., Li, G.R., Nattel, S., 1996. Characterization of an ultrarapid delayed rectifier potassium channel involved in canine atrial repolarization. *The Journal of Physiology* 496, 647–662.
- Zahalak, G.I., 1990. Modeling Muscle Mechanics (and Energetics), in: Winters, J.M., Woo, S.L.Y. (Eds.), *Multiple Muscle Systems: Biomechanics and Movement Organization*. Springer, New York, pp. 1–23.
- Zielinski, M., McGann, L.E., Nychka, J.A., Elliott, J.A.W., 2017. Nonideal solute chemical potential equation and the validity of the grouped solute approach for intracellular solution thermodynamics. *Journal of Physical Chemistry B* 121, 10443–10456.

University Library



MINERVA
ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Pan, Michael

Title:

A bond graph approach to integrative biophysical modelling

Date:

2019

Persistent Link:

<http://hdl.handle.net/11343/230908>

File Description:

Final thesis file

Terms and Conditions:

Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.