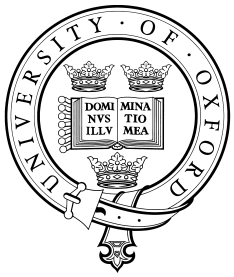
Mathematical Modelling of Drug-Ion Channel Interactions for Cardiac Safety Assessment

Kylie Beattie
Lincoln College



Computational Biology Research Group
Department of Computer Science
University of Oxford

Trinity Term 2015

This thesis is submitted to the Department of Computer Science, University of Oxford, for the degree of Doctor of Philosophy. This thesis is entirely my own work, and, except where otherwise indicated, describes my own research.

Kylie Beattie
Lincoln College

Doctor of Philosophy
Trinity Term 2015

Mathematical Modelling of Drug-Ion Channel Interactions for Cardiac Safety Assessment

Abstract

Unintended drug interactions with ion channels in cardiac cells can alter normal electrical activity in the heart. This can lead to the onset of cardiac arrhythmias which can be potentially fatal. Such arrhythmias are a leading cause of market withdrawal of pharmaceutical compounds. It is therefore important for pharmaceutical companies to identify drugs which may cause adverse changes to cardiac electrical activity as early as possible in the drug development process.

Computational approaches provide great potential for assessing cardiac safety at low cost and high-throughput. Regulatory guidelines for the assessment of cardiac safety are being revised and computational approaches are envisaged to form an integral part of future guidelines. It is therefore important that computational methods proposed for use are both accurate and well validated.

This thesis is focused around the use of mathematical models for cardiac safety assessment. We evaluate an approach which uses information on the extent of drug block of cardiac ion channels to predict changes in cardiac electrical activity in a rabbit-based safety test. A possible limitation of this approach is the representation of drug-ion channel interactions within the mathematical model. We go on to explore the representation of the kinetics of the cardiac potassium channel encoded by hERG (human-Ether-a-go-go Related Gene) within mathematical cardiac electrophysiology models.

We demonstrate how existing mathematical models of hERG channel kinetics can behave very differently when used to simulate the same voltage protocol. We also show the impact that the choice of hERG channel representation used within an action potential model can have on the prediction of drug effects on cardiac activity. We design and perform novel experiments to allow us to construct and validate mathematical models describing hERG channel kinetics.

We demonstrate the potential for the use of computational approaches in cardiac safety assessment. However, at the same time we indicate that models used for this purpose must be carefully constructed and thoroughly validated, as well as illustrating ways this may be achieved. Such considerations are a prerequisite for more predictive computational assessment of cardiac safety.

Contents

1 Introduction **1**
1.1 Motivation 1
1.2 Thesis Structure 5

2 Cardiac Physiology and Drug Safety Assessment **8**
2.1 Cardiac Structure and Function 8
2.2 Cardiac Electrophysiology 10
2.3 Drug-Induced Changes to Cardiac Electrical Activity 19
2.4 Pharmaceutical Cardiac Safety Assessment 21
2.5 Summary 25

3 Mathematical Modelling of Cardiac Electrophysiology **26**
3.1 Mathematical Models of Ion Channel Kinetics 26
3.2 Mathematical Models of Ventricular Cardiomyocyte Electrophysiology . . 49
3.3 Modelling Electrical Propagation in Ventricular Tissue 53
3.4 Mathematical Representations of Drug-Ion Channel Interactions 55
3.5 *In Silico* Approaches for Industrial Pharmaceutical Cardiac Safety Assessment 59
3.6 Summary 63

4 Using Ion Channel Screening Data and Action Potential Simulations to Predict QT Interval Changes in the Rabbit Ventricular Wedge **65**
4.1 Introduction 66
4.2 Methods 67
4.3 Evaluation Results 77
4.4 Limitations 93
4.5 Discussion 94
4.6 Summary 98

5 Voltage Protocol Design for Determining a Model of hERG Channel Kinetics **100**
5.1 Motivation 100
5.2 Methods 116
5.3 Results 132
5.4 Discussion 161
5.5 Summary 164

6 Experimental Recordings Using the Whole-Cell Patch Clamp Technique **165**
6.1 Outline of the Whole-Cell Patch Clamp Technique 165

6.2 Methods . . . 170
6.3 Recording Technique . . . 171
6.4 Voltage Protocols . . . 172
6.5 Experimental Recordings . . . 172
6.6 Quantitatively Comparing Existing hERG Channel Models . . . 185
6.7 Discussion . . . 188
6.8 Summary . . . 189

7 Constructing Cell-Specific Models of hERG Channel Kinetics 190
7.1 Introduction . . . 190
7.2 Summary of Experimental Data Sets . . . 191
7.3 Considerations when Performing Mathematical Model Construction . . . 196
7.4 Methods . . . 199
7.5 Results . . . 206
7.6 Discussion . . . 236
7.7 Summary . . . 238

8 Discussion and Future Research 240
8.1 Summary and Main Contributions . . . 240
8.2 Future Research . . . 247
8.3 Concluding Remarks . . . 249

A Appendix 252
A.1 Mathematical Models of hERG Channel Kinetics . . . 252

B Appendix 271
B.1 Further Combinations of Ion Channel Data Use . . . 271

C Appendix 278
C.1 Numerical Methods . . . 278
C.2 Parameter Optimisation . . . 280

Acknowledgements

First and foremost I would like to thank my supervisors, Gary, David and Jim for their guidance and support and for the amount of time they have dedicated to supervising me over the course of my D.Phil. In particular thanks to Gary for all his help, patience, and never-failing optimism! Thanks to David for his enthusiasm, encouragement and insights throughout. Thanks to Jim for enabling me to undertake the experimental aspects of this work and to spend time at GSK. Thank you also to Yi Cui for his input in the early stages of the project.

I would also like to thank Teun de Boer for his valuable technical guidance and insights into the patch clamp technique; Remi Bardenet for helping me to understand MCMC; and Geoff Williams for his patience while extracting the GSK data used in Chapter 4. Thank you also to Gail Robertson for her assistance in acquiring the cell-line used in the experiments described in Chapter 6.

Thank you to the members of Safety Pharmacology at GSK for making me feel so welcome during my visits and for all the help provided. In particular, thanks to Carol Wilson for teaching me the patch clamp technique and Sara Graham, Sam Turner and Kate Harris for their technical assistance and advice.

Thank you to past and present members of the Computational Biology Group for useful discussions, advice and welcome distractions! In particular thanks to Aidan, AJ, Anna J, James Mb, Lorenz, Ollie, Ross and Sara.

Finally, thanks to my Mum and Dad and Nathan and Samuel for their continued support and lots of laughter over the years!

1

Introduction

1.1 Motivation

Unintended drug interactions with cardiac ion channels can cause changes to normal electrical activity in the heart. These changes can lead to the onset of cardiac arrhythmias which can be potentially fatal. Cardiovascular toxicity is the main safety reason for the discontinuation of development and market withdrawal of pharmaceutical compounds (Valentin, 2010; Shah, 2013). Adverse cardiac side effects have been estimated to be responsible for 45% of drug withdrawals and around 30% of general drug attrition (Laverty et al., 2011). The progression of a drug from the initial discovery stages through to being marketed can take up to 15 years and cost 1–2 billion dollars (Morgan et al., 2011). Market withdrawal of a drug not only means a substantial loss of invested time, money and resources, but more importantly, means that a risk to patient safety has been posed. It is therefore important for pharmaceutical companies to identify compounds which may have adverse cardiac side effects as early as possible in the drug development process.

Computational approaches have great potential for assisting in many aspects of drug development. In this thesis our focus is on their use for pro-arrhythmic cardiac safety assessment. *In silico* methods for assessing drug effects on cardiac activity are inexpensive and can be performed at extremely high-throughput in the very earliest stages of drug development. Computational methods are beginning to receive increased attention from pharmaceutical companies and regulatory agencies, with their potential for routine use during drug devel-

opment currently being explored. However, before their adoption for routine use in cardiac safety assessment, any approach must be thoroughly assessed for its predictive capabilities and limitations. This is achieved through validation; this involves evaluating the ability of the computational method to replicate, not only the experimental data used to construct the mathematical model underpinning the approach, but also additional data which was not used to build the model.

It is not possible for a mathematical model to be able to replicate every aspect of a biological system, nor can a model ever be "fully valid", as we cannot assess the ability of a model to replicate every possible experimental scenario. However, to be genuinely useful in cardiac safety assessment, models must demonstrate the ability to predict the outcome of experiments using protocols outside the range of the protocols used to parameterise them. At present this is typically not the case: models are fitted to all of the (often sparse) experimental data available, and then are often used to investigate possible experimental scenarios far away from those on which they were built and tested. In this thesis we will propose a more comprehensive approach to model validation that will make use of a wide range of experimental data and protocols to investigate the predictive ability of our models.

This is of increasing importance since over the past fifty years of development, a very broad range of biophysical models of cardiac electrophysiology have been developed to describe electrical activity in different cardiac cell types and in different species. As computational power has increased, and more biological detail has been uncovered, mathematical models have become increasingly complex with models composed of many equations and parameters. Within a model of a cardiac cell, mathematical representations of different sub-cellular components are typically constructed using experimental data from many different sources; including from different species and a range of experimental conditions (Niederer et al., 2009).

Simulations of the behaviour of both the whole cell and of different sub-cellular compo-

nents represented in different models can be quite distinct. Some differences in behaviour will certainly be attributable to differences in the data used to construct the models. However, we question whether the range of simulation outputs observed from these models are simply representative of the variability in the data to which each model was trained or whether some differences are a consequence of the model construction process itself. If differences are attributed to the model construction methodology then we will also question whether aspects of the behaviour of the models may not be expected to be observed experimentally. And, in light of differences in the behaviour of different models, which model should be used for making predictions?

A number of subtleties of the model construction process may be overlooked when formulating a new mathematical model. Many models have become so complex or have been constructed in such a way that it is mathematically impossible to determine a unique optimal parameterisation of the model to describe observed experimental data. It may also be the case that there is insufficient information in the experimental data to be able to constrain a model, or perhaps the model structure being used to describe the system is inappropriate. Many of these fundamental issues are often not considered when a new mathematical model is proposed, and yet, after often only minimal validation, the most recent (and often most complex) mathematical model can become widely adopted for scientific investigations. As mathematical models become more complex, the model construction process becomes increasingly complicated, and so the resulting models, while most likely being more biologically detailed, may not necessarily be the most representative of the behaviour of the biological system.

If computational approaches are to be of use during the pharmaceutical drug development process, we must be confident that the mathematical models used are able to replicate the results of the range of experiments they are used to simulate, and not only those from which they were constructed. Consequently, there is a great need to ensure that mathematical models used for predictive purposes are accurate and robust, and that the model construction

process is performed and carefully documented in order to achieve this. We believe such considerations will lead to enhanced predictive ability of the resulting electrophysiology models for use within pharmaceutical drug development.

In this thesis we focus on the use of computational modelling of drug-ion channel interactions for the purpose of cardiac safety assessment. Our research indicates the great potential of computational approaches for facilitating cardiac safety assessment, while also illustrating that caution should be taken when using mathematical models for this purpose. We highlight the necessity of careful model construction to ensure that models are accurate and we also describe the need for thorough model validation. We evaluate and thoroughly validate one computational approach for cardiac safety assessment. We identify one possible limitation of this approach as being the representation of drug-ion channel interactions within the mathematical model. We demonstrate the consequence that the choice of model for the cardiac potassium ion channel encoded by hERG (human-Ether-a-go-go Related Gene) can have on predictions of drug effects on cardiac electrical activity. We then focus on the mathematical representation of hERG channel kinetics, since this ion channel interacts strongly with many pharmaceutical compounds. We demonstrate how the different ion channel models exhibit a range of disparate behaviours when used to simulate different experimental protocols.

We focus in depth on the model construction process for determining robust and accurate models of hERG ion channel kinetics. We illustrate the stages undertaken in the process; from the experimental design that tries to ensure that sufficient information is available for selecting and parameterising the most appropriate model structure, to assessing the information content of different experiments for parameterising existing models, to performing the new experiments, and then finally determining and validating models to describe observed experimental data. In Section 1.2 we outline the thesis structure in more detail.

1.2 Thesis Structure

In **Chapter 2** we outline the biological background of the research presented in this thesis. We describe the biological detail underlying electrical activity in cardiac cells. We outline the typical pharmaceutical drug development process; highlighting the potential role of computational modelling in present and future cardiac safety assessment. We also discuss the focus of existing cardiac safety assessment strategies and motivate the research presented in this thesis in an industrial context.

In **Chapter 3** we introduce the mathematical model descriptions of the biological systems discussed in Chapter 2 which are used throughout the thesis. We describe the mathematical representation of cardiac electrical activity from the scale of ion channels, through to the whole cell and cardiac tissue representations. We also outline the history of the development of these mathematical models. We describe how the effects of drug interactions on cardiac electrical activity can be represented in mathematical models. We discuss some existing computational approaches used for assessing cardiac safety which have been developed in collaboration with pharmaceutical companies.

In **Chapter 4** we evaluate a computational approach which integrates the use of data quantifying the extent of drug block of different ion channels within a mathematical cardiac cell model to predict the results of a rabbit-based cardiac safety screen performed within Safety Pharmacology at GlaxoSmithKline (GSK). Validation is performed using an unprecedented number of compounds to add confidence to the evaluation results. We evaluate the method using different combinations of the available data at different stages of the early drug development process. We identify the combination of data with which the approach demonstrates the highest level of predictive ability. We conclude that while a reasonable level of predictive ability is demonstrated, higher levels would be required before such an approach could be considered as a complete replacement for animal-based cardiac safety testing.

We identify a possible limitation of the computational approach presented in Chapter 4 as being the lack of the representation of drug-ion channel kinetic interactions which may be important for predicting subsequent drug side effects on cardiac electrical activity. Therefore, in the remainder of the thesis we focus on the development of mathematical models describing the kinetics of the cardiac potassium channel hERG, with which a diverse range of drugs have been found to interact strongly.

In **Chapter 5** we describe the design of novel experiments to provide better information for determining the kinetics of the hERG channel. We explore the behaviour of existing hERG models when used to simulate the same experiments. We assess our ability to determine parameters of existing hERG channel models when using data from different combinations of our new experimental protocols. This analysis, based upon simulated data with added noise, then gives an indication of the information content within each of the different experiments for determining parameters of existing models. We use a Bayesian inference approach to determine distributions of estimates for each parameter, providing an indication of parameter uncertainty. We compare the information contained within our new protocols with existing standard and non-standard protocols in the literature. We conclude that the protocols we have designed are theoretically appropriate for determining mathematical models of hERG channel kinetics.

In **Chapter 6** we describe the techniques that we have used to perform the voltage protocols presented in Chapter 5 in patch clamp experiments. We outline the methods and necessary additional considerations required due to the unusual nature of the new protocols. We present our recorded experimental data, and highlight some interesting observations made while collecting the data, in particular focusing on variability between recordings.

In **Chapter 7** we use the experimental data presented in Chapter 6 and the theoretical analysis from Chapter 5 to determine and validate cell-specific models of hERG channel kinetics. We carefully pursue the model construction process, attempting to select the most appro-

priate model structure and parameterise models in such a way as to attempt to ensure that the optimal solution to the parameter optimisation problem is identified. We describe in detail the necessity, and also the challenges encountered when attempting to construct models in this way. We assess qualitatively the variability observed between experiments performed on different cells and compare the parameter estimate distributions obtained when using data from different experiments performed on the same cell.

In **Chapter 8** we provide an overview of the research presented in this thesis. We highlight the main strengths and contributions of this work, as well as the key limitations. We outline numerous potential applications and extensions of the research presented in this thesis, detailing directions for future work.

We provide the mathematical descriptions of the existing hERG channel models considered in **Appendix A**. In **Appendix B** we provide additional results from the evaluation of the *in silico* cardiac safety screen which are not included in Chapter 4. In **Appendix C** we provide further detail on the numerical methods and algorithms used to perform the simulations described in this thesis.

At a time when the nature of cardiac safety assessment for pharmaceutical compounds is evolving, with an increased focus being placed on *in silico* and novel *in vitro* methodologies, the research presented in this thesis explores a number of key issues implicated in the use of computational approaches for cardiac safety assessment. The key message of this thesis is that computational modelling provides great promise for facilitating drug development; however, mathematical models must be carefully constructed and adequately validated for these approaches to achieve maximal success and uptake into pharmaceutical drug development.

2

Cardiac Physiology and Drug Safety Assessment

This chapter introduces key concepts of cardiac anatomy and electrophysiology we will refer to throughout this thesis. We describe drug-induced cardiac side effects and indicate biomarkers which can be used for detecting their incidence. We also outline existing and proposed future guidelines for cardiac safety assessment of new pharmaceutical compounds.

2.1 Cardiac Structure and Function

The heart is a muscle which pumps blood around the body. It is one of the most important organs due to its role in transporting essential supplies of oxygen, glucose and amino acids. Pumping of the blood is achieved though contraction of cardiac muscle cells. Contractions are regulated by electrical impulses propagating through the heart. A diagram of the anatomical structure of the heart is shown in Figure 2.1.

The heart is divided into the left and right sections, each of which has an upper compartment called the *atrium* and a lower compartment called the *ventricle*. Deoxygenated blood flows into the heart through the *vena cava*. Blood first enters the right atrium and then moves into the right ventricle. When the right ventricle contracts, blood is pumped to the lungs, via the *pulmonary artery*, where it is re-oxygenated. Oxygenated blood flows into the left atrium, via the *pulmonary vein*, and passes into the left ventricle. When the left ventricle contracts, oxygenated blood is pumped around the body via the *aorta*. The left side of the

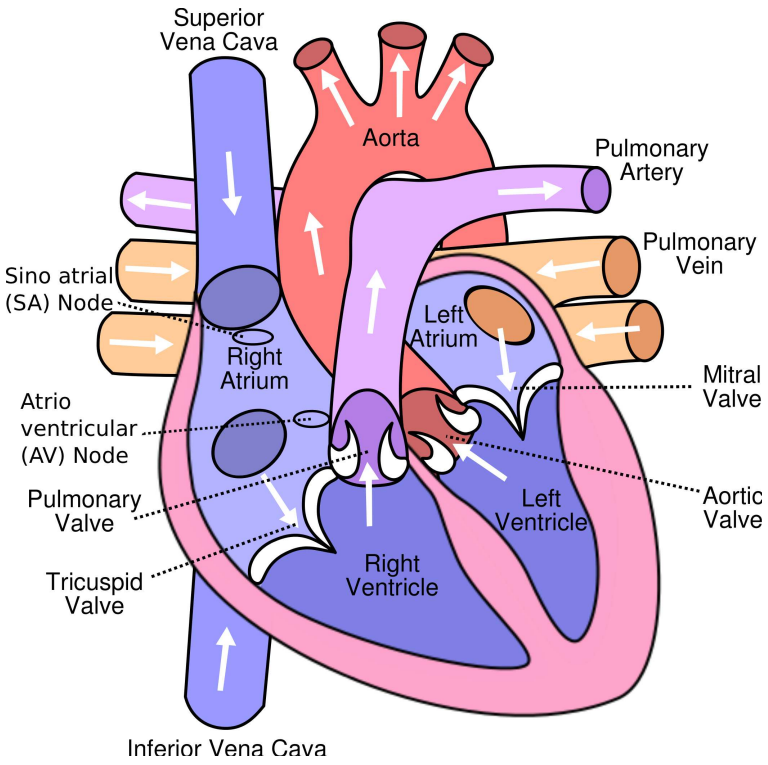


Figure 2.1: Cross-sectional view of the heart. Key components we refer to in this section are indicated. Arrows indicate the direction of blood flow. Figure adapted from *http://commons.wikimedia.org/wiki/File:Diagram_of_the_human_heart.svg*.

heart is thicker than the right side to enable it to undertake this function. For a more thorough review, see (Katz, 2010).

The heart wall is composed of three layers: the *epicardium*, the *endocardium* and the *myocardium*. The myocardium is a thick muscular layer between the endocardium and epicardium. It consists mainly of cardiac muscle cells, called *cardiomyocytes*, and fibroblasts. The epicardium is a thick outer layer of connective tissue and fat. It contains blood vessels, lymphatic tissue and nerve fibres and protects and supplies nutrients to the heart. The myocardial cells nearest the epicardium are termed sub-epicardial cells. The endocardium is a smooth inner layer of the heart which provides a non-adherent surface for blood collection and pumping. There is a layer of connective tissue which contains myocardial cells nearest the endocardium, which are termed sub-endocardial cells. *Purkinje fibres* are also contained within the endocardium. Purkinje fibres are specialised cardiac muscle fibres which facilitate electrical conduction (Katz, 2010).

Cardiomyocytes conduct electrical impulses to coordinate contraction and enable pumping of the heart. Cardiomyocytes are arranged side by side in the myocardium, forming myocardial sheets. This arrangement leads to both electrical and mechanical interaction between cells. Consequently, each heart beat is an electrically and mechanically coupled event (Kohl et al., 2011). Differences in the physiological properties of the cell types which comprise the cardiac wall contribute to the heterogeneity in electrical conduction through the heart.

Electrical impulses propagating through the heart are initiated by self-exciting pacemaker cells located in the *sinoatrial node* at the top of the right atrium. An impulse propagates through the atrium to the *atrioventricular (AV) node*. Electrical stimuli pass from one cardiomyocyte to the next via *gap junctions*. The *bundle of His* conducts the electrical impulse into the ventricle via the *left* and *right bundle branches*, and then through the network of Purkinje fibres. This allows the impulse to propagate to the endocardial surface of each ventricle. A wave of electrical excitation travels from the endocardial surface, through the myocardium and on to the epicardial surface. Cardiac muscle cells contract as a result of this electrical stimulation (Katz, 2010). The sequence of events involved in the propagation of electrical impulses through the heart is illustrated in Figure 2.2.

Electrical activity occurring in the heart over time can be recorded in an *electrocardiogram (ECG)*. The form of typical healthy ECG signals is shown in Figure 2.2. The purple regions highlighted on the ECG trace to the right of each diagram of the heart in Figure 2.2 show the corresponding recording of electrical activity in response to electrical stimulation in the cardiac region highlighted.

2.2 Cardiac Electrophysiology

As described in Section 2.1, propagation of electrical impulses through the heart is achieved through successive electrical stimulation of neighbouring cardiomyocytes. The energy en-

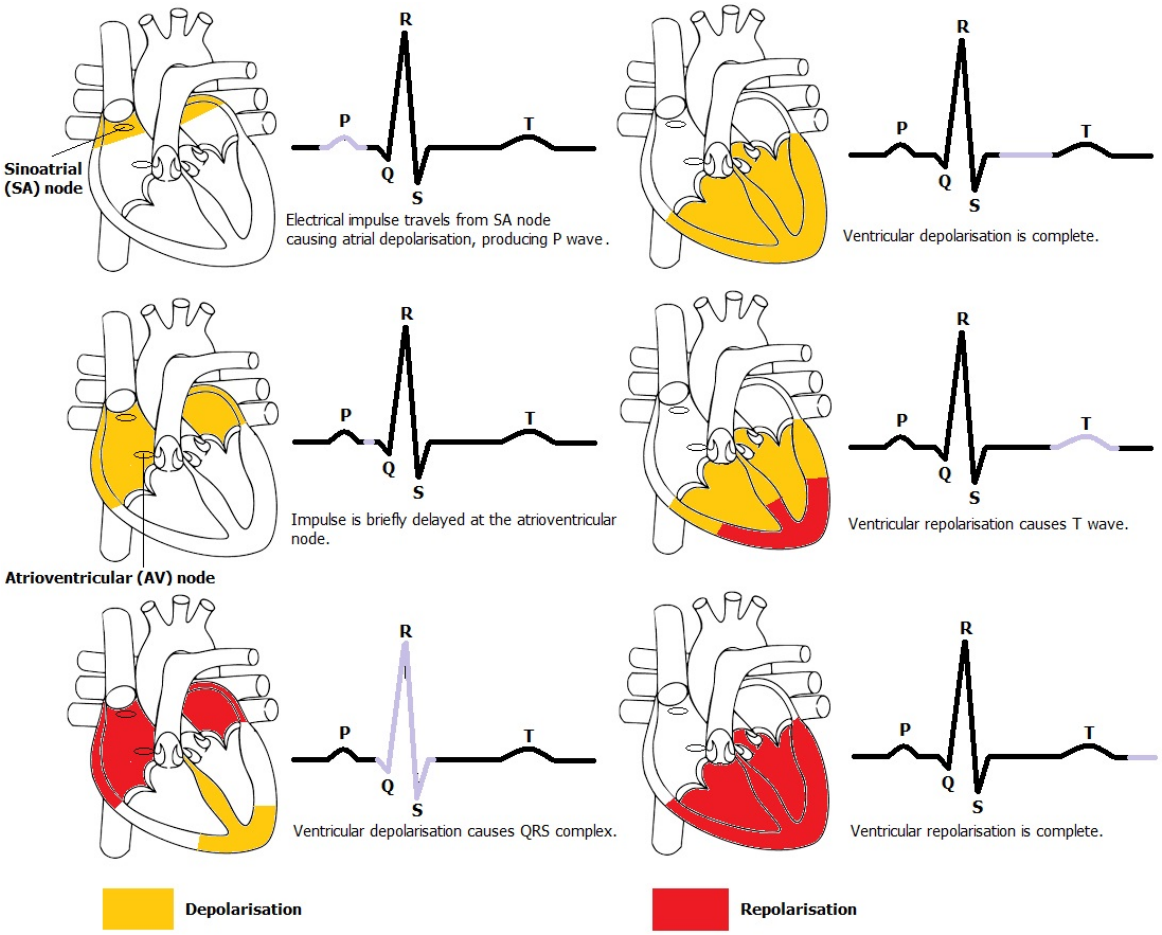


Figure 2.2: Schematic showing the sequence of events occurring during generation and propagation of an electrical impulse through the heart. For each stimulation event, depolarisation (increasing voltage) is highlighted in yellow and repolarisation (decreasing voltage) is highlighted in red in the diagrams of the heart. The corresponding contribution to the ECG recording is highlighted in purple in the ECG trace to the right. The diagram of the heart used in this figure is adapted from *http://commons.wikimedia.org/wiki/File:Heart.svg*.

abling cardiac cells to propagate rapidly electrical signals is created due to maintained electrochemical gradients across the cell membrane both from differences in charge and in ionic concentrations. Ionic concentration differences induce the diffusion of ions from higher to lower concentrations. The flow of ions across the cell membrane is facilitated by *pumps* and *exchangers*. In addition, a third type of membrane protein, *ion channels*, facilitate the passive diffusion of ions. The resulting potential difference across the cell membrane is called the *membrane potential*. Some of these proteins in the cell membrane are voltage-sensitive, and are activated by changes in membrane potential, which enable the movement of ions into, or out of, the cell across the cell membrane. The flux of three key types of ion across the cell

membrane play an important role in determining the membrane potential. These are potassium (K^+), sodium (Na^+) and calcium (Ca^{2+}) ions; chloride (Cl^-) ions are also involved to a lesser extent (Grant, 2009).

2.2.1 Cardiac Action Potential

In response to electrical stimulation, the membrane potential of a cardiomyocyte rapidly increases (*depolarisation*). This change in membrane potential gives rise to the movement of ions across the cell membrane via the voltage-sensitive ion channels, pumps and exchangers described above. This movement of ions across the cell membrane generates membrane currents which further adjust the electrical state of the cell, until the membrane potential decreases (*repolarisation*) and the cell returns to its resting state. The change in membrane voltage of a single cell over time is described as an *action potential* (Carmeliet and Vereecke, 2001). A schematic of the form of a typical action potential, with the contribution of each ionic current to the electrical state of the cell over time, is shown in Figure 2.3.

In the remainder of this section we describe and explain the mechanism of events which take place during a cardiac action potential. In its resting state, a cardiomyocyte has a membrane potential of around -85 mV, termed the *resting potential*, although the value and range varies between species and cell-type. The potential is negative due to the overall concentration of positively charged ions being greater outside the cell than inside. Table 2.1 details the typical extracellular and intracellular concentrations of key ions in mammalian cardiac cells (Molleman, 2003).

Table 2.1 Table of typical intracellular and extracellular ion concentrations in mammalian cardiac cells (Molleman, 2003).

Ion	**Intracellular Concentration Range (mM)**	**Extracellular Concentration Range (mM)**
K^+	130 – 160	4 – 8
Na^+	5 – 20	130 – 160
Ca^{2+}	$5x10^{-4}$ – 10^{-3}	1.2 – 4
Cl^-	1 – 20	100 – 140

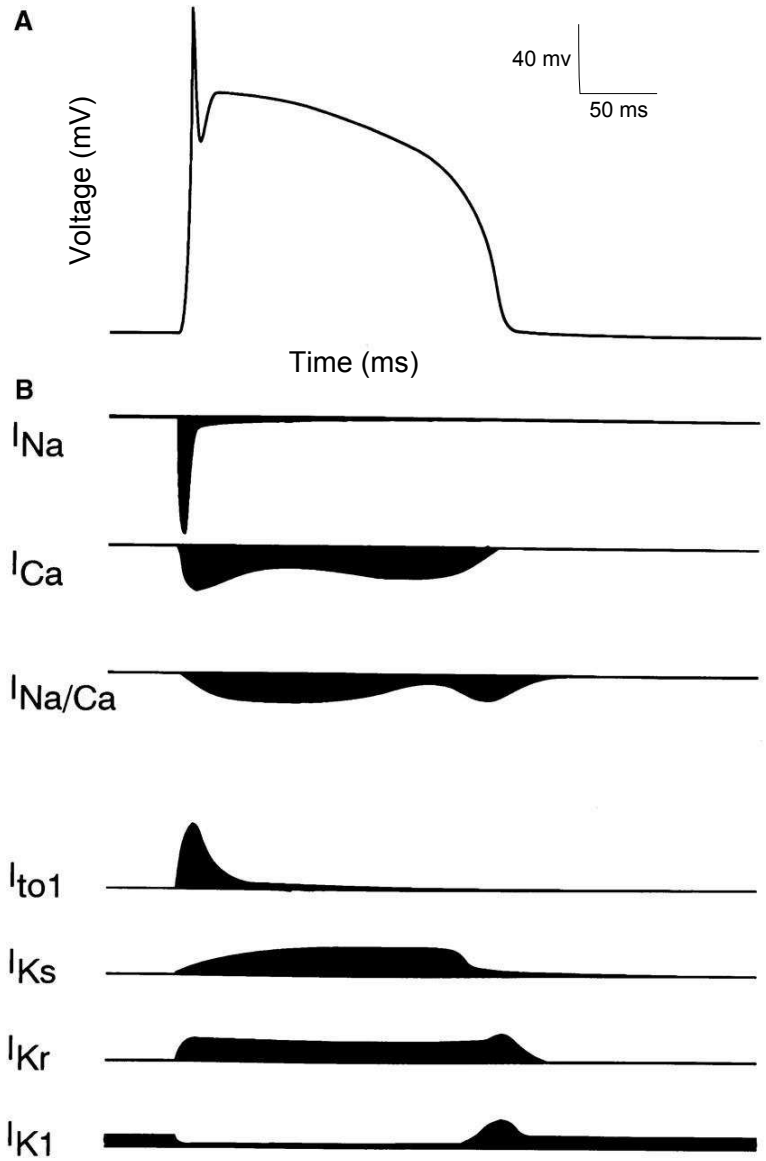


Figure 2.3: Schematic of typical ventricular action potential describing the change in membrane voltage over time is shown in A. In B, the contributions of the key ionic currents involved in generation of the action potential are indicated, with the role of each current described in the text. Downward traces indicate current flowing into the cell and upward traces indicate that the current is flowing out of the cell. Note that the relative current densities of the ion channels are not drawn to scale. Figure adapted from Näbauer and Kääb (1998).

The intracellular concentration of K^+ ions is higher than the extracellular concentration. However, rather than K^+ ions moving down the potassium concentration gradient to the outside of the cell, the negative resting potential of the cell allows for an influx of K^+ ions. The *sodium-potassium exchange pump (I_{NaK})* maintains the potassium concentration gradient

and the *inward rectifying potassium current (I_{K1})* maintains the negative resting potential.

Electrical stimulation causes slight depolarisation of the cell which provokes fast opening of Na^+ channels, allowing sodium ions to flow rapidly into the cell. This further depolarises the cell and the *fast sodium current (I_{Na})* is generated. The rapid rise in membrane potential gives rise to the *upstroke* of the action potential, at the end of which the cell reaches its maximum level of depolarisation (and is at its most positive membrane voltage).

Na^+ channels then become inactivated and the *transient outward current (I_{to})* is activated, causing potassium ions to leave the cell and the peak of depolarisation to be exhibited. This results in a brief repolarisation of the cell. This lower membrane potential is sustained during a *plateau* of the action potential, where a balance between the inward movement of Ca^{2+} ions through the *L-type calcium current (I_{CaL})* channel and the outward movement of potassium ions through the *slow-delayed rectifier potassium current (I_{Ks})* channel is maintained. The I_{NaK} and *sodium-calcium exchanger current (I_{NaCa})* are also involved in maintaining the membrane potential plateau, with the latter playing an important role in removing Ca^{2+} ions from the cell.

The cell repolarises more rapidly as the I_{CaL} channels close and the I_{Ks} channels remain open. This results in a flow of current out of the cell and a subsequent decrease in membrane potential, which triggers the opening of the *rapid delayed rectifier current (I_{Kr})* and I_{K1}-producing channels. The resulting outward, positive current causes the cell to repolarise until it returns to its resting potential. Once the resting potential is reached, the channels that produce I_{Kr} and I_{Ks} close while the I_{K1} channels remain open maintaining the negative resting potential. The cell is then ready to receive electrical stimulation again and for a new action potential to begin (Zaza and Rosen, 2000; Sigg et al., 2010).

2.2.2 Voltage-gated Cardiac Ion Channels

As described in Section 2.2.1, the generation of a cardiac action potential is achieved through a balance of ionic currents flowing into and out of the cell. *Voltage-gated ion channels* in the cell membrane facilitate the regulation of ionic currents. Some channels allow a number of different types of ions to pass while others only allow specific ions to flow through. Changes in membrane potential cause conformational changes in voltage-gated ion channel proteins. All the ion channels we discuss in this thesis are voltage-gated, that is their permeability changes in response to voltage across the cell. Different patterns of voltage changes can be applied to ion channels and behaviour alters accordingly. This is of particular importance in the latter part of this thesis.

A conformational change of a voltage-gated channel in response to a change in voltage can result in the opening (*activation*) of a channel, allowing ions to flow into and out of the cell, or the closing of the channel (by *deactivation* or *inactivation*), preventing the flow of ions across the cell membrane. Hence, ion channels are usually considered to be found in three distinct conformational states: *open*, *closed* or *inactivated*. The open state is the only state in which current is conducted, while inactivated and closed/deactivated states are non-conducting. When the channel is in the inactivated state, there is a brief block of the conduction of ions, during which no current can flow, followed by recovery and resumed current flow through the channel (Aidley, 1996). Ion channel inactivation differs from deactivation as during inactivation the channel closes in the presence of a depolarising voltage, whereas channels deactivate in response to repolarisation. A channel recovers from being in an inactivated state by a period of repolarisation of the cell membrane, and this recovery is necessary for a channel to be able to open again and continue to conduct ions (Hille et al., 2001). One particular voltage-gated cardiac ion channel which is a focus of this thesis, is the *human Ether-a-go-go Related Gene (hERG)* channel. The arrangement of the channel proteins in the different possible conformational states is illustrated in Figure 2.4 (A). In Figure 2.4 ((B) and (C)) we see examples of hERG channel currents recorded in response to

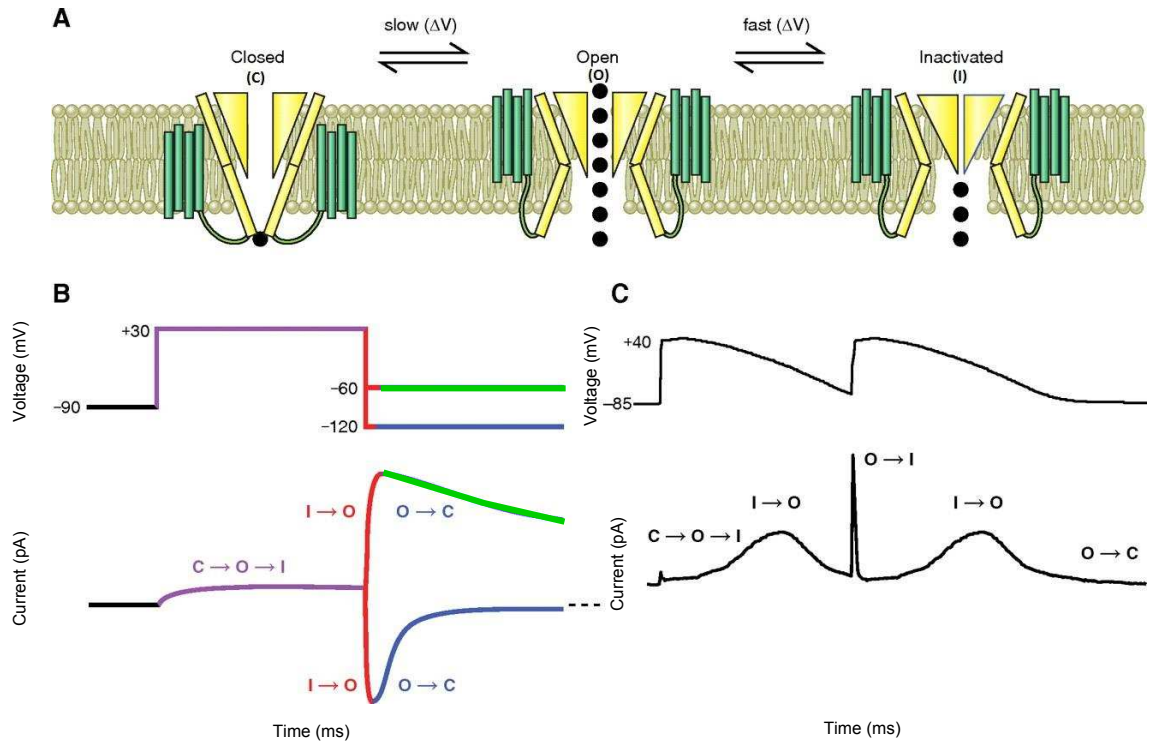


Figure 2.4: Schematic illustrating different conformational states of the hERG channel during gating (A) and the transitions between states occurring, and resulting current flow in response to two different voltage protocols in B and C. Each conformational state diagram (in A) shows two α subunits of the hERG channel as illustrated in Figure 2.5. Figure adapted from Vandenberg et al. (2012).

different patterns of voltage changes.

The hERG gene codes for the protein Kv11.1, which is the primary subunit of the rapid delayed rectifier potassium channel which carries the I_{Kr} current. Four such subunits, termed α subunits form the pore of the I_{Kr} channel. A second type of subunit, termed the β subunit has been identified and is thought to interact with α subunits and to modulate the activity of the I_{Kr} channel (Abbott et al., 1999). The expression of β subunits in ventricular tissue has been found to be low (Pourrier et al., 2003; Lundquist et al., 2005) and a comparison of the currents from hERG channels consisting of just α subunits with those including both α and β subunits, showed that both types of channel were able to recapitulate currents from I_{Kr} channels (Weerapura et al., 2002). Therefore, despite differences observed between currents generated from channels consisting of four α subunits and those from native I_{Kr} channels, the study of Kv11.1 channel currents is well accepted as a representative model of the I_{Kr}

channel. Consequently, the terms 'hERG channel' and 'I_{Kr} channel' are often referred to interchangeably. We will therefore refer to the ion channel carrying the I_{Kr} current, whose primary subunit is encoded by the hERG gene, as the 'hERG channel' for simplicity.

We now focus on the hERG channel to further describe the structure and function of voltage-gated ion channels. The hERG channel has not yet been characterised in full by x-ray crystallography and so a detailed description of its atomic structure is not yet known. However, information about its structure has been inferred by analogy with other ion channels in the same family and from different species, computer modelling, and a range of pharmacological studies (Vandenberg et al., 2012).

Each α subunit of the hERG channel consists of six transmembrane alpha helices which are labelled S1-S6. A schematic of the structure of each subunit and their arrangement in forming the hERG channel is shown in Figure 2.5.

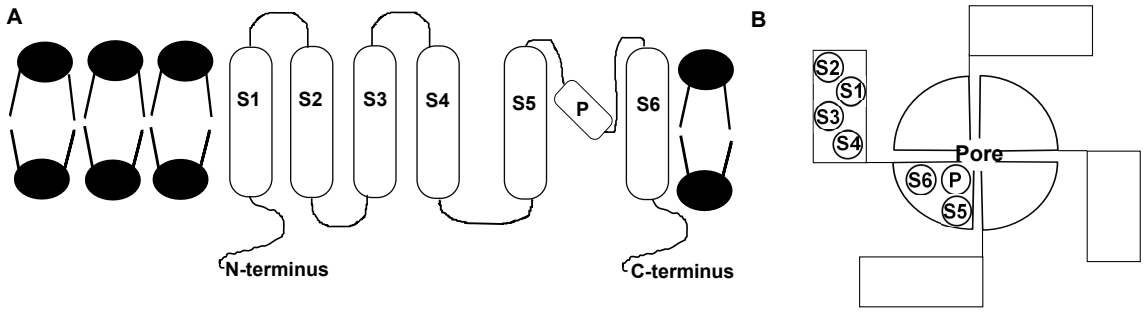


Figure 2.5: Schematic of the structure of the hERG channel and the α subunits of which it is formed. A shows the composition of a subunit of the hERG channel, with the alpha helices of which it is formed labelled S1-6. P represents a pore helix situated between S5 and S6. The subunits are situated within the cell membrane as indicated in the schematic. B is an extracellular view of the hERG channel, showing the arrangement of the four subunits around the central pore. The position of the alpha helices labelled S1-6 in A are indicated in one of the subunits in B. Diagram based on that in Perrin et al. (2008b).

The alpha helix labelled S4 in Figure 2.5 acts as the primary voltage sensor for the channel. It is composed of positively charged amino acids separated by pairs of hydrophobic residues, enabling it to function in this capacity (Sanguinetti and Tristani-Firouzi, 2006).

The description of the hERG channel given here, and as illustrated in Figure 2.5 is for one particular isoform of hERG — the hERG 1a isoform. It was long thought that the I_{Kr} channel was composed of four α subunits encoded by the 1a transcript of the hERG gene. However, a second isoform of hERG has been discovered, termed the hERG 1b isoform. hERG 1b isoforms are identical to the hERG 1a isoforms in the transmembrane and cytoplasmic regions, but the 1b subunit is composed of fewer amino acids, being truncated in the N-terminus. It has been suggested that I_{Kr} channels present in native cardiomyocytes may be composed of combinations of hERG 1a and hERG 1b α subunits (Jones et al., 2004; Sale et al., 2008), which can exhibit different kinetic properties to channels composed of hERG 1a subunits alone (Sale et al., 2008), as well as differences in drug binding affinities (Abi-Gerges et al., 2011). However, in this thesis we focus solely on hERG 1a α subunits, and all subsequent reference to hERG subunits refers to those of the 1a isoform.

Transitions between the open, closed and inactivated states are voltage-dependent in the hERG channel. With an increase in voltage from a negative resting potential, the channel moves from the closed to the open state, but as the voltage becomes increasingly positive, the channel becomes inactivated, restricting the flow of ions. With a decrease in voltage, the channel moves from the inactivated state to the open state, before closing again as the voltage returns to its negative resting potential. An example of this behaviour is illustrated in Figure 2.4 (B). hERG channel activation and deactivation kinetics are slow (of the order of hundreds of milliseconds to seconds) and inactivation and recovery from inactivation are rapid (of the order of milliseconds to tens of milliseconds (Wang et al., 1997; Zhou et al., 1998)). This property makes hERG channel kinetics rather unusual and difficult to characterise. These unusual kinetic properties also underlie the role that hERG plays in cardiac repolarisation (Spector et al., 1996).

In the laboratory, ion channel kinetics can be investigated using the *patch clamp* technique. The voltage across the cell membrane can be controlled while the current that must be sup-

plied to maintain a particular voltage level is recorded. The patterns of voltage changes typically used for such investigations are *voltage-step* protocols. During voltage-step protocols, a particular voltage level is held for a period of time before stepping to a new voltage level. An example of a voltage-step protocol is shown in Figure 2.4 (B), and further examples, typical of those used in patch clamp experiments, are shown in Figure 3.1. The patch clamp technique is described in detail in Chapter 6.

2.3 Drug-Induced Changes to Cardiac Electrical Activity

Pharmaceutical compounds can interact with cardiac ion channels and cause disturbances to normal electrical activity in the heart. Compounds can physically obstruct the ion channel pore or can induce conformational changes in the channel through binding at sites other than the pore. Both types of drug action reduce the flow of ions through the channel, disturbing the normal ionic balance, and consequently altering the electrical activity in cardiomyocytes, changing the heart's susceptibility to arrhythmias (Brennan et al., 2009).

A number of theories have been proposed to describe the mechanism of compound interaction with ion channels. These can be broadly classified into three types: *simple pore block*, *state and voltage-dependent binding theories* (*modulated* and *guarded receptor* theories being examples); and drugs that are thought to act as *allosteric effectors* (Brennan et al., 2009).

In simple pore block theory, the drug molecule is assumed to physically restrict the flow of ions through the channel by binding to, and obstructing, the ion channel pore. The affinity of a drug for a particular ion channel is described by an IC_{50} *value*, which is the concentration at which a 50% reduction of the peak ion channel current is observed, and the *Hill coefficient*, which physically can be interpreted as the number of drug molecules required to block one ion channel. Simple pore block theory assumes that the drug has continuous access to the ion channel pore and that the affinity of the drug molecule for the ion channel is independent of time and membrane voltage (Starmer et al., 1994).

Limitations of simple pore block theory are that drug block is assumed to be immediately in steady state and it does not take into account that drug-ion channel interaction will also be affected by the voltage-dependent behaviour of the channel. State and voltage-dependent drug binding theories attempt to address these limitations. Due to the voltage-dependence of ion channel conformational changes, while voltage and state-dependent binding can occur independently, it is difficult to distinguish between state and voltage-dependence of drug block (Brennan et al., 2009) and so both are often considered simultaneously. Both the modulated and guarded receptor theories assume that drug binding and unbinding to the channel depends on membrane voltage and time. The modulated receptor theory, proposed by Hille (1977), assumes that a drug can bind to the channel when it is in any of the open, closed or inactivated states, but that the binding affinity is different for each state. The guarded receptor theory proposed by Starmer et al. (1984) assumes that a drug binds only when the channel is in specific conformational states, and does so with a fixed affinity. In this case, binding is restricted only by voltage and time-dependent conformation changes of the channel.

The third type of drug binding assumption is that the drug acts as an allosteric effector. Drugs with this mechanism of action are assumed not to block the ion channel directly, but to alter the dynamics of its conformational state changes through binding at a site other than its pore (Brennan et al., 2009).

A wide range of structurally and functionally distinct compounds have been found to have a strong affinity for binding with the hERG channel, most often as an unintended target. It is believed that compounds bind with the hERG channel via the central pore (Spector et al., 1996; Mitcheson et al., 2000) and the hERG channel's promiscuity has been attributed to the large, hydrophobic cavity in its pore. Most compounds block the hERG channel when it is in the open or inactivated states (Ficker et al., 1998) and compounds can only access the binding site in the central pore after the channel has opened (Sanguinetti and Mitcheson, 2005). However, it has not been established whether the open or inactivated states are preferred for

compound binding (Perrin et al., 2008a).

A single compound can interact with a number of different ion channels, some intentionally and others unintentionally. These interactions change the fluxes of different ions across the cell membrane, which may cause changes in cardiac electrical activity in individual cells, and subsequently in cardiac tissue. Alterations to normal cardiac electrical activity, resulting in a disturbed heart beat rhythm, can lead to the onset of potentially fatal cardiac arrhythmias. One such arrhythmia, which has been linked with drug block of the hERG channel is *Torsades de Pointes (TdP)*, which exhibits distinctive characteristics on the ECG, the typical form of which was shown in Figure 2.2. Changes in ECG morphology are used to identify drug-induced changes to cardiac electrical activity. Similarly, in individual cardiomyocytes, changes to the shape of the action potential are used to identify adverse drug effects. The *QT interval* on the ECG and the duration of the action potential in a single cell (*Action Potential Duration (APD)*) are two commonly used biomarkers for the assessment of drug-induced changes to cardiac electrical activity. Figure 2.6 shows a schematic of the ECG and action potential, with the QT interval and APD biomarkers indicated. The correspondence between electrical activity recorded in the action potential and on the ECG is also shown.

2.4 Pharmaceutical Cardiac Safety Assessment

Cardiovascular toxicity remains the main safety reason for the discontinuation of development and withdrawal of drugs from the market (Valentin, 2010). 21.4% of the drugs withdrawn from the market between 1990 and early 2012 were as a result of evidence of QT prolongation and/or association with TdP (Shah, 2013). To address the evident association between hERG channel block, QT prolongation and increased incidence of pro-arrhythmic risk (in particular of TdP), in 2005 two guidance documents were formulated by the International Conference for Harmonisation (ICH); the ICH-S7B[1] and the ICH-E14[2], for the assessment of the cardiac liability posed by new drugs. The ICH-S7B guidelines suggest

[1] http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Safety/S7B/Step4/S7B_Guideline.pdf
[2] http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E14/E14_Guideline.pdf

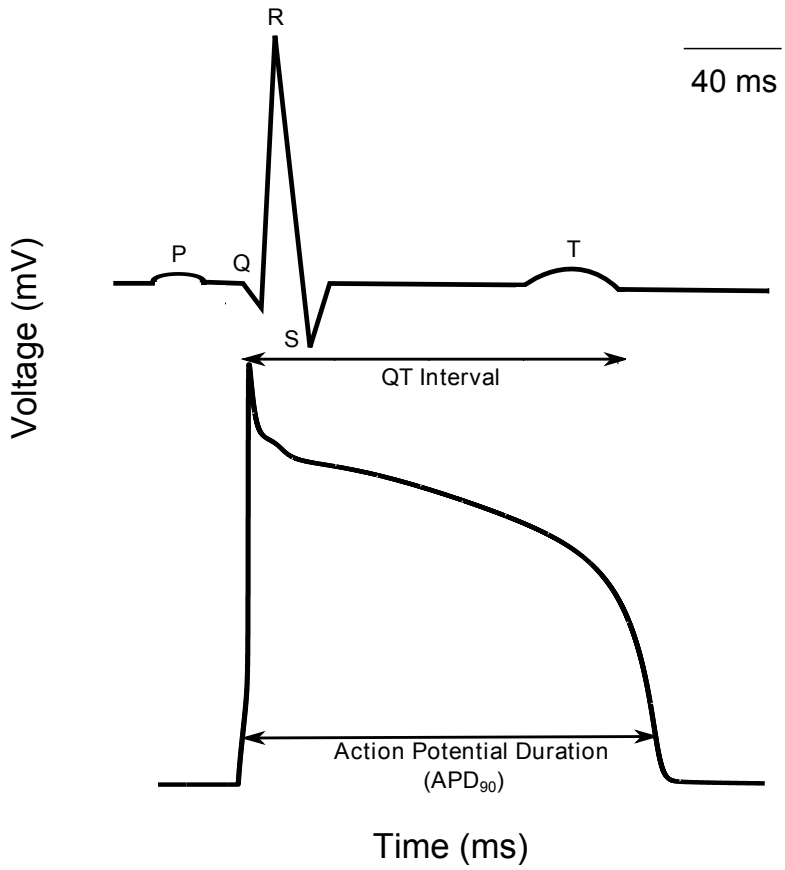


Figure 2.6: Schematic of a typical healthy ECG (top) and action potential (bottom). The P, Q, R, S and T regions of the ECG are labelled and the QT interval (the time from the start of electrical stimulation (at Q) to the return of the ECG to the resting state (at the end of the T wave)) is indicated. On the diagram of the action potential, the Action Potential Duration (APD) is shown. The APD_{90} corresponds to the time between points on the action potential at which the membrane voltage is at 90% of the difference between its minimum and maximum. The assumed correspondence between the QT interval and APD_{90} value is indicated.

that an *in vitro* assay assessing hERG channel block by compounds, and an *in vivo* study investigating drug-induced changes to the QT interval should be performed as standard during drug development. The ICH-E14 document suggests that a human thorough-QT study is performed during the clinical stages of development, which assesses the effect of a drug on the QT interval in humans with the aim of investigating potential pro-arrhythmic liability.

As a consequence of the publication of the ICH guidelines, pharmaceutical companies have developed a range of preclinical assays to evaluate the potential pro-arrhythmic liability of new drug compounds (Pollard et al., 2010). A typical cardiac safety assessment strategy

could take the form of: initial Quantitative Structure Activity Relationship (QSAR) studies to infer properties of biological activity based on chemical structure (described in more detail in Chapter 3.5); ion channel screening using a high-throughput platform; an *in vitro* study, such as measuring the effect of the drug on the QT interval in rabbit cardiac tissue; and an *in vivo* study, such as monitoring the QT interval in a conscious dog telemetry study. A typical cardiac safety assessment pipeline for pharmaceutical compounds is shown in Figure 2.7.

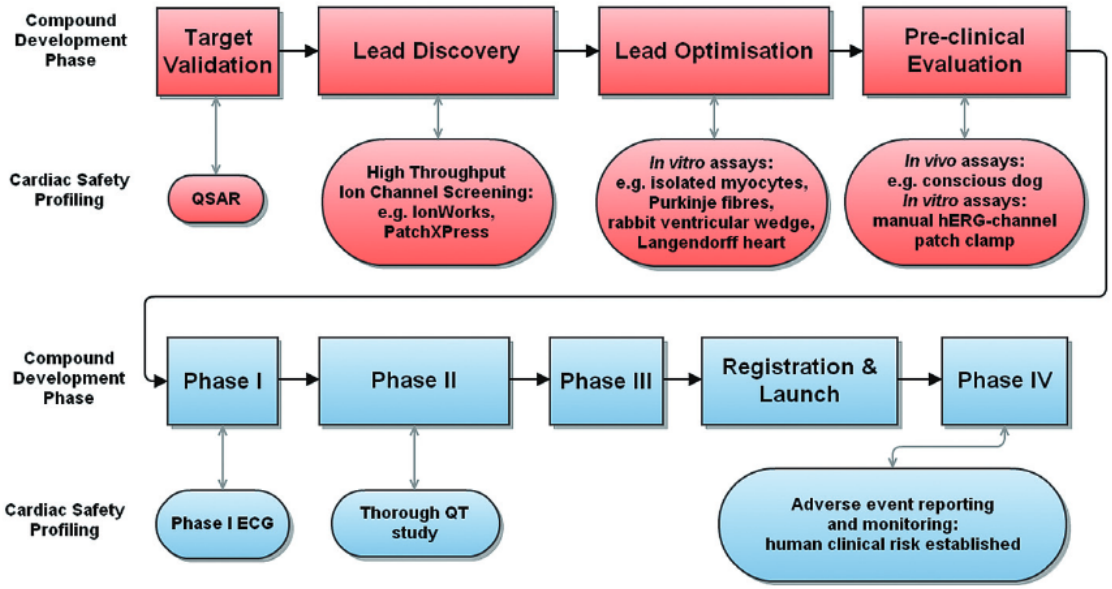


Figure 2.7: Outline of typical cardiovascular safety assessment strategies performed during each stage of pharmaceutical compound development. Stages coloured in red are performed during preclinical development and those in blue during clinical development. Figure taken from Mirams et al. (2012).

The cardiac safety assessment strategies adopted in response to the ICH guidelines have been extremely effective, with no drugs withdrawn from the market due to the occurrence of TdP since their introduction. However, it has been noted that the present approach does not directly assess the main clinical safety concern, which is detecting pro-arrhythmic risk (Sager et al., 2014). The focus of the ICH guidelines on hERG channel block and QT prolongation, which are highly sensitive, but not very specific markers of pro-arrhythmic risk, may have led to the unnecessary discontinuation of development of safe compounds (Sager et al., 2014). hERG channel block alone has been responsible for the discontinuation of as

many as 60% of potential new drugs (De Ponti, 2008) and there are many examples of drugs (such as ranolazine, phenobarbital, and tolterodine (Sager et al., 2014)) which prolong the QT interval (often via block of the hERG channel) at therapeutic concentrations, but are not necessarily pro-arrhythmic.

A recent proposal led by the US Food and Drug Administration (FDA) aims to revise the existing ICH guidelines for cardiac safety assessment. The aim is to reduce the present reliance upon the pro-arrhythmia biomarkers of hERG channel block and QT prolongation in favour of gaining a more mechanistic understanding of the risk of cardiac arrhythmias. The proposed paradigm, entitled the *Comprehensive in vitro Pro-arrhythmia Assay (CiPA)*, aims to combine a new range of pre-clinical assays for the assessment of pro-arrhythmic risk. It is currently anticipated that this will include investigations of compound-ion channel interactions using over-expressed human cardiac ion channels, studies using stem cell derived cardiomyocytes, and *in silico* approaches for modelling drug-induced effects on cardiac electrophysiology. The *in silico* methods are envisaged to integrate compound effects on multiple cardiac ionic currents to reconstruct the ventricular action potential to evaluate pro-arrhythmic risk (Sager et al., 2014). It is hoped that a new combination of recommended assays for cardiac safety assessment will lead to improved identification of drugs posing cardiac safety liabilities and may help to reduce the levels of attrition associated with the existing safety assessment strategies, ultimately enabling more safe drugs to reach the market.

A main initial focus of the CiPA effort, particularly with regard to the use of *in silico* methods, is to assess both the suitability of the use of, and to validate, potential predictive mathematical models. A second question being addressed is whether the mathematical representations of drug ion channel-interactions in existing models of cardiac electrical activity are sufficient. The research presented in this thesis relates to both of these questions; we thoroughly validate one computational approach which integrates compound effects on multiple ion channels to predict the effects on cardiac electrical activity observed in a rabbit-based cardiac safety test. We then explore the mathematical representations of the hERG channel

within models of cardiac electrophysiology and demonstrate the importance of an accurate representation of the behaviour of this channel within cardiac action potential models when predicting drug-effects. We propose and perform novel experiments for determining mathematical models of this channel and illustrate approaches both for assessing the information content of different experiments used for this purpose and for constructing cell-specific mathematical models of hERG channel kinetics.

2.5 Summary

In this chapter we have provided an overview of the structure and function of the heart from the sub-cellular ion channel level through to the whole organ level. We have described the mechanism of electrical activity generation and propagation through the heart, aspects of which we model mathematically in this thesis. We have described how drugs can interact with cardiac ion channels to cause changes to normal electrical activity in the heart. We have reviewed existing strategies for assessing the cardiac liability of drugs during pharmaceutical development and have outlined recent proposals for reforming the existing guidelines.

In Chapter 3 we describe how electrical activity in the heart can be modelled mathematically, through analogy of the electrical activity in a cardiomyocyte to that of an electrical circuit. We outline the physical foundations of this modelling approach and describe the development of the mathematical models of cardiac electrophysiology used. We describe the modelling approaches adopted in this thesis, ranging from those describing the activity of ion channels, to models of the electrophysiology of a single cardiomyocyte and through to models describing an electrical impulse propagating through cardiac tissue. We describe how drug-ion channel interactions can be represented in these models. Finally we elaborate on how such approaches have begun to be explored recently within pharmaceutical companies for the purpose of cardiac pro-arrhythmic safety assessment.

3

Mathematical Modelling of Cardiac Electrophysiology

This chapter introduces the mathematical modelling approaches used to represent different aspects of cardiac electrophysiology in this thesis. The mathematical models used range from descriptions of ion channel kinetics, to representations of the electrophysiology of a ventricular cardiomyocyte, through to how such models are coupled together to describe electrical propagation in cardiac tissue. In addition to describing the mathematical concepts underpinning the work, we also endeavour to provide a brief historical overview of the development of these models. We describe approaches used for mathematically representing drug block of ion channels and we review some recent computational approaches, developed in collaboration with pharmaceutical companies, for cardiac safety assessment of drug compounds.

3.1 Mathematical Models of Ion Channel Kinetics

As described in Chapter 2, electrical activity in the heart is generated and regulated by the flow of ionic currents. Ion channels open and close to control the flow of specific ions to ensure normal electrical activity in the heart. The patch clamp technique, described briefly in Section 2.2.2, and in detail in Chapter 6, has greatly enhanced our ability to investigate electrophysiological properties of ion channels. At the same time, advancing computing power has led to the development of a number of different techniques for determining mathematical

models to describe ion channel kinetics. These range from molecular dynamics simulations, to stochastic representations, and through to continuum modelling approaches. The models we use, detailed in this section, describe ion channels as existing in a number of distinct possible conformational states, with specified kinetic rates defined for changes between each of the states.

The rate at which an ion channel changes from one conformational state to another is termed the *transition rate*. The simplest mathematical model we could propose to describe the kinetics of a single ion channel would have only two distinct kinetic states represented; an *open* state (*O*) and a *closed* state (*C*). If we assume the channel moves from being in the open state to the closed state at a rate of α times per unit time and moves from being in the closed state to the open state β times per unit time, then a kinetic model for the channel could be represented by the state diagram

$$C \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} O \tag{3.1}$$

Transition rates between conformational states may depend on the membrane voltage, as they do for the hERG channel and other voltage-gated ion channels, as described in Section 2.2.2. The transition rates α and β in Equation (3.1) are then functions of voltage (*V*), the exact form of which we discuss later in this section.

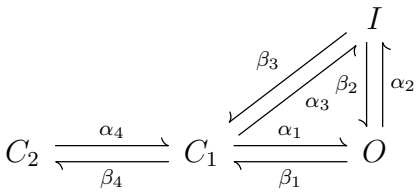
Considering a population of ion channels acting in the way described by the model in Equation (3.1), if we denote the proportion of channels in the open state by *O* and the proportion of channels in the closed state by *C*, then the rate of change of the proportion of channels in each conformational state can be described by the equations

$$\begin{aligned} \tfrac{\mathrm{d}O}{\mathrm{d}t} &= \alpha C - \beta O, \\ \tfrac{\mathrm{d}C}{\mathrm{d}t} &= -\alpha C + \beta O. \end{aligned} \tag{3.2}$$

Noting that the probabilities $O + C = 1$, Equation (3.2) reduces to

$$\frac{\mathrm{d}O}{dt} = \alpha(1 - O) - \beta O. \tag{3.3}$$

Additional conformational states, with the possibility of transitions between any of the states, can be added to this simple model. We can draw a state diagram of the kind shown in Equation (3.1) to describe the model and we can use this to formulate a set of Ordinary Differential Equations (ODEs) of the form shown in Equation (3.2). It is possible to have multiple conformational states of the same type. For example, a model with two closed states (C_1 and C_2), one open state (O) and one *inactivated* state (I) could be proposed, and be represented by the state diagram



(3.4)

α_1, α_2, α_3, α_4, β_1, β_2, β_3 and β_4 are transition rates between the distinct conformational states. Where a "closed loop" arrangement exists in an ion channel state diagram, as it does between the C_1, O and I states in Equation (3.4), the *microscopic reversibility condition* must be satisfied (Tolman, 1938). The microscopic reversibility condition is an addition to the laws of thermodynamics. It is formulated on the basis that the first and second laws of thermodynamics are fulfilled in a cyclic reaction if, during dynamic equilibrium there continues to be movement around the cycle but that this is uniform (Colquhoun et al., 2004). That is to say that according to the microscopic reversibility condition, while during equilibrium there may be no observable changes at a macroscopic level, there is still continuous kinetic activity on a microscopic scale, which is referred to as a dynamic equilibrium. The microscopic reversibility condition therefore states that the frequency of transitions in one direction around the loop must be the same as that in the reverse direction (Colquhoun et al., 2004). This condition ensures that the transition rates are such that one direction around a closed loop is not favoured over another. Consequently, for the example model in Equation (3.4), the following condition must hold:

$$\alpha_1\alpha_2\beta_3 = \alpha_3\beta_2\beta_1. \tag{3.5}$$

In general we denote n distinct conformational states as Y_1,...,Y_n, ordered as the conformational changes occur with increasing voltage. The model is described by a state diagram with

transition rates k_{ij} defining the rate of transition from state Y_i to state Y_j. A set of ODEs of the general form shown in Equation (3.6) can then be formulated to describe the model.

$$\frac{\mathrm{d}Y_j}{\mathrm{d}t} = \sum_{i=1, i\neq j}^{n} (k_{ij}Y_i - k_{ji}Y_j), \tag{3.6}$$

where k_{ij} is the rate of transition from state Y_i to state Y_j.

For a channel to change conformational state it must overcome associated free-energy barriers (Keener and Sneyd, 2010). The rate at which it is able to do this is an exponential function of the height of the potential energy barrier that must be crossed (Eyring, 1935). Therefore, transition rates for voltage-dependent ion channel kinetic models are expressed using the following exponential function of voltage

$$\alpha = A\exp(BV), \tag{3.7}$$

where A and B are constants and V is the membrane voltage. It follows that A must always be positive so that the transition rate is positive. The constant B is positive for forward transition rates (of the form k_{ij} in Equation (3.6)) and negative for backward transition rates (of the form k_{ji} in Equation (3.6)) for transitions from state i to state j. This reflects the relative challenge of moving to different possible conformational states from a given state depending upon the membrane voltage.

3.1.1 Modelling Electrical Current Flow Across the Cell Membrane

Electrical activity across the cell membrane can be modelled through analogy to an electrical circuit as we describe in detail in this section (Hille et al., 2001). As the cell membrane is made of an insulating lipid bilayer, it separates charged particles, and can be modelled as a capacitor. Capacitance is defined as the ratio of charge across a capacitor (Q) to the membrane voltage (V) which is required to hold the charge,

$$C_m = \frac{Q}{V}. \tag{3.8}$$

Current (I), is the rate of flow of charge over time,

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t}. \tag{3.9}$$

Differentiating Equation (3.8) and substituting in Equation (3.9), with the assumption of constant capacitance, the capacitive current (I_{cap}) is described as,

$$I_{cap} = C_m \frac{\mathrm{d}V}{\mathrm{d}t}. \tag{3.10}$$

According to Kirchoff's Current Law, electrical current is conservative, and so the sum of ionic and capacitive currents must be zero. It therefore follows that

$$C_m \frac{\mathrm{d}V}{\mathrm{d}t} + I_{ion} = 0, \tag{3.11}$$

where I_{ion} represents the transmembrane ionic current.

It can be assumed that two factors contribute to the potential difference across the cell membrane: differences in ionic concentrations and the flow of electrical current.

The potential difference due to the ionic concentration differences can be described by the *Nernst equation* (Nernst, 1888), which is defined for any ion type S by:

$$E_S = \frac{RT}{zF} \ln\left(\frac{[S]_o}{[S]_i}\right), \tag{3.12}$$

where R is the gas constant, T is absolute temperature, z is the charge on ion S, F is Faraday's constant, $[S]_o$ is the extracellular concentration of ion S and $[S]_i$ is the intracellular concentration of ion S. The value E_S calculated is termed the *reversal potential* and is the driving force which moves ions across the membrane down their concentration gradient as described in Section 2.2.

The potential difference across the membrane due to flow of electric current of ionic species S, can be described by Ohm's law,

$$V_S = RI_S, \tag{3.13}$$

where R is the channel resistance and I_S is the transmembrane current due to the flow of ion S.

Summing the contributions to changes in the potential from the expressions in Equations (3.12) and (3.13), the total potential difference (V) is expressed as

$$V = RI_S + E_S. \tag{3.14}$$

Equation (3.14), can be rearranged to derive the following expression for the current flowing through ion channel S

$$I_S = g_S(V - Es), \tag{3.15}$$

where g_S is the *membrane conductance* of channel S and is equal to the reciprocal of the resistance R. Conductance is typically specified per unit area of cell membrane and so is equivalent to the maximal conductance of a single ion channel multiplied by the number of open channels per unit area of membrane. Equation (3.15) allows us to calculate the current of each ionic species S flowing in response to any membrane voltage (Hille et al., 2001).

The first quantitative model describing the propagation of an electrical signal was developed by Hodgkin and Huxley (1952). The model describes electrical activity in the squid giant axon. It is assumed that the principal ionic currents are the Na^+ and K^+ currents, and the other ionic currents are represented as one combined current called the *leakage current*. Under these assumptions, Equation (3.11) can be rewritten expressing I_{ion} as a sum of contributions from the Na^+, K^+ and leakage currents (L) to formulate Equation (3.16). I_{app} represents the additional contribution to the total current from a current applied during the experiment.

$$C_m \frac{\mathrm{d}V}{\mathrm{d}t} = -g_{Na}(V - E_{Na}) - g_K(V - E_K) - g_L(V - E_L) + I_{app}. \tag{3.16}$$

3.1.2 Hodgkin-Huxley and Markov Models of Ion Channel Kinetics

In addition to being able to predict key features of excitable behaviour in the squid giant axon, Hodgkin and Huxley's seminal work also indicated important aspects of ion channel gating (Hodgkin and Huxley, 1952). As in Equation (3.16), the conductance of each ion channel is described by a different mathematical formulation. For instance, in the Hodgkin

and Huxley (1952) model, the conductance of the potassium channel is described by

$$g_K = \bar{g_K} n^4, \tag{3.17}$$

where $\bar{g_K}$ is the maximum conductance of the potassium channel and n can be interpreted as the probability that a subunit of the potassium channel is in the open conformation. As described in Section 2.2.2, potassium channels are formed of four subunits, each of which is required to be in the open state for the channel to be open. In the Hodgkin and Huxley (1952) ion channel model formulation it is assumed that the gating of each individual subunit is independent. Hence, the probability of all four subunits being simultaneously in the open state (and consequently of the channel being open) is expressed as n^4. This explanation for the potassium channel conductance in Equation (3.17), however, is a reinterpretation of Hodgkin-Huxley's hypothesis which astonishingly was formulated before it emerged that the opening and closing of ion channels mediated current.

The Hodgkin and Huxley (1952) model for potassium channel kinetics can be represented in a state diagram of the kind shown in Equation (3.1). Four transitions must be represented in the model — for the channel to move from the closed to the open state, each subunit must open in turn, with each transition representing the opening of one subunit. The assumption that the four potassium channel subunits are independent and kinetically identical means that the model can be described by

$$C_4 \underset{\beta}{\overset{4\alpha}{\rightleftharpoons}} C_3 \underset{2\beta}{\overset{3\alpha}{\rightleftharpoons}} C_2 \underset{3\beta}{\overset{2\alpha}{\rightleftharpoons}} C_1 \underset{4\beta}{\overset{\alpha}{\rightleftharpoons}} O \tag{3.18}$$

In Equation (3.18), the subscript of each C corresponds to the number of subunits which are in the closed state. Using this notation, the open state (O) is equivalent to C_0. Only once the channel is open (with all four subunits in the open conformation), will the channel conduct current. In this model, each subunit opens with the transition rate α and closes at a rate β, both of which are voltage-dependent functions. The assumption of independent gating in Hodgkin-Huxley style ion channel model formulations precludes the possibility of, for example, describing the kinetics of a channel in which inactivation of an ion channel

depends on activation of the channel (Armstrong and Bezanilla, 1977; Bezanilla and Armstrong, 1977).

Markov models, which are a generalisation of Hodgkin-Huxley style models, are able to overcome the limitations of the independent gating assumption of Hodgkin-Huxley type models. Markov models are based on the assumption that transitions between different conformational states depend on the present conformation of the channel, but not on previous gating behaviour. Multiple open, closed, and inactivated states can be represented in Markov models and a system of ODEs can be formulated to describe the model in the same way as described for the examples shown at the beginning of this chapter. Markov models enable a more general description of ion channel kinetics to be represented. Solving the system of model ODEs allows us to determine the proportion of channels expected in each conformational state over time in response to an applied voltage protocol, such as in patch clamp experiments, which we describe briefly here and in detail in Chapter 6.

Voltage-step protocols are typically performed in patch clamp experiments for investigating ion channel kinetics. As we have described, an ion channel is able to move from one conformational state to another by overcoming an energy barrier associated with the transition. With the application of a constant voltage during a voltage-step protocol, the distribution of channels in different conformational states reaches an equilibrium. When the voltage is stepped to a new level, the equilibrium is disturbed, and the channel rearranges itself into a new conformation. The duration for which the voltage-step length is held is typically longer than the timescale over which the kinetics of the channel operate, and so the channel should reach an equilibrium state before the next voltage step. Figure 3.1 shows a selection of voltage-step protocols performed in patch clamp experiments for investigating different characteristics of ion channel kinetics. Variations of these forms of voltage-step protocols are used within electrophysiological investigations of hERG channel kinetics.

In Figure 3.2, examples of simulated and experimental current recordings in response to

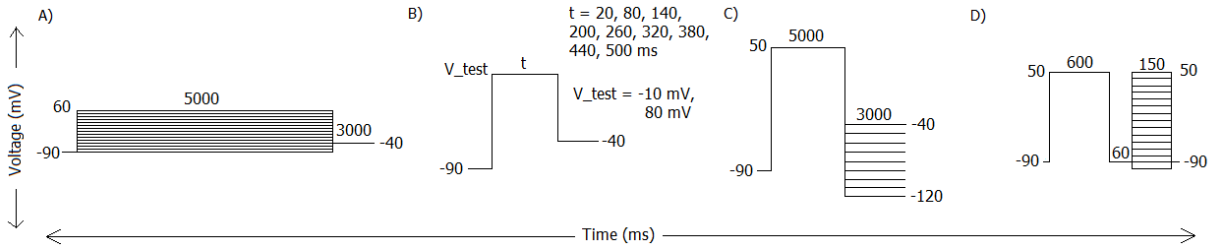


Figure 3.1: Examples of standard voltage-step protocols used for investigating hERG channel kinetics, as used in Bett et al. (2011). Each protocol is used to study a particular aspect of hERG channel kinetics; (A) steady-state activation, (B) activation kinetics, (C) deactivation and (D) inactivation. Each protocol is composed of a series of steps to different voltage levels where the voltage is held before stepping to the next voltage level.

a voltage-step protocol for investigating deactivation kinetics are shown.

Current recordings from the application of a series of voltage-step protocols akin to those shown in Figure 3.1 in patch clamp experiments are used to determine the transition rates in mathematical models of ion channel kinetics. In Section 3.1.3 we overview some different approaches for estimating transition rate parameters in ion channel models using macroscopic current recordings. In addition, we introduce some other concepts and considerations surrounding model parameterisation which are discussed in Chapters 5 and 7.

3.1.3 Determining Transition Rates from Macroscopic Currents

A number of techniques have been employed for determining the transition rate parameters of mathematical models describing ion channel kinetics. These range from using basic optimisation algorithms such as simplex (Nelder and Mead, 1965) or Levenberg-Marquardt (Marquardt, 1963) algorithms (Wang et al., 1997; Mazhari et al., 2001); to approaches using maximum likelihood estimation (Milescu et al., 2005; Moffatt, 2007); to global search methods such as using genetic algorithms, simulated annealing or particle swarm methods (Gurkiewicz and Korngreen, 2007; Wang et al., 2012). Each approach aims to determine parameter sets which allow model simulations to best replicate the observed experimental data. The most common methods use basic optimisation algorithms such as simplex search to fit averaged time constants in the form of exponential functions to summarised experimental

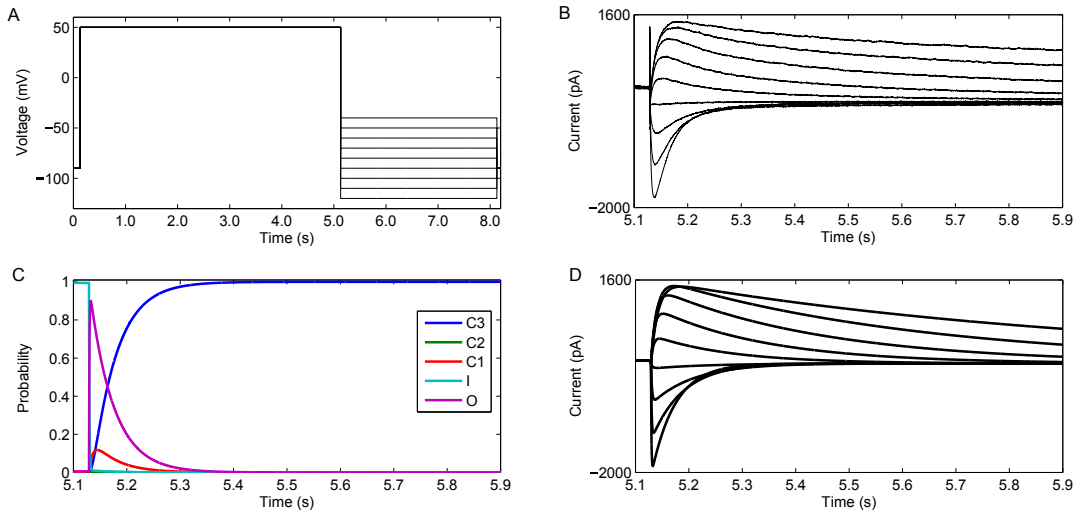


Figure 3.2: Diagram showing a current recording from a patch clamp experiment on hERG expressing cells (B) and simulated current from the Mazhari et al. (2001) hERG channel model (D) in response to application of the deactivation voltage-step protocol shown in A. The Mazhari et al. (2001) model used has 3 closed states (C_3, C_2 and C_1) and one open (O) and inactivated (I) state. C shows the changes in probability of each channel conformational state during one repolarising step (to -100 mV) of the protocol, as determined from solution of the model ODEs.

data obtained from the application of standard voltage-step protocols, of the type shown in Figure 3.1, in patch clamp experiments.

Each of the methods listed require varying amounts of computational resources and have associated advantages and limitations. Most of the approaches, while being able to find a good solution to the optimisation problem, do not guarantee to find a unique optimal solution. In fact, in some cases, manual parameter "tuning" is performed following the application of an optimisation algorithm (Vecchietti et al., 2006). Additionally, ion channel kinetic behaviour is characterised through the application of a number of different voltage protocols (as described in Section 3.1.2), and subsequently, fitting is performed with a number of different experimental data sets. Consequently, there are various ways model fitting might be attempted; for example, the model could be fitted using all of the data at once or different transitions could be fitted to subsets of the data. It is also often the case that all the experimental data is used for determining the parameters of the model, with little or no data left for

validation. So, while a particular model may be able to describe the ion channel behaviour under the protocols to which it was fitted, there is no guarantee that the kinetics predicted by the model under the application of a different protocol, or even the same protocol in a different cell, would be representative of the behaviour observed in experiment. However, it is important that if we want to use a mathematical model to represent the kinetics of an ion channel, it should behave as expected over the range of experimental conditions it is going to be used to simulate.

When developing a new mathematical ion channel model, the model structure (that is the number of states represented in the model and the arrangement of possible transitions between the states) is often assumed (informed by the conclusions of previous studies). The selected model structure is then parameterised to describe relevant experimental observations. However, some quantitative approaches have also been presented for *model selection* between at least two possible model structures by Liu et al. (1996), Wang et al. (1997), Kiehn et al. (1999) and Johnson et al. (1999).

Liu et al. (1996) compared simulation results (when assuming different possible model formulations) with currents observed in experiments to determine both the order of voltage-dependent and voltage-independent transitions during hERG channel activation. They subsequently determined the most appropriate additions to the model structure for it to be able to describe hERG channel inactivation. To determine the order of the transitions during activation they compared the simulation results from two linear models comprised of two closed states and one open state, one in which the voltage-independent transition preceded the voltage-dependent transition and the other where the voltage-dependent transition preceded the voltage-independent one. While they note that the order of these transitions cannot be determined from the time and voltage-dependence of activation data, they observe that the kinetics of deactivation of the drug-sensitive tail currents are strongly voltage-dependent, indicating that the voltage-dependent transition immediately precedes the opening of the channel.

Liu et al. (1996) then consider two alternative model formulations for representing hERG channel inactivation, using the activation model as a basis; in one model inactivation is assumed to be independent of activation, and in the second, inactivation is assumed to be coupled to activation. Each alternative model was then used to simulate hERG channel activation kinetic experiments and significant differences were found between the simulation results when using the two models. They used this analysis to conclude that inactivation of ferret I_{Kr} is most likely coupled to activation, allowing them to select the most appropriate model from the two proposed.

Wang et al. (1997) present a more thorough quantitative approach to select the most appropriate model to describe hERG channel activation. They consider four linear models of increasing complexity. The simplest model is a two state model with one closed state and one open state. Each subsequent model has additional closed states, with the most complex model consisting of four closed states and one open state. They fitted each model structure to data obtained from voltage protocols designed to investigate activation kinetics in patch clamp experiments. Based on consideration of the distribution of the residuals from the fit between simulated and experimental data, they reject model structures which are identified as being unable to describe the observed behaviour. They conclude that a model with three closed states and one open state is the minimal model needed to describe the observed activation kinetics data. They observe that including additional states, as represented in a model with four closed states and one open state, does not show great improvement in the quality of the fit and so based on this analysis were deemed unnecessary. Wang et al. (1997) also reasoned that a voltage-independent step precedes the closed/open transition due to the saturation in the activation rate in the late phase of activation.

Wang et al. (1997) extend their model of hERG channel activation to a full model of hERG channel kinetics by adding an inactivated state to their linear activation model, observing that this model is able to replicate observed experimental data well. Kiehn et al. (1999)

employ a more quantitative approach to determining the nature of the inactivation portion of the full hERG channel model, by considering a range of different possible model structures. They considered a range of models with different numbers of closed states represented and considered the inclusion and exclusion of inactivation transitions from the open state and the final closed state before the open state. They used single-channel recordings and macroscopic measurements to investigate hERG channel inactivation mechanisms. They demonstrate the necessity of the inclusion of an inactivated state to be able to replicate accurately single channel measurements. They conclude that the model able to replicate best observed experimental data includes closed-state inactivation, which was not included in the Wang et al. (1997) model. Through comparison of simulated currents from each proposed model and macroscopic currents measured experimentally, they propose a model with three closed states which includes both open and closed-state inactivation.

Milescu et al. (2005) use an alternative approach for model selection, considering the Akaike (AIC), (Akaike, 1974) and Bayesian (BIC), (Schwarz, 1978) information criteria. These are measures of goodness of fit based on the likelihood of the fit to a particular model, and a penalty for complexity. In general, the simplest possible model formulation which is able to replicate observed experimental behaviour is preferred. Using an example considering macroscopic currents from the acetylcholine receptor, Milescu et al. (2005), employ this method to select a model of the most appropriate size and topology to be able to replicate experimental behaviour.

A concept linked with model selection and parameterisation, is *identifiability*. Identifiability is concerned with whether the parameters of a mathematical model formulation can be uniquely determined from a given experimental data set. While it will be the case that parameters determined from different sets of experimental data will be distinct, due to variability inherent in biological systems, it should not be the case that two distinct sets of model parameters enable the model to behave in exactly the same way. If this is the case, then there is insufficient information in the experimental data being used to determine the model pa-

rameters uniquely. We propose that the design of the experiment should then be reevaluated to enable parameters to be determined uniquely (or the model should be simplified to one in which all the parameters are identifiable).

Milescu et al. (2005) consider parameter identifiability in addition to model selection. They show that the two models with the same conformational states represented, but in a different arrangement, are mathematically distinguishable when using a particular simulation protocol. They do this by fitting the data simulated from one model to the other model and *vice-versa*, and they show that they are unable to find a set of parameters which could generate the same current as the other model. They then go on to assess the ability to determine unique parameter values for each model when using different experimental protocols.

The idea of the identifiability of model parameters being dependent upon the experimental protocol used is further explored by Fink and Noble (2009). They examine the identifiability of parameters from a selection of ion channel models using a sensitivity-based approach. They investigate the identifiability of model parameters when using different voltage-step protocols to demonstrate the influence of the information contained within the available experimental data. In doing this, they raise the question of what are necessary or sufficient data for estimating the parameters of ion channel kinetic models. They compared variations of standard voltage-step protocols of different lengths. As described in Section 3.1.2, standard voltage-step protocols consist of voltage steps which are often held for durations much longer than those over which ion channel kinetics typically occur. In the shortest protocol found by (Fink and Noble, 2009) which allowed identifiable parameters to be determined across a range of models, the duration of time for which the voltage-steps are held are much reduced as compared to in traditional voltage-step protocols. This protocol is shown in Figure 3.3 (A). As the protocol was found to provide a similar level of information for estimating model parameters as the much longer series of standard protocols, this suggested redundancy of information in the typical approach.

Zhou et al. (2009) also use a more efficient combination of standard voltage-step protocols to determine a Markov model describing the effect of the drug verapamil on hERG channel kinetics. They also discuss the challenges of parameter estimation and identifiability. Their model is built using data obtained from native cardiomyocytes. Difficulties are often encountered recording I_{Kr} in native myocytes due to the additional currents from other cardiac ion channels in these cells and the diminishing current observed when using long voltage protocols. Zhou et al. (2009) attempt to overcome these difficulties by using qualitative features from hERG channel experiments in transfected cell lines to help constrain parameter estimation. In this way they ensure that the model is able to reproduce these expected features in addition to the behaviour observed in experiments. The shorter voltage protocol used is composed of variations of standard protocols combined together which cover a wide range of voltages and pulse durations.

Another non-standard voltage protocol for investigating ion channel kinetics was recently proposed by Di Veroli et al. (2013), specifically for elucidating information about drug-ion channel kinetics. This protocol, shown in Figure 3.3 (B), has rapid voltage-steps combining repeated short depolarising steps followed by repolarising steps, allowing for a dynamic assessment of drug binding kinetics.

In Chapters 5 and 7 we consider the design of non-standard voltage protocols for use in patch clamp experiments for determining mathematical models of hERG channel kinetics. In doing this we consider aspects of model parameterisation, model selection, parameter identifiability, and experimental design introduced here. In Section 3.1.4 we review the range of mathematical models which have been proposed to describe hERG channel kinetics.

3.1.4 Mathematical Models of hERG Channel Kinetics

A large number of different models have been proposed to describe the kinetics of the hERG channel. These span both Hodgkin-Huxley and Markov style models, with varying numbers

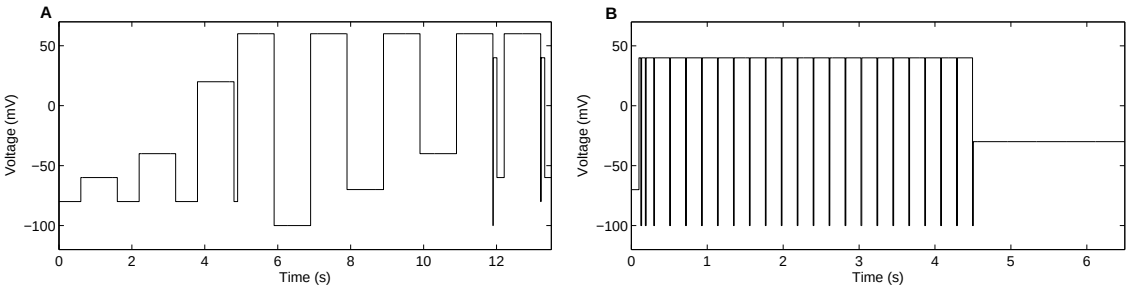


Figure 3.3: Diagram showing non-standard voltage-step protocols used by Fink and Noble (2009) (A) and Di Veroli et al. (2013) (B). Both protocols are much shorter than the standard combinations of voltage-step protocols performed in patch clamp experiments. The protocol proposed by Fink and Noble (2009) was shown to enable a range of transition rate parameters to be recovered uniquely (good identifiability) for a number of models. The protocol proposed by Di Veroli et al. (2013) provides information for determining drug-binding kinetics with the hERG channel.

and arrangements of open, closed and inactivated states represented. The equivalent Markov model structures of the range of existing models, each with different arrangements of open, closed and inactivated states are shown in Figure 3.4.

The earliest mathematical models of hERG channel kinetics, such as the Hodgkin-Huxley style model proposed by Zeng et al. (1995), were based on experimental data obtained from I_K currents, for example from Shibasaki (1987). Due to difficulties in separating out the multiple components of I_K still being encountered at this time, such models are not commonly used today. Courtemanche et al. (1998) and Priebe and Beuckelmann (1998) published later Hodgkin-Huxley style models of hERG channel kinetics and the first Hodgkin-Huxley style model formulated using data from expressed hERG channels was published by Ten Tusscher et al. (2004). Each of these Hodgkin-Huxley style models, in addition to the representations used within the cardiac action potential models published by Lindblad et al. (1996), Nygren et al. (1998), Winslow et al. (1999), Ramirez et al. (2000), Puglisi and Bers (2001), Fox et al. (2002), Seemann et al. (2003), Hund and Rudy (2004), Aslanidi et al. (2009) and Grandi et al. (2010), as well as in subsequent modified versions of these cardiac cell models, all

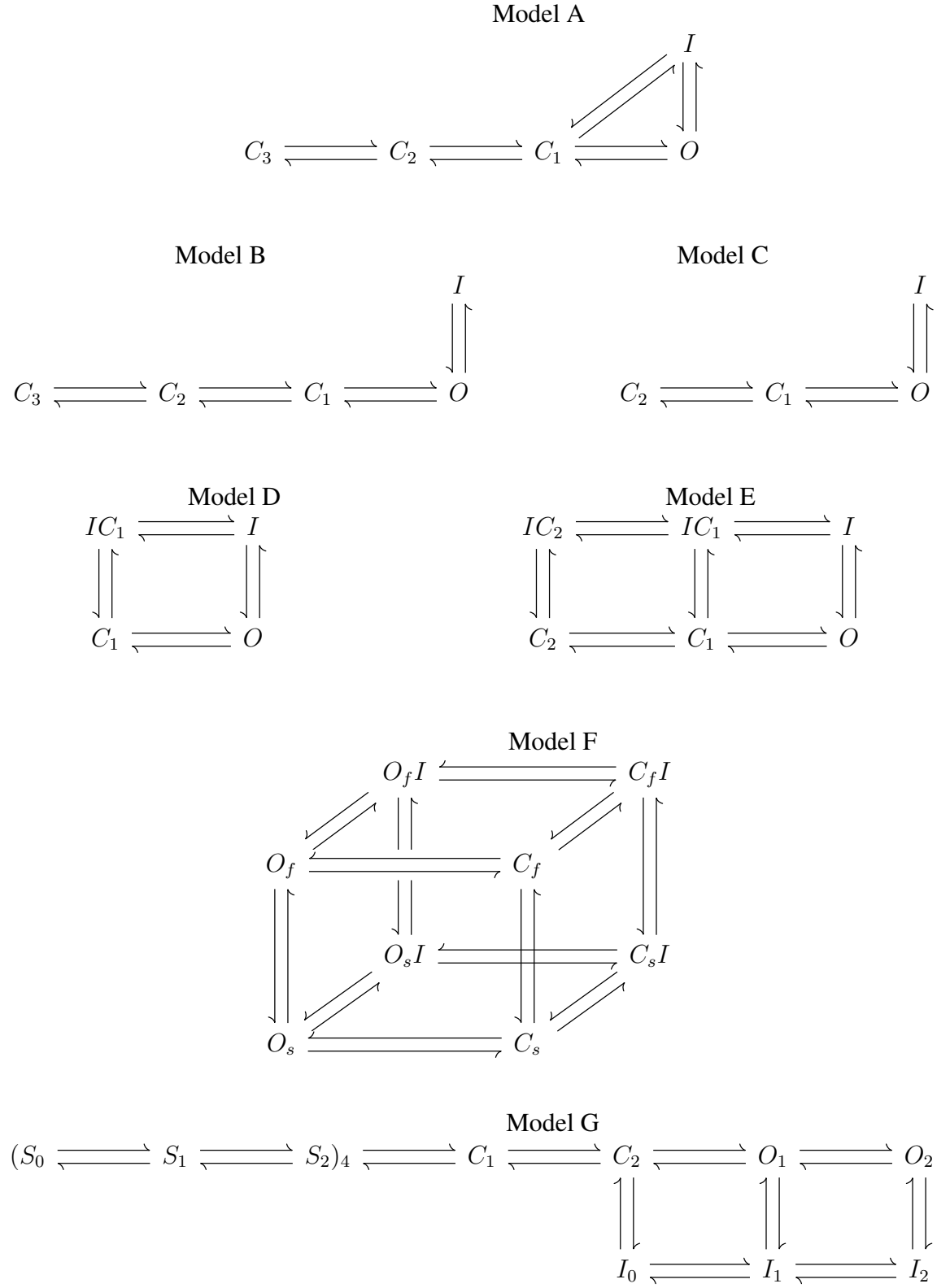


Figure 3.4: Equivalent Markov model structures for the range of existing published hERG channel models. Closed states are represented by C, inactivated states by I and open states by O. The label for each model is referred to in Table 3.1. Model G has a non-standard representation, with Hodgkin-Huxley style gating for four subunits indicated by the power '4' and then a Markov Model for the behaviour of the channel once all four subunits are in state S_2.

share the same hERG channel model structure (model D in Figure 3.4). This structure has been reparameterised in each model formulation to replicate relevant experimental data. A number of Hodgkin-Huxley style models have been suggested for describing hERG channel kinetics with an alternative model structure, which include a description of both fast and slow components of activation kinetics represented in different proportions (model F in Figure 3.4). Such models include those introduced by Noble et al. (1998), Zhang et al. (2000), Kurata et al. (2002), Matsuoka et al. (2003), Inada et al. (2009), O'Hara et al. (2011) and Severi et al. (2012).

Markov model formulations for hERG channel kinetics introduced by Liu et al. (1996), Wang et al. (1997) and Kiehn et al. (1999) have been discussed in detail in Section 3.1.3 and the majority of the subsequent Markov models describing hERG channel kinetics have been based on the model formulations presented by Wang et al. (1997) and Kiehn et al. (1999). The difference in structure of the models proposed by Wang et al. (1997) and Kiehn et al. (1999) is the presence of a transition between the closed and inactivated states in the Kiehn et al. (1999) model which is not present in the Wang et al. (1997) model (model A and B respectively in Figure 3.4). This addition was justified in Kiehn et al. (1999) from analysis of single-channel hERG behaviour. They found that single channels do not always open when depolarised but always open when repolarised. They then attributed this observation to inactivation from a closed state.

Fink et al. (2008) justify their use of the Wang et al. (1997) model structure for investigating the functional role of the hERG channel in the human ventricular action potential by noting that where transitions between closed and inactivated states are represented within other existing hERG channel models, these are defined to be very infrequent and so they consider them to be negligible. Fink et al. (2008) attempt to investigate the possibility of the presence of fast and slow components of deactivation. They do this by fitting a sum of two exponential curves to deactivation voltage-clamp data. They found that while they could always fit the slow deactivation time constant to the data, difficulties were encountered when

fitting the fast time constant, again highlighting the numerical difficulties of parameterising models of this form.

The model structure proposed by Kiehn et al. (1999), which has the structure of model A in Figure 3.4 (including the transition between the closed and inactivated states), was adopted by Clancy and Rudy (2001), Mazhari et al. (2001) and Lu et al. (2001) for use in *in silico* assessments of long QT syndrome and premature stimulation on hERG channel gating. Different parameterisations of the original model structure were used in each case. Lu et al. (2001) attempt to reason why such differences are seen between the parameterisation of each of these models, and in particular to justify why some transition rates in their model formulation are often faster than in the other models. They offer possible explanations to differences including differences in experimental setups, with different cell types and recording temperatures used. In particular, they highlight that in their own approach some recordings were performed at 37°C and others at 35°C. Due to the strong temperature dependence of hERG channel kinetics, as described at the end of this section, this could impact on the model. They also note that the Mazhari et al. (2001) model was fitted to data recorded at 22°C and the parameters were scaled to produce a model capable of describing currents at physiological temperature, again using the approach described at the end of this section.

The model developed by Oehmen et al. (2002) of hERG channel kinetics in the sinoatrial node was based on the hERG Markov model structure proposed by Liu et al. (1996), which has the structure of model C in Figure 3.4. Oehmen et al. (2002) note that this model is the most applicable to their experimental setting. The model was then reparameterised for their purpose of describing rabbit sinoatrial hERG channel kinetics. Di Veroli et al. (2013) propose two models describing hERG channel kinetics, one for the channel at room temperature (model E in Figure 3.4) and one at physiological temperature (model D in Figure 3.4). The Di Veroli et al. (2013) model formulated to describe hERG channel kinetics at room temperature utilises one of the model structures proposed by Liu et al. (1996). However, this model structure was rejected by Liu et al. (1996) for describing hERG channel kinetics in

ferret atrial cells due to the evident coupling between activation and inactivation in their experimental observations which was not encapsulated by this model formulation. The model describing hERG channel kinetics at physiological temperatures is a reparameterisation of the model proposed by Hund and Rudy (2004).

The model proposed by Piper et al. (2003) is distinct from the other formulations described so far in this section. Their Markov model formulation includes two open states, an additional closed state, as well as closed-state inactivation. Another unique aspect of this model is that the transitions between the first two closed states are modelled independently for each of the four subunits. This means that the movement of each subunit between these two closed states is represented directly in the model. The inclusion of two open states in the model, with equal conductance, is based on single-channel recordings from Kiehn et al. (1999) and Zou et al. (1997). Piper et al. (2003) also use the same single-channel data to conclude that there is a high probability of inactivation from the closed state. An analysis of an inactivation-deficient hERG mutation also allowed Piper et al. (2003) to conclude that hERG channel inactivation is coupled to activation. This model is much more complex and requires many additional parameters than the other hERG channel models described in this section.

The behaviour of a subset of the hERG channel models described in this section is compared and contrasted in more detail in Bett et al. (2011). Table 3.1 summarises the details of each of the models described in this section. As well as indicating the corresponding model structure in Figure 3.4 of each of the models, the model type, the number of parameters, and details of the experimental system from which each model was derived are included. In Figure 3.5, we plot the transition rates for each transition present in all the models with Markov model structures A, B or C in Figure 3.4 over a range of voltages.

We can see that the largest differences between transition rates occur at the extremes of the membrane potential range considered with large differences evident between some pa-

Table 3.1 Table summarising details of the structure and experimental system used to derive each of the existing published hERG channel model formulations. The 'Model Type' is 'HH' for Hodgkin-Huxley models and 'MM' for Markov Models, or a hybrid of the two (MM/HH = MM with some identical transition rates). '# Params.' indicates the total number of parameters, including a parameter representing the conductance of the channel. The number of parameters corresponds to the number in the model after the model has been rewritten to be *a priori* identifiable (encompassing two parameters into a single parameter where this is possible, as described in Fink and Noble (2009) and as we elaborate on in Chapter 5). Entries are ordered in chronological order. The corresponding equivalent Markov model structure as shown in Figure 3.4 is also included for each model. (*The transition rates of the Kiehn et al. (1999) model are defined at specific voltages, so for this model there are 8 parameters (and 1 conductance parameter) for each voltage for which the model is defined).

Model	Model Type	# Params.	Experimental Cell Type	Temperature	Structure in Figure 3.4
Zeng et al. (1995)	HH	11	Guinea pig ventricular myocytes	Physiological	D
Lindblad et al. (1996)	HH	11	Rabbit SA, AV, atrial myocytes and rabbit and guinea pig ventricular myocytes	Physiological	D
Liu et al. (1996)	MM	11	Ferret atrial myocytes	Room	C
Wang et al. (1997)	MM	15	*Xenopus* oocytes	Room	B
Courtemanche et al. (1998)	HH	10	Human atrial myocytes	Physiological	D
Noble et al. (1998)	HH	12	Guinea pig ventricular myocyte	Physiological	F
Nygren et al. (1998)	HH	9	Human atrial/rabbit atrial myocytes and *Xenopus* oocytes	Physiological	D
Priebe and Beuckelmann (1998)	HH	9	Human ventricular myocytes	Physiological	D
Kiehn et al. (1999)	MM	9*	*Xenopus* oocytes	Room	A
Winslow et al. (1999)	HH	7	Guinea pig ventricular myocytes	Physiological	D
Ramirez et al. (2000)	HH	13	Canine atrial myocytes	Physiological	D
Zhang et al. (2000)	HH	15	Rabbit sinoatrial node cells	Physiological	F
Clancy and Rudy (2001)	MM	15	Guinea pig ventricular	Physiological	A
Lu et al. (2001)	MM	17	Chinese Hamster Ovary (CHO)	Physiological	A
Mazhari et al. (2001)	MM	17	Human Embryonic Kidney (HEK) 293	Physiological	A
Fox et al. (2002)	HH	10	Canine ventricular myocytes	Physiological	D
Kurata et al. (2002)	HH	18	Rabbit sinoatrial node cells	Physiological	F
Oehmen et al. (2002)	MM	11	Rabbit sinoatrial cells	Physiological	C
Matsuoka et al. (2003)	HH	22	Rabbit pacemaker and guinea pig ventricular myocytes	Physiological	F
Piper et al. (2003)	MM/HH	43	*Xenopus* oocytes	Room	G
Seemann et al. (2003)	HH	7	Human ventricular myocytes	Physiological	D
Hund and Rudy (2004)	HH	11	Canine ventricular myocytes	Physiological	D
Shannon et al. (2004)	HH	11	Rabbit ventricular myocytes	Physiological	D
Ten Tusscher et al. (2004)	HH	13	HEK 293/CHO/*Xenopus* oocytes	Physiological	D
Fink et al. (2008)	MM	15	Human Embryonic Kidney (HEK) 293	Physiological	B
Aslanidi et al. (2009)	HH	8	Canine purkinje cells	Physiological	D
Inada et al. (2009)	HH	20	Rabbit atrioventricular node cells	Physiological	F
Grandi et al. (2010)	HH	12	Human ventricular myocytes	Physiological	D
O'Hara et al. (2011)	HH	19	Human ventricular myocytes	Physiological	F
Severi et al. (2012)	HH	17	Rabbit sinoatrial node cells	Physiological	F
Di Veroli et al. (2013)	MM/HH	13	Chinese Hamster Ovary (CHO)	Room	E
Di Veroli et al. (2013)	HH	17	Human Embryonic Kidney (HEK293) expressing canine ERG	Physiological	D

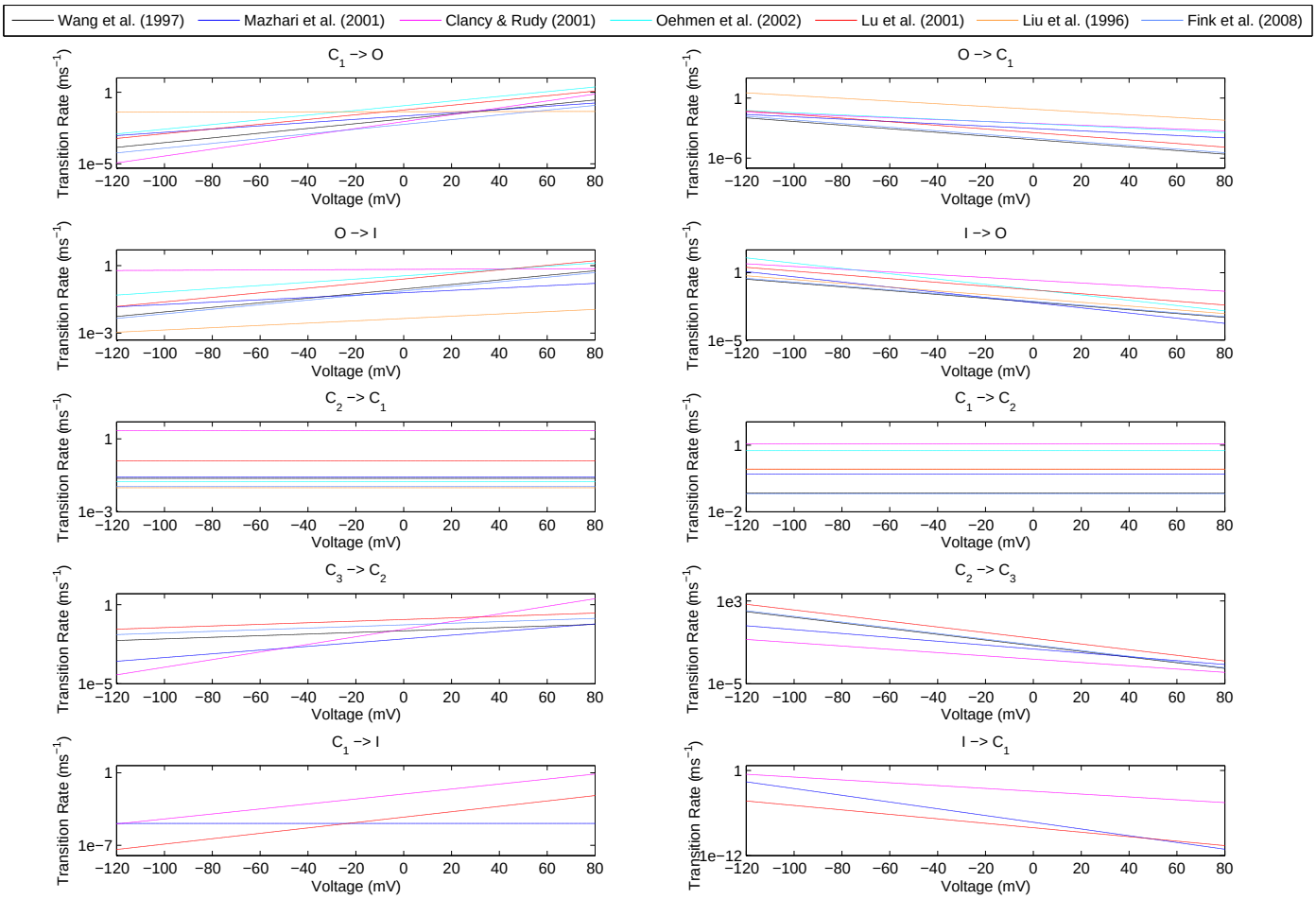


Figure 3.5: Plot of transition rates of each transition in all models in Table 3.1 with structures corresponding to those of model A, B or C in Figure 3.4. It should be noted that the Liu et al. (1996) and Wang et al. (1997) are defined at room temperature and the remainder of the models at physiological temperature. Some differences in transition rates would be expected due to this factor so models defined at the same temperature should be compared with one another.

rameters. The model equations describing each of the models included in Table 3.1 can be found in Appendix A.

A difference between the hERG channel models described in this section, which has been highlighted in Table 3.1, is the temperature for which the models were parameterised. hERG channel kinetics have been shown to be very temperature-dependent (Zhou et al., 1998) and so temperature differences between the models are significant. As the temperature increases, so too do the rates of activation and deactivation, but to a lesser extent than changes to the rates of inactivation and recovery from inactivation (Vandenberg et al., 2012).

A method for adjusting the kinetic rate constants to simulate a model at a temperature different to that at which it was formulated is to scale them using the Q_{10} temperature scaling coefficient. The Q_{10} coefficient is a measure of the rate of change of a biophysical system due to an increase in the temperature of the system by 10°C. The Q_{10} coefficient is calculated by

$$Q_{10} = \left(\frac{\alpha_2}{\alpha_1}\right)^{\frac{10}{T_2 - T_1}}, \tag{3.19}$$

where α_1 and α_2 are the rate constants measured at temperatures T_1 and T_2 respectively (Hille et al., 2001). Vandenberg et al. (2006) note that results recorded at room temperature cannot simply be extrapolated to 37°C using a single temperature scaling coefficient and they provide Q_{10} values for time constants based on experimental data from voltage-step protocols at different temperatures. They estimate Q_{10} values of: 2.1 for activation kinetics, 1.7 for deactivation kinetics, 2.5 for inactivation kinetics, 2.6 for reactivation kinetics and 1.4 for conductance. However, in some examples of hERG channel model formulations, a single Q_{10} scaling factor has been used for all transition rates to make a model formulated on patch clamp data obtained at room temperature appropriate for simulating at 37°C (Mazhari et al., 2001).

3.2 Mathematical Models of Ventricular Cardiomyocyte Electrophysiology

3.2.1 Development of Mathematical Models of Ventricular Cardiomyocytes

Detailed models of ion channel kinetics are included within mathematical models describing electrical activity in individual cardiomyocytes. Building on the seminal work of Hodgkin and Huxley (1952) described in Section 3.1.1, Noble (1960, 1962) showed that the same principles used for describing electrical propagation in the squid giant axon could be applied to describe action potential morphology in cardiac cells and presented the first cardiac computational model of the Purkinje fibre. Leading on from there, the first cardiac model of the ventricular action potential was formulated by Beeler and Reuter (1977). This work built on that of McAllister et al. (1975), who constructed a Purkinje fibre action potential model by extending the earlier work of Noble (1960, 1962). The Beeler and Reuter (1977) model included representations of four distinct membrane currents: a fast inward sodium current, a slow inward current comprised predominately of calcium, a time-dependent inward rectifier potassium current and a time-dependent outward potassium current. Ionic currents in all of these action potential models are described by Hodgkin-Huxley style formulations.

Alongside improvements in electrophysiology experimental techniques, leading to an increase in the quality and quantity of experimental data, the number of mathematical models being proposed to describe cardiac electrophysiology behaviour in a range of experimental settings was also rising rapidly. The first ventricular cell models based on voltage-clamp recordings on isolated myocytes were published by Luo and Rudy (1991) and Noble et al. (1991). The Luo and Rudy (1994) model of the ventricular action potential in the guinea pig was the first model of the ventricular action potential to include dynamic changes in ionic concentrations during the action potential. It also included sarcolemmal Na-K and Ca^{2+} pumps, Na-Ca exchange, Ca^{2+} buffers, calcium-induced calcium release (CICR) and Ca^{2+} cycling. This model forms the foundation of many subsequent ventricular action potential

models. It also marked the beginning of the development of increasingly biophysically detailed computational models of cardiac ventricular myocytes. At the same time, as more details of the mechanisms of ion channel kinetics began to be unravelled, the use of Markov models to represent some ion channel currents within cardiac myocyte models began to be seen (Clancy and Rudy, 1995).

Since the Luo and Rudy (1994) model, a range of models of the ventricular action potential in different species have been published, including: rabbit (Puglisi and Bers, 2001; Shannon et al., 2004; Mahajan et al., 2008); mouse (Bondarenko et al., 2004; Li et al., 2010); rat (Pandit et al., 2001; Wang and Sobie, 2008); guinea pig (Zeng et al., 1995; Noble et al., 1998; Faber et al., 2007); dog (Winslow et al., 1999; Hund and Rudy, 2004; Decker et al., 2009); and human (Priebe and Beuckelmann, 1998; Iyer et al., 2004; Ten Tusscher et al., 2004; Grandi et al., 2010; Carro et al., 2011; O'Hara et al., 2011; Asakura et al., 2014). While we have focused exclusively here on models of the ventricular action potential as these are used in this thesis, there are is also an array of cardiac electrophysiology models describing the behaviour of atrial cells and Purkinje fibres, and models of both healthy and diseased cardiomyocytes have also been developed. A more complete review of the different cardiac electrophysiology models developed can be found in Noble et al. (2012).

The CellML repository (Lloyd et al., 2008) is a database which contains many of the published models of cardiac electrophysiology as well as models of other biological cells. Models in CellML are described in an XML-based markup language, which can be used reliably to convert models into source code in a variety of different programming languages. The mathematical models of cardiac electrophysiology used in this thesis were taken from the CellML repository. Further details about CellML can be found in Appendix C.1.1.

Over time, mathematical models of cardiac electrophysiology have become increasingly more complex, with more details of biological knowledge of different cellular mechanisms and features incorporated. The more recent models contain representations of an increased

number of ionic currents, pumps and exchangers as compared to the earlier models. To illustrate the level of biological detail incorporated within models of cardiac cellular electrophysiology, in Section 3.2.2 we briefly describe the rabbit ventricular myocyte model published by Shannon et al. (2004), which is the main mathematical model used in Chapter 4.

3.2.2 A Mathematical Model of the Rabbit Ventricular Myocyte

Figure 3.6 presents the main biological components represented with the Shannon et al. (2004) model of the rabbit ventricular myocyte.

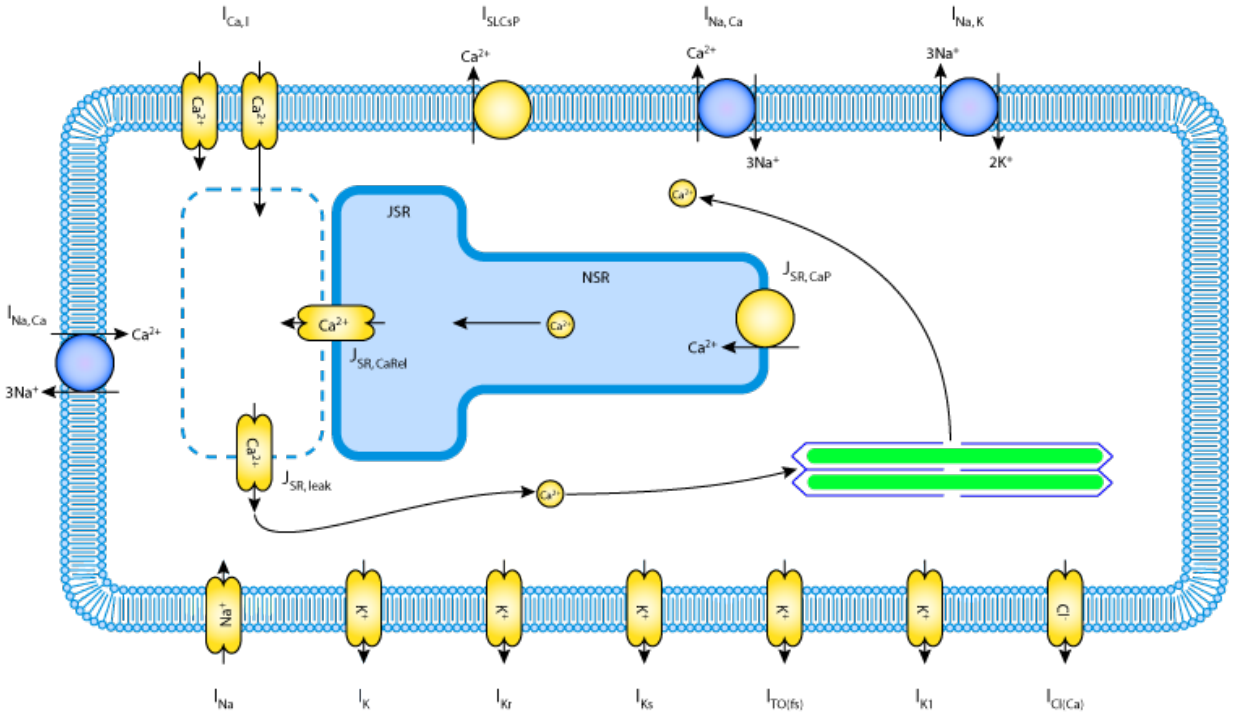


Figure 3.6: Schematic showing the components of the mathematical model of the rabbit ventricular myocyte developed by Shannon et al. (2004). Figure taken from *https://models.cellml.org*. Model components are described in the text.

The Shannon et al. (2004) model builds on the first model of the rabbit ventricular myocyte published by Puglisi and Bers (2001). The model is largely based on the Luo and Rudy (1994) model, with species-specific alterations and additions making the model appropriate

for modelling rabbit ventricular electrophysiology. The model includes representation of 14 transmembrane currents, which are indicated in Figure 3.6. These ionic currents include: fast Na^+, L-type Ca^{2+}, rapid and slow components of the delayed rectifier K^+, inward rectifier K^+, fast and slow transient outward K^+, Ca^{2+}-activated Cl^-, Na-Ca exchanger, Na-K pump, sarcolemmal Ca^{2+} pump, and background Na^+, K^+, and Cl^- currents. A subsarcolemmal compartment is represented in the model in addition to junctional and bulk cytosolic compartments represented in the Puglisi and Bers (2001) model. The subsarcolemmal compartment is the space directly below the cell membrane on the inside of the cell. The junctional cleft is a small space between the cell membrane and the junctional sarcoplasmic reticulum and the bulk cytosolic space comprises a large portion of the remaining intracellular space.

In the Shannon et al. (2004) mathematical model there are 45 state variables, 13 of which describe ion channel gating and the others include variables to represent various ionic concentrations within the different model compartments. The full set of equations describing the model can be found in Shannon et al. (2004). Key features of the Shannon et al. (2004) model include: separate distributions of ion channels within the sarcolemma and the junctional space along with region-specific ionic concentrations; a modified formulation of the Na^+-Ca^{2+} exchange current; and an updated description of intracellular Ca^{2+} handling.

In all of the biophysical cardiac electrophysiology models described, one variable (V) describes the membrane potential. As described in detail in Section 3.1.1, the membrane voltage of the cell over time can be determined by solving

$$C_m \frac{\mathrm{d}V}{\mathrm{d}t} = -\sum I_{ion} + I_{app}, \tag{3.20}$$

where I_{ion} represents the ionic current from each of the channels represented within the model and I_{app} is the stimulus current applied to pace the cell. In Section 3.3 we describe how single cell mathematical models of ventricular cardiomyocytes are used to model the propagation of electrical signals in cardiac tissue.

3.3 Modelling Electrical Propagation in Ventricular Tissue

Mathematical models of single cardiomyocytes, as described in Section 3.2, allow us to model the action potential of individual cardiac cells, but in the heart cells are coupled together creating a tissue through which the electrical waves propagate. Propagation of electrical activity in cardiac tissue can be modelled using both discrete and continuum models. A review of both model types can be found in Clayton et al. (2011). Discrete models are formulated by explicit representation of individual cells within the cardiac tissue being modelled. In continuum models, cardiac tissue is represented mathematically as a syncytium. Using both types of model, a geometry describing the arrangement of the tissue is defined. Such models of cardiac tissue can be formulated as one-, two- or three- dimensional models, with the latter including whole organ and even whole body simulations (Zemzemi et al., 2013). In Chapter 4 we use the simplest form of a continuum tissue model: simulating a one-dimensional line of cells, analogous to a strand of coupled cells, representing the simulation of a cardiac fibre. Therefore, in this section, we restrict our description of cardiac tissue modelling to continuum approaches.

Continuum models of electrical propagation in cardiac tissue are underpinned by Partial Differential Equations (PDEs) in the form of the *bidomain* equations and their simplification, the *monodomain* equations. In the bidomain representation, cardiac tissue is modelled as being composed of intracellular and extracellular domains, with the description of each domain being based on a generalised version of Ohm's law. The monodomain equation is a simplification of the bidomain equation in which it is assumed that the conductivity in the extracellular space is proportional to that in the intracellular space. Both the monodomain and bidomain equations can be solved to determine the electrical activity in the tissue over time. The one-dimensional tissue simulations used in Chapter 4 were solved using the monodomain equation, and so we focus only on the mathematical formulation of this equation here. However, more details about the generalised monodomain and bidomain models can be found in Sundnes et al. (2007).

The monodomain and bidomain equations are based on *cable theory*. In cable theory, a cardiac tissue is represented by an electrically conducting cable. The cable has a membrane potential $V(x,t)$ at each position x along the cable at time t. The cable is assumed to have a resistance R and an axial current I_a which flows down the cable perpendicular to the transmembrane current I_m.

As described in Section 3.1.1 for the single cell, Ohm's law states that,

$$\frac{\partial V}{\partial x} = -RI_a. \tag{3.21}$$

Kirchhoffs laws state that electrical current is conservative, so that any change in current must be due to the transmembrane current and so

$$I_m = -\frac{\partial I_a}{\partial x}. \tag{3.22}$$

Differentiating Equation (3.21) with respect to x and substituting in Equation (3.22) yields

$$\frac{\partial^2 V}{\partial x^2} = -R\frac{\partial I_a}{\partial x} = -RI_m. \tag{3.23}$$

I_m is the sum of both capacitive and resistive currents, as described in Section 3.1.1,

$$I_m = I_{cap} + I_{ion}, \tag{3.24}$$

and the capacitive current is expressed as in Equation (3.10), therefore

$$I_{cap} = C_m\frac{\partial V}{\partial t}. \tag{3.25}$$

Substituting Equations (3.24) and (3.25) into Equation (3.23) completes the derivation of the one-dimensional monodomain equation, to give the expression

$$\frac{\partial V}{\partial t} = -\frac{1}{C_m}I_{ion} + \frac{1}{C_m R}\frac{\partial^2 V}{\partial x^2}. \tag{3.26}$$

Adding the additional contribution of current from pacing of the tissue and rearranging and rewriting this equation using more concise notation yields the expression

$$\chi(C_m\frac{\partial V}{\partial t} + (I_{ion} - I_{app})) = \nabla\cdot(\boldsymbol{\sigma}\nabla V), \tag{3.27}$$

where $\boldsymbol{\sigma}$ is the conductivity tensor, χ is the membrane surface area per unit volume of tissue, I_{app} is the stimulus current applied to pace the tissue, and I_{ion} is the flow through the cell membrane per unit area, as determined from the solution of the mathematical model of the single cardiac cell.

Within cardiac tissue simulations, we could represent all cells as being of the same type, in which case we would describe the tissue as being *homogenous*. However, different cell types can also be represented within different sections of the tissue being simulated, to represent the arrangement of cells in the heart wall, as described Section 2.1. Such a tissue simulation is referred to as being *heterogeneous*. In practice, to incorporate different cell types within tissue simulations, electrophysiological properties in the mathematical model used in different regions of the tissue are altered, corresponding to the different cell types being represented.

3.4 Mathematical Representations of Drug-Ion Channel Interactions

3.4.1 Conductance Block using Simple Pore Block Theory

As described in Section 2.3, drug block of ion channels can be described by the simple pore block theory, which assumes that a drug physically obstructs the pore of the channel by an amount depending on the affinity of the drug for the particular ion channel. The affinity of a drug for an ion channel is described by its IC_{50} value and Hill coefficient, which correspond to the concentration at which 50% block of current flowing through an ion channel is achieved and the number of drug molecules required to block one ion channel respectively, as defined in Section 2.3. Simple pore block theory can be represented within mathematical models of ventricular cell electrophysiology by scaling the conductance of the ion channel j by a factor b defined by

$$b = \frac{1}{\left(1 + \left(\frac{[IC_{50}]_j}{[D]}\right)^n\right)}, \tag{3.28}$$

where $[D]$ is the concentration of drug, $[\text{IC}_{50}]_j$ is the IC_{50} value for channel j and n is the corresponding Hill coefficient. The conductance of any channels with which the drug is known to interact can be scaled in this way. For example, a value of b = 0.2 corresponds to a 20% reduction of the maximal conductance and so a 20% block of the channel by the drug. Simulation of ventricular cell models with drug block represented in this way allows the expected effects of different drug concentrations on the action potential to be examined. This approach for representing drug block within cardiac electrophysiology models has been used, for example, by Starmer et al. (1994), Starmer et al. (1995), Bottino et al. (2006), Mirams et al. (2011), Obiol-Pardo et al. (2011), Davies et al. (2012), Zemzemi et al. (2013) and Mirams et al. (2014). We also use this method for representing drug block in Chapter 4.

Starmer et al. (1994) and Starmer et al. (1995) used the simple pore block assumption to investigate the mechanisms by which a potassium channel-blocking drug suppressed responses of cardiac cells to premature stimuli and also to investigate the anti-arrhythmic effects of potassium channel block.

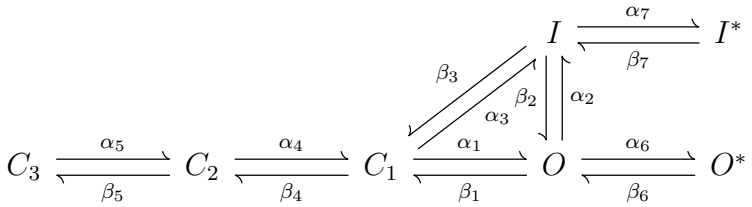
Zemzemi et al. (2013) present a multiscale modelling study investigating drug-induced effects on ventricular electrical activity from the ion channel level through to cellular electrical activity and up to measurement of body surface potentials. They use a three-dimensional model of the human body in order to simulate the effects of drugs on the ECG measured on the surface of the thorax. Drug block is simulated using the simple pore block theory for the hERG, fast sodium and L-type calcium channels.

Drug block of different ion channels can affect the morphology of the action potential in distinct ways due to the role the different ionic currents play in generating and maintaining the action potential. For example, as I_{Kr} is important in action potential repolarisation, block of the hERG channel has been linked with prolongation of the action potential duration and consequently QT prolongation.

When simulating ion channel block effects on the action potential, due to the feedback mechanisms existing between the different ionic currents within cardiac cells, it is important to incorporate information about drug interactions with multiple ion channels. For example, the drug ranolazine, has a strong affinity for blocking hERG, but due to its interactions with additional channels, it does not appear to induce extensive QT prolongation, nor to be pro-arrhythmic (Noble and Noble, 2006). We provide a further demonstration of this in Section 4.3.2.

3.4.2 Markov Model Describing Drug-Ion Channel Interactions

Another way to represent drug block within mathematical cardiac cell models is to extend directly the Markov model representation of ion channel kinetics with drug blocked states. In this case, all the states in which the drug is thought to be able to bind to the channel have an additional blocked state represented. As an example, if a drug was assumed to block a channel when it was in the open or inactivated states, then the following Markov model could be proposed to describe the drug-ion channel interactions.



(3.29)

Here, the states marked with asterisks represent drug block states. In Equation (3.29), α_6 and α_7 are equal to the association rates in the open and inactivated states respectively multiplied by the drug concentration. These rates are termed k_{on} rates and describe the rate of binding of the drug to the open and inactivated states respectively. Similarly, β_6 and β_7 are the rates of unbinding from the open and inactivated states respectively, and are termed k_{off} rates. The affinity of the compound for the channel can be represented by the ratio of the k_{on} and k_{off} rates, in the same way that the affinity was described by the IC_{50} value when using the simple pore block assumption for drug binding. For each state with which the drug is thought to interact, the system of ODEs describing the ion channel kinetic model in the absence of drug block can be extended to incorporate the additional drug blocked states. The system

of ODEs describing the model in Equation (3.29) would be formulated using mass action equations of the form shown in Equation (3.6).

Such a model represents drug block under the assumption of state-dependent binding. Such a representation overcomes the limitation of the simple pore block theory which assumes that drug block of the channel is not dynamic and is immediately in steady state. The Markov model approach for representing drug block also has the ability to describe important drug effects such as rate- and use-dependence that cannot be represented by methods relying solely upon the IC_{50} value. Markov model representations of drug block can be incorporated directly within mathematical cardiac cell models and can be simulated with different drug concentrations to examine the effect of the drug on the action potential morphology. This approach has been used to represent drug block in Brennan et al. (2007), Clancy et al. (2007), Zhou et al. (2009), Moreno et al. (2011), Saiz et al. (2011), Di Veroli et al. (2013), Romero et al. (2014) and Hill et al. (2014).

Clancy et al. (2007) present an approach for investigating drug-induced effects on cardiac electrical activity using Markov model formulations of the sodium channels in normal and a mutated state associated with long-QT syndrome, ventricular arrhythmias and sudden cardiac death. They investigate the effects of interactions of the compounds mexiletine and lidocaine with the normal and mutated representations of sodium channels.

Moreno et al. (2011) also formulated a Markov model of the sodium channel. They used the modulated and guarded receptor theories, as well as considering allosteric effects, to extend their model to be able to replicate the behaviour of the channel in the presence of a selection of anti-arrhythmic drugs. They then incorporated the formulation into a computational model of the human ventricular myocyte to predict the effects of the drugs on normal human ventricular electrical activity. By investigating the vulnerable period in the presence of drugs, during which electrical stimulation could initiate arrhythmias in cardiac tissue, they were able to quantify drug-induced changes in arrhythmic risk.

Romero et al. (2014) systematically altered transition rates in the Fink et al. (2008) hERG channel Markov model to investigate the impact of altered hERG channel kinetics on the susceptibility to acquired long QT syndrome. In using the approach presented, they attempt to unravel the kinetic properties of I_{Kr}-blocking drugs that lead to an increased risk of long QT syndrome.

Hill et al. (2014) report the first instance of directly measured kinetics of drug binding and unbinding to the hERG channel for the example compound clozapine. They demonstrate that there are at least two kinetically distinct components to drug block and unblock, from which they are able to conclude that clozapine binds to the hERG channel when it is in open and inactivated states. They therefore propose a model with open and inactivated state binding to the hERG channel (as in Equation (3.29)), and describe clozapine interaction with the hERG channel in acquired long QT syndrome.

In Section 3.5 we discuss some *in silico* approaches that have been developed in collaboration with pharmaceutical companies for drug cardiac safety assessment.

3.5 *In Silico* Approaches for Industrial Pharmaceutical Cardiac Safety Assessment

Mathematical models of cardiac electrophysiology are becoming increasingly more detailed, and are beginning to be validated more thoroughly to assess their suitability for use during drug cardiac safety assessment. These *in silico* approaches have the potential to offer an efficient, ethical and very low cost approach for aiding the assessment of drug cardiac safety, and have begun to receive increased attention from the pharmaceutical industry in recent years. The unrestrictive nature of the ICH guidelines allowed pharmaceutical companies to explore new *in vivo*, *in vitro* and *in silico* methodologies for assessing the effects of drugs on cardiac electrical activity (Darpo, 2010; Picard et al., 2011). In light of the recent CiPA

proposals outlined in Section 2.4, this interest is likely to increase further in the coming years.

One *in silico* method commonly employed within pharmaceutical companies during early pharmaceutical development is Quantitative Structure Activity Relationship (QSAR) modelling. QSAR modelling uses information about the chemical structure of compounds to infer properties of their biological activity (Inanobe et al., 2008). For instance, a QSAR model could be used to predict the expected IC_{50} value of a compound based on its molecular structure through comparison to other compounds with similar chemical features. This allows the IC_{50} value of a compound for different ion channels to be estimated before beginning to synthesise it.

Pharmaceutical companies have also begun to explore the use of mathematical models of cardiac electrophysiology for predicting drug-induced changes to electrical activity in the heart. Bottino et al. (2006), in collaboration with Novartis, proposed a computational approach to predict the pro-arrhythmic risk biomarkers of APD and QT interval prolongation and transmural dispersion of repolarisation. To do this they estimated IC_{50} values of compounds for a range of channels using a reverse engineering approach to determine the information from an action potential recorded from a Purkinje fibre. This technique allowed the block of ion channels beyond the small number typically screened within pharmaceutical companies to be represented in the cardiac electrophysiology model when predicting drug-induced effects.

Mirams et al. (2011), in collaboration with GlaxoSmithKline, used concentration-effect data quantifying drug block of the hERG, late sodium and L-type calcium channels to predict the TdP risk category of a drug, using the classification system proposed by Redfern et al. (2003). Compounds were classified based on the maximum APD_{90} value from simulations at a range of drug concentrations. The method, evaluated using 31 pharmaceutical compounds, demonstrated a good level of predictive ability, much improved to that of the existing TdP risk indicator based on the compound's hERG IC_{50} value. This also confirmed that consideration of drug-ion channel interactions beyond solely considering block of the hERG channel

is necessary for predicting TdP risk.

Davies et al. (2012) at AstraZeneca, evaluated the ability of an *in silico* model to predict drug effects on the canine action potential in isolated myocytes using concentration-effect data for five cardiac ion channels; hERG, I_{Na}, I_{CaL}, I_{to}, and I_{Ks}. They also attempted to predict associated inter-individual variability and this is the first example of the use of an ensemble of models to gain an understanding of how drug effects may vary within a population. They evaluated their approach by comparing simulated and experimental measurements of the APD_{90} value for a set of 53 compounds. The method was found to demonstrate a good level of predictivity, with 78% sensitivity and 58% specificity for predicting whether a compound will cause greater than 10% prolongation or shortening in APD_{90}. When evaluating the approach on its ability to predict simply whether a compound will cause greater than 10% change in APD_{90}, the method demonstrated 82% sensitivity and 79% specificity.

Britton et al. (2013), working with Janssen Pharmaceuticals, also outlined an approach for predicting variability in action potential measurements based on construction of an ensemble of models. They calibrated their population of models describing action potential behaviour in the rabbit Purkinje fibre using the ranges of key features exhibited in experimental action potential recordings. They also indicated how this approach could be used to predict variability of drug-induced effects on the APD.

Di Veroli et al. (2013), in collaboration with AstraZeneca, compared the two approaches of representing drug block outlined in Section 3.4 for predicting drug-induced changes to the APD. Their electrophysiological experiments investigating drug-ion channel interactions made use of the novel voltage protocol described in Figure 3.3 (B), allowing for an improved assessment of drug-binding kinetics over time. They compared the predictions of drug effects on the canine action potential when assuming the simple pore block theory of drug block and when modelling drug block kinetics via a Markov model representation. They concluded from their study that neglecting details of drug-binding kinetics in mathematical

electrophysiology models when simulating drug effects may lead to an underestimation of drug-induced changes to the action potential. Di Veroli et al. (2014) extended this work to demonstrate further the importance of considering kinetics of drug-ion channel interaction for cardiac safety assessment, demonstrating that pro-arrhythmic liability depends on the mode of action and kinetics of binding of a drug, not solely on the hERG channel potency as determined by the IC_{50} value.

A strength of the work presented by Di Veroli et al. (2013) is that the methodology described is appropriate for use on the IonWorks high-throughput ion channel screen, and so, combined with automated post processing, it could demonstrate extremely high-throughput capabilities. However, further investigation of the ideas they presented would be of great interest. Firstly, a comparison of the use of different Markov model structures for representing the kinetics of the hERG channel would be interesting. The model representations used for both the room temperature and physiological temperature simulations of hERG channel kinetics assume that inactivation and activation are not coupled. This is opposite to the conclusion of Liu et al. (1996) and Wang et al. (1997), and was counter to the assumptions underpinning the development of a number of other models describing hERG channel kinetics. In addition, the model structure used for simulating hERG channel kinetics at physiological temperature is much simpler than that used for representing the kinetics at room temperature due to the limitation of not being able to distinguish between closed states with the data used for parameterising the model. It may therefore be of interest to consider the impact of this further, perhaps extending the experimental data set used to determine the model. The use of the same rates of drug binding and unbinding are also assumed in the two models formulated at different temperatures. Di Veroli et al. (2013) justify this by stating that IC_{50} values are comparable at the two temperatures (Kirsch et al., 2004) and by attributing observed differences in kinetics at room and physiological temperatures to temperature-dependent effects at the level of the whole channel rather than at the level of drug-ion channel interaction. However, further comparison of using this approach, as well as reparameterising the rates at different temperatures, again using additional experimental data at physiological temper-

atures, would be of interest. Additionally, the methodology has been presented for three example compounds, an extension to a larger number of compounds, and evaluation of these results, would also be valuable.

Elkins et al. (2013) quantified the variability associated with concentration-effect curves fitted to recordings from a high-throughput ion channel screen for control compounds used at AstraZeneca and GlaxoSmithKline. The authors investigated how variability in the ion channel screening data affects predictions made using *in silico* action potential models and demonstrated how confidence intervals could be placed on such model predictions.

Mirams et al. (2014), working with GlaxoSmithKline and AstraZenca, evaluated an *in silico* approach of using ion channel screening data to predict the results of the human thorough-QT study, which is suggested by the ICH E14 to be performed as standard during clinical development. Using the conductance block representation of drug block of ion channels, these authors simulated the effects of drug block on the human APD for 34 compounds. Mirams et al. (2014) used the methodology described in Elkins et al. (2013) to quantify the uncertainty in the model predictions due to variability in the ion channel screening assay. While demonstrating a proof-of-concept method for predicting the results of the human thorough-QT study, they also indicate possible refinements which could lead to improved predictivity of such an approach.

In the light of the CiPA proposal objectives, it is expected that *in silico* approaches for the assessment of drug-cardiac safety will increasingly be explored and adopted by pharmaceutical companies.

3.6 Summary

In this chapter we have introduced the mathematical foundations of the models used in this thesis for describing electrical activity in the heart. We have discussed mathematical rep-

resentations of cardiac electrophysiology from the kinetics of ion channels to describing electrical activity in individual cardiomyocytes and through to propagation of electrical signals in cardiac tissue, all of which are utilised in this thesis. The physical foundations and assumptions associated with each modelling approach have been outlined. We have also included an historical perspective on the development of different mathematical representations of cardiac electrophysiology. In addition, we have described how drug-ion channel interactions can be represented within mathematical models of cardiac electrophysiology. We have reviewed some *in silico* approaches, developed in collaboration with pharmaceutical companies, for the assessment of drug-induced changes to electrical activity in the heart.

In Chapters 5 – 7 we focus on mathematical representation of hERG ion channel kinetics. In Chapter 5 we design novel voltage protocols and evaluate their suitability for use for determining mathematical models of hERG channel kinetics. We assess the ability of novel protocols to determine parameters of existing hERG channel models introduced in this chapter. In Chapter 6 we describe the techniques used to perform these experiments using the whole-cell patch clamp technique and in Chapter 7 we use the experimental data collected to construct and validate cell-specific models of hERG channel kinetics.

Before focusing specifically on the representation of hERG channel kinetics, in Chapter 4 we present an evaluation of an *in silico* approach for predicting changes in cardiac electrical activity observed in an animal-based cardiac safety assay. This computational approach integrates information about the extent of drug block of cardiac ion channels using data from high-throughput screens within mathematical action potential models to assess changes in cardiac electrical activity.

4

Using Ion Channel Screening Data and Action Potential Simulations to Predict QT Interval Changes in the Rabbit Ventricular Wedge

Measurement of the QT interval on the ECG with different drug concentrations is one of the main biomarkers used for assessing cardiac drug safety. In this chapter we evaluate a computational approach for predicting the results of a rabbit-based cardiac safety assay performed at GlaxoSmithKline. The method uses data quantifying the degree of block of different ion channels by a compound which comes from one of two different high-throughput ion channel screens or QSAR, a statistical model. Data quantifying compound block of the hERG, fast sodium, L-type calcium and KCNQ1 channels are used within simulations using a biophysical model of the rabbit ventricular myocyte, to predict the QT interval expected in the rabbit wedge assay in response to different drug concentrations. Using a range of performance metrics, we compare the predictive ability of the approach using data from different ion channel screening technologies, considering the effect of including compound interactions with different numbers of ion channels, and performing both single cell and one-dimensional tissue simulations. The approach demonstrates a good level of predictive ability and is suitable for very high-throughput use in early drug development. This approach is now routinely used within Safety Pharmacology at GSK for screening compounds prior to performing the rabbit wedge assay. Integrating ion channel

data from routine screens to infer results of animal-based tests in this way, provides a cost- and time-effective cardiac safety screen. Most of the content of this chapter has been published in Beattie et al. (2013), with some text included in the publication also appearing in this chapter.

4.1 Introduction

As discussed in Section 2.3, drug interactions with cardiac ion channels can cause disturbances to normal electrical activity in the heart. In particular, such interactions can lead to prolongation of the QT interval on the ECG (Grant, 2009), which has been linked with potential cardiac safety liabilities, and the incidence of arrhythmias. As the present regulatory guidelines require an assessment of the effects of a compound on the QT interval, it is important for pharmaceutical companies to be able to detect QT prolonging compounds as early, and at as low a cost, as possible during drug development. A recommendation of the International Committee for Harmonization (ICH) guidelines is that the drug block of the hERG channel is assessed, due to the link between hERG channel block and QT prolongation. However, drug interactions with additional ion channels, such as the sodium and L-type calcium channels, and the I_{Ks} current-carrying channel KCNQ1/minK, are also often routinely screened by some pharmaceutical companies.

In Section 3.5, we outlined a number of *in silico* approaches for cardiac safety assessment which have been developed in collaboration with pharmaceutical companies. Such approaches have great potential for alleviating the high levels of attrition associated with current drug development practices; as low cost, high-throughput methods for aiding drug cardiac safety assessment. Each of these approaches use information about drug block of cardiac ion channels; typically hERG, NaV1.5, CaV1.2 and KCNQ1 channels, to predict the changes in action potential morphology at different drug concentrations. Each of these approaches have been shown to demonstrate a reasonable level of predictive ability. In this chapter we evaluate another such *in silico* approach, using high-throughput ion channel

screening data to predict changes in QT interval observed in the rabbit left-ventricular wedge cardiac safety screen, which is performed during drug development at GSK. In Chapter 1 we alluded to the importance of validation of computational approaches used for predictive purposes. This evaluation includes an unprecedented number of compounds, to thoroughly assess the predictive capacity of our *in silico* screen. We compare the predictive ability of the approach using data from different ion channel screens and using different combinations of the available data. We also compare the use of simulations of electrical activity in single cells with a one-dimensional tissue representation. While much of the content of this chapter has been published in Beattie et al. (2013), we include additionally here a comparison of the use of two different mathematical models of rabbit ventricular myocyte electrophysiology to select the most appropriate for use for the *in silico* approach. We also include some additional observations made, in particular with regard to the correspondence between predictions using single cell and one-dimensional tissue simulations as described on Page 90 and a discussion on the convergence of performance metrics with increasing number of compounds as described on Page 95.

4.2 Methods

In this study we use the experimental data obtained from the high-throughput ion channel screens at GSK to simulate the effects of drug block of different ion channels by different compounds and then compare the simulation results to the experimental recordings obtained at GSK where the compounds are applied in the rabbit wedge preparation. All the experimental work in this chapter was undertaken by our collaborators at GSK.

As described in Section 3.4, the amount by which a compound blocks a particular channel can be quantified by an IC_{50} value. The IC_{50} value indicates the concentration at which a drug is expected to block peak current from a channel by 50%. When a compound is screened for ion channel block, the degree of block at different concentrations is examined. This is determined by fitting a curve through the data quantifying the ion channel block (as

determined by the reduction in peak current) at different drug concentrations. The form of this curve is a Hill curve; as described by Equation (4.1)

$$b = \frac{1}{\left(1 + \left(\frac{[IC_{50}]_j}{[D]}\right)^n\right)}, \tag{4.1}$$

where, b is the current blocking factor of the channel due to the effect of the drug, $[D]$ is the compound concentration, $[\mathrm{IC}_{50}]_j$ is the IC_{50} value for channel j and n is the corresponding Hill coefficient. Fitting of the Hill coefficient was constrained in the optimisation algorithm so that a value in the experimentally expected range (between 0.5 and 5) was obtained. We wrote a C++ implementation of the Nelder and Mead (1965) optimisation algorithm to fit the data for this purpose. We can use the Hill curve and interpolate or extrapolate to determine an estimate of the IC_{50} value. A typical set of data outputs from a high-throughput screen for one compound for one channel is shown in Figure 4.1.

In practice, where more than a 50% block of an ion channel was observed during screening with the PatchXpress and IonWorks/FLIPR assays, the IC_{50} value and Hill coefficient (Hill, 1910) were calculated automatically and used directly as the inputs for simulation. Details of the methodology used to do this are described in the supplementary material of Beattie et al. (2013). Where a 50% block was not observed at the compound concentrations applied during ion channel screening, we fitted concentration-effect data acquired from the assays with a curve of the form in Equation (4.1).

4.2.1 Ion Channel Screening Data

We call the data quantifying the extent of block of each channel at different compound concentrations concentration-effect data. We use data from two ion channel screens commonly used within pharmaceutical companies: the Molecular Devices PatchXpress 7000A assay and the Molecular Devices IonWorks Quattro assay. The PatchXpress 7000A (referred to as the PatchXpress assay) was used to screen hERG, NaV1.5 and CaV1.2 channels and the Molecular Devices IonWorks Quattro (referred to as the IonWorks assay) was used to screen

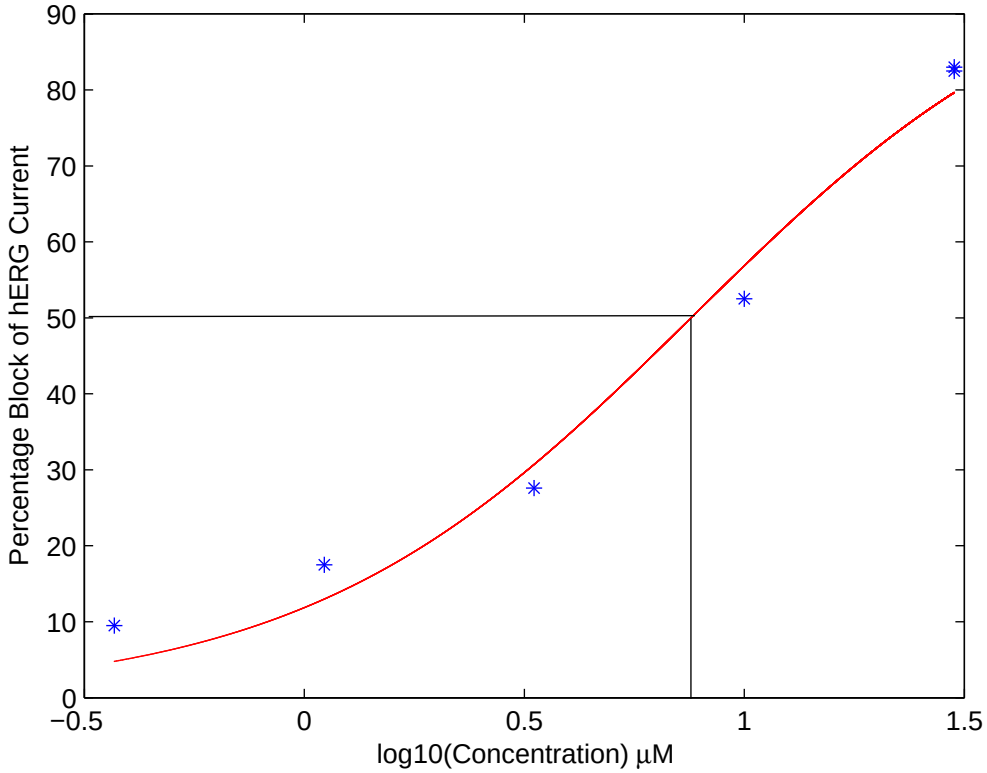


Figure 4.1: Example dose response curve fitted to concentration-effect data for the hERG channel from the PatchXpress screen for compound GSK1 (We rename the compounds for purpose of anonymity throughout this chapter). Interpolation is used to determine an IC_{50} value, that is the concentration at which a 50% block of the channel is observed, from the fitted curve as indicated.

hERG, NaV1.5 and KCNQ1 channels. The Molecular Devices FLuorescence Imaging Plate Reader (FLIPR) assay was used to screen the CaV1.2 channel to be used alongside data from the IonWorks assay. The concentration-effect data acquired in each screen quantifies the change in current in response to the application of different compound concentrations.

Both the PatchXpress and IonWorks assays perform parallel data acquisition for multiple doses of different compounds at the same time. Plates with discrete wells containing different compound concentrations are used in each assay. The PatchXpress assay plates have 16 such wells for 16 simultaneous recordings. The PatchXpress screen allows for the addition of multiple drug concentrations to each cell, enabling concentration response curves to be generated from one cell. The IonWorks assay uses population patch clamp technology

which enables the recording of up to 64 cells in parallel in each well and the machine allows recordings from either 96 or 384 well plates. The measured current from each well in the IonWorks assay is then an average of the ensemble of cells. The PatchXpress assay produces higher-quality data than the IonWorks Quattro assay, but has a lower-throughput capability. The quality of PatchXpress ion channel screening data is not as good as manual patch clamp, but the PatchXpress screen offers a two- to ten-fold increase in throughput. The IonWorks screen offers up to a four hundred-fold increase in throughput capabilities as compared to manual patch clamp (Fermini and Priest, 2008).

Human Embryonic Kidney-293 (HEK-293) and Chinese Hamster Ovary (CHO) cells are used for ion channel screening performed by electrophysiologists at GSK. Full details of the cell culture and preparation process, the solutions prepared for ion channel screening and controls can be found in the supplementary material of Beattie et al. (2013). Voltage protocols applied to obtain IC_{50} value estimates from the IonWorks and PatchXpress assays and the method of fluorescence measurement used in the FLIPR assay are also outlined in the supplementary material of Beattie et al. (2013). Concentration-effect data acquired from each of the assays were used to determine IC_{50} values for each channel, as described above.

In addition to data acquired from the PatchXpress and IonWorks/FLIPR assays, IC_{50} values were also acquired from QSAR models for the hERG, NaV1.5 and CaV1.2 channels. The QSAR models were created by computational chemists at GSK using an automated modelling software application (QSAR Workbench (Accelrys Inc.)). Models based upon Support Vector Machines (SVM) (Burbidge et al., 2001) were used for the hERG and NaV1.5 channels, and a random forest classifier (Svetnik et al., 2003) was used for the CaV1.2 channel. The output from the SVM was a predicted IC_{50} value for each compound. The output for the CaV1.2 channel was a predicted class which was then equated with an IC_{50} value. Data obtained from the QSAR model already take the form of an IC_{50} value and so were also used directly. However, as only an IC_{50} value was recorded in the QSAR model, the Hill coefficient was always assumed to be 1.

4.2.2 Aims

Our aims were to evaluate a computational approach of predicting QT prolongation and QT shortening in the rabbit wedge cardiac safety assay when using different combinations of available ion channel screening data. We wanted to do this to establish the best use of available experimental data and computational approach for achieving the highest ability to predict the results observed in experiment. Our evaluation included; 77 compounds with data from the PatchXpress assay, 121 with data from the IonWorks/FLIPR assay, and 372 compounds with predicted IC_{50} values from the QSAR model. Predictive capacity of the computational approach is compared by:

1. assuming drugs interact with just the hERG channel or with multiple channels (hERG, NaV1.5, CaV1.2 and KCNQ1);

2. parameterising the drug block model with just the IC_{50} value (and assuming the Hill coefficient is 1) or both the IC_{50} value and Hill coefficient;

3. using IC_{50} values obtained from data from the IonWorks/FLIPR and PatchXpress ion channel assays and the QSAR statistical model; and

4. performing single cell and one-dimensional tissue simulations.

4.2.3 Simulating Cardiac Electrical Activity

Two biophysical models of the rabbit ventricular myocyte were considered for use in this evaluation: the Shannon et al. (2004) and Mahajan et al. (2008) models. Both models were modified to incorporate compound interactions with the hERG, NaV1.5, CaV1.2 and KCNQ1 channels. Compound block was represented using the simple pore block mechanism, as described in Section 3.4.1, by scaling the maximal conductance of each channel by the factor in Equation (4.1). The ionic current of each channel is then expressed as

$$I_j = g_j(1-b)O(V-E_j), \tag{4.2}$$

where g_j is the maximum conductance of channel j, O is the probability of the channel being in the open state, V is the membrane potential and E_j is the reversal potential of the channel.

The Shannon et al. (2004) model was derived from the Puglisi and Bers (2001) rabbit ventricular myocyte model and is described in detail in Section 3.2.2. The Mahajan et al. (2008) model is a modified version of the Shannon et al. (2004) rabbit ventricular myocyte model. The Mahajan et al. (2008) model includes an updated description of I_{CaL}, with a Markov formulation replacing the Hodgkin-Huxley formulation in the Shannon et al. (2004) model. It also includes a new description of intracellular calcium handling. To incorporate these new features into the Mahajan et al. (2008) model, the ventricular myocyte model was tuned by the authors to replicate the experimentally measured APD and intracellular calcium transient at rapid pacing rates. This may consequently explain the observations by Romero et al. (2011) that while both the Shannon et al. (2004) and Mahajan et al. (2008) models are able to reproduce many characteristics of rabbit electrophysiology, the Shannon et al. (2004) model is able to describe behaviour more accurately at low pacing rates, whereas the Mahajan et al. (2008) model is better able to replicate observed behaviour at faster pacing rates. In Section 4.3.1 we perform a preliminary comparison of predictions made using both the Shannon et al. (2004) and Mahajan et al. (2008) models to determine which mathematical model to use for the full evaluation.

Pacing was initiated in the Shannon et al. (2004) model by applying a stimulus current of magnitude 9.5 $\mu A/\mu F$ for a duration of 5 ms. In the Mahajan et al. (2008) model, pacing was initiated by applying a stimulus of magnitude of 15 $\mu A/\mu F$ for a duration of 3 ms. The magnitude and duration of these pacing protocols (and the units they are expressed in) correspond to the square wave stimuli as defined in the models' CellML representations, which are taken from the original published models. Pacing was continued at 0.5 Hz, corresponding to the pacing frequency performed in the rabbit ventricular wedge experiments. The stimulus was applied at 0.5 Hz until a steady state was reached. Steady state is defined to be when the square root of the sum of the squared differences of the state variables at the

start of successive paces is less than 10^{-6}. The action potential was determined by solving

$$\frac{\mathrm{d}V}{\mathrm{d}t} = -\frac{1}{C_m}\left(\sum_j I_j + I_{stim}\right), \tag{4.3}$$

where V is the membrane voltage, C_m is the membrane capacitance, I_j represents the current from each channel j represented in the model, and I_{stim} is the stimulus current applied to pace the cell. We record the APD_{90} and its percentage change relative to control at each concentration.

The one-dimensional simulation setup is analogous to a line of coupled cells, as described in Section 3.3, forming a 1 cm long homogenous fibre. Simulation nodes were spaced 0.1 mm apart. The monodomain Partial Differential Equation (PDE), as described in Equation (3.27) was solved, using a time step of 0.01 ms and an adaptive ODE time step for Equation (4.3), to determine the voltage across the tissue over time. To assess convergence of the solution with our choice of PDE time step, we compared the conduction velocity across the tissue and the calculated steady state APD_{90} value of the last node in the chain in both control and drug block conditions when using a range of different time steps. Our selected time step produced APD_{90} values and conduction velocities which were consistent with those obtained when finer PDE time steps were used.

A stimulus of 250,000 μ A/cm^2 was applied at one end of the fibre for 5 ms in order to initiate propagation. The fibre was paced at 0.5 Hz until steady state was reached, with steady state being defined to be when the difference in the APD_{90} at the central node on successive time steps was less than 0.1 ms. The intracellular conductivity was set to 0.174 S/m (Bishop et al., 2010). A pseudo-ECG, as defined in Gima and Rudy (2002), was determined by measuring the membrane potential through a virtual electrode situated at the central node of the one-dimensional chain. A QT interval duration was estimated from the simulated pseudo-ECG, defined as being from the time of application of the stimulus to the minimum of the T-wave on the pseudo-ECG (noting that the T-wave is inverted on the ECG of a homogeneous one-dimensional chain as compared to a heterogeneous simulation). An

example simulation of a pseudo-ECG in control conditions indicating how the QT interval is measured is shown in Figure 4.2. Simulations were performed at the drug concentrations tested in the rabbit wedge assay, as well as at lower concentrations of 0.001 μM, 0.01 μM, 0.1 μM and 1 μM where these were not included in the experimental assay. The percentage change in QT interval at each drug concentration as compared to the control was recorded.

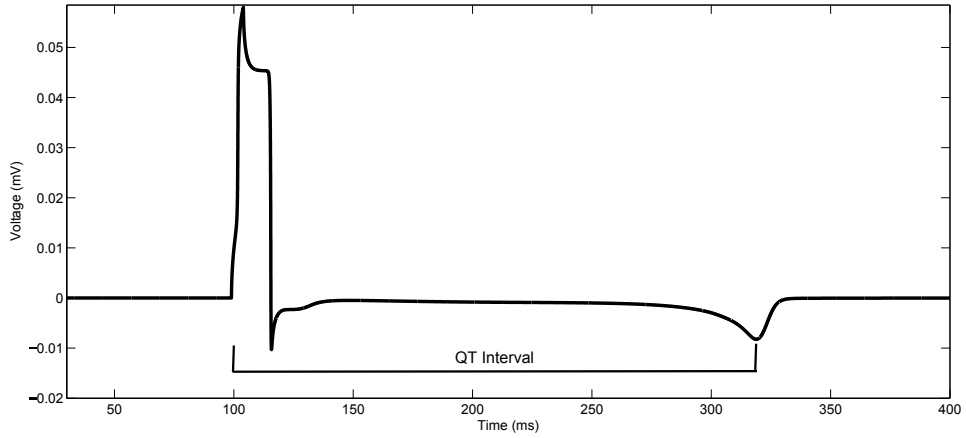


Figure 4.2: Simulation of pseudo-ecg using Shannon et al. (2004) rabbit ventricular myocyte model in control conditions, with QT interval as measured in the one-dimensional tissue simulations indicated.

XML representations of the Shannon et al. (2004) and Mahajan et al. (2008) rabbit action potential model were taken from the CellML repository (Lloyd et al., 2008) and PyCML (Garny et al., 2008) was used to translate the CellML format into C++ code. Simulations were performed using the Chaste library (Mirams et al., 2013), with PETSc PDE solvers (Balay et al., 2012) and the CVODE package for ODEs (Hindmarsh et al., 2005), based upon adaptive time stepping with numerical differentiation formulae, using relative and absolute tolerances of 10^{-6} and 10^{-8} respectively (Cooper et al., 2014). Chaste, CellML and CVODE are described in more detail in Appendix C.1.

4.2.4 Rabbit Ventricular Wedge Assay

Data measuring QT prolongation in response to different drug concentrations was acquired from the rabbit left-ventricular wedge assay performed by safety pharmacologists at GSK. Full details of the assay be found in Liu et al. (2006). Describing the preparation briefly: female rabbits were sedated with 6 mg/kg xylazine (intramuscular administration), anticoagulated with 800 U/kg heparin (i.v.) and anaesthetised with ketamine (3035 mg/kg, i.v.), or with pentobarbital (50 mg/kg, i.v.). The left circumflex or anterior descending branch of the coronary artery of the excised rabbit heart was cannulated and perfused in cardioplegic solution. A transmural left-ventricular wedge was dissected and placed in a tissue bath and arterially perfused with Tyrode's solution. After approximately 1 hour of equilibration in the bath at a stimulation frequency of 1 Hz, the stimulation frequency was reduced to 0.5 Hz for 5 min of stabilisation where the baseline ECG was measured. The preparations were then returned to a stimulation frequency of 1 Hz and perfused with Tyrode's solution containing a test compound. For each test compound concentration, the preparation was perfused for approximately 30 min at a frequency of 1 Hz followed by 5 min at a frequency of 0.5 Hz, where again the ECG was recorded.

4.2.5 Statistical Evaluation

For evaluation of the *in silico* approach, the simulated percentage changes in APD_{90}/QT interval were compared to the percentage changes in the QT interval recorded in the rabbit left-ventricular wedge assay. Where multiple experimental repeats were performed, an average of the assay repeats was used for comparison. A number of statistical measures were used to assess the ability of the approach to predict both QT prolongation and shortening in the rabbit ventricular wedge assay, which we describe below.

Experimental and simulation results for each drug used in the evaluation were independently classified into one of two categories. When evaluating the ability to predict QT prolongation, if at any experimental drug concentration, the experimental/simulation result exhibited more

than a 10% change in the recorded QT interval/APD_{90} value, the compound was classified as a prolonger of the QT interval in the experimental/simulation result as appropriate. Otherwise, the compound was classified as a non-prolonger. For the prediction of QT shortening, if at any concentration tested the percentage change in APD_{90} value/QT interval was less than -10%, the compound was classed as a shortener. Otherwise the compound was classified as a non-shortener. 10% was chosen as the threshold for defining prolongation and shortening as GSK Safety Pharmacologists advised that such a change in QT interval is a cause for concern and warrants further investigation during pharmaceutical development.

Results classified in this way can be presented in a classification matrix, and a range of performance-related metrics can be calculated to assess the agreement between experiment and simulation results. The *overall accuracy* is the proportion of compounds for which the experimental and simulation classification categories coincide respectively. The *sensitivity* and *specificity* represent the proportion of true positive and true negative results correctly identified. The *positive* and *negative predictive values* describe how likely a positive or negative simulation result is to be true. The formulae used for calculating each of these metrics are included in Table 4.1, as also calculated in Fawcett (2006). We calculate these metrics when assessing the ability of the approach described for predicting both QT prolongation and shortening in the rabbit left-ventricular wedge experiments.

An additional metric used for comparing the predictive ability is the *Kappa value* (κ) (Cohen, 1968). The Kappa value is a measurement of the agreement between simulation and experimental results while also compensating for chance agreement, as determined by the distribution of experimental results across each category. The Kappa value is calculated using the equation

$$\kappa = \frac{p_a - p_e}{1 - p_e}, \tag{4.4}$$

where p_a is the proportion of compounds where simulation and experimental classification categories coincide and p_e is the proportion of compounds where agreement is expected by chance. p_e is calculated by multiplying the proportion of compounds with simulation results

in a given category by the proportion of compounds whose experimental results are in that category and summing these results for all categories. A value between 0 and 0.2 represents poor agreement, 0.2–0.4 represents fair agreement, 0.4–0.6 moderate agreement, 0.6–0.8 good agreement and 0.8–1 very good agreement (Cohen, 1968).

95% confidence intervals are included for each metric calculated. The Kappa value is assumed to be normally distributed when the data set being considered is sufficiently large, as in our evaluation. Intervals of size 1.96 times the standard error above and below the calculated value can then be taken. The standard error is calculated as described by Cohen (1968). Confidence intervals on the accuracy, sensitivity, specificity, positive and negative predictive values are obtained using Wilson's score for binomial proportions, as described in Wilson (1927).

4.3 Evaluation Results

Before presenting the evaluation results from this study, our first step was to select a biophysical model of rabbit ventricular myocyte electrophysiology to be used for performing the evaluation.

4.3.1 Selecting a Biophysical Rabbit Ventricular Myocyte Model

To decide which mathematical model of the rabbit ventricular myocyte to use for the purpose of evaluation, we compared the predictive ability of the approach when using the Shannon et al. (2004) and Mahajan et al. (2008) rabbit ventricular myocyte models. Figure 4.3 shows simulated percentage change in APD_{90} at a range of drug concentrations for a compound which we will call GSK2, when using the Shannon et al. (2004) and Mahajan et al. (2008) models, as compared to the QT percentage change recorded in experiment.

The classification matrices in Table 4.1 presents the statistical metrics calculated when assessing the ability to predict QT prolongation using the Shannon et al. (2004) and Mahajan

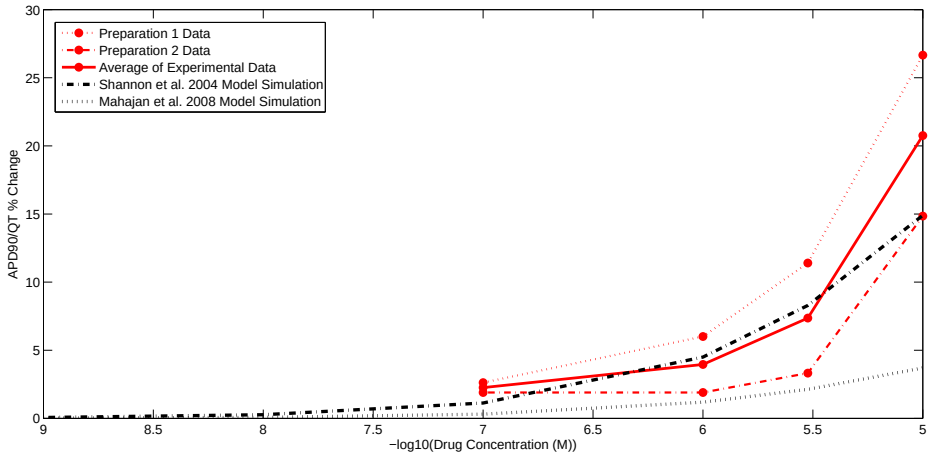


Figure 4.3: Comparison of experimental and simulation results when using the Shannon et al. (2004) and Mahajan et al. (2008) models to simulate the change in APD_{90} at different concentrations of compound GSK2.

et al. (2008) models. In this evaluation, data from the PatchXpress ion channel screen was used, it was assumed that drugs could interact with the hERG, NaV1.5 and CaV1.2 channels and both IC_{50} values and Hill coefficients were used to parameterise the drug block model in Equation (4.1).

The classification matrices presented in Table 4.1 demonstrate the inability of the Mahajan et al. (2008) model to predict large QT prolongation, with no compounds predicted to induce more than a 10% change in the APD_{90} value from the simulation results. This is also evident for compound GSK2 shown in Figure 4.3. This apparent inability of the Mahajan et al. (2008) model to predict QT prolongation could be explained by the differences in I_{Kr} conductance between the Shannon et al. (2004) and Mahajan et al. (2008) models. The maximal I_{Kr} conductance is $0.03\ mS/\mu F$ in the Shannon et al. (2004) model and 0.0125 $mS/\mu F$ in the Mahajan et al. (2008) model. The reduced I_{Kr} conductance in the Mahajan et al. (2008) model formulation translates into a prediction of a lesser change in APD_{90} with I_{Kr} block. The full set of metric values calculated when comparing the use of the Shannon et al. (2004) and Mahajan et al. (2008) models for the prediction of QT prolongation and QT shortening are shown in Table 4.2.

Table 4.1 Classification matrices presenting categorical classification of simulation and experimental results when assessing the ability to predict QT prolongation when using the Shannon et al. (2004) (top) and Mahajan et al. (2008) (bottom) ventricular myocyte model. The results were obtained when using PatchXpress assay data, assuming compounds can interact with multiple channels and parameterising the drug block model with both the IC_{50} value and the Hill coefficient for a set of 77 compounds. The accuracy, sensitivity, specificity, and positive and negative predictive values are calculated as indicated.

Shannon et al. (2004) Model	Simulation Prolonger	Simulation Non-Prolonger	
Experimental Prolonger	**18**	**7**	Sensitivity = $\frac{18}{18+7}$ = **72.0%**
Experimental Non-Prolonger	**10**	**42**	Specificity = $\frac{42}{42+10}$ = **80.8%**
	Positive Predictive Value = $\frac{18}{18+10}$ = **64.3%**	Negative Predictive Value = $\frac{42}{42+7}$ = **85.7%**	Accuracy = $\frac{18+42}{18+7+10+42}$ = **77.9%**

Mahajan et al. (2008) model	Simulation Prolonger	Simulation Non-Prolonger	
Experimental Prolonger	**0**	**25**	Sensitivity = $\frac{0}{0+25}$ = **0.0%**
Experimental Non-Prolonger	**1**	**51**	Specificity = $\frac{51}{51+1}$ = **98.1%**
	Positive Predictive Value = $\frac{0}{0+1}$ = **0.0%**	Negative Predictive Value = $\frac{51}{51+25}$ = **67.1%**	Accuracy = $\frac{0+51}{0+25+1+51}$ = **66.2%**

Table 4.2 Prolongation and shortening metric values calculated (with 95% confidence intervals indicated) for evaluation of the ability to predict QT prolongation and QT shortening in the rabbit ventricular wedge experiments when using the Shannon et al. (2004) and Mahajan et al. (2008) rabbit ventricular myocyte models. Ion channel screening data from the PatchXpress assay was used, and drugs were assumed to interact with the hERG, NaV1.5 and CaV1.2 channels and both IC_{50} values and Hill coefficients were used in the drug block representation.

Model	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
Shannon et al. (2004) (QT Prolongation)	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**77.9** *(67.5-85.7)*	**0.51** *(0.31-0.72)*	**72.0** *(52.4-85.7)*	**80.8** *(68.1-89.2)*	**64.3** *(45.8-79.3)*	**85.7** *(73.3-92.9)*
Mahajan et al. (2008) (QT Prolongation)	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**66.2** *(55.1-75.8)*	**-0.03** *(-0.35-0.30)*	**0.00** *(0.00-13.3)*	**98.1** *(89.9-99.7)*	**0.00** *(0.00-79.3)*	**67.1** *(55.9-76.6)*
Shannon et al. (2004) (QT Shortening)	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**75.3** *(64.6-83.6)*	**0.24** *(-0.06-0.54)*	**33.3** *(16.3-56.3)*	**88.1** *(77.5-94.1)*	**46.2** *(29.3-63.9)*	**81.3** *(70.0-88.9)*
Mahajan et al. (2008) (QT Shortening)	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**79.2** *(68.9-86.8)*	**0.21** *(-0.14-0.55)*	**16.7** *(5.84-39.2)*	**98.3** *(91.0-99.7)*	**75.0** *(11.8-98.5)*	**79.5** *(68.8-87.1)*

The results presented in Table 4.2 further illustrate the reduced ability of the Mahajan et al. (2008) model to predict QT prolongation. When using the Mahajan et al. (2008) model for simulating the action potential, the Kappa value for predicting QT prolongation has a negative value which indicates that categorisation by chance would be superior to using this approach. The Mahajan et al. (2008) model performs better when predicting QT shortening, with some metrics calculated being higher than those calculated when using the Shannon et al. (2004) model. However, overall, an improved level of predictive ability is demonstrated when using the Shannon et al. (2004) model.

Following this preliminary evaluation we proceeded to use the Shannon et al. (2004) model to perform a full evaluation comparing the predictive performance when using different combinations of the available ion channel screening data and when performing both single cell and one-dimensional tissue simulations, the results of which are presented in Section 4.3.2.

4.3.2 Evaluation of Assay Predictivity

We now present the full evaluation results with each performance metric calculated when using different combinations of the available ion channel screening data when using the Shannon et al. (2004) biophysical model for performing simulations. Table 4.3 presents the performance metric values calculated when predicting QT prolongation in the rabbit wedge assay when considering different uses of the available ion channel screening data. Table 4.4 presents the performance of the assay when predicting QT shortening.

Table 4.3 Prolongation metric values calculated (with 95% confidence intervals indicated) for evaluation of the assay's ability to predict QT prolongation in the rabbit ventricular wedge experiments when considering the different uses of the available ion channel data described using the Shannon et al. (2004) ventricular myocyte model. Entries are ordered alphabetically by the data type used in the simulations.

Case Number	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
1	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value *(121 Compounds)*	**77.7** *(69.5-84.2)*	**0.50** *(0.33-0.67)*	**73.0** *(57.0-84.6)*	**79.8** *(70.0-87.0)*	**61.4** *(46.6-74.3)*	**87.0** *(77.7-92.8)*
2	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill coefficient *(121 Compounds)*	**75.2** *(66.8-82.0)*	**0.45** *(0.28-0.62)*	**70.3** *(54.2-82.5)*	**77.4** *(67.4-85.0)*	**57.8** *(43.3-71.0)*	**85.5** *(75.9-91.7)*
3	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value *(121 Compounds)*	**77.7** *(69.5-84.2)*	**0.51** *(0.34-0.67)*	**75.7** *(59.9-86.6)*	**78.6** *(68.7-86.0)*	**60.9** *(46.5-73.6)*	**88.0** *(78.7-93.6)*
4	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value *(77 Compounds)*	**72.7** *(61.9-81.4)*	**0.48** *(0.29-0.67)*	**96.0** *(80.5-99.3)*	**61.5** *(48.0-73.5)*	**54.5** *(40.1-68.3)*	**97.0** *(84.7-99.5)*
5	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value *(77 Compounds)*	**74.0** *(63.3-82.5)*	**0.41** *(0.18-0.63)*	**60.0** *(40.7-76.6)*	**80.8** *(68.1-89.2)*	**60.0** *(40.7-76.6)*	**80.8** *(68.1-89.2)*
6	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**77.9** *(67.5-85.7)*	**0.51** *(0.31-0.72)*	**72.0** *(52.4-85.7)*	**80.8** *(68.1-89.2)*	**64.3** *(45.8-79.3)*	**85.7** *(73.3-92.9)*
7	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value, One-Dimensional Simulation *(77 Compounds)*	**71.4** *(60.5-80.3)*	**0.41** *(0.20-0.62)*	**76.0** *(56.6-88.5)*	**69.2** *(55.7-80.1)*	**54.3** *(38.2-69.5)*	**85.7** *(72.2-93.3)*
8	QSAR Data, Multiple Channel Block *(372 Compounds)*	**55.4** *(50.3-60.3)*	**0.17** *(0.08-0.27)*	**91.4** *(86.0-94.8)*	**27.6** *(22.0-34.0)*	**49.3** *(43.7-55.0)*	**80.6** *(70.0-88.0)*

Table 4.4 Shortening metric values calculated (with 95% confidence intervals indicated) for evaluation of the assay's ability to predict QT shortening in the rabbit ventricular wedge experiments when considering the different uses of the available ion channel data described using the Shannon et al. (2004) model. Entries are ordered alphabetically by the data type used in the simulations. The missing entry is due to the inability to calculate the metric due to there being no compounds in the relevant categories for its computation.

Case Number	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
1	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value *(121 Compounds)*	**69.4** *(60.7-76.9)*	**0.13** *(-0.10-0.37)*	**29.0** *(16.1-46.6)*	**83.3** *(74.3-89.6)*	**37.5** *(24.7-52.3)*	**77.3** *(68.0-84.5)*
2	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill coefficient *(121 Compounds)*	**71.1** *(62.4-78.4)*	**0.23** *(0.02-0.45)*	**41.9** *(26.4-59.2)*	**81.1** *(71.8-87.9)*	**43.3** *(30.0-57.8)*	**80.2** *(70.9-87.1)*
3	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value *(121 Compounds)*	**69.4** *(60.7-76.9)*	**0.13** *(-0.10-0.37)*	**29.0** *(16.1-46.6)*	**83.3** *(74.3-89.6)*	**37.5** *(25.0-51.9)*	**77.3** *(68.0-84.5)*
4	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value *(77 Compounds)*	**76.6** *(66.0-84.7)*	**0.00** *(-0.40-0.40)*	**0.00** *(0.00-17.6)*	**100** *(93.9-100)*	**-**	**76.6** *(66.0-84.7)*
5	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value *(77 Compounds)*	**76.6** *(66.0-84.7)*	**0.26** *(-0.04-0.56)*	**33.3** *(16.3-56.3)*	**89.8** *(79.5-95.3)*	**50.0** *(31.8-68.2)*	**81.5** *(70.4-89.1)*
6	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**75.3** *(64.6-83.6)*	**0.24** *(-0.06-0.54)*	**33.3** *(16.3-56.3)*	**88.1** *(77.5-94.1)*	**46.2** *(29.3-63.9)*	**81.3** *(70.0-88.9)*
7	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value, One-Dimensional Simulation *(77 Compounds)*	**72.7** *(61.9-81.4)*	**0.22** *(-0.06-0.51)*	**38.9** *(20.3-61.4)*	**83.1** *(71.5-90.5)*	**41.2** *(26.5-57.6)*	**81.7** *(70.1-89.4)*
8	QSAR Data, Multiple Channel Block *(372 Compounds)*	**79.8** *(75.5-83.6)*	**0.10** *(-0.09-0.28)*	**14.1** *(7.58-24.6)*	**93.5** *(90.2-95.8)*	**31.0** *(26.1-36.5)*	**84.0** *(79.7-87.5)*

The metric values presented in Tables 4.3 and 4.4 aid in identifying the most advantageous use of the available ion channel data to maximise the predictive ability of this approach. Comparisons of further combinations of the use of available ion channel screening data can be found in Appendix B. The highest valued performance metrics, when predicting both QT prolongation and shortening, were found when using data from the PatchXpress assay, and when assuming that drugs interact with multiple ion channels (case 6 in Table 4.3 and case 5 in Table 4.4). There is an improved ability to predict QT prolongation when using information about both the IC_{50} value and the Hill coefficient in the drug block representation (case 6) as opposed to just the IC_{50} value (assuming the Hill coefficient is 1) (case 5). QT prolongation is predicted with 77.9% accuracy and QT shortening is predicted with 76.6% accuracy; according to the European Centre for the Validation of Alternative Methods (ECVAM) (Genschow et al., 2002), these levels of accuracy correspond to good levels of predictive ability. Kappa values of 0.51 for predicting QT prolongation and 0.26 for predicting QT shortening were also found, corresponding to moderate and fair agreement respectively (Cohen, 1968).

Multiple Ion Channel Interactions

To illustrate the impact of including drug interactions with different combinations of ion channels on the simulated action potential, Figure 4.4 shows the simulated action potential using data from the IonWorks/FLIPR assays when including interactions with just the hERG channel, with the hERG, NaV1.5 and CaV1.2 channels and with the hERG, NaV1.5, CaV1.2 and KCNQ1 channels when simulating the effects of drug block for a compound that we will call GSK3.

In Figure 4.4, the arrows indicate the direction of change of the APD with increasing drug concentration. We can see that when considering hERG channel block alone the APD_{90} is predicted to increase with increasing drug concentrations for this particular compound. Considering interactions with the NaV1.5 and CaV1.2 channels in addition leads to a prediction of APD shortening with increasing drug concentrations. Considering additional interactions

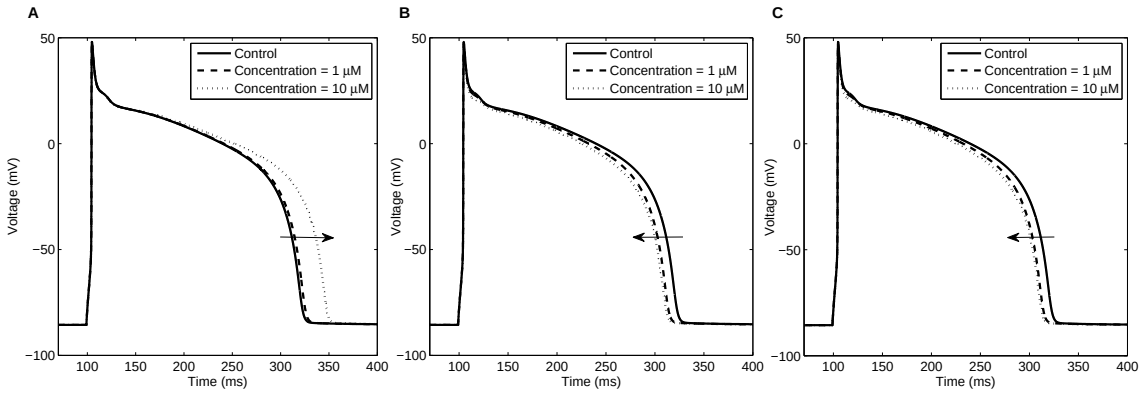


Figure 4.4: Comparison of simulated action potentials with different drug concentrations when including drug interaction with different numbers of channels by the compound GSK3 using data from the IonWorks/FLIPR assays and using information about both the IC_{50} value and Hill coefficient. In A it is assumed that GSK3 interacts with only the hERG channel. In B it is assumed that GSK3 interacts with the hERG, NaV1.5 and CaV1.2 channels and in C it is assumed that GSK3 interacts with the hERG, NaV1.5, CaV1.2 and KCNQ1 channels. Arrows indicate the direction of the change in APD with increasing drug concentration.

with the KCNQ1 channel causes little change to the simulated action potential, with a decrease in APD with increasing drug concentrations but to a slightly lesser extent than when considering drug interaction with just the hERG, CaV1.2 and NaV1.5 channels.

When using data from all three ion channel screening assays, improved predictive ability was demonstrated when assuming drugs interact with multiple channels (hERG, NaV1.5, CaV1.2 (and KCNQ1 for the IonWorks assay)) as opposed to just the hERG channel. Using additional ion channels demonstrates increased accuracy, as well as an improved balance between sensitivity and specificity. Increased predictive ability when including interactions with additional channels is most apparent when predicting QT shortening. When assuming compounds interact only with the hERG channel (case 4), a Kappa value of 0 is obtained. This indicates the same level of agreement as expected if classification had been performed by chance. A sensitivity value of 0% is observed for QT shortening. By assuming that drugs interact solely with the hERG channel, the model is unable to predict shortening, as block of the hERG channel always leads to APD_{90}/QT prolongation, with no additional ion channel

interactions to counterbalance this effect.

Inclusion of drug interactions with the KCNQ1 channel, in addition to the hERG, fast sodium and L-type calcium channels, when using IonWorks/FLIPR assay data leads to only a small improvement in predictive ability. Comparing cases 1 and 3, for the prediction of QT prolongation, the inclusion of interactions with the KCNQ1 channel yields no increase in accuracy, and a 2.7% improvement in the sensitivity (75.7% as compared to 73.0%). Whilst only a slight improvement, this is concordant with the increased importance of interactions with the KCNQ1 channel that may be manifested when drugs block the hERG channel (consequently leading to increased APD prolongation).

We also compare calculated performance metrics when predicting QT prolongation and shortening using a subset of 44 drugs for which data were available for all channels from all three ion channel assays and the rabbit wedge assay. The evaluation results for this subset of compounds are presented in Appendix B. Many of the same trends observed for the full compound sets were also observed for this subset of common compounds.

Including Hill Coefficients in Drug Block Representation

The predictive ability of the approach was also compared when using information about the IC_{50} value and Hill coefficient as compared to just the IC_{50} value in the parameterisation of the drug-block representation. The effect of this difference on simulation results was only subtle. Using PatchXpress assay data and assuming that drugs interact with multiple ion channels (cases 5 and 6) in Table 4.3, an improved ability to predict QT prolongation is observed when parameterising the drug block model with both the IC_{50} value and Hill coefficient (case 6 as compared to case 5 in Table 4.3). However, a decreased ability to predict QT shortening is seen (cases 5 and 6 in Table 4.4). When using IonWorks/FLIPR data, the opposite is seen. In these cases, the majority of the performance metric values being compared fall within each other's 95% confidence intervals, indicating that, where

differences are seen, these are not large.

Comparing the Use of Data from Different Ion Channel Screens

The assay with the highest level of predictivity for both QT prolongation and shortening is PatchXpress, which produces the highest quality and lowest throughput data used in this evaluation. When predicting QT prolongation, a Kappa value which indicates moderate agreement is obtained and an accuracy of 70-80% is obtained when using PatchXpress data (cases 4-7 in Table 4.3). A Kappa value which demonstrates fair agreement is obtained for the prediction of QT shortening and assuming drug interaction with multiple ion channels (cases 5-7 in Table 4.4), which is an improvement to when using IonWorks/FLIPR assay data (cases 1 and 3) or QSAR model data (case 8). When using IonWorks/FLIPR data and parameterising the drug block model with both the IC_{50} value and Hill coefficient (case 2), fair agreement is exhibited for the prediction of QT shortening. For predicting QT prolongation, and when assuming drugs interact with multiple ion channels, a similar level of predictive ability is seen when using IonWorks/FLIPR data (case 1), as when using PatchXpress assay data (case 6). An accuracy of 77.7% as compared to 77.9% and a sensitivity and specificity of 73.0% and 79.8% as compared to 72.0% and 80.8%, respectively, are obtained. This suggests it is possible to use IonWorks/FLIPR assay data in the *in silico* assay without a marked loss in accuracy. At GSK this data is available earlier in the drug development process than PatchXpress ion channel screening data.

Using the QSAR model as the source of ion channel data, the Kappa values calculated when predicting both QT prolongation and shortening indicate a small increase in predictive ability over chance alone (case 8 in Tables 4.3 and 4.4). However, when predicting QT shortening, the predictive ability is not much lower than that demonstrated when using IonWorks/FLIPR assay data (case 8 compared to cases 1 and 3 in Table 4.4). For all cases in Table 4.4, where IonWorks/FLIPR and PatchXpress assay data were used, the Kappa values calculated all fall within the 95% confidence interval calculated on the Kappa value when QSAR model data

is used. However, for the prediction of QT prolongation, the performance of the model when using QSAR data is markedly worse than when using data from the other assays. It can be seen that the Kappa values obtained in cases 1-7 do not fall within the 95% confidence interval of the Kappa value from evaluation of the QSAR model, illustrating this large difference. With further development and improvements to the QSAR model, there could be potential for its use as an ion channel data source for such an *in silico* approach, in the very earliest stages of drug development process. This could enable chemists to identify the potential for QT prolongation, and modify and optimise the design of compounds appropriately, before beginning compound synthesis.

Single Cell vs. One Dimensional Tissue Simulations

In Tables 4.3 and 4.4 we compared the predictivity of the approach when performing single cell simulations (case 5) as compared to one-dimensional tissue simulations (case 7), when using PatchXpress data, parameterising the drug block model with just the IC_{50} value (assuming the Hill coefficient is 1), and assuming drugs interact with multiple ion channels. An example of such a simulation for one compound, showing the change in the form of the pseudo-ECG from the one-dimensional tissue simulation, along with a comparison of change in the simulated action potential using a single cell model is shown in Figure 4.5.

The metrics presented in Tables 4.3 and 4.4 indicate that a reduced predictive ability is demonstrated when predicting both QT prolongation and shortening when performing one-dimensional simulations as compared to single cell simulations (an accuracy of 71.4% as compared to 74.0% for QT prolongation and an accuracy of 72.7% as compared to 76.6% and Kappa value of 0.22 as compared to 0.26 for QT shortening). However, the same Kappa value of 0.41, corresponding to moderate agreement, is yielded in both cases when predicting QT prolongation, and a higher sensitivity (and consequently lower specificity) is associated with the use of one-dimensional simulations in the prediction of both QT prolongation and shortening. The metric values calculated when performing one-dimensional simulations fall

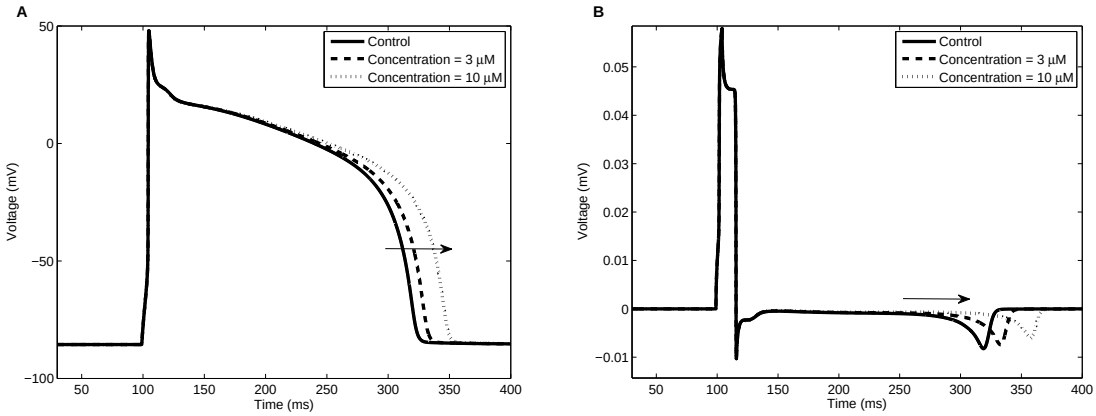


Figure 4.5: Simulated action potential (A) and pseudo-ECG (B) at different drug concentrations using single cell and one-dimensional tissue simulations respectively. Arrows indicate the direction of the APD90/QT interval calculated with increasing drug concentration.

within the 95% confidence intervals of the metrics calculated when performing single cell simulations, suggesting that the difference in the predictive ability demonstrated between the two methods is not large.

The measurement of the QT interval on the pseudo-ECG in these simulations was made using the minimum of the T-wave as a reference point for measurement due to the difficulty in precisely defining the end point of the T-wave. However, we also compared the results when identifying a point at the end of the T-wave for measurement. This point was located by searching for the time, after the minimum of the T-wave, at which the difference between two successive points on the pseudo-ECG was close to zero, approximating the point at which the gradient of the T-wave had returned to the resting level. We found that using this alternative measure for the QT interval calculation had no effect on the performance metrics calculated for predicting QT shortening and only a small effect on the metrics calculated for predicting QT prolongation.

We next investigated the correlation between the quantitative predictions made using single cell simulations and one-dimensional tissue simulations. The results obtained using PatchX-

press data using the IC_{50} value to parameterise the drug block model are presented in Figure 4.6. The one-dimensional tissue results correspond to those obtained when the QT interval is determined by measuring to the end of the T-wave.

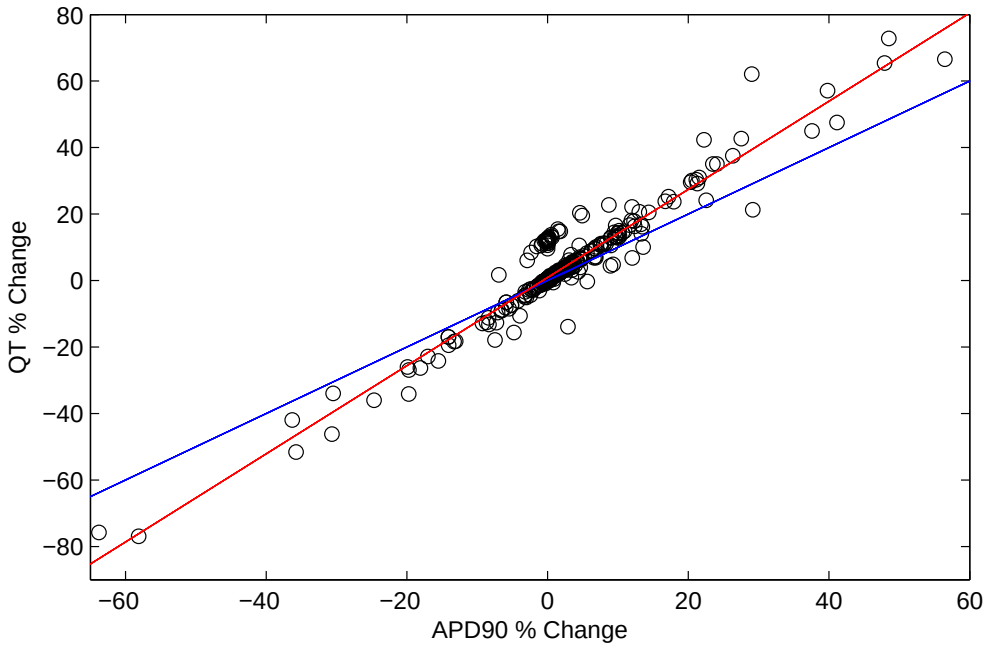


Figure 4.6: Plot of simulated APD_{90} vs QT interval for every concentration tested in each compound. Each circle marks the simulated APD_{90} and QT intervals at a particular drug concentration tested. The blue line indicates the identity line, if all points were to lie on this identity line then the predictions would be the same from both single cell and one-dimensional tissue simulations The red line indicates the line of best fit through the data which has a gradient of 1.34.

Good agreement between the simulated APD_{90} and QT interval changes is evident, with a ratio of 1:1.34 between change in APD_{90} and change in QT interval. This possibly suggests that a rescaling of the APD_{90} value yielded from single cell simulations may enable improved predictions of drug effects on the QT interval. In general it can be seen that the predictions obtained from the one-dimensional simulations tend to be slightly more extreme than those from the single cell simulations, with elevated levels of prolongation or shortening predicted as compared to the single cell simulations.

4.3.3 Evaluation of the Accuracy of Predictions

The performance metrics in Tables 4.3 and 4.4 allow an evaluation of the qualitative ability of the approach to predict QT "prolongation" or "shortening". Figure 4.7 allows us to gain an appreciation of the quantitative accuracy of these predictions. In Figure 4.7, the results of case 6, when using PatchXpress data and information about both the IC_{50} value and Hill coefficient, are presented. The figure includes compounds for which at least one of the simulation or experimental APD_{90} value/QT interval measurements demonstrates more than a 10% change at one of the drug concentrations tested as compared to the control. For these compounds, the percentage change in APD_{90}/QT interval and corresponding concentrations tested are interpolated to determine the concentration at which a 10% (for results classified as demonstrating prolongation) or -10% (for results classified as demonstrating shortening) change in APD_{90} value/QT interval is expected (EC_{10} value). The EC_{10} values from the experimental and simulation results for each compound are plotted against one another in one of the quadrants 1 to 4 marked in Figure 4.7.

The half log unit lines included in Figure 4.7 indicate the margin of error commonly associated with the ion channel assays used and so can be thought to provide an indication of the boundaries of acceptable error. Where less than a 10% change was observed in either the experimental or simulation results, but more than a 10% change was observed in the other result, the EC_{10} values for compounds are plotted on the relevant axis. Most compounds can be seen to be plotted in quadrants 1 and 3, indicating qualitative agreement between simulation and experimental results. Most of these points are also plotted within, or close to, half log unit boundaries, further indicating quantitative accuracy of the predictions. Few compounds are plotted in quadrants 2 and 4, which would indicate that simulation and experimental predictions were opposite (one being shortening and the other prolongation). Points plotted along the inner axes of Figure 4.7 are located mostly along the central and inner portions of the axes (closest to the origin). This indicates that where a disagreement between the classification of experimental and simulation results occurs (with one predicting that the

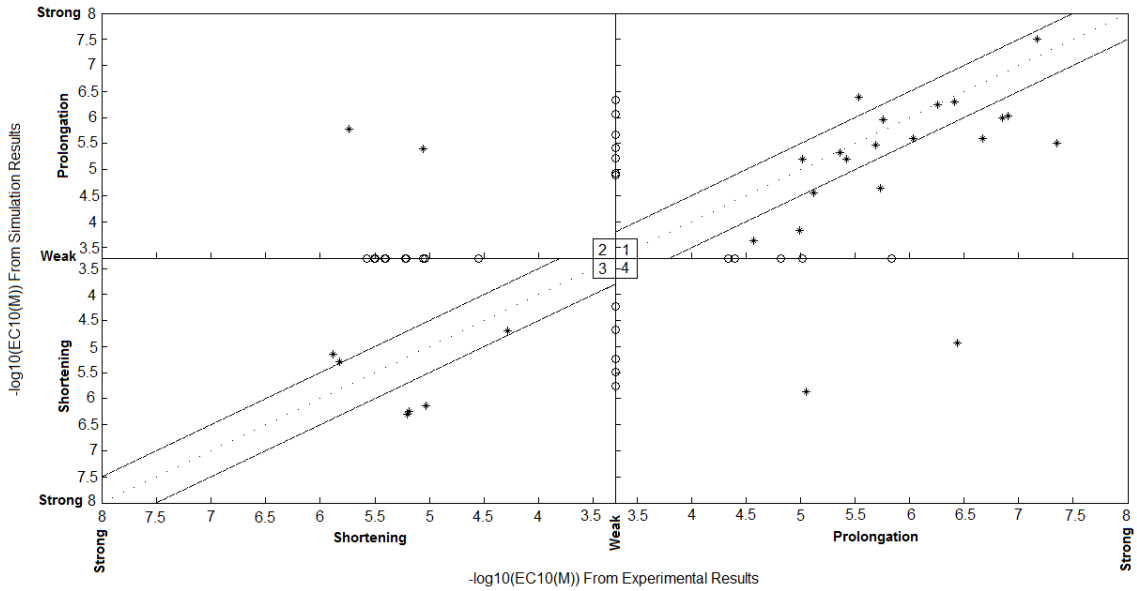


Figure 4.7: Plot of concentrations (EC_{10} values) at which a 10% change in APD_{90} value/QT interval is expected, as interpolated from simulation and experimental results, for each compound. The results from case 6, using PatchXpress data, are presented. Compounds exhibiting more than 10% change in their experimental or simulation results after drug administration, as compared to the control measurement, at at least one test concentration are included. Points plotted with an asterisk (*) are compounds for which more than a 10% change in the QT interval length (as compared to the control measurement) is exhibited in both simulation and experimental results. These points are plotted in one of four quadrants according to the classification of the simulation and experimental results: both experimental and simulation results show prolongation (quadrant 1), both show shortening (quadrant 3), experimental results show shortening and simulation results show prolongation (quadrant 2), simulation results show shortening and experimental results show prolongation (quadrant 4). Points plotted with a circle (o) are compounds for which only one of the experimental and simulation results show more than a 10% change in the QT interval as compared to the control. For the remaining experimental or simulation result, the percent change in QT interval is between -10% and 10% (showing "no effect") at all concentrations tested. This result is assumed to have an EC_{10} value corresponding to the maximum concentration tested in the rabbit wedge assay amongst all the compounds which show "no effect" (with the percent change in APD_{90} or QT interval between -10% and 10%), which has a value of 500 μM. Half log unit lines which correspond to the error commonly associated with the ion channel assays are included as an indication of the range of accepted error.

compound will induce more than a 10% change in the QT interval and the other not), it is often the case that the EC_{10} value is low, corresponding to a relatively weak action. Thus, whilst a qualitative agreement is not evident using the binary classification system, there is still a degree of quantitative agreement between the simulation and experimental results.

4.4 Limitations

The results presented in Section 4.3 assess the ability of the *in silico* approach described in this chapter to predict the results of the rabbit left-ventricular wedge assay using integrated concentration-effect data for four cardiac ion channels (hERG, NaV1.5, CaV1.2 and KCNQ1). While the level of predictive ability demonstrated is reasonable, elevated levels of predictive ability would be necessary before this approach could be considered as a replacement for such an animal-based assay. A number of limitations of the approach presented here could be addressed in order to improve the predictive ability. Compound block is represented in the mathematical model using the simple pore block theory, as describe in Section 2.3. While this is a commonly adopted approach, and is easy to represent using readily available high-throughput ion channel screening data, it is a simple representation and lacks the ability to describe complex drug-ion channel interactions, as discussed in Section 3.4. The incorporation of a Markov model drug block formulation, which could represent the kinetics of drug-ion channel interactions would thus be desirable, as has recently been explored by Di Veroli et al. (2013). This is the motivating factor for the focus of the research presented in the remainder of this thesis and the limitation of this approach which we decide to address first.

The comparison of the use of single-cell and one-dimensional tissue simulations in Section 4.3.2 was performed to acknowledge that differences inherent in tissue may not be fully captured by single cell models. This initial comparison demonstrated a small increase in sensitivity when performing one-dimensional simulations as compared to single cell simulations. Other limitations which could be addressed include exploring the use of higher dimensional simulation setups, more representative of the experimental conditions of the rabbit left-ventricular wedge assay, and the resulting predictive ability evaluated. We could also consider the addition of a perfusing bath to the simulation setup. The inclusion of interactions with additional ion channels, as well as further increasing the number of compounds used for evaluation would add further weight to findings and conclusions drawn from this

evaluation. Additionally, the variability associated with high-throughput screens can be considerable. The uncertainty associated with high-throughput ion channel screens has been quantified in Elkins et al. (2013), where a method is proposed for estimating the subsequent confidence in simulation outputs when using such data.

Variability is also apparent across the screening platforms depending upon the cell type used and the voltage protocol used. The ion channel screening methods detailed in the supplementary material of Beattie et al. (2013), demonstrate such variations between the methods adopted which may also have contributed to differences in predictive ability observed when using data acquired from different ion channel screens. In addition, the ion channel screens are performed at room temperature. If experiments were instead to be performed at physiological temperatures, any temperature-dependent effects would be eliminated and an increase in agreement between simulation and experimental results may be observed.

4.5 Discussion

We have evaluated the ability of an *in silico* approach which integrates concentration-effect data from high-throughput ion channel assays to predict the results of the rabbit left-ventricular wedge assay. The predictive ability of the approach was found to be good, particularly for predicting QT prolongation, which is the main focus of present cardiac safety assessment guidelines. The approach is also able to predict QT shortening, albeit with lower levels of sensitivity demonstrated. Furthermore, the approach has been shown to demonstrate good quantitative agreement with experiment, as discussed in Section 4.3.3.

A substantial number of drug compounds have been included for evaluation; 77 compounds with data from the PatchXpress assay, 121 with data from the IonWorks/FLIPR assay, and 372 compounds with predicted IC_{50} values from the QSAR model. The unprecedented number of compounds used for evaluation is the main strength of this work, as we illustrate through exploring the calculated performance metric values when including increasing num-

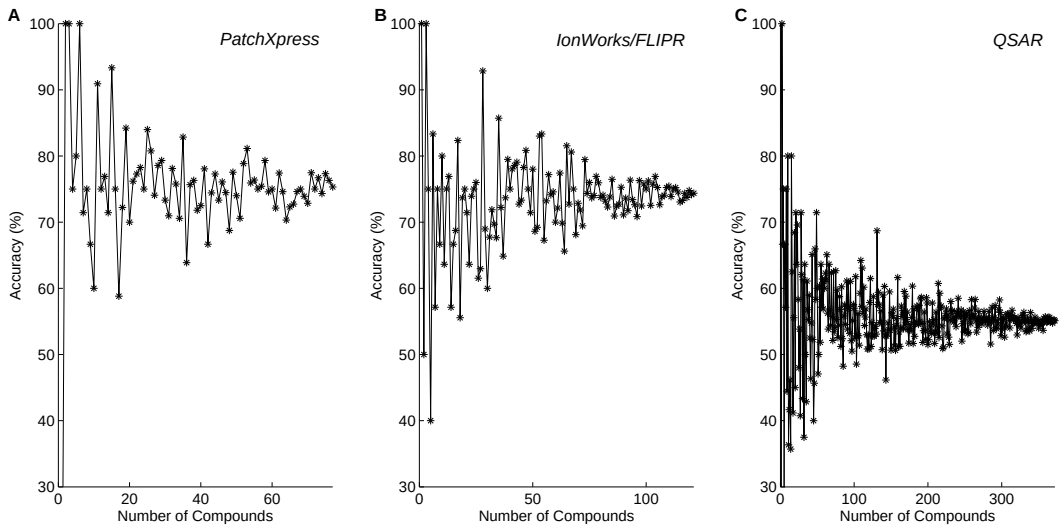


Figure 4.8: Plot showing convergence of calculated accuracy metric when including increasing numbers of compounds for evaluation when using data from each of the three assays. In A results using PatchXpress assay data are shown, in B data from the IonWorks/FLIPR assay were used and in C results using data from the QSAR model are shown.

bers of compounds for evaluation, an example of which is shown in Figure 4.8.

It can be seen when using data from all three ion channel screening assays, with the inclusion of increasing numbers of compounds for evaluation, the accuracy metric value calculated begins to converge. When considering lower numbers of compounds, as is typically used for such an evaluation, the accuracy metric can still be seen to be fluctuating, and so conclusions based upon few compounds may not be as representative of the true predictive ability of an *in silico* approach. However, with the numbers of compounds used for evaluation in the approach described in this chapter we have an increased confidence in the evaluation results obtained and the conclusions drawn from these. Another point to note is that the results as presented in Figure 4.8 would clearly depend on the order the compounds were used to calculate the metric, particularly when only few compounds are included, before the metrics converge; this illustrates the necessity of including more than a handful of compounds for evaluation.

In the evaluation presented in this chapter we have identified the most advantageous use of

the available ion channel screening data. The highest level of predictivity was demonstrated using PatchXpress assay data as the source of concentration-effect data, assuming drugs interact with multiple ion channels (hERG, NaV1.5 and CaV1.2), and fitting the drug block model with both the IC_{50} value and Hill coefficient (case 6). Performing one-dimensional tissue simulations in place of single cell simulations did not appear to greatly enhance the predictive ability of the model. Using ion channel data for the KCNQ1 channel from the IonWorks/FLIPR assay, the predictivity of the model was compared when assuming drugs interact with the KCNQ1 channel in addition to the hERG, NaV1.5 and CaV1.2 channels (case 3 as compared to case 1). A slight improvement in the predictive ability was observed; however, in the compound set considered, there were few compounds exhibiting very strong block of the KCNQ1 channel (with the highest pIC_{50} value observed being 5.6). Inclusion of interactions with the KCNQ1 channel would likely have more of an influence on the level of predictivity demonstrated by the assay if the compound set had included more compounds with a strong affinity for this channel. On the basis that evaluation results are no worse with the inclusion of interactions with the KCNQ1 channel, if available, it would probably be advantageous to include these interactions.

The observation of improved predictive ability when assuming drugs interact with multiple ion channels is concordant with the findings of Mirams et al. (2011) and Obiol-Pardo et al. (2011). Improved predictivity when using PatchXpress assay data is as expected, due to this being the best quality ion channel screening data available for the study. However, it is encouraging to see a similar predictive ability when using IonWorks/FLIPR data, which is acquired earlier in the drug development process, at a lower cost and a higher throughput. This indicates the potential for the use of this *in silico* assay early in the drug development process, with the approach described in this chapter being suitable for very high-throughput use and has the potential to complement and reduce the use of the rabbit left-ventricular wedge assay. Our approach is now in routine use at GSK. It is used prior to performing the rabbit wedge assay to prioritise compounds for animal screening and to help establish the dose range to be tested in the experiment. Where compounds of a similar chemical class

show similar ion channel properties, and the simulation agrees with the rabbit wedge experiment for one of these compounds, GSK no longer run the rabbit wedge assay for the other similar compounds.

The level of predictive ability demonstrated when performing one-dimensional tissue simulations and single cell simulations is comparable, suggesting that the use of single cell simulations, which have lower computational cost, may be sufficient for this purpose. The investigation of the correlation between single cell and one-dimensional tissue simulation results shown in Figure 4.6 also raises the question as to whether a simple scaling of single cell simulation results could provide an improved prediction of tissue level simulation results.

Low predictive ability is demonstrated when using IC_{50} values from the QSAR model. This indicates the value of using direct measurement of ionic current inhibition from the PatchXpress and IonWorks screens. However, some predictive capacity is demonstrated by the QSAR method, with metrics indicating levels of predictive power above those achieved by chance alone, and this is improved when including interactions with multiple ion channels. Thus, with further enhancements of the QSAR model used, there may be future potential for its use with this assay. Such improvements could include replacing the class prediction for the CaV1.2 channel with an actual IC_{50} value as is the case with the hERG and NaV1.5 channels. Using only data from one test system could also decrease the subsequent data variability.

The effect on simulation predictions when parameterising the drug block model with just the IC_{50} value (and assuming a Hill coefficient of one), as has often previously been the method adopted (Bottino et al., 2006; Mirams et al., 2011), compared to using both the IC_{50} value and Hill coefficient has also been explored. The effects of including Hill coefficients in the drug block parameterisation are not conclusive. Improvements in predictivity were seen in some cases but reductions in others: in all cases the differences were only slight. Further investigation, comparing the impact of the inclusion of Hill coefficients in the drug

block model, dependent on the quantity and quality of concentration-effect data available for fitting would clarify this further.

4.6 Summary

In this chapter we have evaluated the ability of an *in silico* approach using concentration-effect data from high-throughput ion channel screens to predict the results of the rabbit left-ventricular wedge cardiac safety assay. The approach has been evaluated comparing the use of different combinations of available ion channel screening data, including: assessing the effects of incorporating drug block with different numbers of ion channels, comparing the use of data from different ion channel screening assays, as well as using information about both the IC_{50} values and Hill coefficients as compared to just IC_{50} values in the model representing drug block. Additionally we have compared the use of both single cell and one-dimensional simulations for predicting experimental results. These comparisons have allowed us to identify the best use of the available ion channel screening data for predicting the results of the rabbit ventricular wedge assay and the approach was found to demonstrate a good level of predictive ability. Such simulations are now performed routinely during safety assessment at GSK prior to conducting the rabbit wedge assay and this evaluation provides an indication as to the confidence with which the results should be interpreted. The *in silico* screen has reduced the number of rabbits used for testing, however, the current level of predictive ability is not sufficiently high that such an approach could be considered as a complete replacement for an animal-based assay.

We have identified some limitations of the approach described in this chapter, which could be addressed to attempt to improve the predictive ability. One such limitation is the use of the simple pore block theory for representing drug block. A Markov model representation of drug block may allow for an improved and more accurate description for drug-ion channel interactions and consequently improved predictions of drug-induced effects on cardiac electrical activity. In the remainder of this thesis we focus on determining the most appropriate

model for accurately describing the kinetics of the hERG channel, considering aspects of experimental design, model selection and parameter identifiability. In Chapter 5 we design and evaluate novel experiments for exploring hERG channel kinetics in addition to further demonstrating the need to determine an accurate representation for this channel for use in action potential models. We assess parameter identifiability of existing hERG channel kinetic models when using these new protocols. In Chapter 6 we describe how we perform these novel protocols in experiment and in Chapter 7 we use the novel protocols to construct and validate cell-specific models of hERG channel kinetics, as well as illustrating the value of considering cell-specific models.

5

Voltage Protocol Design for Determining a Model of hERG Channel Kinetics

In this chapter we focus on the representation of hERG channel kinetics within mathematical action potential models. We explore the behaviour of existing hERG channel models under a range of standard voltage-step and action potential clamp protocols. We also demonstrate the impact that the choice of hERG channel representation within an action potential model can have on the simulated action potential in both rabbit and human ventricular myocyte models. We describe novel voltage protocols, designed with the intention of aiding the selection of the most appropriate structure and parameterisation of models describing hERG channel kinetics. Using a Bayesian inference approach, we investigate the accuracy of parameter estimation within existing hERG channel models using standard voltage-step protocols, non-standard voltage protocols used in previous studies, and our novel protocols. Our protocols are shown theoretically to allow accurate estimation of many parameters across a range of existing model structures. In addition, the new protocols are able to do this in a much shorter time period than a traditional series of voltage-step protocols.

5.1 Motivation

In chapter 4 we evaluated an *in silico* approach for predicting the results of an animal-based cardiac safety assay using high-throughput ion channel screening data acquired in the early stages of drug development. As outlined in Section 4.4, there are a number of limitations of

the approach taken in Chapter 4 that could be addressed to attempt to improve its predictive ability. One such limitation is that drug block is represented within the mathematical model via simple pore block theory, as described in Section 3.4.1. Di Veroli et al. (2013) demonstrated that neglecting the kinetics of drug block may lead to an underestimation of the true effects of drugs on the action potential. Simple pore block theory is not able to describe the complexities of drug-ion channel interaction. In addition, the aims of the CiPA initiative, as described in Section 2.4, are to enable a more mechanistic understanding of pro-arrhythmic risk. As mathematical models are intended to contribute to these aims it would be advantageous for more detailed representations of drug-ion channel interactions to be included within these models, provided that this leads to enhanced predictive ability.

Consequently, the rest of this thesis focuses on the mathematical representation of ion channel kinetics. We focus in particular on the cardiac potassium channel encoded by hERG due to its promiscuity for block by a diverse range of drug compounds, the critical role it plays in action potential repolarisation, and the strong regulatory focus on hERG due to the link between block of this channel, QT interval changes on the ECG, and consequently pro-arrhythmic risk.

5.1.1 Behaviour of hERG Channel Models Using Standard Voltage Protocols

As described in Section 3.1.2, typically a series of voltage-step protocols are performed in patch clamp experiments to study the kinetics of ion channels in both control conditions and in the presence of drug compounds, using the methods described in Chapter 6. The standard protocols consist of a series of steps to different voltages which are held for a time period on the order of seconds before stepping to a different voltage. This process is then repeated a number of times with the total duration of each protocol being between tens of seconds to over a minute. The length of time necessary to perform the full series of protocols is then several minutes. It is assumed that activation and inactivation are independent in Hodgkin-Huxley style models. It is therefore typical for parameters describing activation in the model

to be determined from fitting to a protocol designed specifically for investigating activation kinetics, and inactivation parameters also from dedicated protocols, etc.

Parameters in Markov and Hodgkin–Huxley style models are often determined through calibration to experimental data using optimisation algorithms to minimise the difference between model predictions and observed data. As detailed in Table 3.1, a number of different models have been proposed to describe the kinetics of the hERG channel. Many of these are reparameterisations of existing models which were published with some evidence of being able to replicate characteristics of the experimental data to which the model was fitted. However, the majority of models were constructed using all of the experimental data recorded with none left for validation. Consequently, models can behave quite differently under voltage protocols different to those with which they were constructed. In addition, it is now thought that hERG channel inactivation and activation may not be independent (Liu et al., 1996). Consequently, methods for parameterising models may need to be re-evaluated, as such dependence would not be reflected by fitting model parameters independently to data from different protocols.

Typically, voltage protocols composed of a number of steps to different voltage levels are used to investigate hERG channel kinetics. Each protocol has a particular arrangement of voltage steps for investigating different aspects of hERG channel gating; activation, inactivation and deactivation. It is typically these types of voltage protocols which are used for determining the parameters of new mathematical models describing hERG channel kinetics. As described in Section 3.1, existing hERG channel models consist of a number of open, closed and inactivated states, with parameters describing the transition rates between each of the states. In addition, each model has a parameter describing the conductance of the channel which is used when calculating the current flowing through the channel; we recall that the current is calculated by

$$I_{Kr} = g_{Kr} O (V - E_{Kr}), \tag{5.1}$$

where g_{Kr} is the maximal conductance, O is the probability of the channel being open and $(V - E_{Kr})$ is the driving force due to the difference in the applied voltage (V) and the reversal potential (E_{Kr}). In Figures 5.1-5.4 we explore the behaviour of simulated currents from a selection of existing hERG channel models in response to different standard voltage-step protocols. These standard voltage-step protocols correspond to those considered in Bett et al. (2011), and are of a similar form to the protocols used to formulate the majority of existing hERG channel models. Simulations were performed using a conductance parameter normalised to 1 in each model. We also show the probabilities of the channel being in each state at each time point in the protocol for the Wang et al. (1997) model. In Figure 5.1 (A), voltage-step protocols are shown, with each step in the protocol labelled and the section of the probability plot corresponding to each step indicated in Figure 5.1 (B). In the current diagrams in Figures 5.1 ((C)-(F)), the currents corresponding to each voltage-step in the protocol are overlaid. The voltage-step protocols in Figures 5.2–5.4 are presented in a similar way.

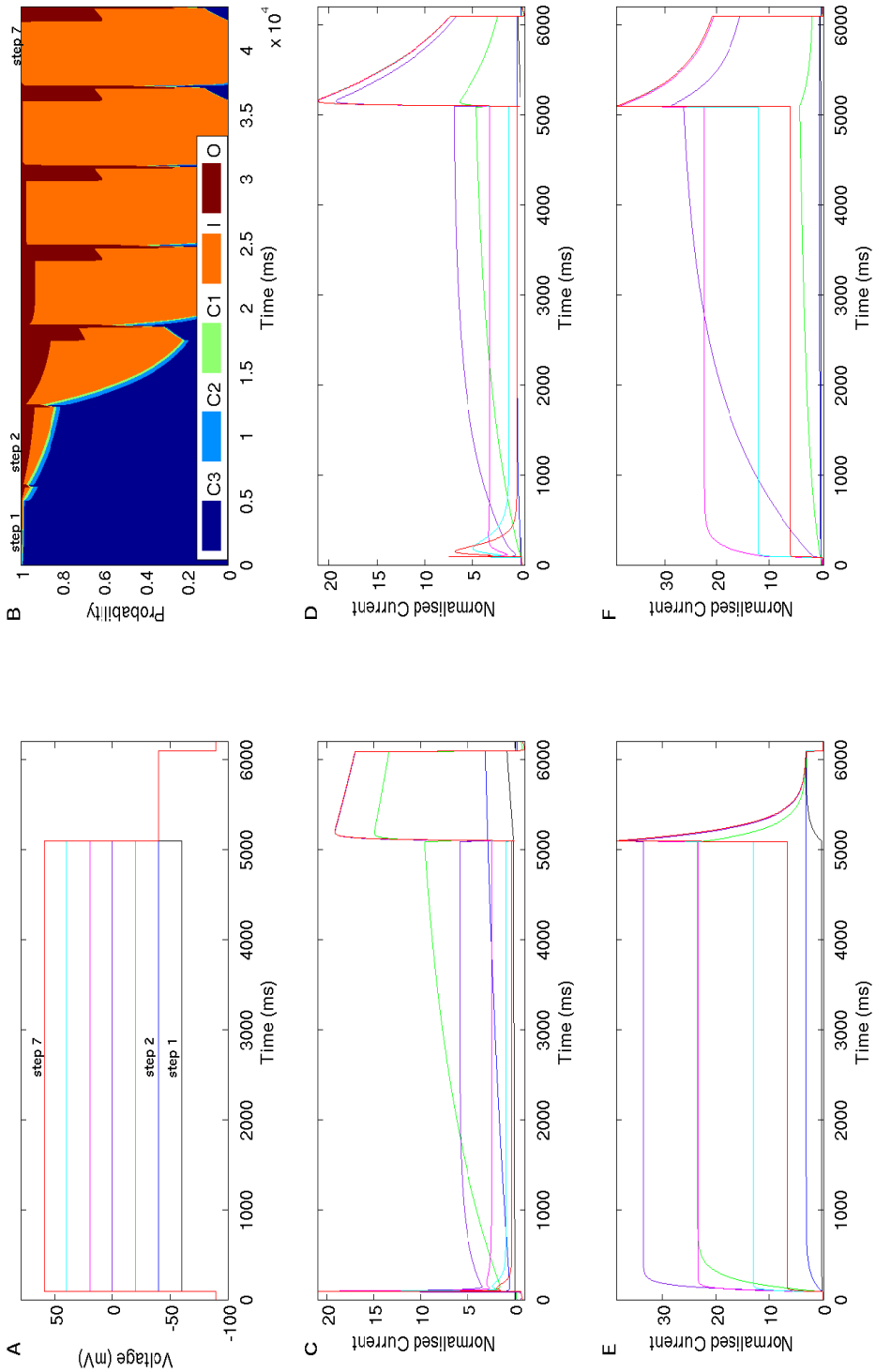


Figure 5.1: Simulated current traces in response to the standard voltage-step protocol shown in A which is intended to investigate steady-state activation kinetics. B shows the probability of the channel being in each state represented in the Wang et al. (1997) model at each time point in the protocol. Simulated current traces in C-F correspond to those from the Wang et al. (1997), Mazhari et al. (2001), Zeng et al. (1995) and O'Hara et al. (2011) hERG channel models respectively. Each colour in the voltage-step protocol in A corresponds to the current shown in the same colour in plots C-F.

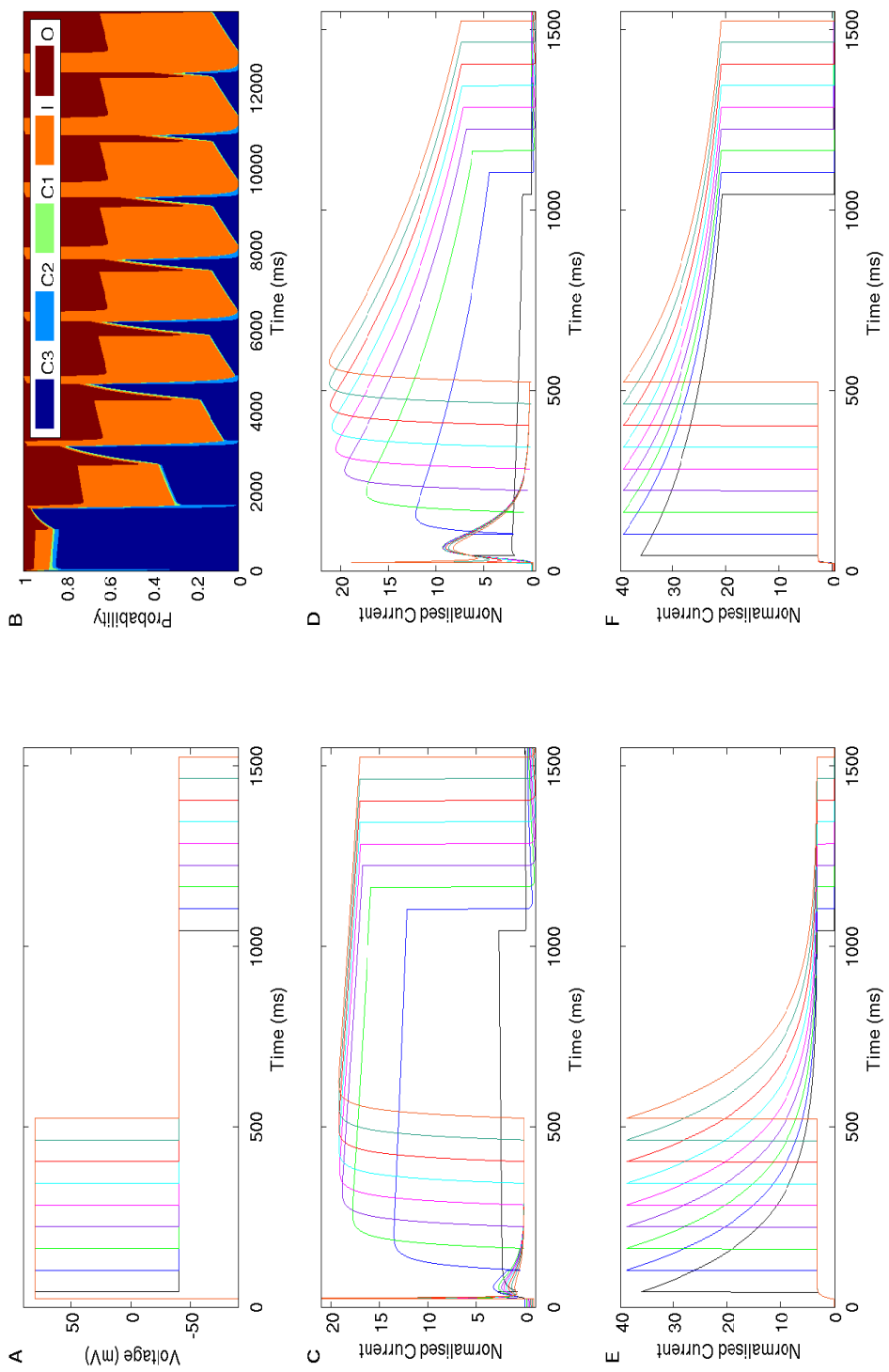


Figure 5.2: Simulated current traces in response to the standard voltage-step protocol shown in A which is intended to investigate activation kinetics. B shows the probability of the channel being in each state represented in the Wang et al. (1997) model at each time point in the protocol. Simulated current traces in C-F correspond to those from the Wang et al. (1997), Mazhari et al. (2001), Zeng et al. (1995) and O'Hara et al. (2011) hERG channel models respectively. Each colour in the voltage-step protocol in A corresponds to the current shown in the same colour in plots C-F.

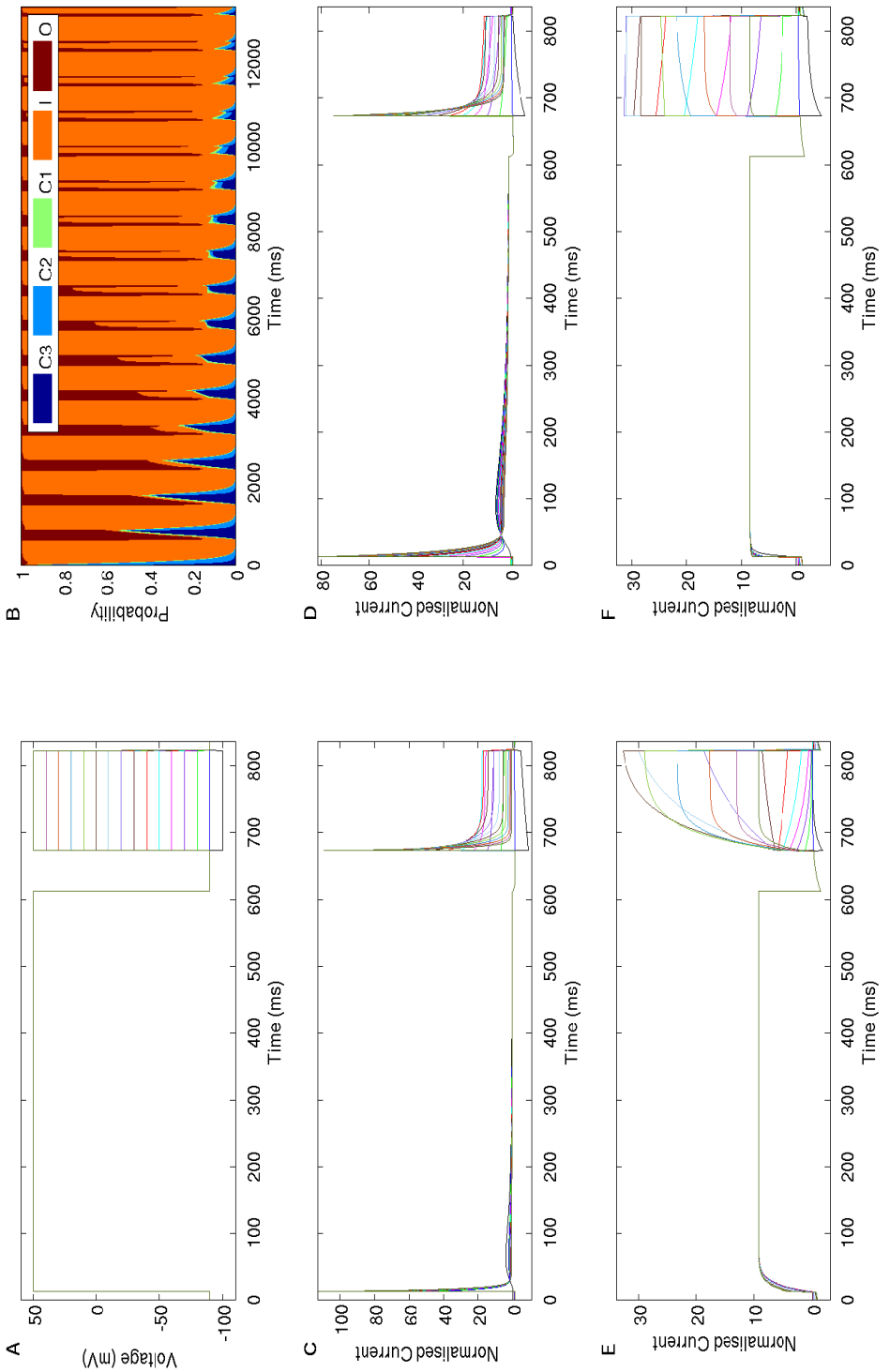


Figure 5.3: Simulated current traces in response to the standard voltage-step protocol shown in A which is intended to investigate inactivation kinetics. B shows the probability of the channel being in each state represented in the Wang et al. (1997) model at each time point in the protocol. Simulated current traces in C-F correspond to those from the Wang et al. (1997), Mazhari et al. (2001), Zeng et al. (1995) and O'Hara et al. (2011) hERG channel models respectively. Each colour in the voltage-step protocol in A corresponds to the current shown in the same colour in plots C-F.

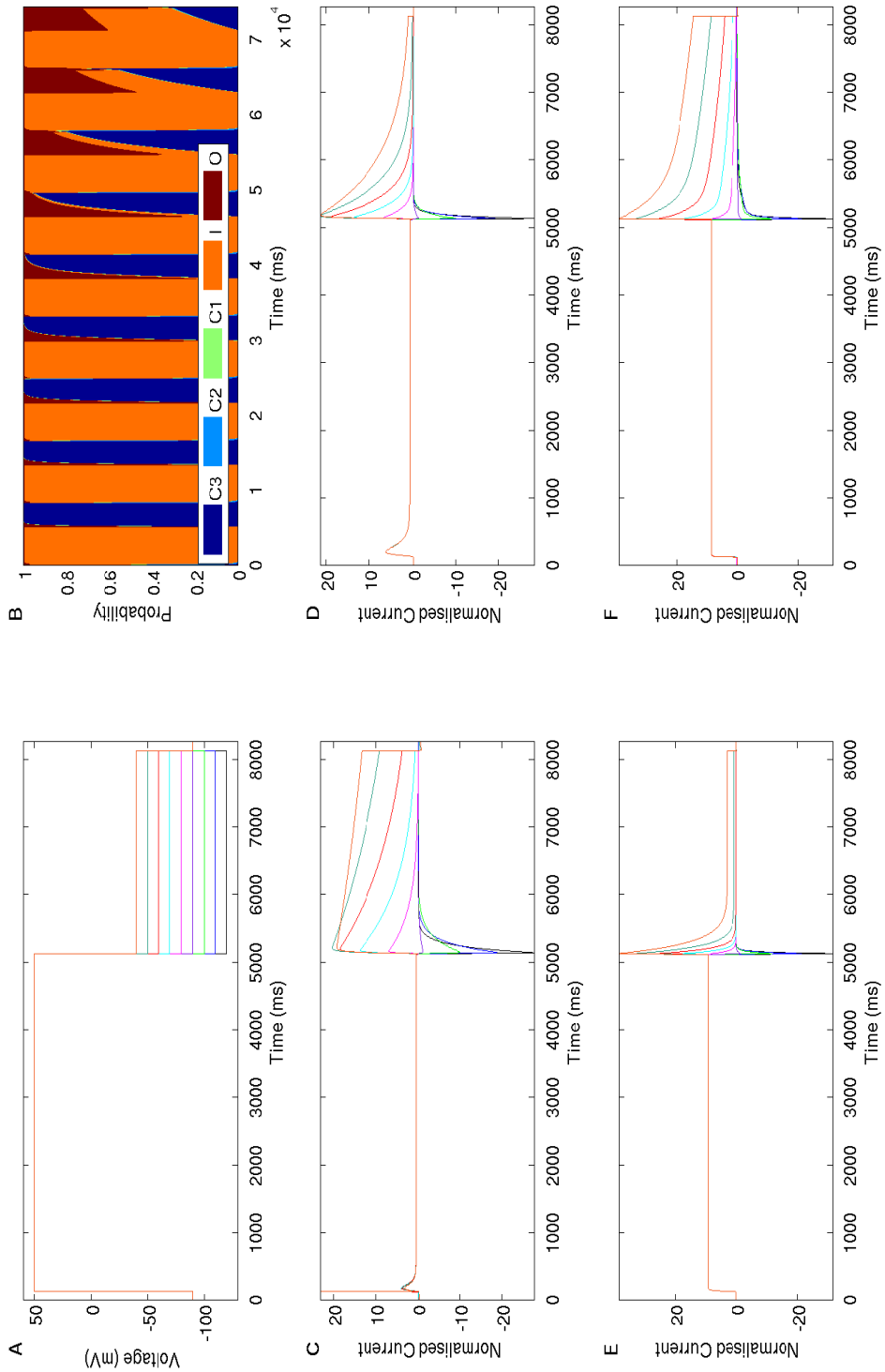


Figure 5.4: Simulated current traces in response to the standard voltage-step protocol shown in A which is intended to investigate deactivation kinetics. B shows the probability of the channel being in each state represented in the Wang et al. (1997) model at each time point in the protocol. Simulated current traces in C-F correspond to those from the Wang et al. (1997), Mazhari et al. (2001), Zeng et al. (1995) and O'Hara et al. (2011) hERG channel models respectively. Each colour in the voltage-step protocol in A corresponds to the current shown in the same colour in plots C-F.

In Figures 5.1-5.4, very different behaviour in the simulated current is evident from each of the models under each protocol. The voltage-step protocols simulated here are variations of those with which the models were originally constructed, and each model showed some agreement with experimental data from protocols similar to those considered here during the calibration process. However, a wide range of different behaviours are predicted under these protocols. Some differences would be expected due to different experimental conditions and natural variability in the cells from which experimental recordings were taken. It is, however, important to note that of the models considered in Figures 5.1-5.4, the Wang et al. (1997) model was formulated to describe the kinetics of the hERG channel at room temperature, whereas the other models describe kinetics at physiological temperatures. Therefore, some differences in behaviour would be expected due to the temperature-dependence of the hERG channel.

Beyond differences due to experimental factors, other possible explanations for this wide variation lie in differences in the mathematical structure of the model and its parameterisation. As many of the models have not been validated with non-training data, there is the possibility that there was insufficient information in the training data to fully characterise the behaviour of the channel, or that the model parameters are not a unique fit to the data. In this latter case, the model may be able to describe the data to which it was fitted but might not behave as expected when simulating protocols different to those with which it was calibrated. It would be of interest to investigate whether such observed differences in model behaviour are due to variability in experimental conditions or other factors related to the construction of the model.

The probability plots presented in Figures 5.1-5.4 (B) allow us to observe the different states that the channel is expected to be in at different times during the protocol. As expected, during the course of the protocol designed to investigate inactivation kinetics, the channels spend the majority of time in the inactivated state, as seen in Figure 5.3. Likewise, the channels spend an increased amount of time in the closed states during the protocol designed to

investigate deactivation kinetics, as seen in Figure 5.4. However, the blocks of time during which the channel behaves in a similar way and remains in the same conformational states are of note in the probability plots. As was discussed in Section 3.1.3, and explored in particular by Fink and Noble (2009), the question arises as to whether there are redundant sections of typical voltage-step protocols during which no additional information is gained; for example, parts of the protocol where voltage-steps are held for a number of seconds. Reducing the length of the protocols, allowing data to be collected in a shorter period of time, could enhance the possibility of collecting training and validation data from the same cell and to enable an assessment of cell-to-cell variability. The short experiment length would also allow multiple current recordings from the same cell in different experimental conditions, enabling more consistent exploration of drug-dependent or temperature-dependent kinetics, for example.

Another point to note is the form of the standard voltage-step protocols. Typically, these bear little resemblance to the electrical activity experienced in cardiac cells, which is in the form of action potentials. Once again, without validation or assurance that a globally optimal solution and unique fit of the model parameters has been achieved, there is no guarantee that a model which is able to replicate the response to the standard voltage-step protocols will be able to adequately describe the response to an action potential. In Figures 5.5 and 5.6 we explore the behaviour of a range of hERG channel models when simulated with action potential clamp protocols. The protocol in Figure 5.5 is representative of electrical activity in physiological conditions and the protocol in Figure 5.6 in pathophysiological conditions. The 10 second voltage protocols were created by pacing the Grandi et al. (2010) human ventricular myocyte model at a pacing rate of 2 Hz for the physiological protocol and a faster pacing rate of 3.9 Hz for the pathophysiological protocol, a rate at where alternating action potential behaviour was demonstrated.

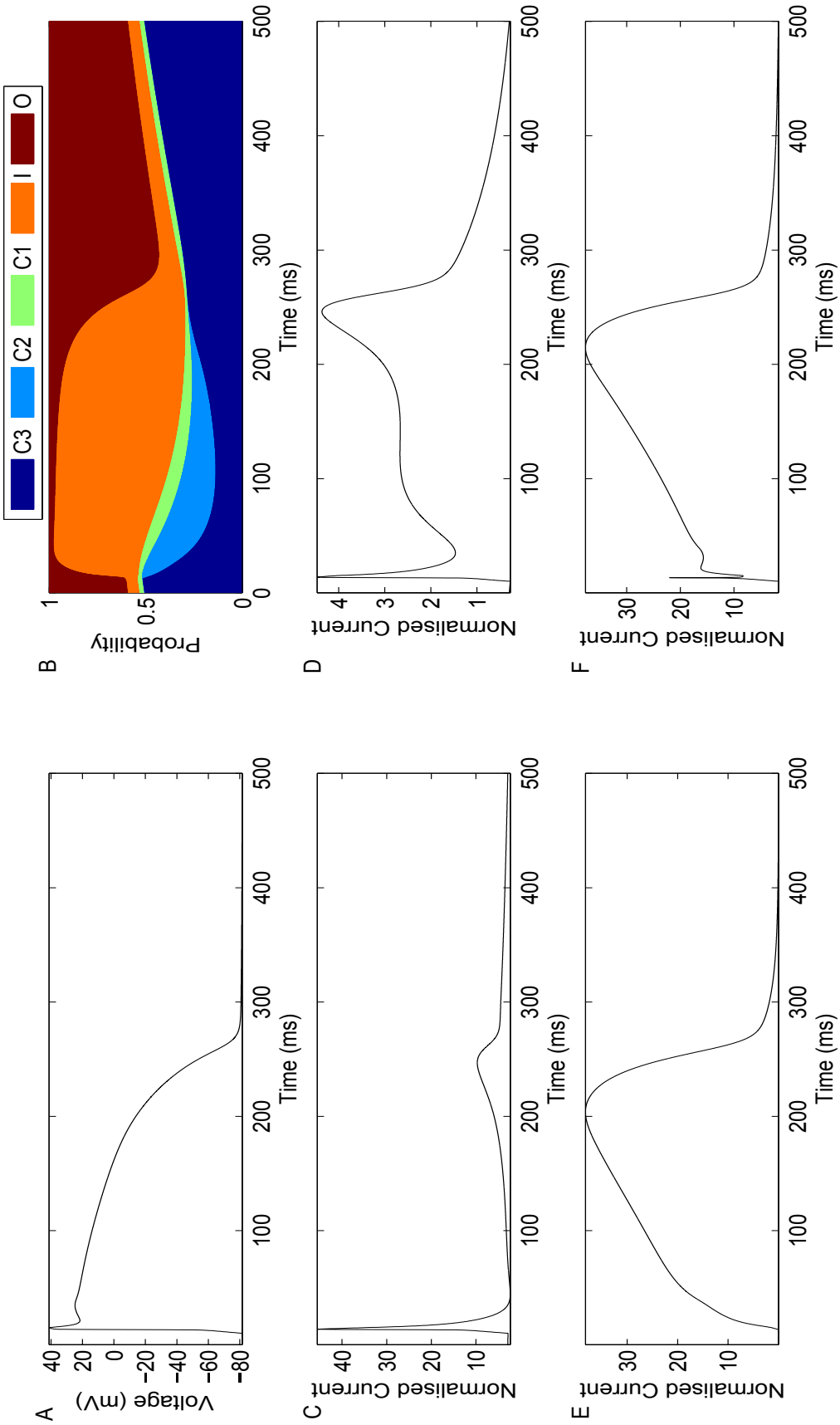


Figure 5.5: Simulated current traces in response to the voltage protocol shown in A which was created by pacing the Grandi et al. (2010) human action potential model at 2 Hz. The final 500 ms of the 10 second protocol is shown. B shows the probability of being in each state represented in the Wang et al. (1997) model at each time point in the protocol. Simulated currents in C-F correspond to those from the Wang et al. (1997), Mazhari et al. (2001), Zeng et al. (1995) and O'Hara et al. (2011) hERG channel models respectively. Simulations were performed using a conductance parameter normalised to 1 in each model, and corresponding normalised current plotted.

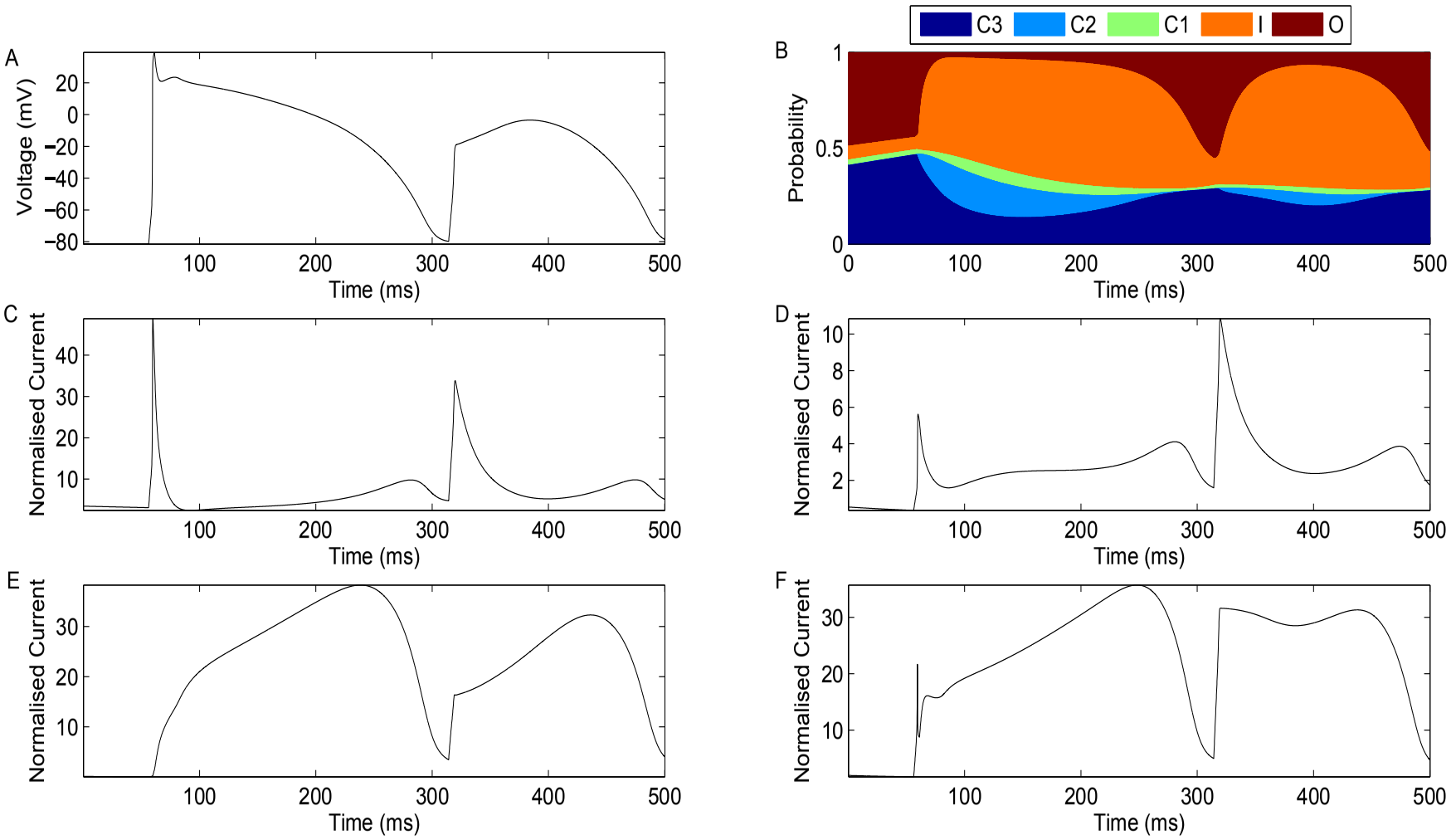


Figure 5.6: Simulated current traces in response to the voltage protocol shown in A which was created by pacing the Grandi et al. (2010) human action potential model at 3.9 Hz. The final 500 ms of the 10 second protocol is shown. B shows the probability of being in each state represented in the Wang et al. (1997) model at each time point in the protocol. Simulated currents in C-F correspond to those from the Wang et al. (1997), Mazhari et al. (2001), Zeng et al. (1995) and O'Hara et al. (2011) hERG channel models respectively. Simulations were performed using a conductance parameter normalised to 1 in each model, and corresponding normalised current plotted.

The differences observed when simulating the models with action potential protocols as shown in Figures 5.5 and 5.6 are of concern. Depending on which hERG channel model representation is used within an action potential model, there could be quite different I_{Kr} currents simulated, which could consequently impact on the morphology of the simulated action potential. In Section 5.1.2 we investigate the potential impact the choice of hERG channel representation within a biophysical action potential model could have on the simulated action potential when predicting drug-induced effects on cardiac electrical activity.

5.1.2 Impact of hERG Channel Model Representation on Simulated Action Potentials

Vandenberg et al. (2012) examined the impact of the hERG channel model used within the ten Tusscher and Panfilov (2006) human action potential model on the simulated action potential. They demonstrated that the choice of hERG channel model can have a large effect on the duration and shape of the simulated action potential. Here, we assess the impact of the choice of hERG channel representation when simulating action potential models with hERG channel block. We incorporated a selection of different hERG channel models (which were formulated to describe kinetics at physiological temperature) within the Shannon et al. (2004) rabbit ventricular myocyte action potential model (used for the evaluation presented in Chapter 4) and the recent human action potential model developed by O'Hara et al. (2011).

Recalling that the hERG channel current is calculated according to Equation (5.1), hERG channel models were incorporated within the action potential models by calculating a value for g_{Kr} which made the control steady state APD_{90} value the same with the new hERG channel model incorporated as with its original hERG channel model formulation. The cell was then paced at 1 Hz to steady state, which is defined to be when the square root of the sum of the squared differences of the state variables at the start of successive paces is less than 10^{-6}. The resulting steady state action potential simulated from each action potential model with different hERG channel model formulations embedded in the presence of 90% hERG channel block is shown in Figure 5.7.

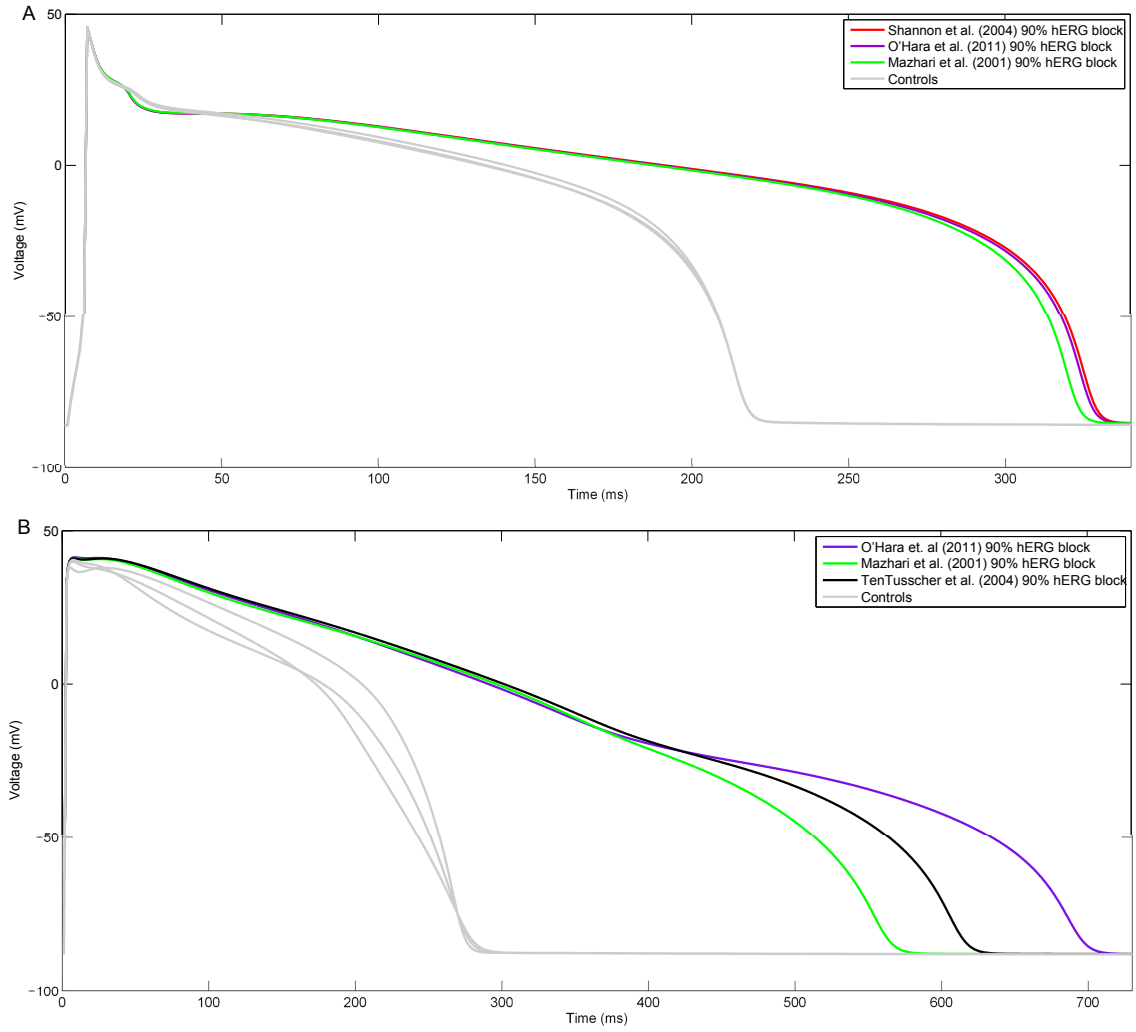


Figure 5.7: Simulated steady state action potential from the Shannon et al. (2004) rabbit ventricular myocyte model in A and the O'Hara et al. (2011) human ventricular myocyte model in B with 90% hERG channel block conditions with different I_{Kr} model formulations embedded within the myocyte model. The grey lines show the control action potentials for each model.

As seen in Figure 5.7, and Vandenberg et al. (2012), the choice of hERG channel model representation can have a great impact on the simulated action potential, both in control conditions, and to a greater extent when simulating the effects of drug block. While replacing the choice of hERG channel model representation in the Shannon et al. (2004) model leads to smaller differences in the simulated action potential than in the O'Hara et al. (2011) model (Figure 5.7 (A)), there is still a range of almost 10 ms difference in the action potentials which would translate into an almost 5% difference in the calculated change in APD_{90} as compared to the control. Such a difference could have a great impact on results using the classification system based on a 10% change in APD_{90} as we used in Chapter 4.

The impact of the choice of hERG channel model representation on the simulated action potential is more obvious when using the O'Hara et al. (2011) model, with a range of APD_{90} values of over 100 ms as shown in Figure 5.7 (B). This highlights the need to ensure that hERG channel kinetics are represented accurately within action potential models, particularly when assessing drug-induced effects. This result is particularly important as the present arrhythmic risk biomarkers are mostly based upon changes in action potential repolarisation time (Mirams et al., 2011).

A final point to note is that a number of the existing hERG channel model representations contain *a priori* unidentifiable parameters. This means there is no possible unique parameterisation of the model due to the model formulation. For example, any model with transition rates expressed in the form

$$A \exp(B(V + C)), \tag{5.2}$$

where A, B and C are constant parameters and V is the applied voltage protocol is *a priori* unidentifiable as Equation (5.2) can be rewritten as

$$D \exp(BV) \tag{5.3}$$

where $D = A \exp(BC)$. A and C can be jointly varied to ensure that D still has a unique

value. Models which are *a priori* unidentifiable can be rewritten so that all parameters are *a priori* identifiable, as indicated in Appendix A.

Following the work presented in Chapter 4, our original intention was to investigate the mathematical representations of drug-ion channel interactions and assess the impact of including a more detailed kinetic representation of drug-induced changes to cardiac electrical activity. However, an initial review of the existing hERG channel model representations, and the exploration of their behaviour as presented so far in this chapter, highlighted that the first step towards achieving this would be to decide upon an accurate and robust model of hERG channel kinetics in the absence of any drug compound. This could be an existing hERG channel model, a reparameterisation of an existing model, or an entirely different model structure.

We approach this problem using ideas of model selection and considering parameter identifiability. Our aim is to construct a model of the least complexity possible which is able to adequately describe behaviour observed in a range of experiments. We want to select an *a priori* identifiable model structure and then attempt to ensure that the parameters of the model are in practice identifiable when using the proposed experimental data to calibrate them.

To determine models of hERG channel kinetics in this way, we designed a range of voltage protocols for use in patch clamp experiments. The rationale behind the design of each of the protocols was to aid in model selection and parameterisation. Motivated by the ideas introduced in Fink and Noble (2009), with regard to the redundancy of information in standard voltage-step protocols, we wanted to explore the use of more dynamic protocols to see if these would enable equivalent amounts of information to be gathered in a shorter time. Shorter protocols would allow the complete characterisation of kinetics from a single cell, and the collection of validation data. More thorough investigations of other effects in a single cell would become possible, such as drug-interactions and temperature-dependence.

In the remainder of this chapter we describe the rationale behind the design of a selection of voltage protocols suitable for use in patch clamp experiments. We explore how existing models behave under these protocols and attempt to assess the identifiability of parameters in existing hERG channel models under these new protocols. To do this we simulate data with added noise, and attempt to identify voltage protocol combinations which may be suitable for the purpose of retrieving the original parameters within a specified tolerance. We conclude this chapter by comparing the identifiability of model parameters using the new protocols with the protocol proposed by Fink and Noble (2009), and a standard set of four voltage-step protocols (as shown in Figures 5.1-5.4).

5.2 Methods

In this section we describe the design of a range of voltage protocols proposed for use in patch clamp experiments to determine models of hERG channel kinetics. We also describe the statistical methods used to assess the identifiability of model parameters when using each of the voltage protocols.

5.2.1 Voltage Protocol Design

The protocols we designed took the form of sums of sine waves. Sine wave protocols were chosen because of the ability to incorporate a range of frequencies appropriate for investigating kinetics which occur over a range of timescales (as described in Section 2.2.2). The dynamic nature of the protocols also allows for the investigation of kinetics away from steady state, in contrast to traditional voltage-step protocols which focus primarily on investigating ion channel kinetics in steady state. Sine wave protocols have previously been utilised in the field of electrochemistry (Sher et al., 2004) and also for exploring the behaviour of the Shaker potassium channel (Kargol et al., 2004).

Each of the sine wave protocols was composed of a sum of three sine waves and took the

general form

$$V = A_1 \sin(2\pi f_1 t) + A_2 \sin(2\pi f_2 t) + A_3 \sin(2\pi f_3 t). \tag{5.4}$$

where A_1, A_2 and A_3 are amplitudes and f_1, f_2 and f_3 are frequencies. We chose to include three sine waves in the protocol formulation so that we could incorporate three different frequencies, roughly corresponding to those of hERG channel activation, inactivation and deactivation, as we will go on to describe later in this section. The protocols that we designed fall into two categories: protocols designed with the rationale of aiding model selection and protocols designed to provide information for model parameterisation. The sine wave amplitudes and frequencies were optimised to minimise the objective function defined in the design of each protocol, which we now describe in detail.

Protocols Designed for Model Selection

The rationale behind the design of the first protocols was to aid with the selection of the most appropriate model structure to describe hERG channel kinetics. As described in Section 3.1.3, many recent studies have utilised the hERG channel representations proposed by Wang et al. (1997) and Mazhari et al. (2001). Both of these models consist of 3 closed states, one inactivated state and one open state. The structural difference between the two models is the presence of a transition between the closed state (proximal to the open state) and the inactivated state in the Mazhari et al. (2001) model, which is absent in the Wang et al. (1997) model. Bett et al. (2011) compared the behaviour of a range of hERG channel models and concluded that models which were able to describe observed experimental data which included this transition had a transition rate of effectively zero and so was effectively negligible and that models which did not include this transition were also able to replicate the data. The same was concluded by Fink et al. (2008).

However, assuming that both are possible model formulations for describing hERG channel kinetics, we endeavoured to design a voltage protocol to maximise the difference in the current predicted from the Wang et al. (1997) and Mazhari et al. (2001) models. The in-

tention behind this method of design was that if the currents predicted by different models were maximally different; this should aid selection between the two models when comparing simulations with experimental data.

The objective function used to design this protocol was the negative of the absolute difference between simulated current traces from each model

$$Z = -\sum_{t=1}^{N} |\mathbf{I}_{\text{Wang}}(t) - \mathbf{I}_{\text{Mazhari}}(t)|, \tag{5.5}$$

where $\mathbf{I}_{\text{Wang}}$ is the simulated current from the Wang et al. (1997) model and $\mathbf{I}_{\text{Mazhari}}$ is the simulated current from the Mazhari et al. (2001) model. This objective function was then minimised for each of the random starting protocol parameters to determine the optimal amplitudes and frequencies for the protocol as defined in Equation (5.4)[1]. The parameters for this protocol, referred to as the *sine wave 1 protocol*, as well as the other sine wave protocols described in this chapter, are shown in Table 5.1. The resulting voltage protocol is shown in Figure 5.8.

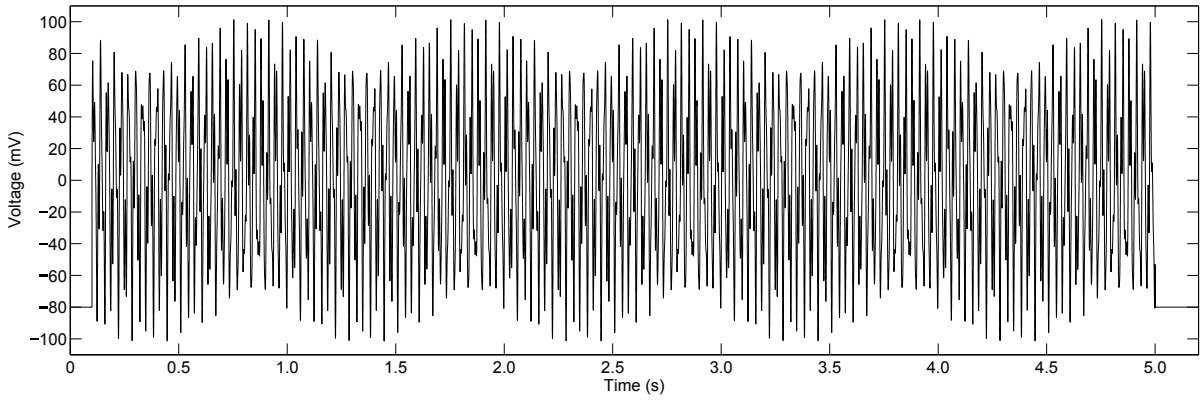


Figure 5.8: *Sine wave 1 protocol* designed to maximise the difference in simulated currents from the Wang et al. (1997) and Mazhari et al. (2001) hERG channel models.

We applied the same approach to design a protocol which maximised the difference in cur-

[1]It should be noted that for the design of this protocol, two additional time delay parameters were included in the protocol formula. That is this protocol is defined by $V = A_1 \sin(2\pi f_1(t - 100)) + A_2 \sin(2\pi f_2 t) + A_3 \sin(2\pi f_3(t - 100))$, but as this later was found to be unnecessary, these additional time delay parameters were removed for the design of subsequent protocols.

Table 5.1 Voltage protocol parameters defining each of the sine wave protocols which are of the form described in Equation (5.4).

	Amplitudes (mV)			**Frequencies** (ms^{-1})		
Protocol	A_1	A_2	A_3	f_1	f_2	f_3
Sine wave 1	57.0	28.0	18.0	0.031	0.08	0.112
Sine wave 2	-25.3	99.7	-4.20	0.00007	0.003	0.16
Sine wave 3	-51.0	71.0	17.0	0.00008	0.004	0.005
Sine wave 4	99.9995	9.8626	0.1378	0.0032	0.0128	0.2161

rent output from each of the Wang et al. (1997), Mazhari et al. (2001) and Di Veroli et al. (2013) hERG channel models after the more recent model was published. In this case the objective function used to determine the protocol parameters was

$$Z = -\frac{1}{3}\left(\sum_{t=1}^{N} \left| \mathbf{I}_{\text{Wang}}(t) - \mathbf{I}_{\text{Mazhari}}(t) \right| + \sum_{t=1}^{N} \left| \mathbf{I}_{\text{Wang}}(t) - \mathbf{I}_{\text{Diveroli}}(t) \right| + \sum_{t=1}^{N} \left| \mathbf{I}_{\text{Mazhari}}(t) - \mathbf{I}_{\text{Diveroli}}(t) \right| \right), \tag{5.6}$$

where $\mathbf{I}_{\text{Wang}}$, $\mathbf{I}_{\text{Mazhari}}$ and $\mathbf{I}_{\text{Diveroli}}$ are the simulated currents from the Wang et al. (1997), Mazhari et al. (2001) and Di Veroli et al. (2013) models respectively. The parameters defining this protocol are shown in Table 5.1 and the voltage protocol, referred to as the *sine wave 2 protocol*, is shown in Figure 5.9.

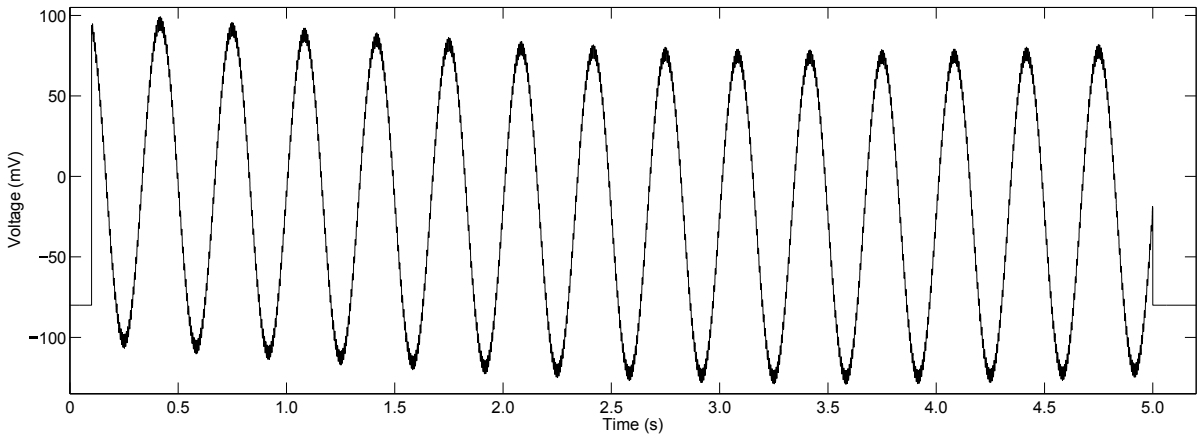


Figure 5.9: *Sine wave 2 protocol* designed to maximise the difference in simulated currents from the Wang et al. (1997), Mazhari et al. (2001) and Di Veroli et al. (2013) hERG channel models.

Protocols Designed For Parameter Estimation

The motivation behind the design of the final two sine wave protocols was to design voltage protocols to provide information to determine accurately transition rate parameters within the hERG channel models. One protocol was designed with an objective function which attempted to cause the Wang et al. (1997) hERG channel model to spend an equal proportion of time in each of its N conformational states. The objective function used was

$$Z = \sum_{i=1}^{N} |\frac{1}{N} - \bar{\mathbf{S}}_\mathbf{i}|, \tag{5.7}$$

where $\bar{\mathbf{S}}_\mathbf{i}$ is the average probability of being in model state i represented in the Wang et al. (1997) model over the whole protocol. The 1/N within the objective function represents the ideal scenario that over the course of the protocol there is an equal chance of finding the channels in each of the states represented in the model. The resulting optimised protocol parameters are included in Table 5.1 and the voltage protocol, referred to as *sine wave 3 protocol*, is shown in Figure 5.10.

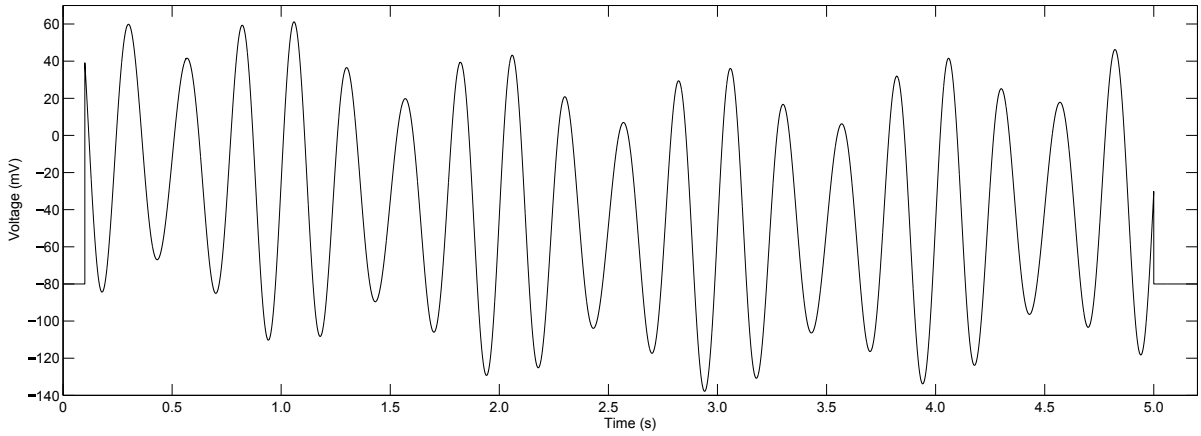


Figure 5.10: *Sine wave 3 protocol* designed for equal proportions of time to be spent in each of the five states represented within the Wang et al. (1997) hERG channel model.

The final protocol was designed to make all the deactivation and activation transition rates in the Wang et al. (1997) hERG channel model as close to equal as possible over the duration

of the whole protocol. The objective function used to achieve this was

$$Z = -\sum_{i=1}^{6} \frac{\sum_{t=1}^{M} T_i(t)}{\max_i\{\sum_{t=1}^{N} T_i(t)\}}, \tag{5.8}$$

where $\mathbf{T_i}$ are the transition rates of the form $\alpha \exp(\beta V)$ at each time in the protocol for each of the six transitions between closed states and the open state within the Wang et al. (1997) hERG channel model. We excluded the transitions between the open and inactivated states in the design of this protocol, as from preliminary exploration with the *sine wave 1 protocol* it appeared that this protocol provided good information about the inactivation transition (as we will demonstrate later in this chapter) and we envisaged using a combination of the sine wave protocols. The parameters for this protocol are shown in Table 5.1 and the protocol, referred to as *sine wave 4 protocol*, is shown in Figure 5.11.

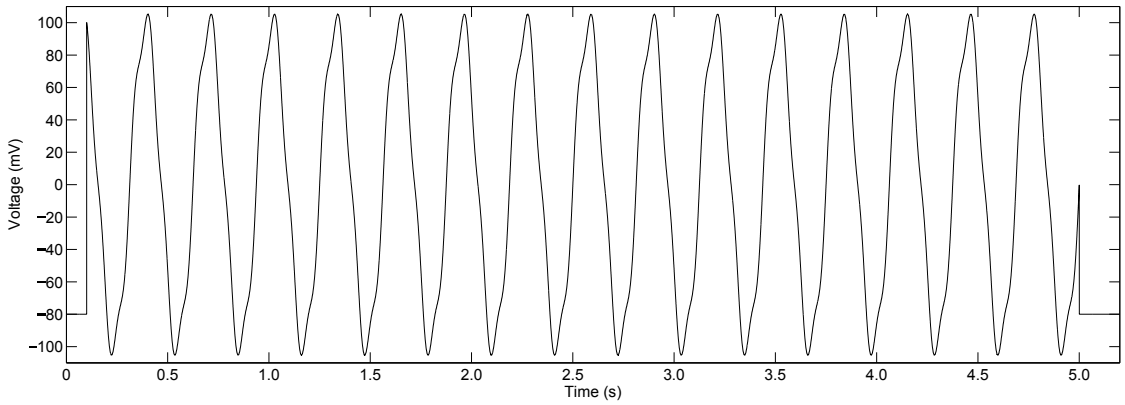


Figure 5.11: *Sine wave 4 protocol* designed for equal transition rates during activation and deactivation in the Wang et al. (1997) hERG channel model.

In the design of each of the protocols described in this section, ODEs were solved using Matlab's `ode15s` solver. To design each protocol, the specified objective function was minimised using a simplex search algorithm (Lagarias et al., 1998) using the `fminsearch` function in Matlab (MatLab, 2012) to determine a set of parameters to define each voltage protocol. The optimisation process was repeated with 100 random starting parameter sets and a tolerance of 10^{-3} on the objective function value was used for termination of the optimisation process.

The random initial starting values for the simplex algorithm in each case were chosen to encompass the range of frequencies observed in the kinetics of the hERG channel. To allow the range of different voltages typically seen in voltage-step protocols to be represented, random initial values for the amplitudes A_1, A_2 and A_3 were scaled by 70, 35 and 15 mV respectively to form the random starting amplitudes. Similarly, the random initial frequencies f_1, f_2 and f_3 were multiplied by 5×10^{-5}, 5×10^{-3} and 5×10^{-1}, to cover the range of possible frequencies of transition in the hERG channel from the time scales of milliseconds to seconds, as described in Section 2.2.2. The length of each sine wave protocol was designed to be 5.2 seconds in total, which included an initial 100 ms section at the holding potential of -80 mV and a final 200 ms section where the voltage is again held at the holding potential.

The optimisation was constrained so that the objective function was penalised during optimisation if ($|A_1| + |A_2| + |A_3|$)> 140 mV. This was to ensure that 140 mV would be the maximum voltage magnitude in the protocol (when all three sine waves were in phase), as voltages above this level are rarely seen in nature and are not expected to be tolerated by cells. We also penalise the objective function so that none of the magnitudes of the individual sine wave amplitudes exceed 100 mV. In addition, we stipulate that $0 < f_1 < f_2 < f_3$ to ensure that we are not searching mathematically equivalent parameters, that is that the protocol parameterisations are identifiable.

5.2.2 Implementation

All simulations described in this chapter were performed in Matlab. Each of the ion channel models detailed in Appendix A were written in C using Matlab Mex functions (MatLab, 2012). The model ODEs were solved using CVODE (Hindmarsh et al., 2005), with a combination of backward differentiation formulae and Newton's method (a more detailed description of CVODE can be found in Appendix C.1.3). Absolute and relative tolerances for the ODE solution were set to 10^{-8}. All of the protocols we consider begin with a period at holding potential where we would expect all channels to be closed and for no current to

flow. Our initial conditions for all Hodgkin-Huxley style models correspond to having zero probability of the channels being in activated/inactivated states. In Markov models, we assign a probability of 1 to the channel being in the closed state furthest from open state and a zero probability of being in the remaining closed, open or inactivated states.

5.2.3 Bayesian Inference

We adopt a Bayesian inference approach to assess how accurately parameters of hERG channel models are estimated when using simulated current traces corresponding to the voltage protocols designed in Section 5.2.1. This approach allows us to determine distributions of estimates for each parameter when fitting proposed models to data from each protocol. In this chapter we consider an ideal scenario, using simulated data with added noise and in Chapter 7 we will apply the same technique using experimental data.

For some observed current recordings which we will denote $\mathbf{y}$, we can infer the probability of different choices of model parameters $\boldsymbol{\theta}$. Bayes' rule underpins the approach taken and is expressed as

$$P(\boldsymbol{\theta}|\mathbf{y}) = \frac{P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\mathbf{y})}. \tag{5.9}$$

$P(\boldsymbol{\theta}|\mathbf{y})$ is the probability that the parameters of the model are $\boldsymbol{\theta}$ after observing the experimental data $\mathbf{y}$ and this is termed the *posterior* probability. $P(\mathbf{y}|\boldsymbol{\theta})$, is the probability of observing the experimental data $\mathbf{y}$ given a model parameterised with parameters $\boldsymbol{\theta}$. $P(\boldsymbol{\theta})$ encapsulates our beliefs about $\boldsymbol{\theta}$ before observing any experimental data and is termed the *prior*. By defining a prior, we are able to emphasise parameter values which we know to be more probable than others, using any information we know beforehand. Where we have no information about the parameter values, we can use a *uniform prior* over a large but finite range. If we know that parameters must fall within a certain range then we can define a uniform distribution across this range as the prior. $P(\mathbf{y})$ is a normalising term which is the integral of all possible probabilities $P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})$ and ensures that the probabilities $P(\boldsymbol{\theta}|\mathbf{y})$

sum to 1. In practice, this normalising term is calculated by

$$P(\mathbf{y}) = \int P(\mathbf{y}|\boldsymbol{\theta})P(\boldsymbol{\theta})d\boldsymbol{\theta}. \tag{5.10}$$

As more and more experimental data are acquired, the estimate of our beliefs about $\boldsymbol{\theta}$ can therefore be updated according to Bayes' rule. A Bayesian inference approach to parameter estimation combines beliefs about the parameters in the *prior distribution* with density $P(\boldsymbol{\theta})$ with the *likelihood* $P(\mathbf{y}|\boldsymbol{\theta})$ to determine the *posterior distribution* $P(\boldsymbol{\theta}|\mathbf{y})$ of $\boldsymbol{\theta}$. The posterior distribution then gives us information about the possible distribution of each of the model parameters in $\boldsymbol{\theta}$, rather than simply determining the single most likely set of parameters, as with a maximum likelihood approach.

For the ion channel models we are considering, we are interested in the posterior distribution of the transition rate parameters ($\boldsymbol{\theta}$). Our observed data are the macroscopic current trace $\mathbf{y}$, with data recorded at times $t = [0, \cdots, T]$. The likelihood (L) of a set of parameters values $\boldsymbol{\theta}$ given the observed data $\mathbf{y}$ is equal to the probability of observing $\mathbf{y}$ given parameter values $\boldsymbol{\theta}$, that is

$$L(\boldsymbol{\theta}|\mathbf{y}) = P(\mathbf{y}|\boldsymbol{\theta}). \tag{5.11}$$

Bayes rule in Equation (5.9) can then be rewritten in terms of the likelihood as

$$P(\boldsymbol{\theta}|\mathbf{y}) \propto P(\boldsymbol{\theta})L(\boldsymbol{\theta}|\mathbf{y}). \tag{5.12}$$

When the assumed prior is a uniform distribution, inferences can be made based just on the likelihood as $P(\boldsymbol{\theta})$, as defined in Equation (5.10), is constant. The priors we use for the results presented in this chapter are uniform priors on each parameter value. Parameters which have to be positive, such as conductance values and certain transition rate parameters (as described in Section 3.1) have priors of $[0, 10^{20}]$. Likewise, parameters identified as being required to be negative (as described in Section 3.1) have priors of $[-10^{20}, 0]$. Parameters in the Hodgkin Huxley style models which represent proportions have a prior of $[0, 1]$. In Hodgkin-Huxley style models, where transition rate terms cannot be obviously identified to impose priors on individual parameters, the prior is formulated such than any parameter set

causing a time constant to become negative at any voltage in the protocol range, is given an acceptance probability of 0. For parameters identified as voltages in these models, a prior of [-200,200] was used.

We assume that the errors at each time point are independent and the process determining the "true" current is deterministic. This is a substantial assumption, however, this is deemed to be acceptable in this situation when there is a substantial amount of data, as is discussed in Milescu et al. (2005) and Moffatt (2007). Therefore, we proceed with the assumption of independence in the current recordings and so the conditional probability of observing the whole current trace recorded from time 0 to time T given the model parameter set $\boldsymbol{\theta}$ is

$$L(\boldsymbol{\theta}|\mathbf{y}) = \prod_{t=0}^{T} P(y_t|\boldsymbol{\theta}). \tag{5.13}$$

In addition we assume that the experimental noise is independently and normally distributed with a mean of 0 and variance of σ^2. Therefore, it follows that each observation y_t is drawn from a normal distribution with a mean equal to the model prediction $f(x_t, \boldsymbol{\theta})$ and variance σ^2. The likelihood is then expressed as

$$L(\boldsymbol{\theta}|\mathbf{y}) = \prod_{t=0}^{T} \mathcal{N}(y_t|f(x_t, \boldsymbol{\theta}), \sigma^2) = \prod_{t=0}^{T} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(- \frac{(y_t - f(x_t, \boldsymbol{\theta}))^2}{2\sigma^2} \right). \tag{5.14}$$

The aim is to find the parameter set $\boldsymbol{\theta}$ which maximises this likelihood expression. As the likelihoods associated with experimental data sets with recordings at a large number of time points, calculated using Equation (5.14), consist of the products of many numbers less than one, which may be very small, it is often more convenient to consider the *log-likelihood* instead. Maximising the likelihood is equivalent to maximising the log-likelihood, or minimising the negative log-likelihood. In terms of the log-likelihood, Equation (5.13) can be expressed as

$$\ln(L(\boldsymbol{\theta}|\mathbf{y})) = \sum_{t=0}^{T} \ln P(y_t|\boldsymbol{\theta}). \tag{5.15}$$

Therefore, Equation (5.14) becomes

$$\ln(L(\boldsymbol{\theta}|\mathbf{y})) = -\frac{1}{2} \sum_{t=0}^{T} \ln(2\pi\sigma^2) - \frac{1}{2} \sum_{t=0}^{T} \frac{(y_t - f(x_t, \boldsymbol{\theta}))^2}{\sigma^2}. \tag{5.16}$$

In order to explore the posterior distribution $P(\boldsymbol{\theta}|\mathbf{y})$ we still need to calculate the normalising integral in Equation (5.10). It is often not possible to calculate Equation (5.10) directly, and so *Markov Chain Monte Carlo (MCMC)* algorithms, as we now describe, can be used. These are iterative sampling methods which allow us to sample from the posterior without needing to solve the normalising integral in Bayes' rule.

Markov Chain Monte Carlo

MCMC algorithms are stochastic methods which create a Markov Chain of samples that represent the posterior distribution. A *proposal distribution* is defined, from which samples are taken. This distribution depends only on the present parameter set and not on any previous sets and so satisfies the Markovian property. A range of different MCMC methods have been proposed, details of which can be found in Robert and Casella (2013). In this thesis, the MCMC method we use is the *Metropolis-Hastings algorithm*, which we now describe in detail.

In the Metropolis-Hastings algorithm, candidate parameter sets $\boldsymbol{\theta}_{cand}$ are created from a proposal distribution $q(\boldsymbol{\theta}_{cand}|\boldsymbol{\theta}_i)$ that depends only on the previously accepted parameter set $\boldsymbol{\theta}_i$. We use a multivariate normal distribution as our proposal distribution, which is a symmetric distribution. Any candidate model parameter set $\boldsymbol{\theta}_{cand}$ is compared to the present parameter set $\boldsymbol{\theta}_i$ by calculating the ratio of the likelihood of the two parameter sets. This ratio is then used to determine whether the proposed parameter set should be accepted as part of the MCMC chain, as we describe below. As the proposal distribution is symmetric, that is that $q(\boldsymbol{\theta}_{cand}|\boldsymbol{\theta}_i) = q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{cand})$, then the ratio used to compare the proposed parameter set $\boldsymbol{\theta}_{cand}$ with the current parameter set $\boldsymbol{\theta}_i$ is expressed as

$$\frac{P(\boldsymbol{\theta}_{cand}|\mathbf{y})q(\boldsymbol{\theta}_i|\boldsymbol{\theta}_{cand})}{P(\boldsymbol{\theta}_i|\mathbf{y})q(\boldsymbol{\theta}_{cand}|\boldsymbol{\theta}_i)} = \frac{P(\boldsymbol{\theta}_{cand}|\mathbf{y})}{P(\boldsymbol{\theta}_i|\mathbf{y})} = \frac{L(\boldsymbol{\theta}_{cand}|\mathbf{y})P(\boldsymbol{\theta}_{cand})}{L(\boldsymbol{\theta}_i|\mathbf{y})P(\boldsymbol{\theta}_i)}. \tag{5.17}$$

As we are assuming uniform priors on each of our parameters, $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_{cand}$ will either be constant within the range of the uniform distribution we assume for the prior, or 0 if outside

the range of the prior. Therefore, Equation (5.17) reduces to considering the *likelihood ratio*

$$\frac{L(\boldsymbol{\theta}_{cand}|\mathbf{y})P(\boldsymbol{\theta}_{cand})}{L(\boldsymbol{\theta}_i|\mathbf{y})P(\boldsymbol{\theta}_i)} = \frac{L(\boldsymbol{\theta}_{cand}|\mathbf{y})}{L(\boldsymbol{\theta}_i|\mathbf{y})}. \tag{5.18}$$

If the candidate parameter set has a greater likelihood than the present point in the chain then it will be accepted and added to the Markov chain ($\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_{cand}$). Otherwise, the parameter set may still be accepted with a probability equal to the likelihood ratio in Equation (5.18). Therefore, in the Metropolis Hastings algorithm, a proposed parameter set generated from the multivariate normal distribution is accepted with a probability

$$\alpha = \min\left\{\frac{L(\boldsymbol{\theta}_{cand}|\mathbf{y})}{L(\boldsymbol{\theta}_i|\mathbf{y})}, 1\right\}. \tag{5.19}$$

However, as we have discussed, for convenience we use log-likelihoods in place of likelihoods when calculating the ratio in Equation (5.19). Also to note is that if the proposed parameter set contains any parameters outside the range of their prior, they are immediately rejected with an acceptance probability of 0 and the previous parameter set $\boldsymbol{\theta}_i$ is added to the Markov chain ($\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i$). Similarly, if the proposed parameter set is not accepted because of a low probability of acceptance then $\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i$.

A number of modifications based on this basic Metropolis-Hastings algorithm have been proposed to enhance its computational efficiency. For MCMC to be able to efficiently explore the posterior distribution, a number of parameters have to be tuned to ensure that the chain converges as quickly as possible. One such choice to be made, which is crucial for the performance of the algorithm (Andrieu and Thoms, 2008), is the selection of the proposal distribution. We choose to sample from a multivariate normal distribution with a mean of $\boldsymbol{\theta}_i$ and a covariance of $s\Sigma$ where Σ is a covariance matrix and s is a scaling factor. When sampling from the proposal distribution, if parameter sets are proposed which are far away from $\boldsymbol{\theta}_i$ (that is in the case where the covariance term $s\Sigma$ is large) then proposed parameter sets will be rarely accepted and the Markov chain will not explore the distribution efficiently. On the other hand, if parameter sets very close to $\boldsymbol{\theta}_i$ are proposed (that is where the covariance term $s\Sigma$ is small) then there will be a high rate of acceptance of points and the chain will

only explore the posterior distribution very slowly as it will take many iterations to move away from $\boldsymbol{\theta}_i$.

A MCMC chain that is able to efficiently explore the posterior distribution is described as *well mixed*. It has been suggested that an acceptance rate of around 23.4% for parameter sets of dimension greater than 2 is close to optimal for a well mixed MCMC chain (Gelman et al., 1996). The acceptance rate can be controlled by adjusting the scaling factor s in the covariance term $s\Sigma$ in the proposal distribution. If the acceptance rate is too high, the scaling factor can be increased so that parameter sets further away from the current set are proposed, resulting in more sets being rejected. Similarly, if the acceptance rate is too low then the scaling factor can be decreased so that the chain explores the region closer to the present parameter set, resulting in higher levels of acceptance.

We can attempt to maintain an acceptance rate of around 23.4% by adapting the scaling factor during sampling. After an initial period with no adaptation, during which the scaling factor is set to 1, we can scale the covariance matrix in the proposal by a factor s which is updated according to the formula

$$\log(s_{i+1}) = \log(s_i) + \gamma_{i+1}[\alpha(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{cand}) - \bar{\alpha}_*], \tag{5.20}$$

where $\bar{\alpha}_*$ is the optimal acceptance rate for which we use 23.4%, $\alpha(\boldsymbol{\theta}_i, \boldsymbol{\theta}_{cand})$ is the calculated probability of acceptance for the proposed parameter set $\boldsymbol{\theta}_{cand}$ based on the last accepted point $\boldsymbol{\theta}_i$ as defined in Equation (5.19). γ_{i+1} takes the form of

$$\gamma_{i+1} = \frac{1}{(i+1)^{0.7}}. \tag{5.21}$$

This form for γ_{i+1} ensures that as $i \to \infty$ the scaling factor tends to a constant. The exponent in Equation (5.21) can take any value in the interval $(\frac{1}{2}, 1]$, with a higher value meaning that the scaling factor becomes constant more rapidly (Bardenet, 2013).

Another practical challenge of Metropolis-Hastings is determining a suitable choice for the

covariance matrix for the multivariate normal proposal. This selection often requires some manual tuning before the algorithm is able to run efficiently. It is also possible to update the covariance matrix in the proposal adaptively using the covariance of the MCMC chain after each iteration. It should be noted that this means that the proposed parameter set depends not only on the previous accepted parameter set, but additionally on all the other parameter sets in the Markov chain, and therefore no longer satisfies the Markovian property. However, while not strictly Markov, this method has been shown to satisfy the required properties for MCMC methods (Haario et al., 2001).

The algorithm includes an initial period of no adaptation, during which some points should have been accepted so that the chain has a non-zero covariance. After this initial period, the covariance of the MCMC chain for each parameter is calculated after every iteration, and the covariance matrix updated accordingly. We define an initial covariance to be a diagonal matrix with the diagonal entries proportional to the initial parameter estimates scaled by a factor identified to allow the chain to quickly begin with some acceptance (we choose this manually so that initial acceptance rate is roughly in the range of 20-40%). We then update the covariance matrix using a recursion formula (Andrieu and Thoms, 2008)

$$\boldsymbol{\Sigma}_{i+1} = \boldsymbol{\Sigma}_i + \gamma_{i+1}[(\boldsymbol{\theta}_{i+1} - \boldsymbol{\mu}_i)(\boldsymbol{\theta}_{i+1} - \boldsymbol{\mu}_i)^T - \boldsymbol{\Sigma}_i], \tag{5.22}$$

$$\boldsymbol{\mu}_{i+1} = \boldsymbol{\mu}_i + \gamma_{i+1}(\boldsymbol{\theta}_{i+1} - \boldsymbol{\mu}_i),$$

where, $\boldsymbol{\Sigma}_i$ is the covariance and $\boldsymbol{\mu}_i$ is the mean at iteration i in the MCMC chain, and γ_{i+1} is as defined in Equation (5.21). In practice, the MCMC algorithm we use updates both the scaling factor s and the covariance matrix $\boldsymbol{\Sigma}$ after each iteration of the algorithm, as described in Andrieu and Thoms (2008) and Bardenet (2013).

The start of the MCMC chain, during which the adaptation process is taking effect, and before the chain has converged, is referred to as the *'burn in'* period. This section of the chain is discarded to ensure that the resulting distribution is representative of the posterior distribution. For the results presented in this chapter we run the MCMC algorithm for a total

of 200,000 iterations and discard the first 50,000 iterations as a burn in period. The length of the burn in period can vary from chain to chain, but the period we chose was selected after observing sufficient time for the adaptation process to take effect and for the chain to converge. For complex problems, pre-MCMC optimisation is required to find a starting point for the MCMC chain, as MCMC algorithms are only efficiently able to explore parameter space locally. We discuss strategies to determine starting points for MCMC chains in Section 7.4.2, as well as further details of assessing MCMC convergence when the true underlying parameter values are not known in Section 7.4.3. The length of the burn in period also depends heavily on the quality of the initial guess for the MCMC algorithm. For the results presented in this chapter we start MCMC chains from the true parameters values of the simulated traces, which represents an ideal scenario. Pseudo-code for the Metropolis-Hastings algorithm used is shown in Algorithm 5.12.

for i = 1:N **do**
  Sample $\boldsymbol{\theta}_{\mathbf{cand}} \sim \mathcal{N}(.|\boldsymbol{\theta}_{\mathbf{i-1}}, s_{i-1}\Sigma_{i-1})$
  Sample $u \sim \mathcal{U}_{(0,1)}$
  Calculate Acceptance ratio:
  $\alpha = \min(1, \frac{L(\theta_{cand})}{L(\theta_{i-1})})$
  if $u < \alpha$ **then**
    $\theta_i = \theta_{cand}$
  else
    $\theta_i = \theta_{i-1}$
  end if
  $\boldsymbol{\Sigma}_i = \boldsymbol{\Sigma}_{i-1} + \gamma_i[(\boldsymbol{\theta}_i - \boldsymbol{\mu}_{i-1})(\boldsymbol{\theta}_i - \boldsymbol{\mu}_{i-1})^T - \boldsymbol{\Sigma}_{i-1}]$
  $\boldsymbol{\mu}_i = \boldsymbol{\mu}_{i-1} + \gamma_i(\boldsymbol{\theta}_i - \boldsymbol{\mu}_{i-1})$
  $\log(s_i) = \log(s_{i-1}) + \gamma_i[\alpha - \bar{\alpha}_*]$
end for=0

Figure 5.12: Adaptive Metropolis Hastings Algorithm. γ_i is as defined in Equation (5.21). L is the log-likelihood, calculated as described in Equation (5.16). θ_0 is the initial starting parameter set for the MCMC chain; in this chapter this corresponds to the true simulated parameter values for each model and in Chapter 7 is determined from pre-MCMC parameter optimisation. s_0 and Σ_0 are set by the user upon initiating the algorithm to achieve an initial acceptance rate of approximately 20-40% as described in the text. $\bar{\alpha}_*$ corresponds to the optimal acceptance rate of 23.4%. In practice, we wait for an initial period (250 iterations of the algorithm for the results presented in this chapter and 2500 iterations for the results presented in Chapter 7) before we begin adapting the covariance matrix to allow for a range of parameter sets to be accepted before adapting according to the covariance of the chain. The scaling factor of the covariance matrix is also set to 1 when the adaptation begins.

To assess whether the voltage protocols designed as described in Section 5.2.1 would provide sufficient information for determining parameters of existing models of hERG channel kinetics we first studied parameter distributions recovered using the MCMC approach described using simulated data with noise added. In order to estimate a representative level of noise to be added to the simulated traces, we performed the *sine wave 1 protocol* in a patch clamp experiment (as we describe in detail in Chapter 6) and based our estimate of noise on this recording. During the initial 100 ms section of the protocol, the voltage is held at the holding potential of -80 mV (a voltage at which all channels would be expected to be closed and so a zero current recorded). A histogram of the current recordings during this 100 ms interval of the protocol is shown in Figure 5.13. The standard deviation of a normal distribution ($\mathcal{N}(0, \sigma)$) was fitted to the current in this interval to estimate the noise on the trace, which took a value of σ=18.2869 pA. This estimate is in the middle of the typical range of estimates seen in repeats of this protocol.

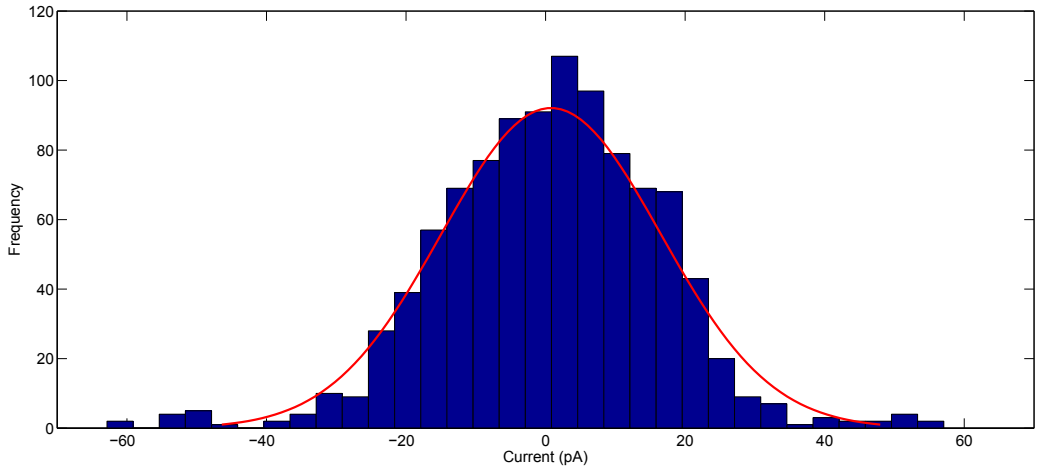


Figure 5.13: Histogram of the current recordings from the first 100 ms interval of *sine wave 1 protocol* where the voltage is held at -80 mV (and the channels expected to be closed and so the current to be 0) from one experimental trace. The red line shows a normal distribution ($\mathcal{N}(0, \sigma)$) fit to the data.

It can be seen in Figure 5.13 that the noise is approximately normally distributed about a

mean of 0 pA, which confirms the appropriateness of our choice of normally distributed noise in the likelihood function (Equation (5.14)). In order to add noise to simulated data across the range of different hERG channel model representations, which have a range of different current amplitudes in response to the same protocol (due to differences in open probabilities over the course of the protocol in different models), we added noise of a magnitude of 0.3% of the peak amplitude of the current trace when the model is simulated with a conductance parameter with the value 1, corresponding to the noise level estimated from this test experimental trace.

Each of the existing models are formulated with currents expressed in a variety of different units, with a wide range of conductance values based on the variable experimental data from which they were estimated. For consistency, we have scaled the conductance to 1 in every model for the simulated results presented in this chapter. However, this still means that the range of current magnitudes varies between simulations from different models. So in order to directly compare the behaviour of different models when we present them together, we normalise the simulated traces by scaling each trace by a factor which minimises the absolute difference between each simulated trace and one reference trace. This simply amounts to selecting an appropriate conductance parameter for each model so that the magnitudes of the simulated traces are comparable.

5.3 Results

5.3.1 Simulated hERG Current with Novel Protocols

We first explored the behaviour of the existing hERG channel models under the novel voltage protocols described in Section 5.2.1. We compare differences in the simulations from a selection of different models and the probabilities of being in each state throughout each protocol (Figures 5.14-5.21). Throughout this chapter, where we refer to particular states or parameters within models, the labelling of these corresponds to that shown for the particular model in Appendix A. It should be noted that the majority of the existing hERG channel

models describe hERG channel kinetics at physiological temperature. In the following plots we have indicated the models formulated to describe kinetics at room temperature by dashed lines. For all simulations a temperature of 22°C was used for calculation of the reversal potential in the Nernst Equation (Equation (3.12)) as this is the temperature at which the experiments described in Chapter 6 are performed.

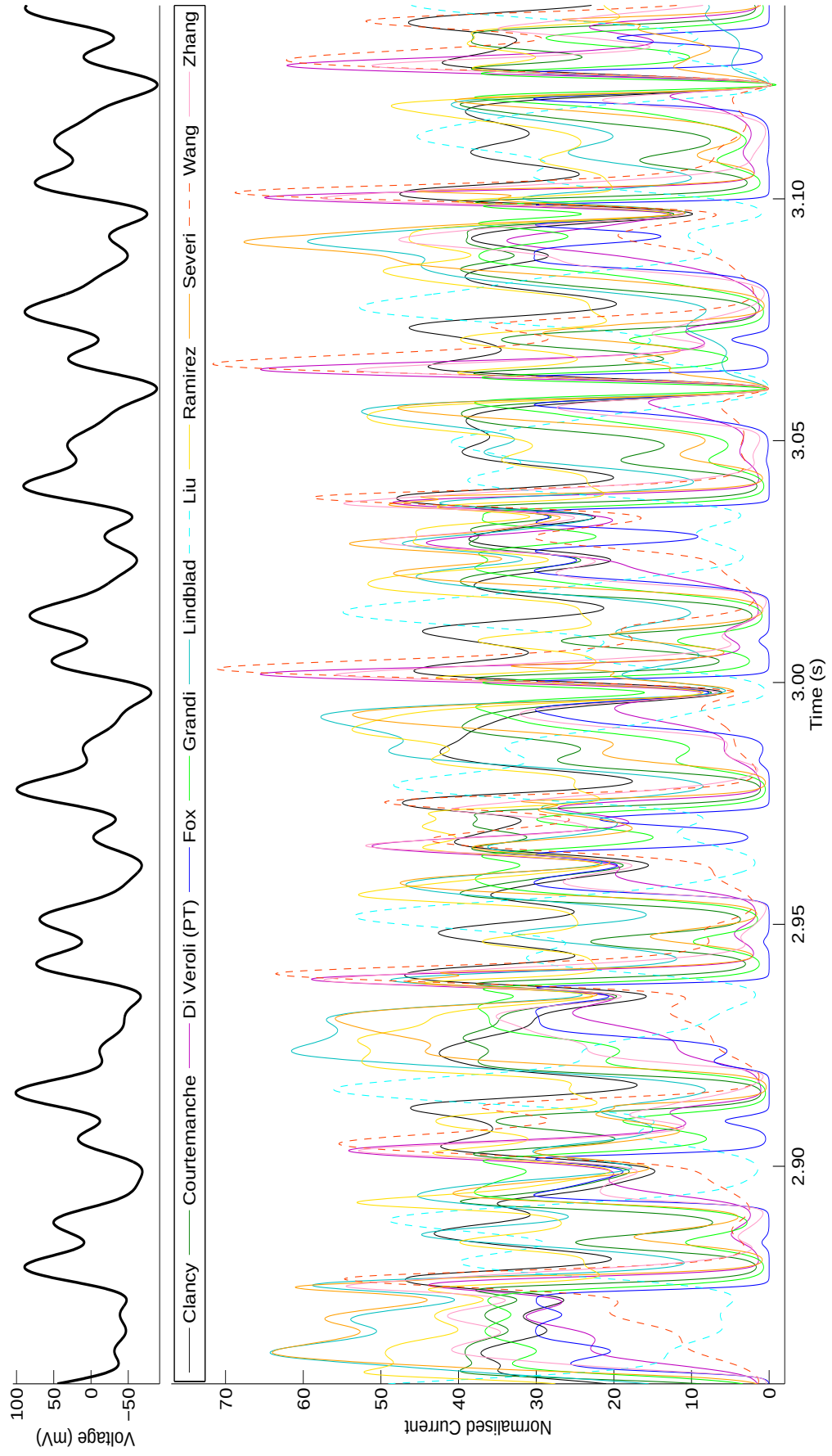


Figure 5.14: Simulated current traces using the *sine wave I protocol* from a selection of existing hERG channel models. The section of voltage protocol is shown above the corresponding simulated currents. Dashed lines indicate models formulated to describe hERG channel kinetics at room temperature.

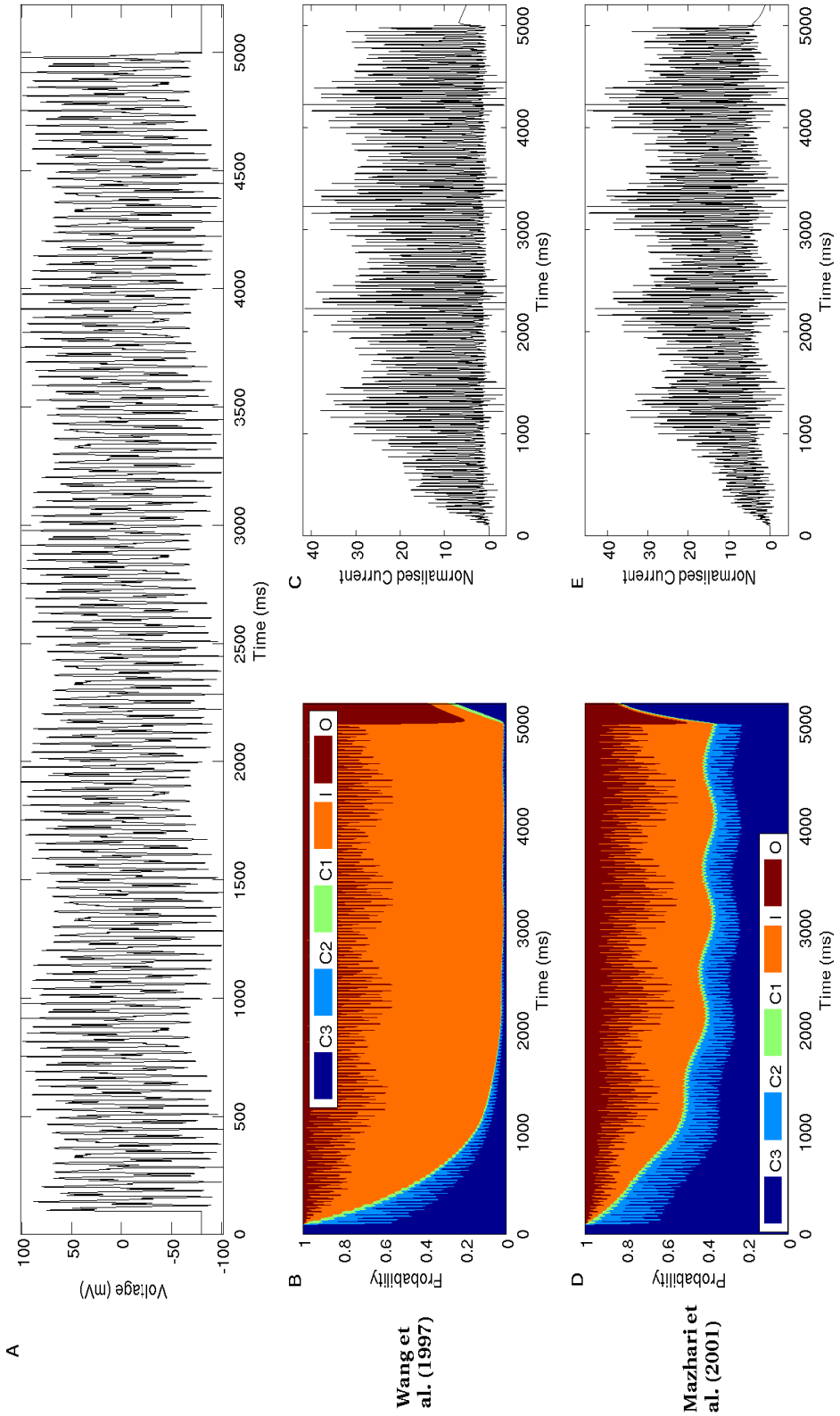


Figure 5.15: A shows the *sine wave I protocol* which was designed to maximise the difference between the Wang et al. (1997) and Mazhari et al. (2001) models. In B and C the probabilities of the channel being in each state represented in the Wang et al. (1997) model and the simulated current from this model over the duration of the protocol are shown. Similarly the simulated probabilities and current from the Mazhari et al. (2001) model are shown in D and E.

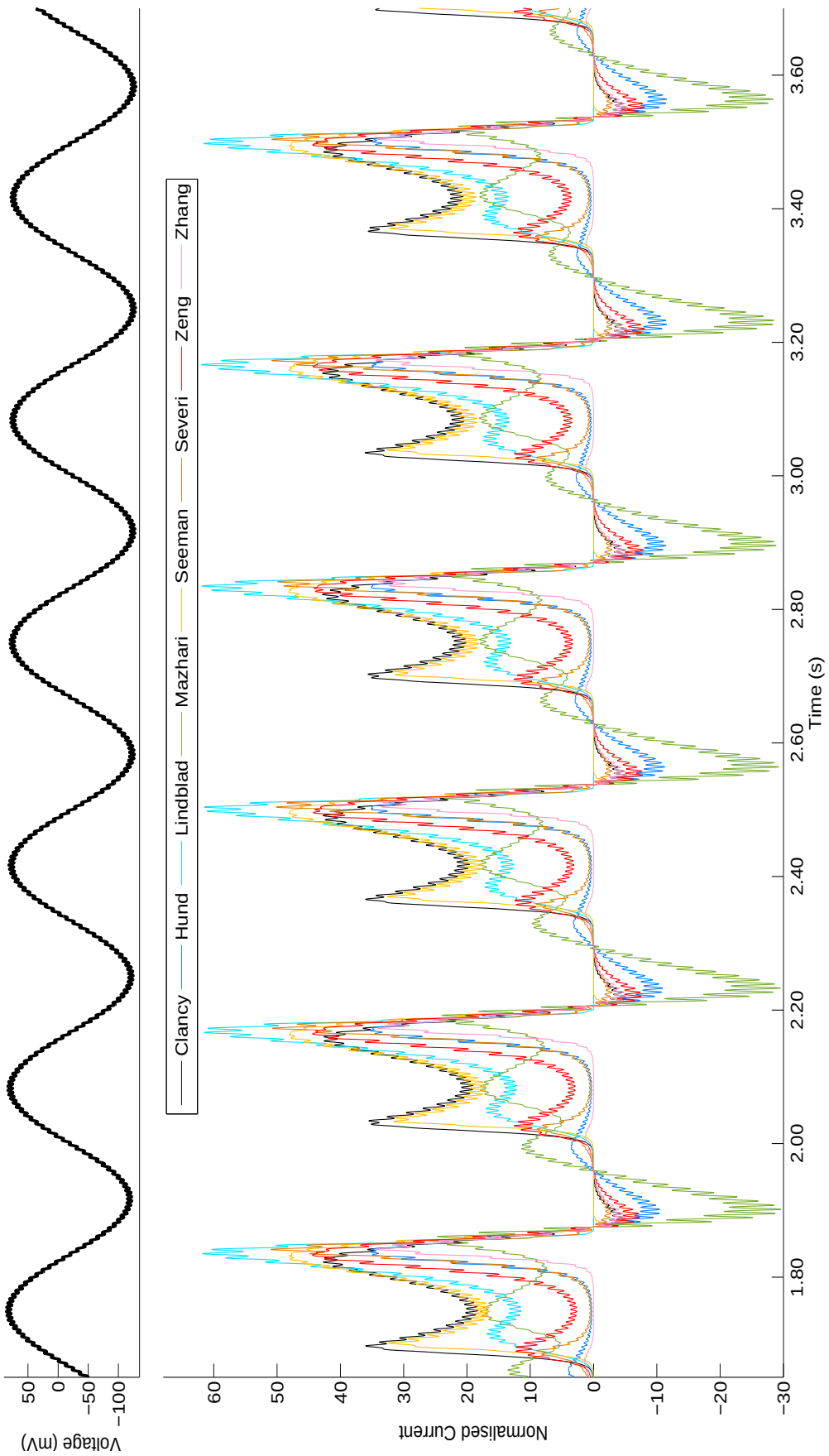


Figure 5.16: Simulated current traces using *sine wave 2 protocol* from a selection of existing hERG channel models. The section of voltage protocol is shown above the corresponding simulated currents.

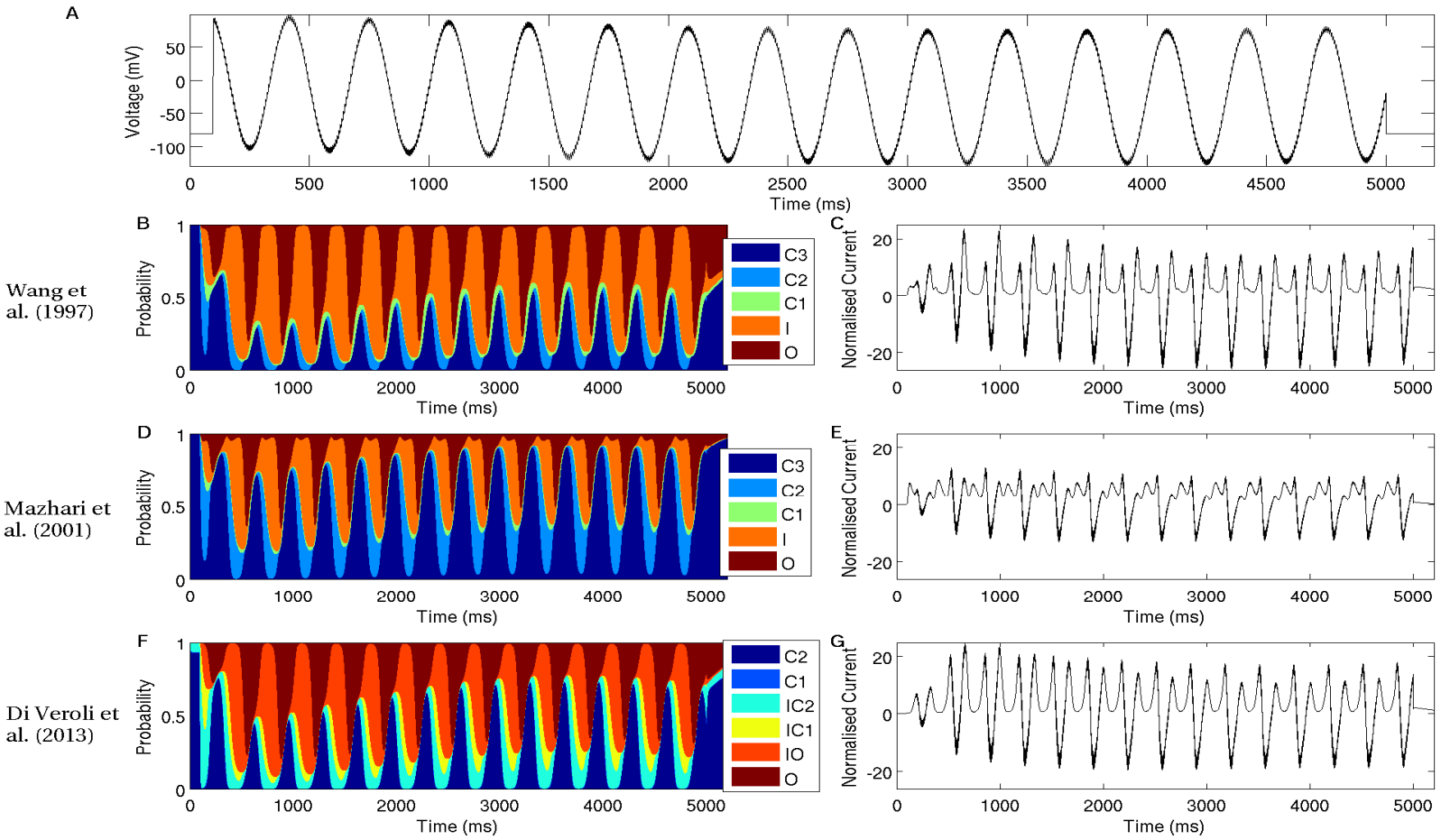


Figure 5.17: A shows *sine wave 2 protocol* which was designed to maximise the difference between the Wang et al. (1997), the Mazhari et al. (2001) and Di Veroli et al. (2013) models. In B and C the probabilities of the channel being in each state represented in the Wang et al. (1997) model and the simulated current from this model over the duration of the protocol are shown. Similarly the simulated probabilities and current from the Mazhari et al. (2001) model are shown in D and E and those from the Di Veroli et al. (2013) room temperature model in F and G.

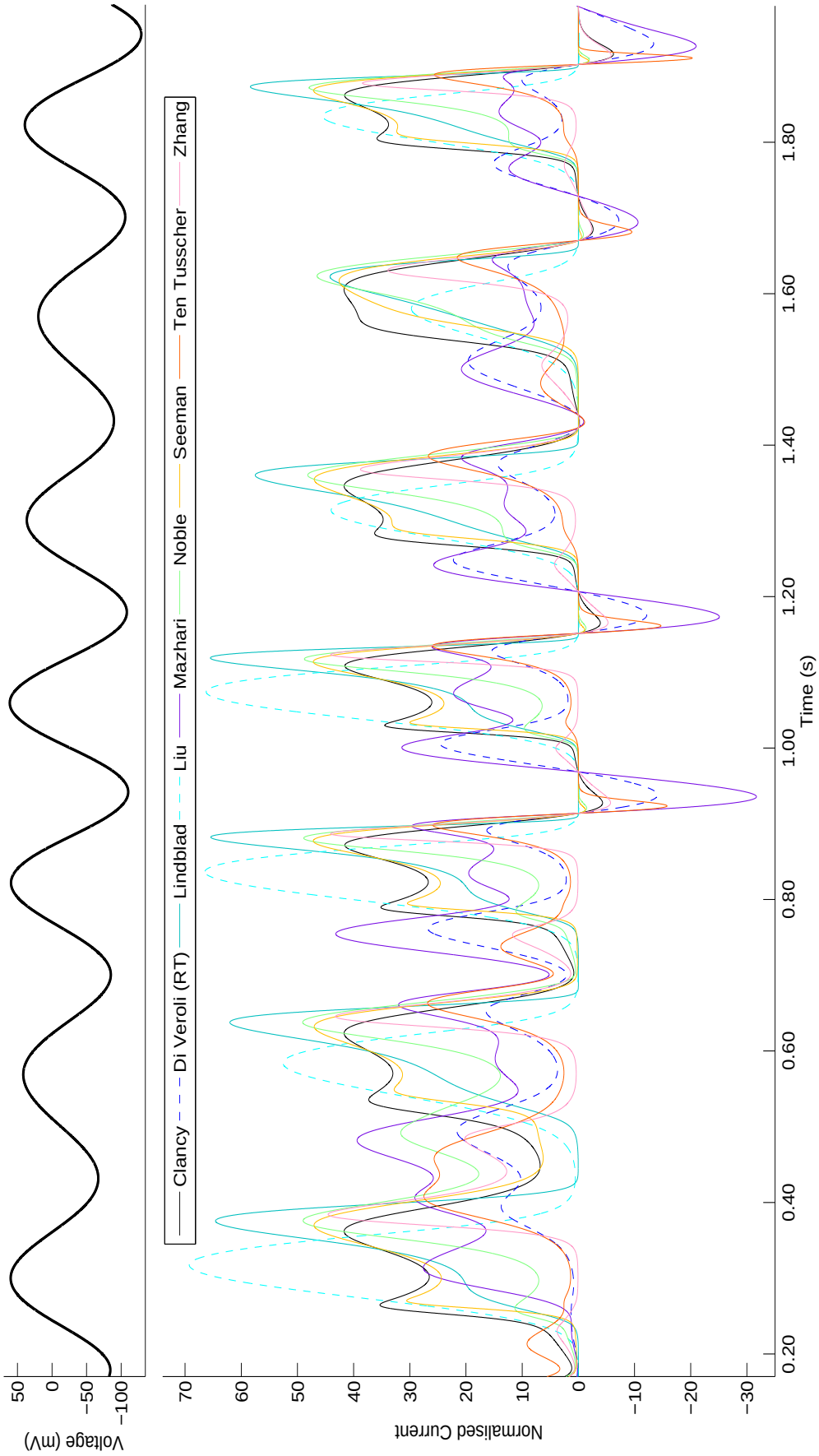


Figure 5.18: Simulated current traces using *sine wave 3 protocol* from a selection of existing hERG channel models. The section of voltage protocol is shown above the corresponding simulated currents. Dashed lines indicate models formulated to describe hERG channel kinetics at room temperature.

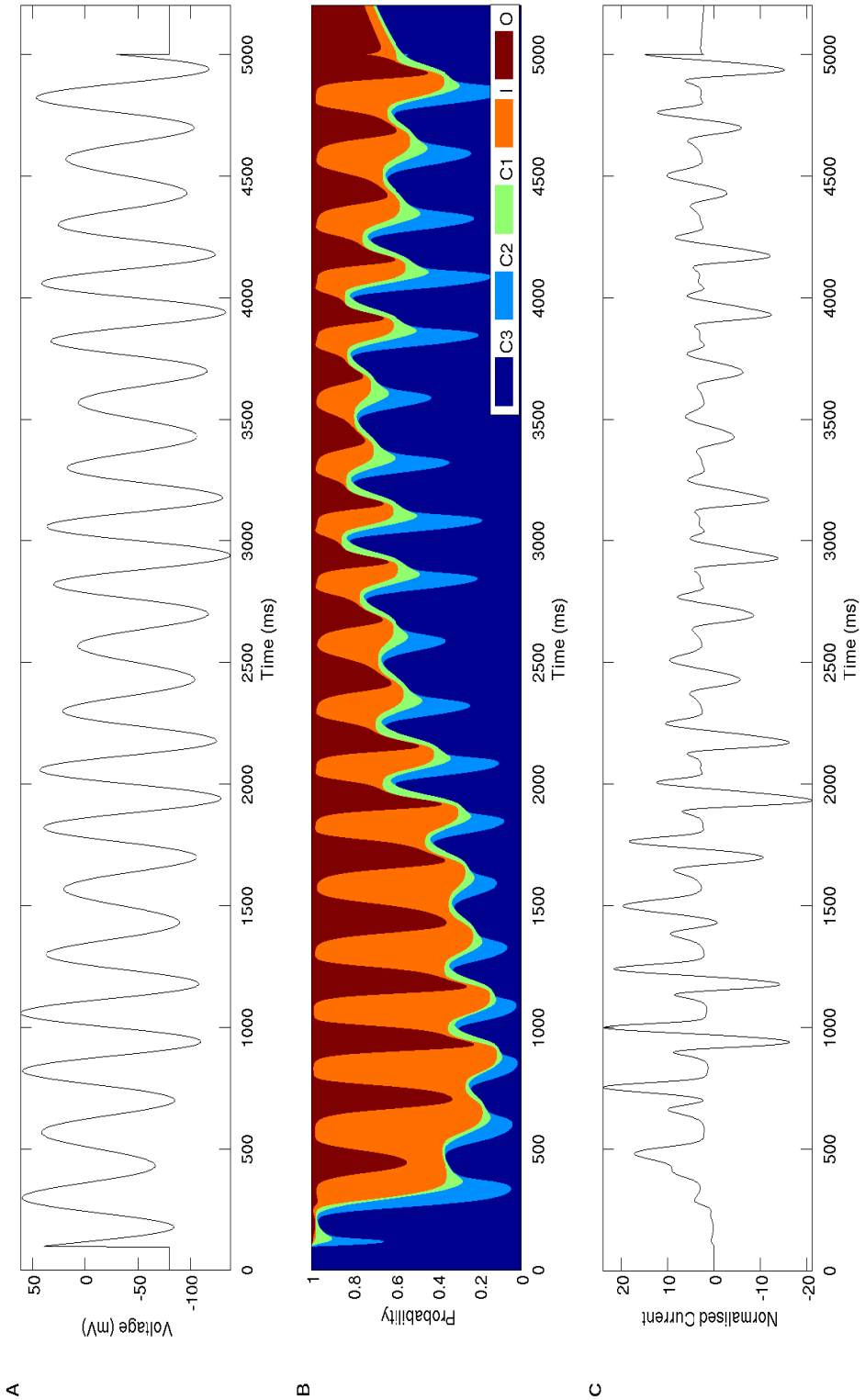


Figure 5.19: A shows *sine wave 3 protocol* which was designed with the aim of the channel spending an equal proportion of time in each of the conformational states represented within the Wang et al. (1997) model. In B and C the probabilities of the channel being in each state represented in the Wang et al. (1997) model and the simulated current from this model over the duration of the protocol are shown.

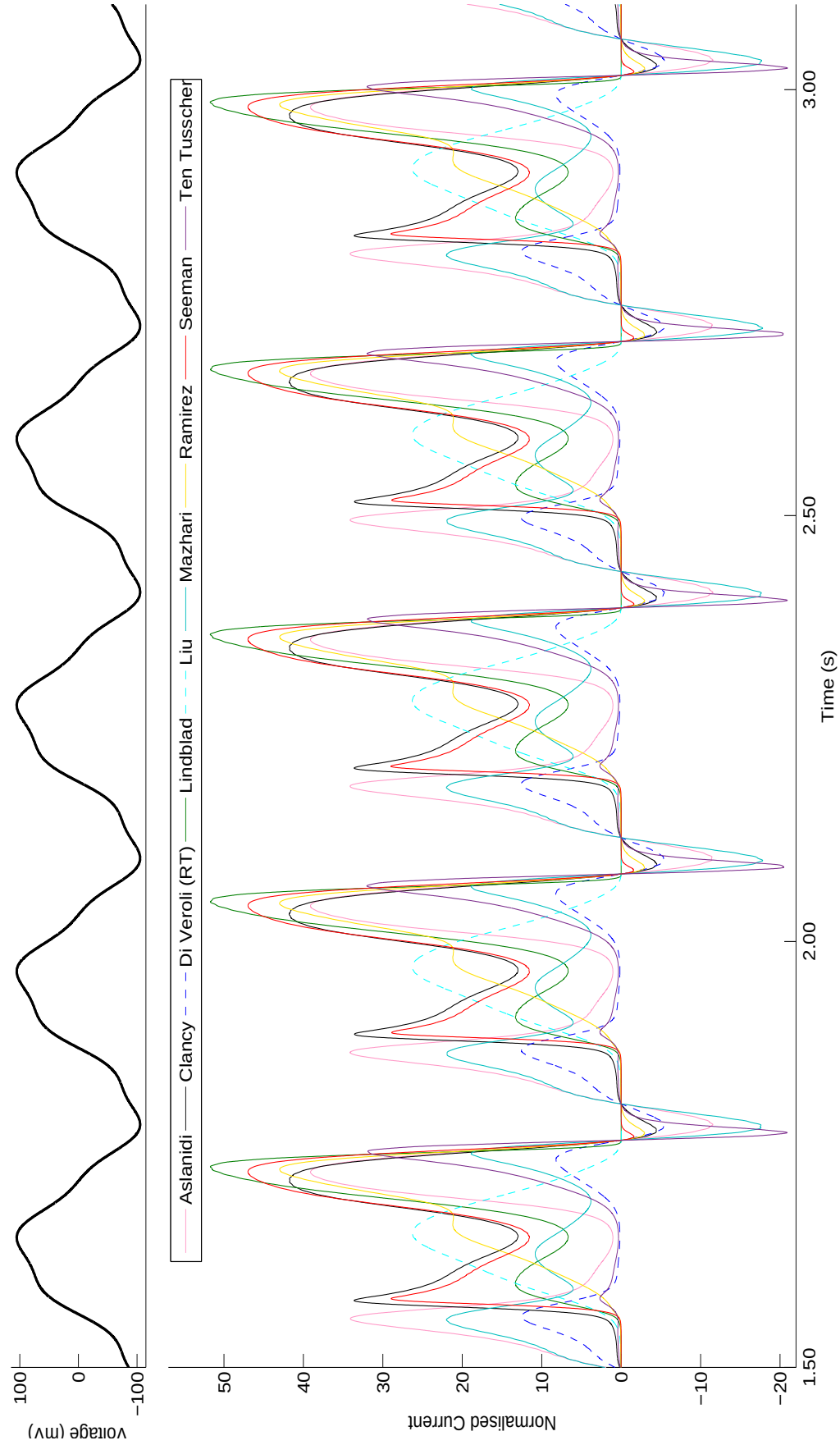


Figure 5.20: Simulated current traces using *sine wave 4 protocol* from a selection of existing hERG channel models. The section of voltage protocol is shown above the corresponding simulated currents. Dashed lines indicate models formulated to describe hERG channel kinetics at room temperature.

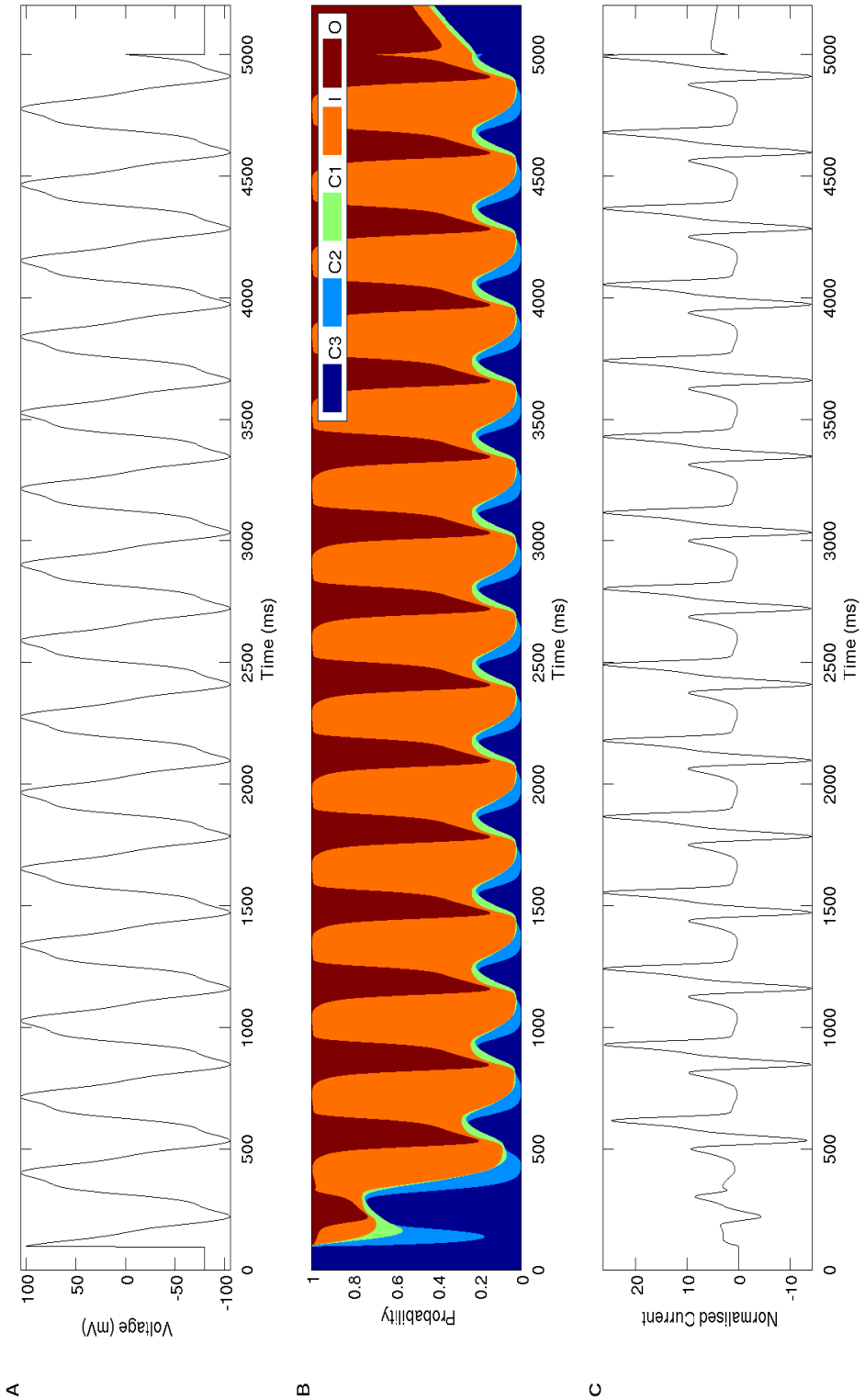


Figure 5.21: A shows *sine wave 4 protocol* which was designed with the aim of the channel having equal rates of transition between each of the conformational states represented within the Wang et al. (1997) model. In B and C the probabilities of the channel being in each state represented in the Wang et al. (1997) model and the simulated current from this model over the duration of the protocol are shown.

A number of interesting observations can be made from the results presented in Figures 5.14-5.21. The first is the differences in behaviour observed between the simulated current traces from different models when simulated with the new protocols as compared to the standard voltage-step protocols as shown in Figures 5.1-5.4. This is particularly apparent with the *sine wave 1 protocol*. Also of interest are the plots of probabilities of being in each state represented in the models over the duration of the protocols. Clear differences are evident between the behaviour of each of the models over the duration of the protocol in the probability plots for the protocols designed to maximise the difference between the current output from the Wang et al. (1997) and Mazhari et al. (2001) models in Figure 5.15, and also between the Di Veroli et al. (2013) model as in Figure 5.17. This allows us to see which aspects of hERG channel gating each protocol is interrogating.

Also of note is the *sine wave 3 protocol*, which was designed for the channel to spend an equal proportion of time in each of the states represented within the Wang et al. (1997) model. The extent to which this has been achieved is evident from the probability plot presented in Figure 5.19. While it is clear that less time is spent in the C_1 and C_2 states as compared to the other states, more time is seen to be spent in the C_1 and C_2 states than is seen in any of the other protocols. This is particularly noticeable for the standard voltage-step protocols, where the time spent in the C_1 and C_2 states is barely visible on the probability plots presented in Figures 5.1-5.4. Across the range of sine wave protocols, it is evident from the probability plots that much more time is spent in different states with evidence of more transitions between states than with both our new voltage-step protocol and the standard voltage-step protocols. This is particularly apparent with the *sine wave 1 protocol*, whose probability plot looks the most different to the standard voltage-step protocols with many more transitions between states due to the higher frequency sine waves represented within the protocol.

A further point to note is that the *sine wave 4 protocol* does not appear to be providing much new information after the first second of the protocol, with repeated behaviour after

this point. We may therefore expect there not to be as much information in this protocol as for the other protocols.

From this initial exploration of the behaviour of the existing hERG channel models when simulated with the novel protocols, we have ascertained that the new protocols appear to have achieved the first aim: the protocols induce a wider range of changes between states in a much shorter period of time than standard protocols. The particular intentions behind the design of each protocol appear to have been achieved in some capacity, as compared to the other protocols. However, there are a number of ways the design of these protocols could be improved further, as we discuss in Chapter 8. Each novel protocol appears to induce a variety of different behaviours from the existing models, which we hope will aid model selection. Our next step is to assess how well we expect to be able to determine the transition rate parameters in existing hERG channel models when using currents recorded in response to these new protocols.

5.3.2 Assessing the Identifiability of Voltage Protocols

In an attempt to assess the accuracy with which we would expect parameters of different hERG channel models to be determined when applying the new protocols, we analysed the posterior distributions of recovered parameters from simulated experimental data with noise (of 0.3% of the peak amplitude of the current trace added, as described in Section 5.2.3) after starting a MCMC chain at the true parameter values (as in each published model). This represents an ideal scenario, the best starting point for the MCMC chain, is the true simulated parameter set, and so allows us to infer the best possible accuracy with which we would expect to be able to recover parameters when using data from different voltage protocols with a comparable amount of noise to experiments.

To compare all of the different existing hERG channel models, including those which have been identified as being *a priori* unidentifiable, we reformulated such models so that all the

parameters were *a priori* identifiable, as described in Appendix A. Through comparison of the recovered parameter distributions for each model we could then infer details of the information content contained within each of the protocols. In addition, examination of the pairwise correlations between different parameters in the MCMC chains provides further information about identifiability of parameters, with strong correlations between parameters indicating possible identifiability issues if the parameter distribution is not narrow.

For each model presented in Table 3.1 (excluding the Winslow et al. (1999) and Noble et al. (1998) for which there were some numerical issues when simulating the models with some of the new protocols with parameters other than those defined in the original publications), and for each combination of the new protocols, we calculated the posterior distributions via MCMC. Figure 5.22 shows the parameter values sampled in the MCMC chain plotted against one another in the off-diagonal plots and the posterior distribution of each parameter on the diagonals. The true simulated parameter values are also indicated. An example of a model for which the parameters were found to be very accurately recovered when using the *sine wave 1 protocol* is shown in Figure 5.22.

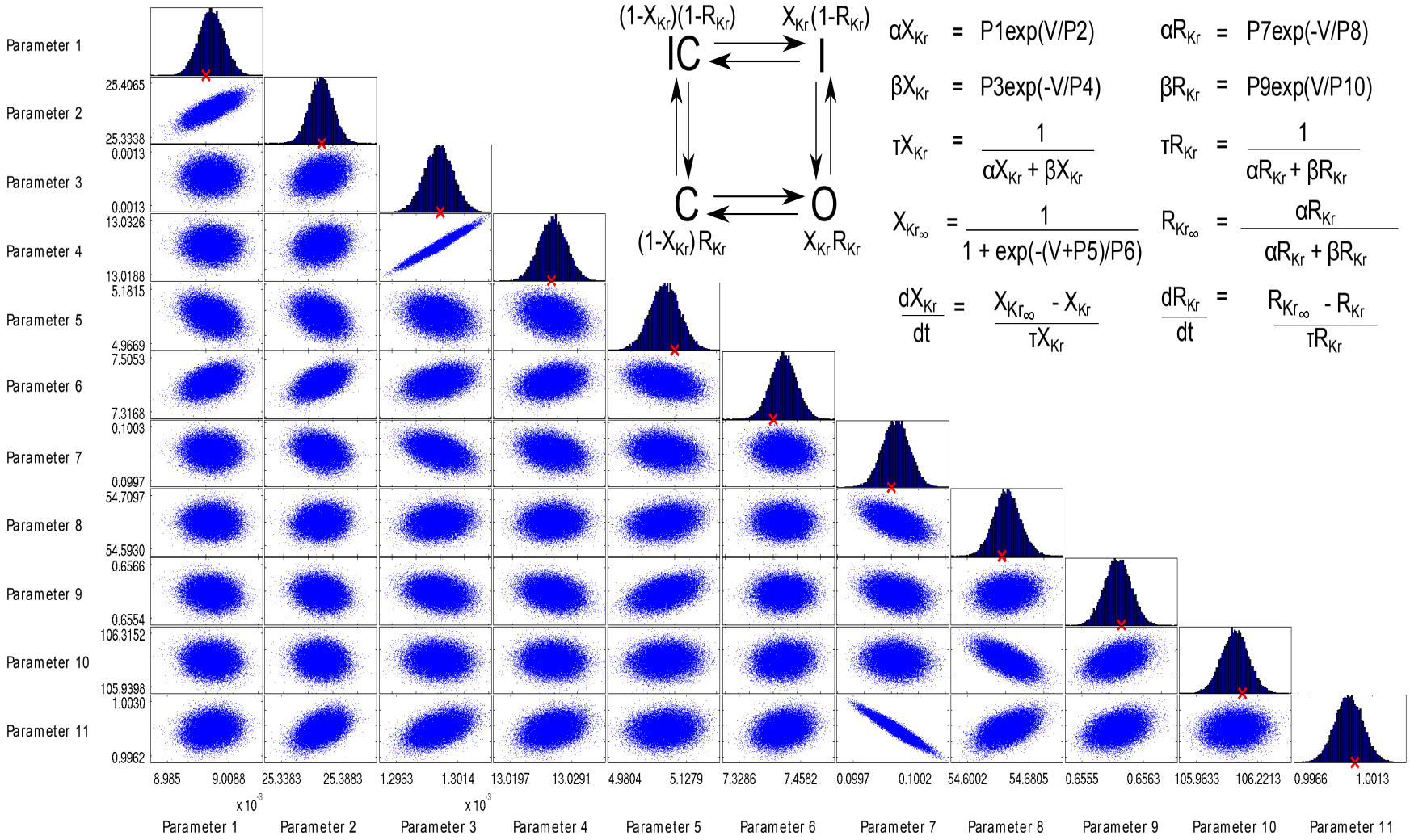


Figure 5.22: Matrix plot of the sampled MCMC chains for each parameter in the Lindblad et al. (1996) model when using the *sine wave 1 protocol*. Plots of the sampled parameters are plotted pairwise on the off-diagonals and histograms of the posterior distribution for each parameter are plotted on the diagonal. Parameter numbers correspond to those indicated in the model shown inset, with parameter 11 corresponding to the conductance value. Red crosses indicate true simulated parameter values for each model parameter. Results are based on MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in.

In Figure 5.22, we see that the posterior parameter distributions for each of the parameters are very narrow, with the true simulated parameter lying close to the mean of the distribution in each case. There are some strongly correlated parameters evident in this example, for example parameters 3 and 4 and parameters 7 and 11, but in both cases the posterior distributions are so narrow that this is not of concern. Mostly uncorrelated pairwise parameter plots and narrow normally distributed posterior estimates centered about the true simulated parameter value are observed. This result indicates good identifiability of the parameters in the Lindblad et al. (1996) model when using data generated by the *sine wave 1 protocol*.

In another example, where we consider the identifiability of the parameters in the Lu et al. (2001) model when using the *sine wave 1 protocol*, as shown in Figure 5.23, we see evidence of a possible unidentifiable parameter as in the case of parameter 15.

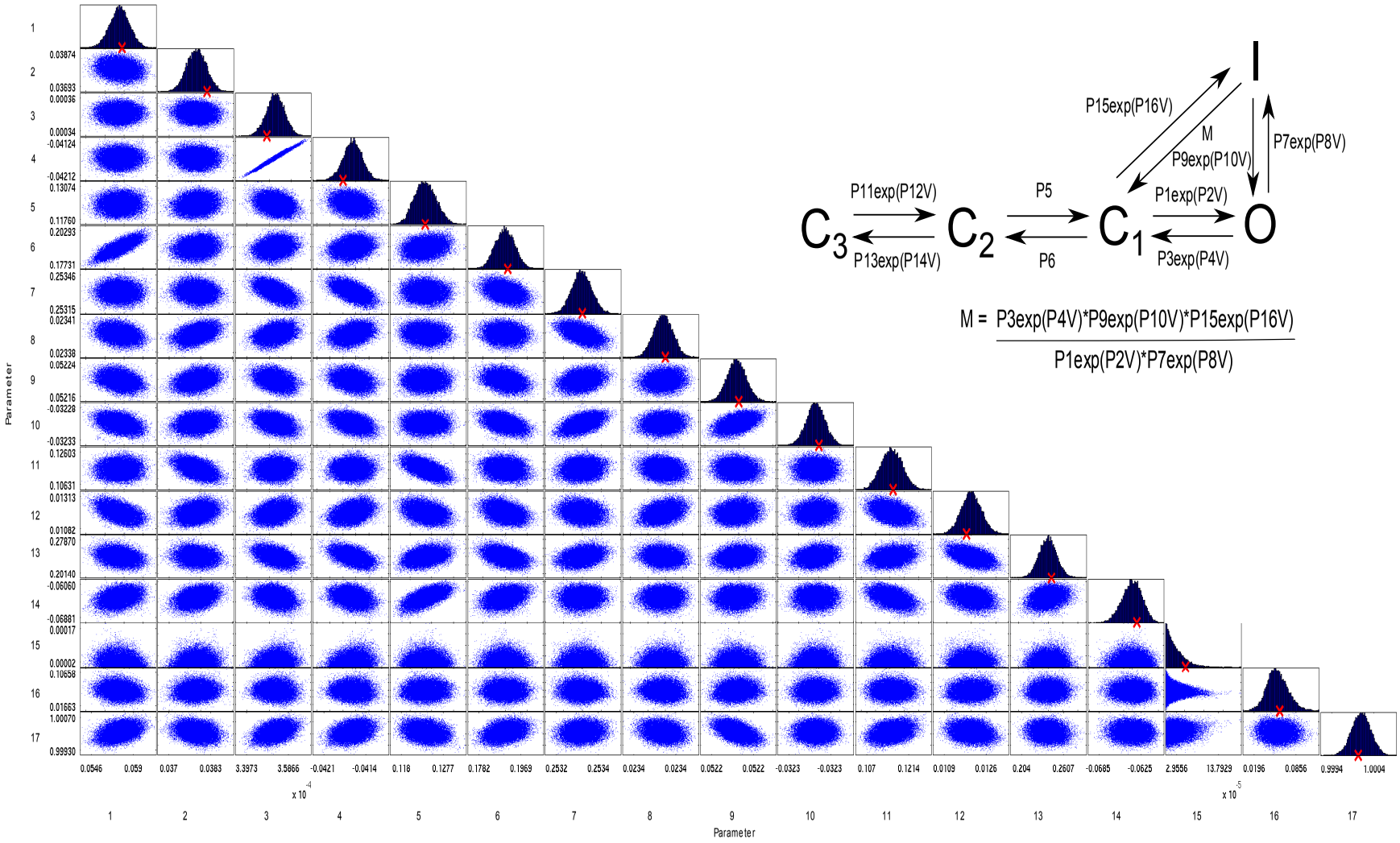


Figure 5.23: Matrix plot of the sampled MCMC chains for each parameter in the Lu et al. (2001) model when using the *sine wave 1 protocol*. Plots of the sampled parameters are plotted pairwise on the off-diagonals and histograms of the posterior distribution for each parameter are plotted on the diagonal. Parameter numbers correspond to those indicated in the model shown inset, with parameter 17 corresponding to the conductance value. Red crosses indicate true simulated parameter values for each model parameter. Results are based on MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in.

In Figure 5.23, we again see that most of the parameters in the Lu et al. (2001) model are identifiable when using the *sine wave 1 protocol*. However, for parameter 15 we see that this can take a wide range of different values, with the posterior distribution no longer being normally distributed. It is restricted to being positive by its prior but is seen to take a wide range of positive values, and the probability is high at the edge of the prior. The apparent unidentifiability of this parameter is due to its small size. The parameter has such little influence on the behaviour or the model when simulated with this particular protocol that it can take a wide range of values and this does not affect the simulated current enough to warrant that parameter set being rejected. This suggests that if a parameter is having such little impact on the model behaviour that it cannot be reasonably estimated, then it could be excluded from the model and a model of lesser complexity could be proposed instead for this particular protocol.

Another example where unidentifiable model parameters are seen is shown in Figure 5.24 where the O'Hara et al. (2011) model is simulated with the *sine wave 1 protocol*.

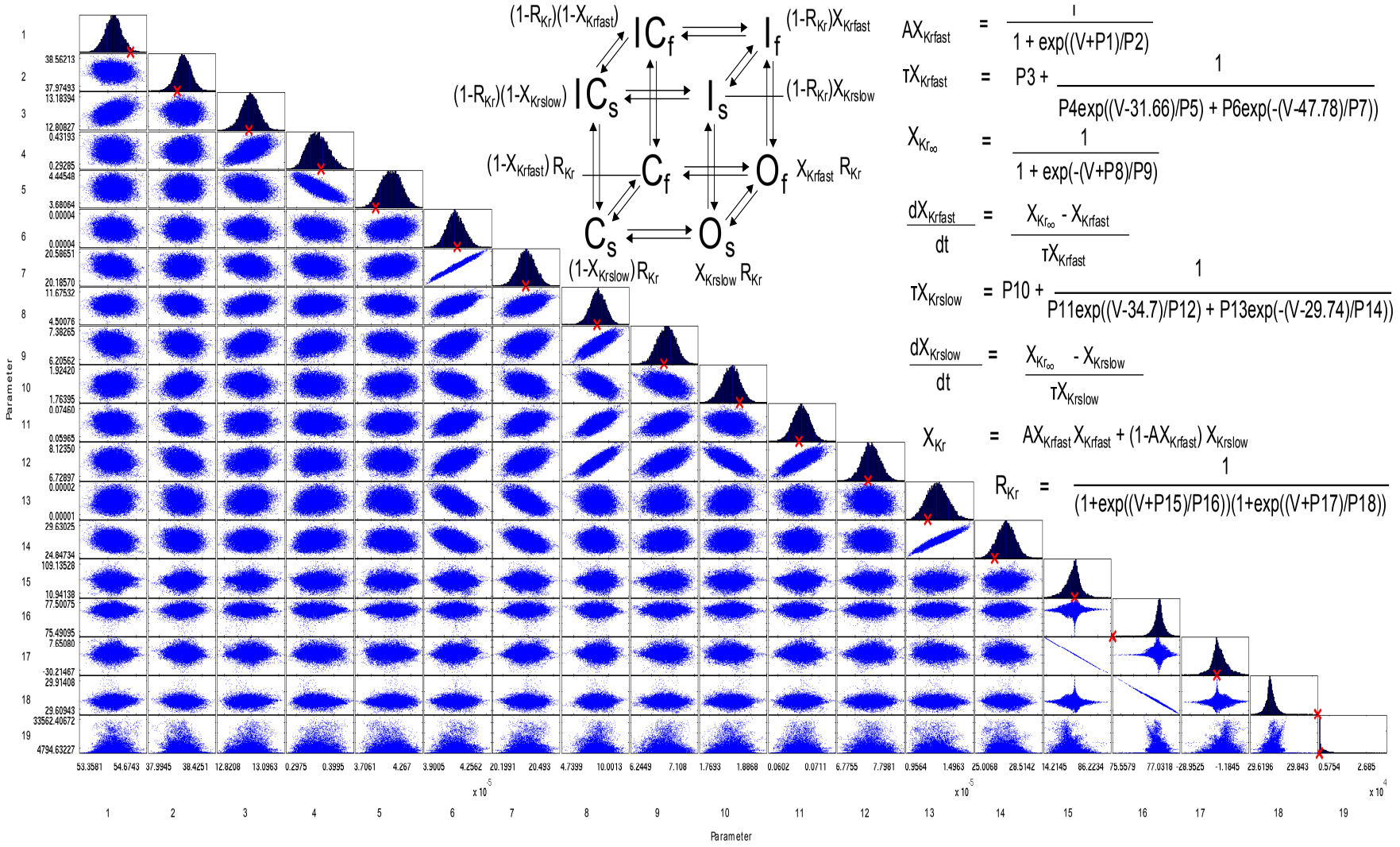


Figure 5.24: Matrix plot of the sampled MCMC chains for each parameter in the O'Hara et al. (2011) model when using the *sine wave 1 protocol*. Plots of the sampled parameters are plotted pairwise on the off-diagonals and histograms of the posterior distribution for each parameter are plotted on the diagonal. Parameter numbers correspond to those indicated in the model shown inset, with parameter 19 corresponding to the conductance value. Red crosses indicate true simulated parameter values for each model parameter. Results are based on MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in.

In Figure 5.24, more correlated parameter estimations are seen than in the examples in Figures 5.22 and 5.23. The posterior parameter distributions are also much wider than those in the previous examples and, by the end of the burn in period, some parameters have moved some distance away from their true parameter values, meaning that the model has found a new region of parameter space which has almost the same likelihood as the original parameter set. In this particular example, the conductance value (parameter 19) and the parameters defining the time-independent transition from the closed to the inactivated state (parameters 15-18) are the most unidentifiable when using *sine wave 1 protocol*. As the conductance parameter is multiplied by this time-independent transition rate when calculating the current, it appears that the conductance parameter encapsulates this transition term (which varies only with voltage). The conductance parameter becomes very large and eliminates the impact of the inactivation transition term from the model behaviour, meaning that the parameters in this term can take a range of values without affecting the behaviour of the model and the likelihood of the proposed parameters. There is therefore clearly insufficient information in *sine wave 1 protocol* to constrain and estimate all the parameters in a model with the same structure as that in the O'Hara et al. (2011) model.

The information content of each voltage protocol for determining the parameters of each model can be compared. A voltage protocol is better suited for determining the parameters of a kinetic model if the recovered parameter distribution has a mean close to the true simulated parameter value and if the width, or standard deviation, of the recovered distribution is small. We assessed the identifiability of all model parameters when using the *sine wave 1 protocol* in this way. For each MCMC chain we calculated the Mean Absolute Percentage Error (MAPE) in the parameter estimates as compared to the true simulated parameter values using MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in. The MAPE is calculated using the formula

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N} \left|\frac{\boldsymbol{\theta}^* - \boldsymbol{\theta}_i}{\boldsymbol{\theta}^*}\right| \times 100\%, \tag{5.23}$$

where $\boldsymbol{\theta}^*$ is the true simulated parameter set, $\boldsymbol{\theta}_i$ is the parameter set accepted in the MCMC

chain at iteration i and N is the number of samples in the chain. It should be noted that for some model/protocol combinations, some model parameters showed strong dependencies and had not fully converged to the new model parameters with higher likelihood (even after starting at the true model parameters). Such an example of this behaviour was seen in the O'Hara et al. (2011) model as we just described. However, this is still reflected in the MAPE score as it compares each point in the chain with the true underlying parameter values. In Figure 5.25 we present the MAPE calculated for each parameter in each model when using *sine wave 1 protocol*.

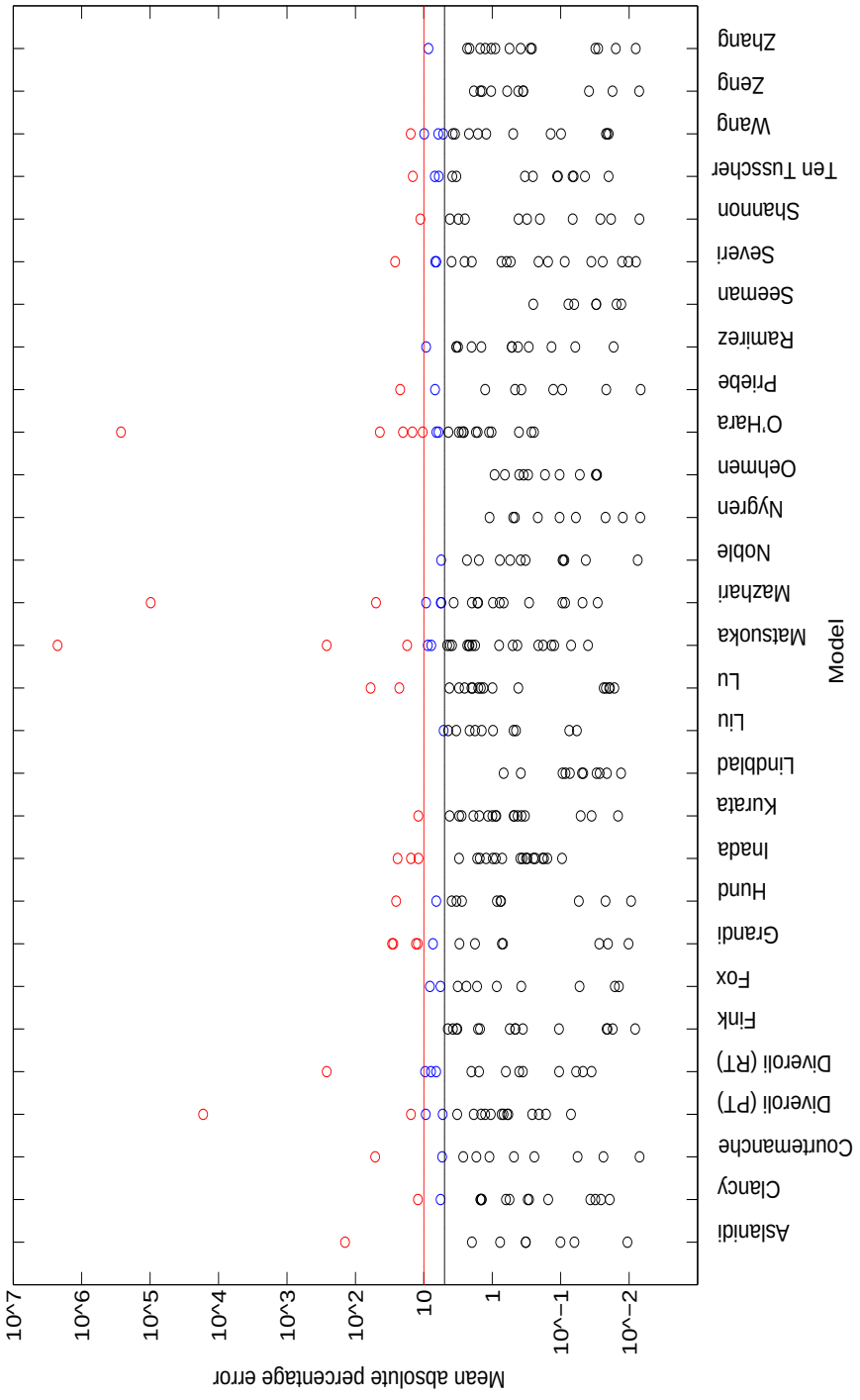


Figure 5.25: Dot plot of MAPE of each parameter in each existing hERG channel model when the *sine wave I protocol* is applied. The black and red lines mark where the MAPE is 5% and 10% of the parameter value respectively. Red dots represent parameters with a MAPE greater than 10%, blue between 5 and 10% and black less than 5%.

From Figure 5.25 it can be seen that for some models, the MAPE for all the parameters is less than 5%, indicating good parameter identifiability for these models with *sine wave 1 protocol*. The MAPE for most parameters in all models is less than 10% but there are some models with a few parameters with MAPE greater than 10%. However, in the majority of cases the reasons for this can be explained mostly by the presence of very small parameter values or due to one parameter growing so large that it dominates another term in the model allowing the parameters in that term to take on a range of values, as was described with the O'Hara et al. (2011) model. Therefore it is encouraging to see that even with the 5.2 second *sine wave 1 protocol*, it allows us to determine the parameters of the majority of the existing hERG channel models relatively accurately. Another point to note is that the MAPE, will be somewhat larger for smaller parameters and so in some cases outlying values could be a consequence of very small parameters present in the model. This is of note due to differences in identifiability of models with the same model structure, for example the Zeng et al. (1995), Hund and Rudy (2004) and Shannon et al. (2004) models.

We compared the identifiability of model parameters when applying different combinations of the voltage protocols. To present the results for all the voltage-protocol combinations, a model parameter was defined as being identifiable if the MAPE was less than 5%. This is a subjective choice but is chosen merely as a guide to the appropriateness of different protocols for determining model parameters. In Figure 5.26 we summarise these results by indicating the percentage of parameters in each model found to be identifiable based on this indicator when using different combinations of protocols.

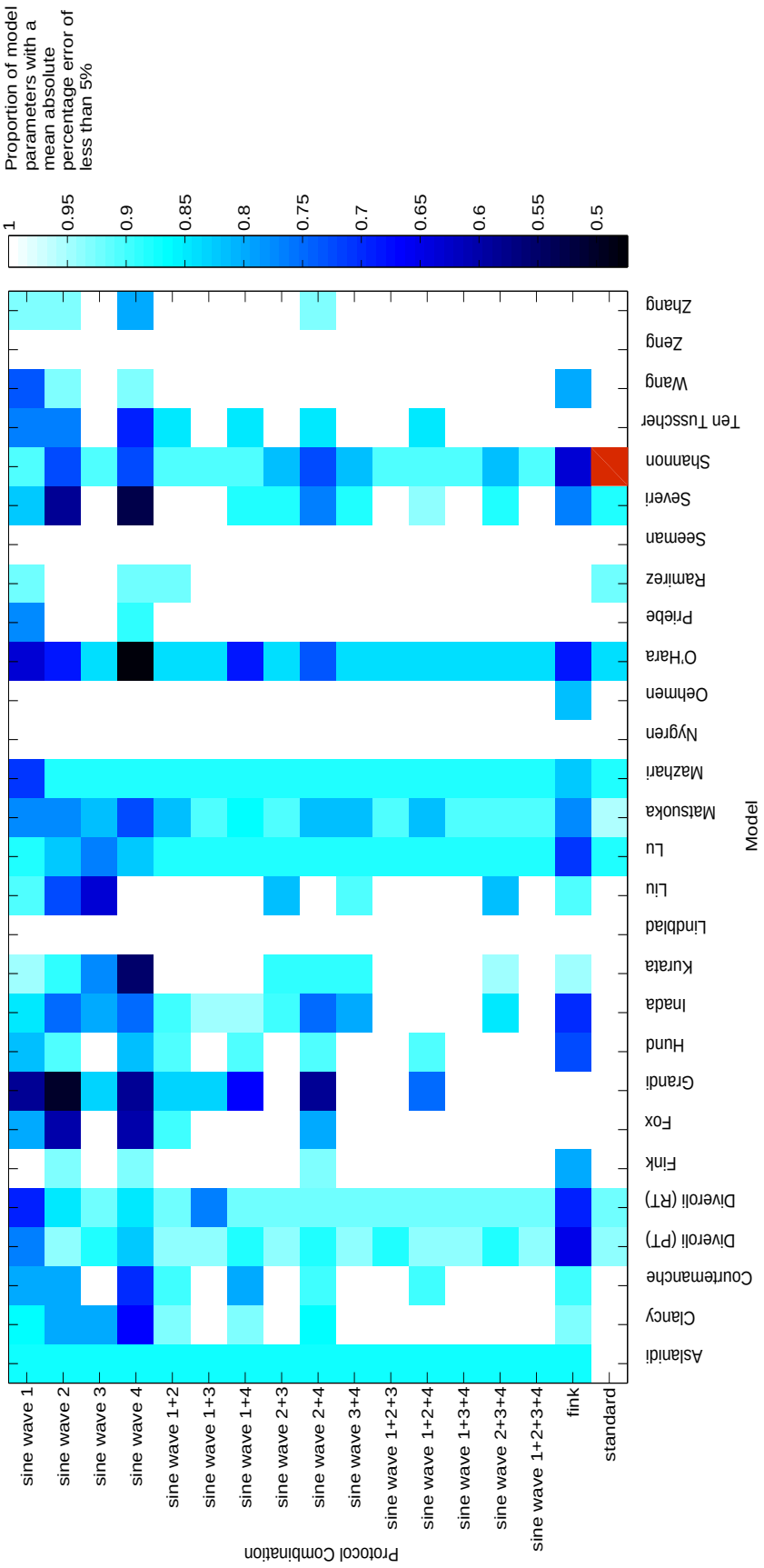


Figure 5.26: Plot of percentage of identifiable parameters in each hERG model when using different voltage protocol combinations. An identifiable parameter was defined to be one for which an MAPE of less than 5% was calculated from an MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in. Numerical errors were encountered when simulating the Shannon et al. (2004) model standard voltage-step protocol intended to investigate inactivation kinetics and so the parameter identifiability for this model under the standard protocols is not included in this figure, as indicated by the shaded red entry.

We see that for some models in Figure 5.26, all, or almost all, of the model parameters are identifiable with every protocol combination, such as the Zeng et al. (1995), Lindblad et al. (1996), Nygren et al. (1998) and Seemann et al. (2003) and Oehmen et al. (2002) models when written in an *a priori* identifiable form. Other models, such as the O'Hara et al. (2011) and Ten Tusscher et al. (2004) models appear to have a large number of unidentifiable parameters when using many of the different protocol combinations. We can see however that the best single sine wave protocol, with which models had the most identifiable parameters was the *sine wave 3 protocol*. As increasing numbers of sine wave protocols were added to the protocol combination, in general more model parameters became identifiable. The protocol combinations for which most parameters across the range of models are identifiable is using the *sine wave 1, 2 and 3 protocols* together (using the *sine wave 4 protocol* in addition increases the parameter identifiability in one model only marginally). The shortest protocol combination with which many model parameters are identifiable is the combination of the *sine wave 1 and 3 protocols*. The results presented in Figure 5.26, suggest that we will be able to use these sine wave protocols to determine models of hERG channel kinetics, with evidence of sufficient information contained within the protocols to allow parameters across a range of different existing hERG channel models to be identifiable. We will use the information presented in Figure 5.26 to help us to identify suitable protocol combinations for fitting to experimental data in Chapter 7.

We now compare the accuracy with which we expect to obtain parameter estimates using the new protocols against standard voltage protocols. We compare the posterior parameter distributions obtained when using the protocol designed by Fink and Noble (2009) which has been shown to allow some parameters of existing hERG channel models to be locally identifiable. We also compare the use of the four standard voltage-step protocols as described in Section 5.1.1. Figure 5.27 shows the parameter distributions for each parameter in the Wang et al. (1997) model when using each of these protocols as compared to using the *sine wave 1 protocol*. The *sine wave 1 protocol* was chosen for comparison here as it is the protocol we use for comparing variability between cells in Chapter 7 so we wanted to assess and compare

its ability for estimating parameters of the Wang et al. (1997) model before being used for this purpose.

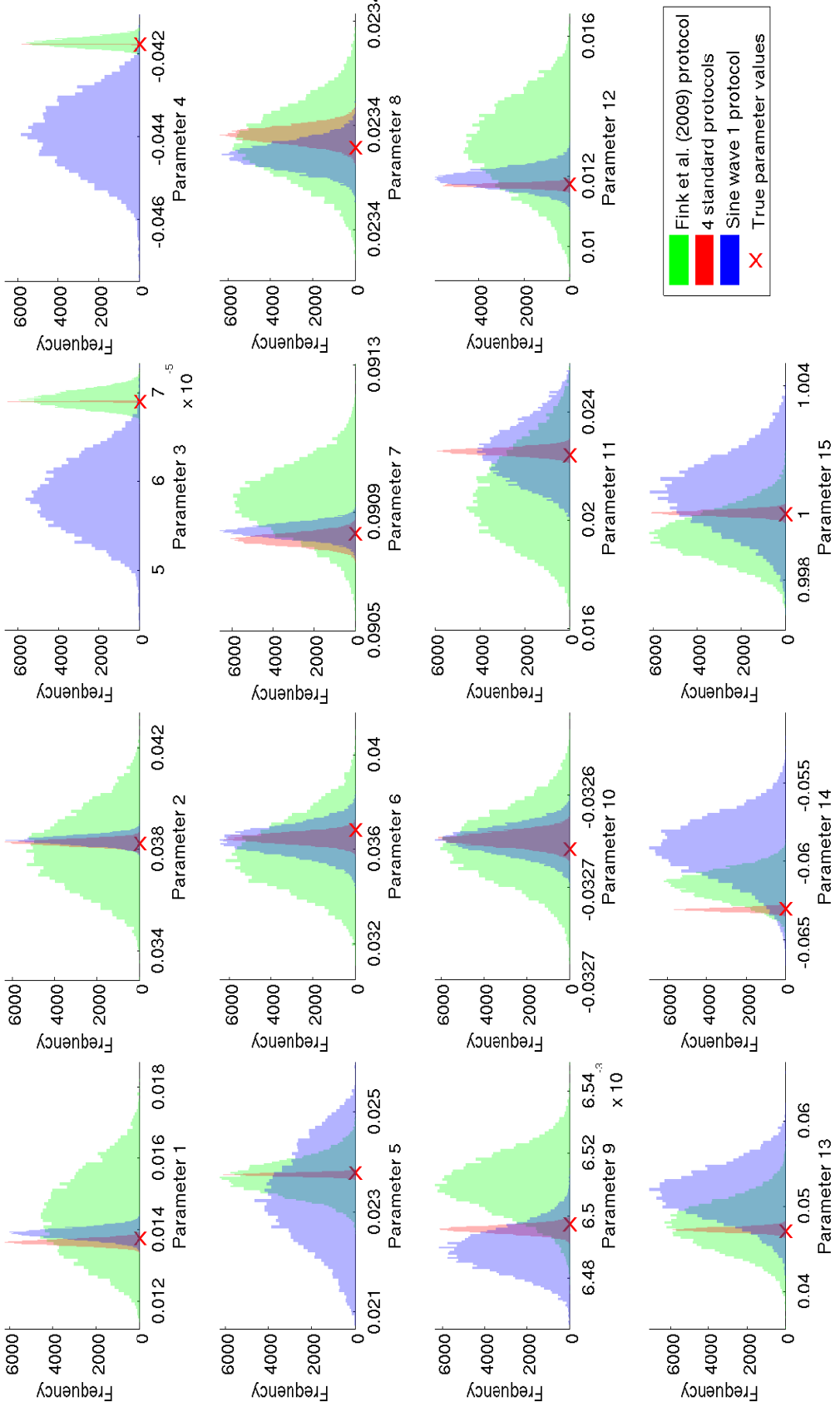


Figure 5.27: Comparison of posterior parameter distributions for the parameters of the Wang et al. (1997) model when using the combination of 4 standard voltage-step protocols, the Fink and Noble (2009) protocol and the *sine wave 1 protocol*. Red crosses indicate true simulated parameter values for each model parameter. Results are based on MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in.

We can see from Figure 5.27 that the set of 4 standard voltage-step protocols allows the parameters of the Wang et al. (1997) model to be estimated more accurately than when using the *sine wave 1 protocol*. For many of the model parameters, the *sine wave 1 protocol* allows parameters to be estimated more accurately than the Fink and Noble (2009) protocol. The noticeable exceptions to this are parameters 3 and 4. This observation indicates that the *sine wave 1 protocol* does not provide as much information for parameterisation of these parameters as it does for other parameters in the model. Based on this, we anticipate that a combination of protocols may be more suitable for accurately determining all model parameters for this particular structure. This is something we will investigate further in practice in Chapter 7. In Figure 5.28 we compare the parameter distributions using *sine wave 1, 2 and 3 protocols* with the Fink and Noble (2009) protocol; and in Figure 5.29 we compare the combination of *sine wave 1, 2 and 3 protocols* (one of our protocol combinations which allows more parameters to be identified as identifiable) with the 4 standard voltage-step protocols.

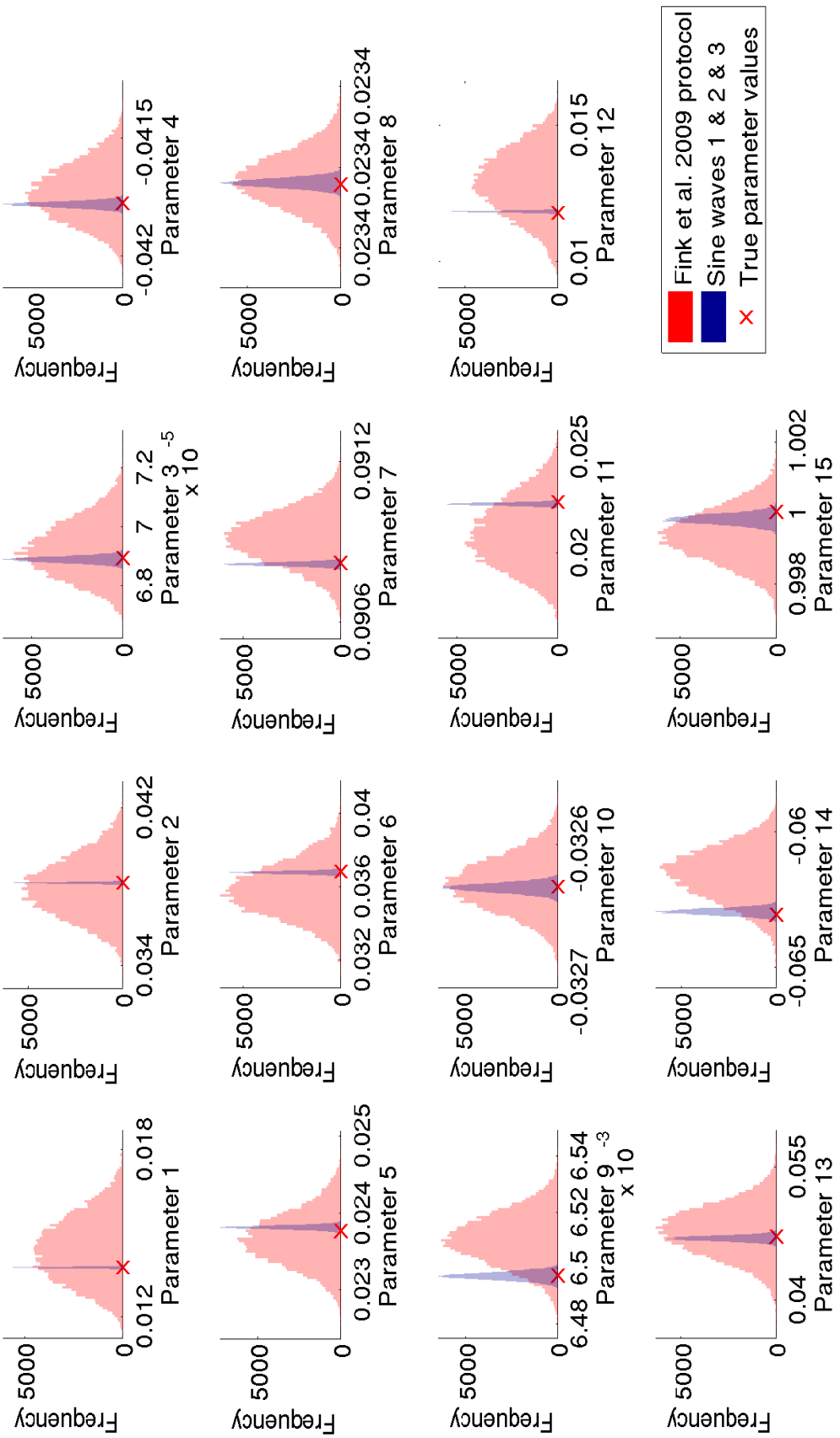


Figure 5.28: Comparison of posterior parameter distributions for the parameters of the Wang et al. (1997) model when using the Fink and Noble (2009) protocol and the *sine wave 1, 2 and 3 protocols*. Red crosses indicate true simulated parameter values for each model parameter. Results are based on MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in.

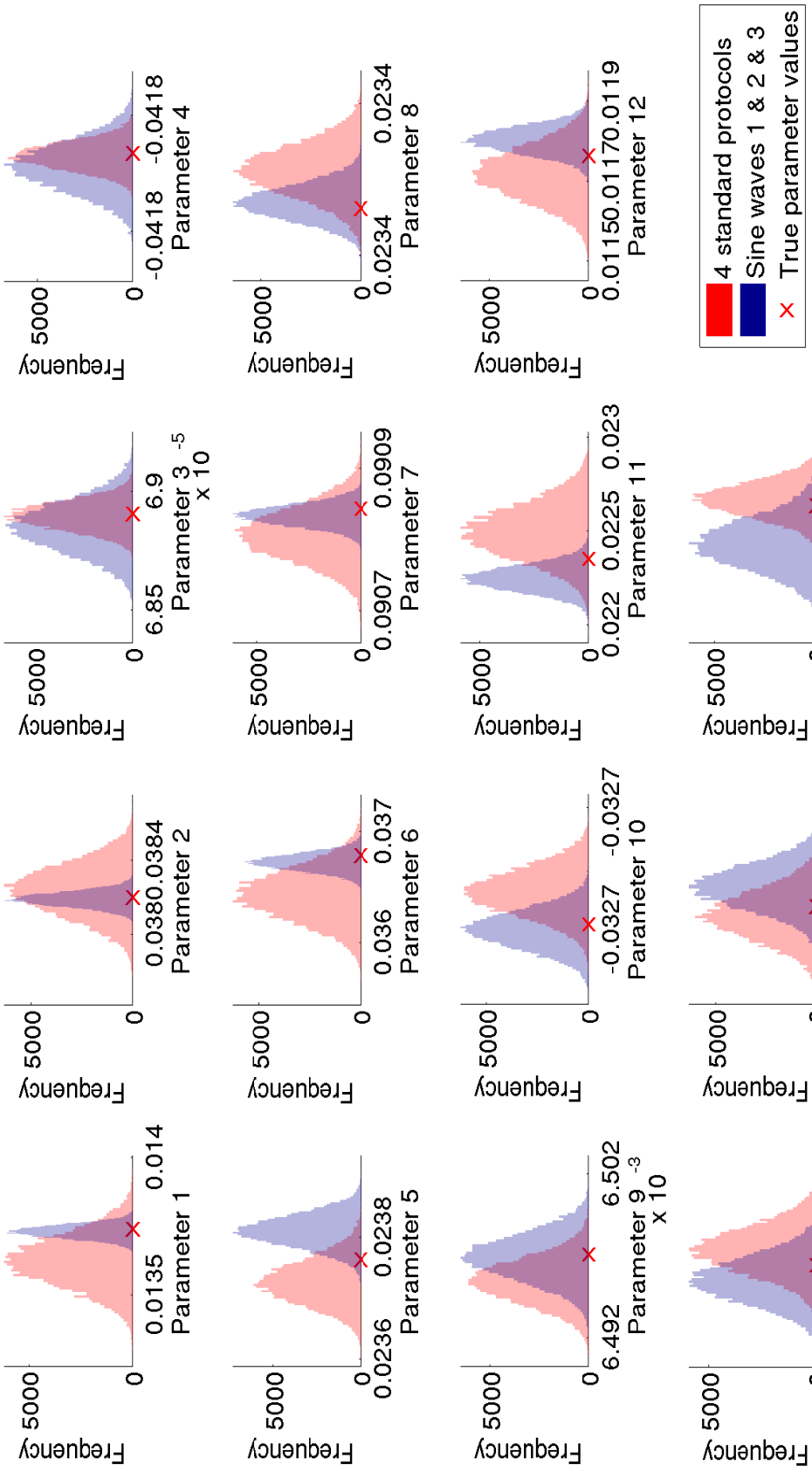


Figure 5.29: Comparison of posterior parameter distributions for the parameters of the Wang et al. (1997) model when using the combination of 4 standard voltage-step protocols and *sine wave 1, 2 and 3 protocols*. Red crosses indicate true simulated parameter values for each model parameter. Results are based on MCMC chains of length 200,000 with the first 50,000 entries discarded as burn in.

In Figure 5.28 we see that we are able to estimate all the parameters in the Wang et al. (1997) model much more accurately when using our combination of 3 sine wave protocols than the Fink and Noble (2009) protocols. However, note that the parameter distributions for the estimates using the Fink and Noble (2009) protocol are relatively narrow for many parameters, but the combination of 3 sine wave protocols provides increased accuracy. The Fink and Noble (2009) protocol has a length of 13.5 seconds and the combination of 3 sine wave protocols has a combined length of 15.6 seconds, so a much higher level of accuracy is observed for a comparable length of protocol.

In Figure 5.29, we compare the ability to estimate parameters of the Wang et al. (1997) model when using the set of three sine wave protocols (1, 2 and 3) and the series of standard voltage-step protocols representative of those which are typically used for determining transition rates in ion channel kinetic models. For most parameters we see a comparable, or increased, level of accuracy in the parameter estimates for each parameter when using the combination of three sine wave protocols as compared to using the 4 standard voltage-step protocols. We are also able to achieve this with a protocol of the length of 15.6 seconds as opposed to a combined length of 2 minutes 25 seconds with the set of standard protocols. We also see an improvement in the accuracy with which parameters 3 and 4 are estimated as compared to when using just the *sine wave 1 protocol* as shown in Figure 5.27; this indicates that additional information contained within the *sine wave 2 and 3 protocols* is helping to identify appropriate transition rate parameters for the backwards transition between the open state and the most proximal closed state in particular.

5.4 Discussion

In this chapter we have explored the behaviour of the range of existing hERG channel models when simulated with a wide variety of different voltage protocols, ranging from standard voltage-step protocols typically performed in patch clamp experiments, to physiological and pathophysiological action potential protocols, to a range of novel non-standard voltage pro-

tocols which we have designed for establishing hERG channel kinetics. We have seen that models can behave very differently to one another even when simulated with protocols very similar to those with which they were constructed. They are also seen to predict a variety of different responses to patterns of electrical stimulation which would be expected in cardiac cells — action potential protocols. We have also shown how differences in the hERG channel representation within an action potential model can propagate through to cause very different behaviour when predicting drug-induced effects on the action potential. It would therefore be of value to be able to determine whether such observed differences are due to differences in the experimental conditions, natural variability in the cells used to generate data to construct the models; or perhaps whether the results are due to the selection of an inappropriate model, the result of not fully optimised parameterisation, lack of identifiability of parameters, or a lack of validation of the proposed model.

From the analysis of the behaviour of the novel protocols presented in this chapter, a number of observations have been made. The first is that the *sine wave 4 protocol* does not appear to provide as much information as the remaining protocols—as is evident from the probability plot in Figure 5.21, with repeated behaviour of transitions between states after the first 1000 ms. This observation is also concordant with the results presented in Figure 5.26, where we see that fewer parameters are identifiable within the existing models with this protocol than with the other single sine wave protocols applied. It also follows that the addition of this protocol to any other protocol combination would not add much additional information for purposes of determining models of hERG channel kinetics.

The next observation is that the small oscillations in the simulated current in the *sine wave 2 protocol*, as can be seen in Figure 5.16, may not be expected to be easily differentiated from noise during experiment, however, we will explore whether this is the case when we perform this protocol in experiment as described in Chapter 6.

As seen in Figures 5.27 and 5.26, the combination of standard voltage-step protocols as

proposed by Fink and Noble (2009) does reasonably well at providing information for determining model parameters of existing models in 13.5 seconds. Fink and Noble (2009) recognise that this is not an optimised version of the protocol, but provides information with little redundancy. Although not presented in this thesis, we also designed a 9 second voltage-step protocol to maximise the difference between the currents of the Wang et al. (1997) and Mazhari et al. (2001) models. The optimisation problem is more complex; as it is required to optimise both the amplitudes of the voltage steps and the duration for which they are held. However, it would be of interest to explore optimising a voltage-step protocol more thoroughly and compare with the results from the sine wave protocols presented here.

In this chapter we have presented a range of novel protocols designed to facilitate selection of hERG channel model structures and to aid parameterisation. The short length of the protocols and dynamic nature should allow us to gather more information about hERG channel kinetics in a shorter time. This efficiency should enable the collection of more data for the purposes of validation and further interventions. We have compared the ability to estimate parameters across the full range of existing hERG channel models when using different protocol combinations. We have demonstrated that we were able to estimate parameters when using a combination of three sine wave protocols more accurately than when using the Fink and Noble (2009) protocol and, for many parameters, as well as, or better than using the standard series of voltage-step protocols. This exercise allowed us to identify for which parameters in which models there is insufficient information in the protocols to estimate. In addition, it allowed us to identify suitable protocol combinations for fitting experimental data to select and parameterise models of hERG channel kinetics as we will describe in Chapter 7. The results presented in this chapter, provide promise that our novel protocols may be appropriate for use for this purpose.

5.5 Summary

In this chapter we have explored the behaviour of the range of existing hERG channel models when simulated with different voltage protocols. We have presented novel sine wave protocols which aim to explore hERG channel kinetics more efficiently than standard voltage-step protocols. The protocols were designed with the intention of aiding in the selection or parameterisation of a mathematical model of hERG channel kinetics. We have investigated the accuracy with which we expect to be able to estimate parameters of existing hERG channel models when using combinations of the new protocols. We have done this using a Bayesian inference approach which allows us to observe the distribution of possible values for each parameter estimate. This analysis represents a best case scenario as the data used is simulated with added noise, but provides promise for the appropriateness of such protocols for determining a model of hERG channel kinetics. The use of such protocols offers a large number of possibilities for further investigation of hERG channel kinetics using data recorded from the same cell, which may not be possible with a series of protocols that take several minutes to perform.

The results presented in this chapter indicate the potential of the new protocols designed for use in determining hERG channel kinetics. They also provide a guide for selecting the most appropriate combination of protocols for this purpose. In Chapter 6 we perform the protocols described in this chapter in patch clamp experiments using cell line cells expressing hERG and in Chapter 7 we use the experimental data generated together with the results presented in this chapter to select and calibrate models of hERG channel kinetics.

6

Experimental Recordings Using the Whole-Cell Patch Clamp Technique

In this chapter we describe the experimental techniques used to record currents flowing through hERG-expressing cells when applying the voltage protocols described in Chapter 5. We outline the whole-cell patch clamp setup used and describe the electrical theory underpinning it. We discuss some additional considerations and checks that we performed due to the non-standard nature of the sine wave protocols. We also describe some interesting observations made while collecting the experimental data, in particular focusing on the variability observed between recordings of the voltage protocols in different cells. We conclude by quantitatively evaluating the ability of existing models to describe our experimental recordings.

6.1 Outline of the Whole-Cell Patch Clamp Technique

In this chapter we describe the whole-cell patch clamp technique used to perform the voltage protocols described in Chapter 5 in cell-line cells expressing hERG channels. I learnt how to perform the patch clamp technique and subsequently performed all the experiments described in this chapter. As introduced in Section 2.2.2, the patch clamp technique enables investigation of the kinetics of voltage-gated ion channels by allowing the measurement of current flow in response to different applied voltage protocols. The principle basis of this technique is that a glass microelectrode is applied to the surface of a cell to isolate a small

section ("patch") of the membrane. The voltage across this patch is held steady ("clamped") by a feedback amplifier so that the current flowing across the membrane can be measured. There are a number of different variations of the patch clamp technique, which are described in Molleman (2003). The method we use is *whole-cell patch clamp* which allows currents to be recorded from an entire cell. The equivalent circuit diagram of the whole-cell patch clamp setup is shown in Figure 6.1.

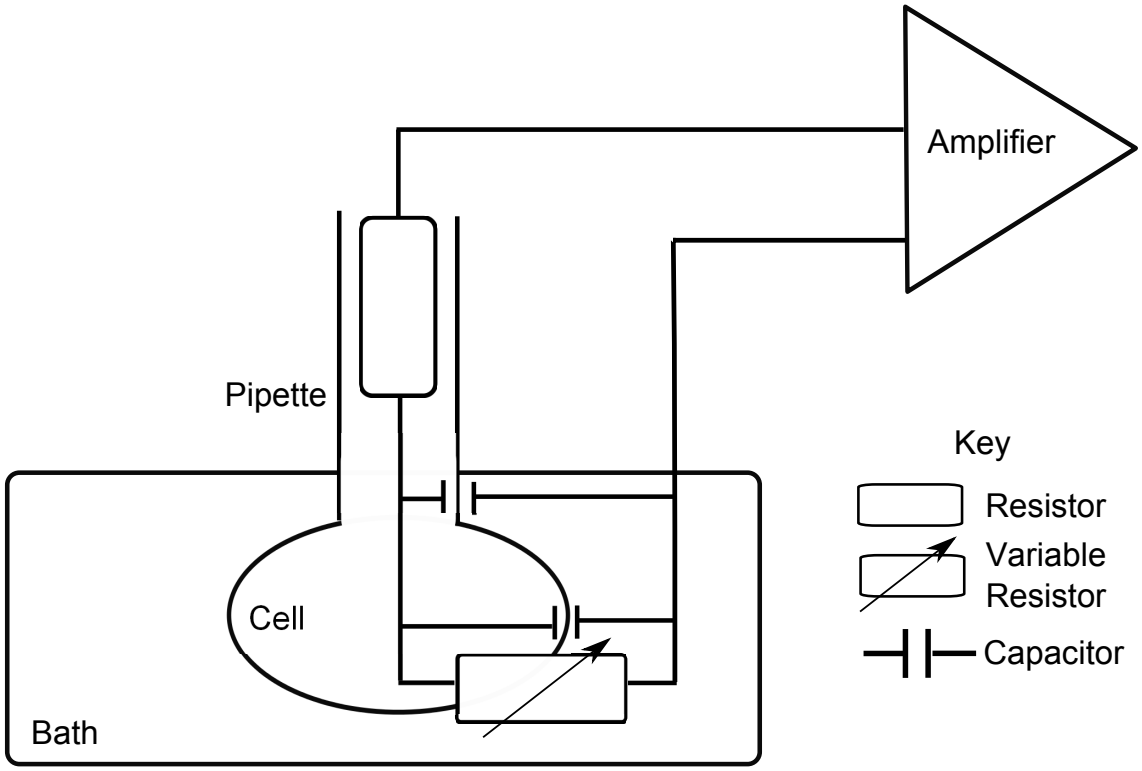


Figure 6.1: Equivalent circuit diagram of the whole-cell patch clamp technique. Points of resistance and capacitance in the circuit are indicated. The role of each component is indicated in the diagram is discussed in detail in the text.

We now describe the methodology and each of the components of the whole-cell patch clamp setup in detail and outline considerations taken when performing this technique. The amplifier is the key component of the setup — it contains the necessary circuitry both to clamp the voltage at the desired level and to inject current; as we describe in more detail throughout this section.

When performing patch clamp experiments, the first step is to use an inverted microscope to locate an appropriate cell on a cover slip which lies in a bath placed on the microscope stage. A grounded electrode in the bath sets the outside of the cell membrane at a potential of 0 mV. To obtain a patch clamp recording from the cell, a glass pipette containing an electrode is sealed against the cell membrane of the selected cell. The patch of membrane under the pipette is then disrupted to enable electrical access to the interior of the cell — this is achieved by applying suction to the cell using a syringe. Pipettes are made of borosilicate glass and fire polished to smooth the tips; this prevents destruction of the cell when contact is made and enables a high resistance seal to be formed with the cell membrane.

The pipette is filled with a conducting salt solution whose composition resembles that of the cytoplasm of the cell and therefore the primary ionic species added to the solution is potassium, as is concordant with Table 2.1. Similarly, the bath is filled with solution which mimics the extracellular environment of the cell, which in the case of our cell-line cells corresponds to that of the culture medium. Therefore, the main ionic components of the bath solution are sodium and chloride. The pipette is moved using a micromanipulator to guide its tip towards the cell. The manipulator is placed on a vibration isolation table to prevent unwanted movement of the electrode. The coarse setting on the micromanipulator is used to lower the pipette into the bath solution. The pipette is moved towards the cell using the microscope, iteratively switching the focus between the cell and the pipette tip. The pipette tip is then moved very close to the cell using the fine micromanipulator setting. Suction is applied via a syringe to establish a seal of at least one gigaohm between the cell membrane and the pipette. The cell membrane is then ruptured, placing the cell cytoplasm and the pipette solution (and therefore the electrode) in direct contact. The gigaohm seal resistance allows electrical isolation of the currents flowing across the membrane.

The electrode in the pipette is then electrically connected to the inside of the cell membrane and the ground electrode in the bath to the outside. This configuration allows the potential across the cell membrane to be measured. In voltage clamp mode, the membrane potential

is compared to an input voltage protocol. If there is a difference between the two voltage levels then the amplifier injects a current into the cell via the electrode as part of a negative feedback loop. The injected current required to equate the input and measured membrane potentials therefore corresponds to the current flowing across the membrane (Hamill et al., 1981; Molleman, 2003). The current from the electrode in the pipette is low pass filtered and sent to an analogue-to-digital converter and recorded using data acquisition software.

The pipette electrode is a chloride-coated silver electrode which electrically connects the pipette solution and facilitates conductance of current in both directions along the wire. The grounded bath electrode is also a chloride-coated silver wire. We add a chloride layer to the silver wire by immersing it in bleach. Before patch clamp recordings can be successfully made using this setup, a number of adjustments have to be made to optimise the quality and accuracy of the voltage-clamp of the membrane, as we describe in Section 6.1.1.

6.1.1 Series Resistance and Capacitive Current Compensation

As evident in the circuit diagram in Figure 6.1, an applied voltage encounters resistances in the pipette as it is applied to the cell. The sum of the pipette resistance and all additional resistances between the two silver chloride electrodes which are in series with the membrane is termed the *series resistance*. Series resistance affects the time constant for charging of the patch membrane and can lead to a discrepancy in the measured membrane potential and the true potential difference across the membrane when a current flows. This will consequently introduce artefacts in the recorded currents. The effect of series resistance should be minimised for successful patch clamp recordings to be made.

The patch clamp amplifier allows series resistance to be compensated, with analog electronics, via a positive feedback circuit. The level of compensation is set using dials on the amplifier. The voltage error estimated as being caused by this resistance is compensated for by proportionally increasing the voltage applied (Sakmann and Neher, 2009).

Series resistance compensation will fail if any currents in addition to the injected current are measured. The glass pipette in the experimental setup, in addition to the cell membrane, acts as a capacitor. As the applied voltage changes, the capacitance in the circuit will change. A sudden large change in voltage would provoke a large capacitive current response. These currents could then result in errors in the series resistance compensation feedback circuit. However, capacitive current can also be compensated using the amplifier controls. Fast and slow capacitive currents are compensated separately. We initially compensate for pipette capacitance, before the cell membrane is disrupted and added to the electrical circuit. Adjusting the fast controls compensates for the majority of the pipette capacitance, but a small amount of the total capacitive current is eliminated using the slow controls.

The cell membrane capacitance depends on the series resistance, and as these are in series, they are iteratively compensated for using their respective amplifier controls. A perfect patch clamp recording would have 100% series resistance compensation, however, due to the nature of the compensation circuitry and its propensity to oscillations, 80% compensation is considered good (Molleman, 2003; Molecular-Devices, 2008). It is particularly important to perform adequate series resistance and capacitance compensation when applying the novel sine wave protocols due to the more rapid changes in voltage compared to standard voltage-step protocols.

However, the compensation methods described in this section are not perfect and to further eliminate remaining artefacts in current recordings, the method of *leak subtraction* can be employed. This method reduces capacitance currents and also helps to reduce the effect of an additional leak current which is due to an imperfect seal between the pipette and the cell membrane. The method we use is P/N leak subtraction as described by Bezanilla and Armstrong (1977). In this method, N repeats of a scaled down version of a voltage protocol are repeated before the protocol is performed. The scaled version of the protocol (scaled by a factor of 1/N) is performed centered around a very negative potential (holding potential in

our case) to ensure that the scaled protocol has voltages all within a range that is too negative to cause activation of ion currents. Currents recorded in the N scaled versions of the protocol are then summed together and subtracted from the current recorded in the full unscaled voltage protocol to produce the P/N leak subtracted current recording. It is often not necessary to perform P/N leak subtraction with standard voltage-step protocols, but due to the nature of the voltage changes in the sine wave protocols we found this to be beneficial.

There are a number of sources of noise which may impact on patch clamp recordings. Noise can come from power cables in the setup, micromanipulator motors, background electrical activity generated in the electronic components of the recording system and other devices in the experimental setup. Experiments are performed within a Faraday cage to reduce electromagnetic interference from other electrical devices in the proximity of the setup. In our recordings we do not apply any software filtering of the recorded currents, however additional filtering could be applied to reduce the effects of any characterised noise on the recordings.

6.2 Methods

6.2.1 Cell Culture

hERG transfected HEK 293 cells were obtained from the University of Wisconsin. These cells are held in cryogenic storage at GSK and also maintained in culture. All cell culture and maintenance was performed by laboratory technicians at GSK. The cells were continuously maintained in (and passaged using) minimum essential medium supplemented with 10% foetal bovine serum, 1% non essential amino acids, 1% sodium pyruvate, 1% penicillin/streptomycin and 0.4 mg/mL geneticin. Trypsin-Ethylenediaminetetraacetic acid and phosphate buffered solution were used to split the cells. For the patch clamp recordings, cells were plated onto glass coverslips in 6 well (35 mm^2) plates (containing 3 mL medium without geneticin) at an appropriate density to allow selection of isolated cells. The plates were stored in an incubator at 37°C with 5% CO_2.

6.2.2 Electrophysiology Solutions

The bath solution was composed of: NaCl (137 mM), KCl (4 mM), $MgCl_2$ (1 mM), HEPES (10 mM), glucose (10 mM), and $CaCl_2$ (1.8 mM). The pH of the solution was adjusted to 7.4 with NaOH. Borosilicate glass micropipettes were pulled using a DMZ-Universal Puller and were fire polished and filled with pipette solution containing: KCl (130 mM), $MgCl_2$ (1 mM), HEPES (10 mM), EGTA (5 mM), and MgATP (5 mM). The solution pH of the solution was adjusted to 7.2 with KOH. All experiments were performed at room temperature ($\approx$22°C). Using this temperature and the composition of the bath and pipette solutions, a reversal potential of -88.5 mV was calculated using the Nernst equation formula in Equation (3.12).

For experiments where different concentrations of dofetilide were used, the required doses were obtained by serial dilution. Dofetilide was first dissolved in Dimethyl sulfoxide (DMSO) before being added to the bath solution to produce the required concentrations.

6.3 Recording Technique

Current recordings were made using an Axopatch 200B amplifier in whole-cell patch clamp mode. Data acquisition was performed using pClamp10 software (Molecular-Devices, 2008). A CV 203BU amplifier headstage and a Digidata 1440A were used. The amplifier headstage is attached to the electrode holder and provides an interface between the amplifier and the cell. The Digidata sends and receives signals from the headstage amplifier. A Sutter micromanipulator was used for positioning of the microelectrode. The current signal was sampled at a rate of 10 kHz. 75 – 85% series resistance compensation was applied and data was either 1 or 5 kHz low pass Bessel filtered by the hardware. Whole-cell capacitance compensation was applied electronically. P/N leak subtraction was applied using the pClamp10 software (Molecular-Devices, 2008). P/8 leak subtraction was applied for the sine wave protocols and P/4 leak subtraction for the voltage-step protocols.

6.4 Voltage Protocols

We performed the voltage protocols described in Chapter 5 and the standard voltage-step protocols shown in Figure 3.1 (of the form described in Bett et al. (2011)). The standard voltage-step protocols were created directly using the pClamp10 software and the non-standard protocols were first created as text files and then converted to `.abf` stimulus files to make corresponding `.pro` protocol files in the pClamp10 software. P/N leak subtraction was performed using the pClamp10 software; we set the holding potential and the N value in the stimulus tab of the protocol options. To make a series of successive recordings using different protocols on the same cell, the "Sequencing Keys" tool was utilised. To do this a `.sks` file was made detailing the order of the sequence of protocols to be performed once triggered to start recording current.

6.5 Experimental Recordings

6.5.1 Recording Multiple Protocols in the Same Cell

In 6 different cells, we performed the full series of standard protocols described in Section 6.4, followed by the full set of novel protocols described in Chapter 5 and then repeated the full sequence of standard protocols again. Standard voltage-step protocols were performed at the start and the end of the recording to allow us to ensure that the cell was not adversely affected by applying the non-standard protocols — it provided a means to check that the cell was still behaving in the same way after the application of these protocols as it was before. This approach also provided us with a set of validation data in the form of data from the standard voltage-step protocols which could be used to evaluate model calibration after constructing models using data obtained from applying the non-standard protocols on the same cell (as we describe in Chapter 7). A range of repeats of subsets of the different voltage protocols were performed on different cells. Example current recordings obtained when performing the full sequence of standard and non-standard protocols described in one cell are shown in Figures 6.2 and 6.3.

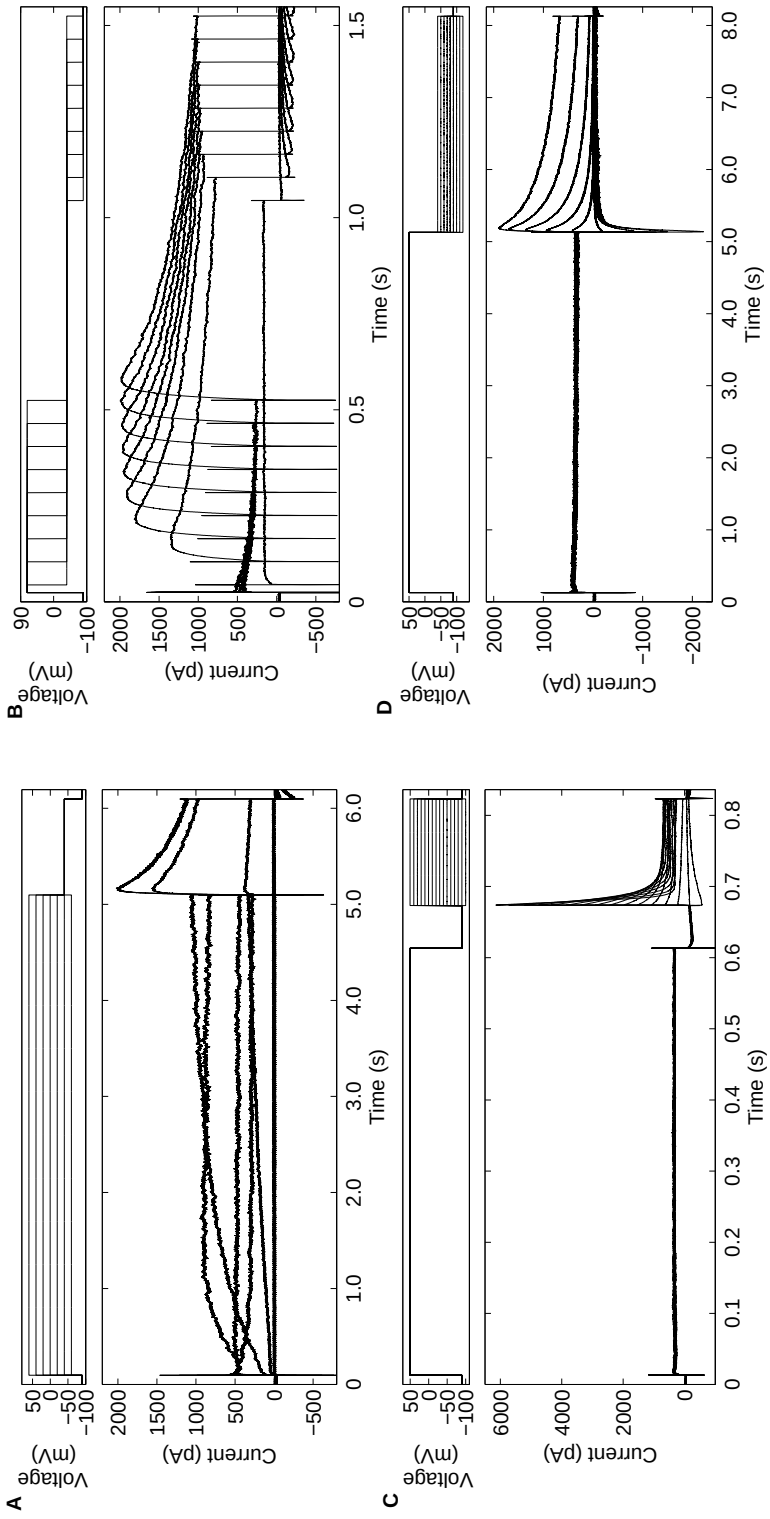


Figure 6.2: Currents recorded when applying the series of standard voltage-step protocols in one cell. The protocols investigate specific aspects of hERG channel gating: steady state activation in A, activation kinetics in B, inactivation kinetics in C and deactivation kinetics in D. The recordings shown have been 1 kHz low pass Bessel hardware filtered and have not been corrected with P/N leak subtraction. Capacitive artefacts are evident in the recordings at points where there is a large step in voltage.

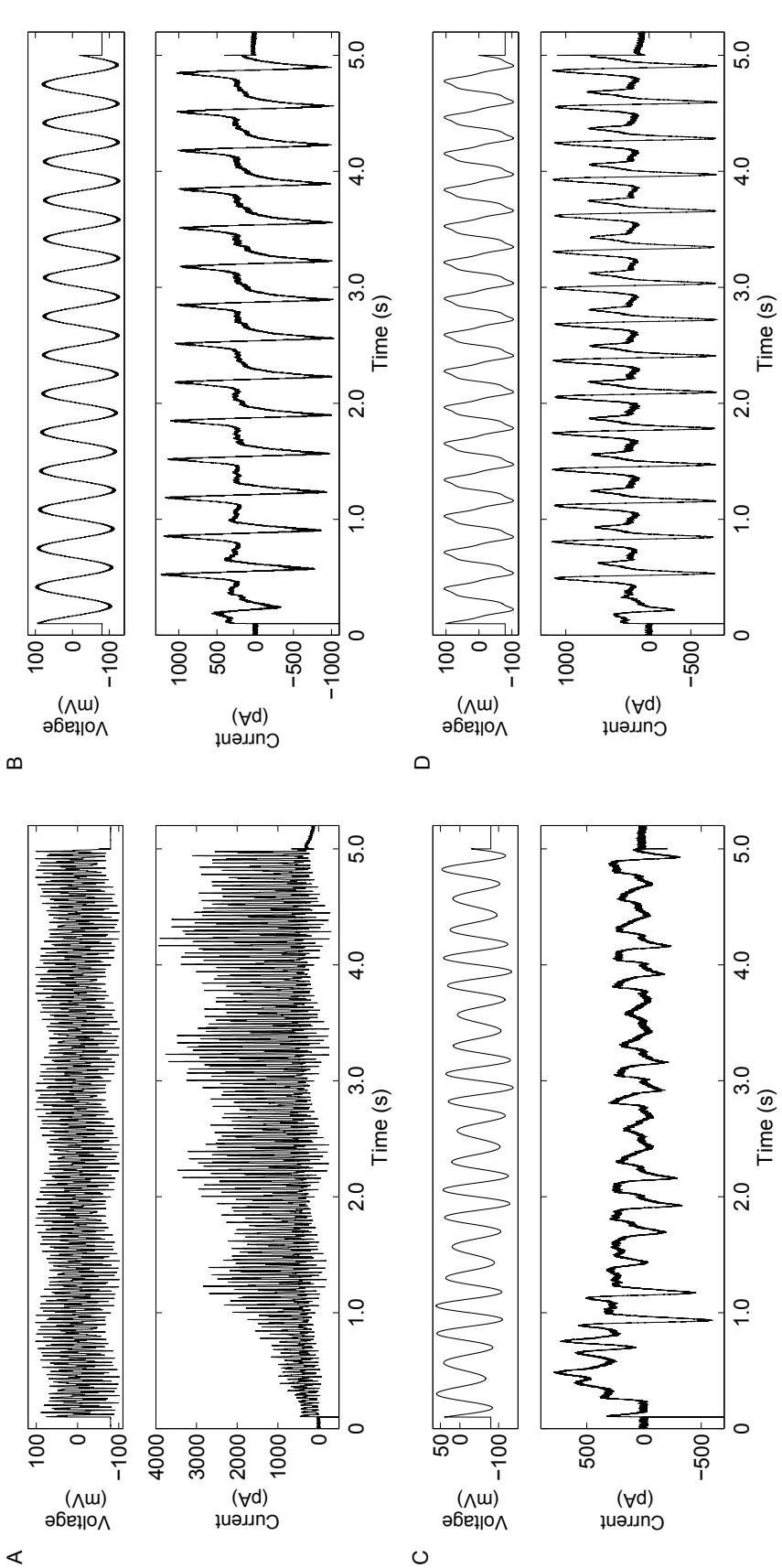


Figure 6.3: Recorded currents from each of the non-standard voltage protocols in one cell. The corresponding protocols applied to produce each current recording are: *sine wave 1-4 protocols* in A-D respectively. The sine wave protocols have been corrected with P/8 leak subtraction.

6.5.2 Preliminary Observations

While repeating patch clamp experiments in different cells, some variability in the recordings was apparent. In many cases just a difference in the magnitude of the current was observed due to the presence of different numbers of channels in each cell, and very similar characteristics of the current recorded were seen. However, in other cases the behaviour observed appeared more distinct. An example of different behaviour observed in two cells in response to the standard voltage-step protocol for investigating steady-state activation kinetics is shown in Figure 6.4.

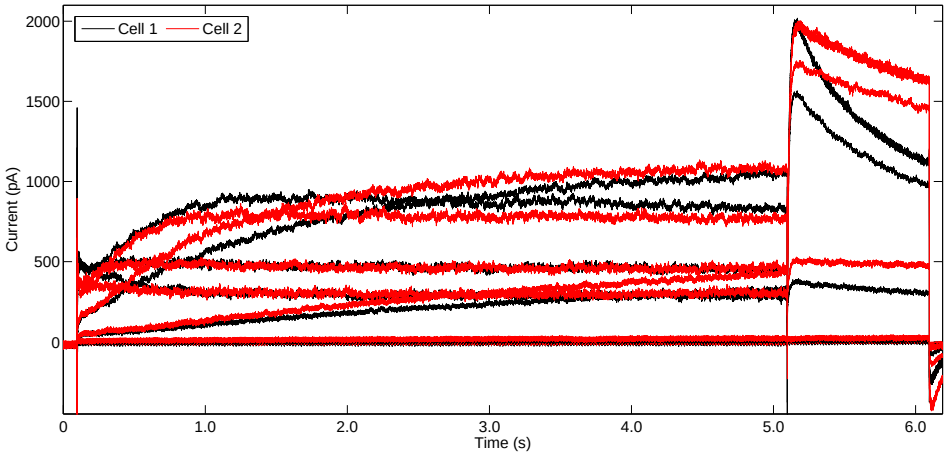


Figure 6.4: Current recorded in two different cells when applying the standard voltage-step protocol for investigating steady-state activation kinetics. Clear differences in the time scale of the recovery from activation in the two cells are apparent.

We also observed some differences in the behaviour in different cells when applying the new protocols. Figure 6.5 shows the currents recorded in two different cells when applying the *sine wave 1 protocol*, which was designed to maximise the difference in current between the Wang et al. (1997) and Mazhari et al. (2001) models. We have overlaid the simulated current traces from these two models (scaling the conductance parameter so as to allow comparison with experimental data by minimising the absolute difference between each trace, using the cell 2 trace as a reference trace but leaving all of the other parameters as in the published

models). It can be seen that the experimental recording in cell 1 behaves in a similar way to the current simulated by the Mazhari et al. (2001) model and the experimental recording in cell 2 behaves in a similar manner to the Wang et al. (1997) model.

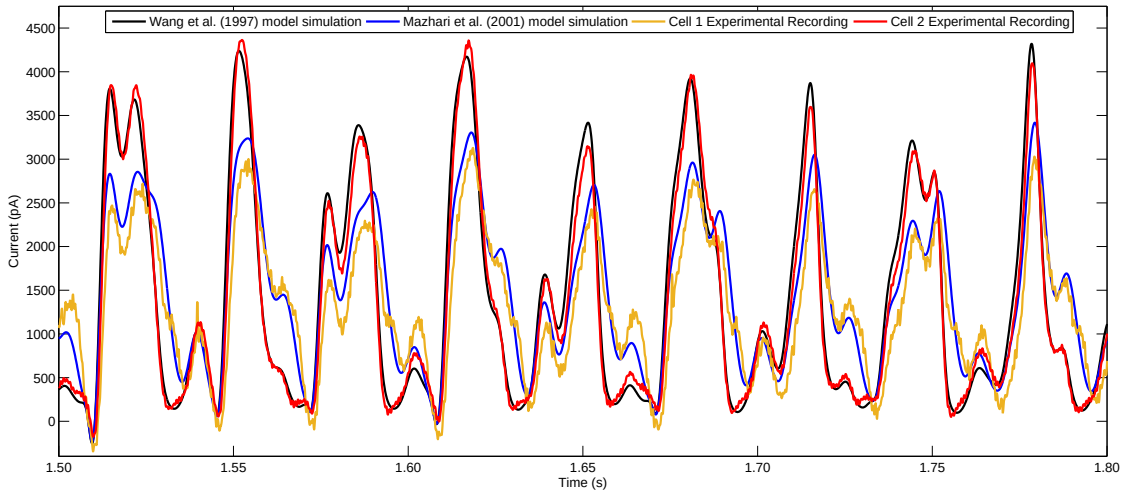


Figure 6.5: Experimental recordings from two different cells using the *sine wave 1 protocol* designed to maximise the difference in current between the Wang et al. (1997) and Mazhari et al. (2001) hERG channel models. Experimental recordings are compared to simulated currents from Wang et al. (1997) and Mazhari et al. (2001) models. The conductance parameter in each model has been scaled to minimise the absolute difference between each trace and the reference trace.

The structural difference between the Wang et al. (1997) and Mazhari et al. (2001) model formulations is the presence of closed-state inactivation from the closed state proximal to the open state. Upon observing that the behaviour of some cells is best described by the Mazhari et al. (2001) model with closed-state inactivation present and in other cells is best described by the Wang et al. (1997) model where this transition is absent, we hypothesised that perhaps closed-state inactivation may occur in some cells but not in others. To investigate this further we simulated the Wang et al. (1997) model with the addition of closed-state inactivation (as formulated in the Mazhari et al. (2001) model) and the Mazhari et al. (2001) model with closed-state inactivation removed. These simulations are shown in Figure 6.6.

Figure 6.6 shows that there is very little difference between the simulated traces of both the

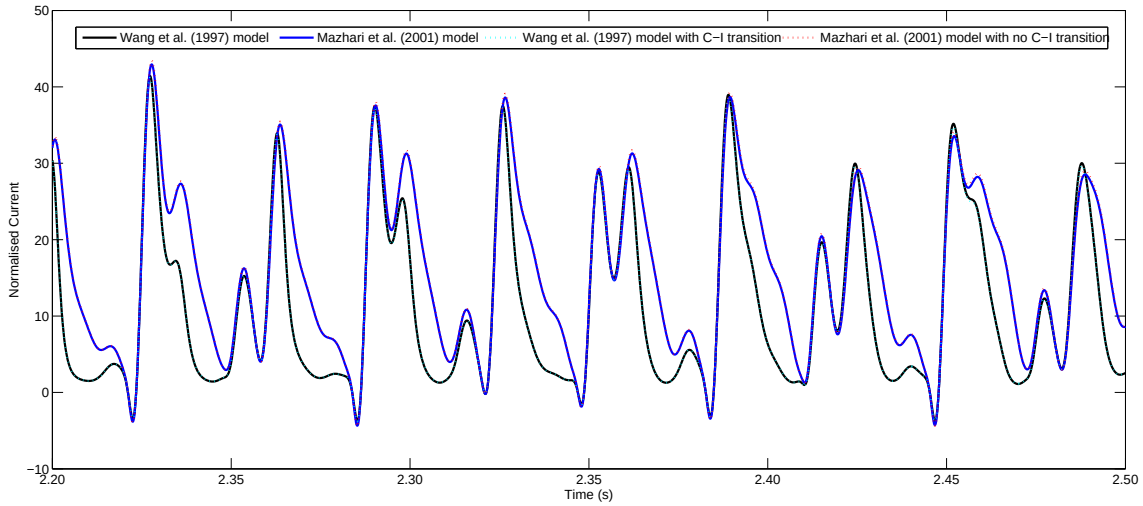


Figure 6.6: Comparison of simulated current traces when applying the *sine wave 1 protocol* when using the Wang et al. (1997) model, the Mazhari et al. (2001) model and modified versions of these two models with closed-state inactivation removed from the Mazhari et al. (2001) model and added to the Wang et al. (1997) model. The conductance parameter is given a value of 1 in each model simulation. We see that the traces from the original Wang et al. (1997) model and the Wang et al. (1997) model including the closed-inactivated transition (rates from the Mazhari et al. (2001) model) are extremely similar, traces are almost indistinguishable. Likewise, there are only slight differences between the simulated traces from the Mazhari et al. (2001) model and the Mazhari et al. (2001) model without the closed-inactivated transition, again traces are almost indistinguishable.

Wang et al. (1997) and Mazhari et al. (2001) models with and without the inclusion of closed-state inactivation in each model. We conclude that the main difference in the behaviour of the Wang et al. (1997) and Mazhari et al. (2001) models is due to the parameterisation of the other transitions in the model and not due to the presence or absence of the closed-state inactivation transition. This observation is consistent with the argument presented by others, including Fink et al. (2008); that the closed-state inactivation transition present in the Mazhari et al. (2001) model is negligible and perhaps unnecessary. It should also be noted that while the Mazhari et al. (2001) model was formulated to describe hERG channel kinetics at physiological temperature, as the experiments described in this chapter were performed at room temperature, the temperature used for calculation of the reversal potential in the model simulations is 22 °C and the model parameters are the same as those in the published physiological temperature model.

6.5.3 Variability in Experimental Recordings

With some evident variability between currents recorded in different cells when the *sine wave 1 protocol* is applied (see Figure 6.5), we performed many repeats of the *sine wave 1 protocol* in different cells to enable us to investigate this observation further. Figure 6.7 shows the recorded currents when applying the *sine wave 1 protocol* in 40 different cells. In Figure 6.7 we scale the traces from each experiment to account for differences in cell conductance to allow direct comparison of experimental traces. We scale the conductance parameter by minimising the absolute differences between each point on a trace and a reference trace in each experimental data set. We also compare the variability evident in the experimental recordings with the range of simulated traces from existing models.

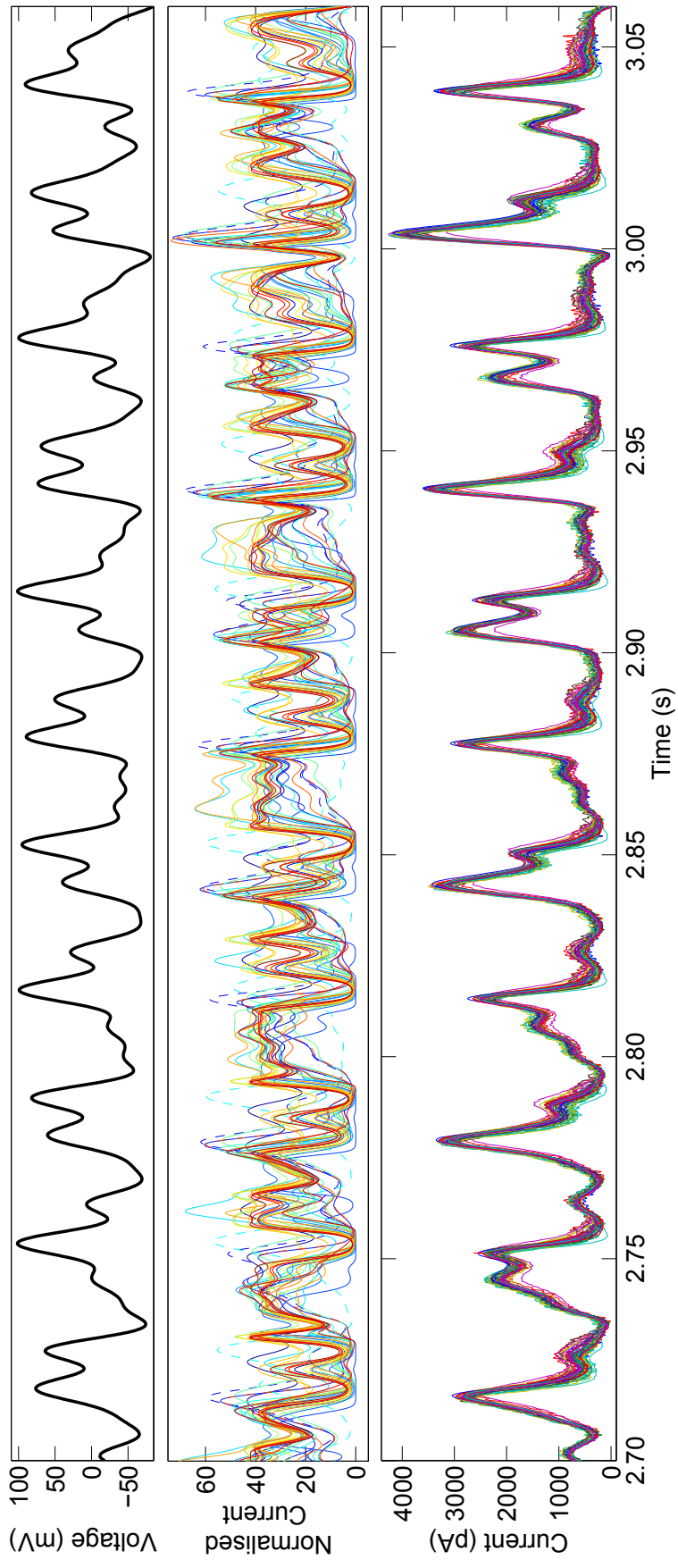


Figure 6.7: Section of the current traces recorded from 40 different cells when applying the *sine wave 1 protocol* in patch clamp experiments. The section of voltage protocol is shown in the top panel, simulations from the range of existing hERG channel models in the middle panel and recordings from 40 cells in the bottom panel. Simulated and experimental traces have been linearly scaled to minimise the absolute difference between each trace and a reference trace. Some subtle differences in the behaviour in different cells can be observed, but these are not as striking as the differences observed between simulated currents from the different models.

Some subtle differences in the behaviour of the current recorded in the different cells are evident in Figure 6.7. However, the variability observed is not as striking as that observed when comparing simulated currents from the range of published mathematical hERG channel models. In Section 6.6 we quantify the differences between experimental and simulated traces to assess the ability of different models to replicate our experimental observations. In Chapter 7 we use the experimental data presented here to further investigate and characterise the variability in these experimental recordings quantitatively.

We also observed some differences in behaviour when repeating each of the other non-standard voltage protocols in different cells. In Figure 6.8 we show the currents recorded in different cells when performing each of the *sine wave 2-4 protocols*.

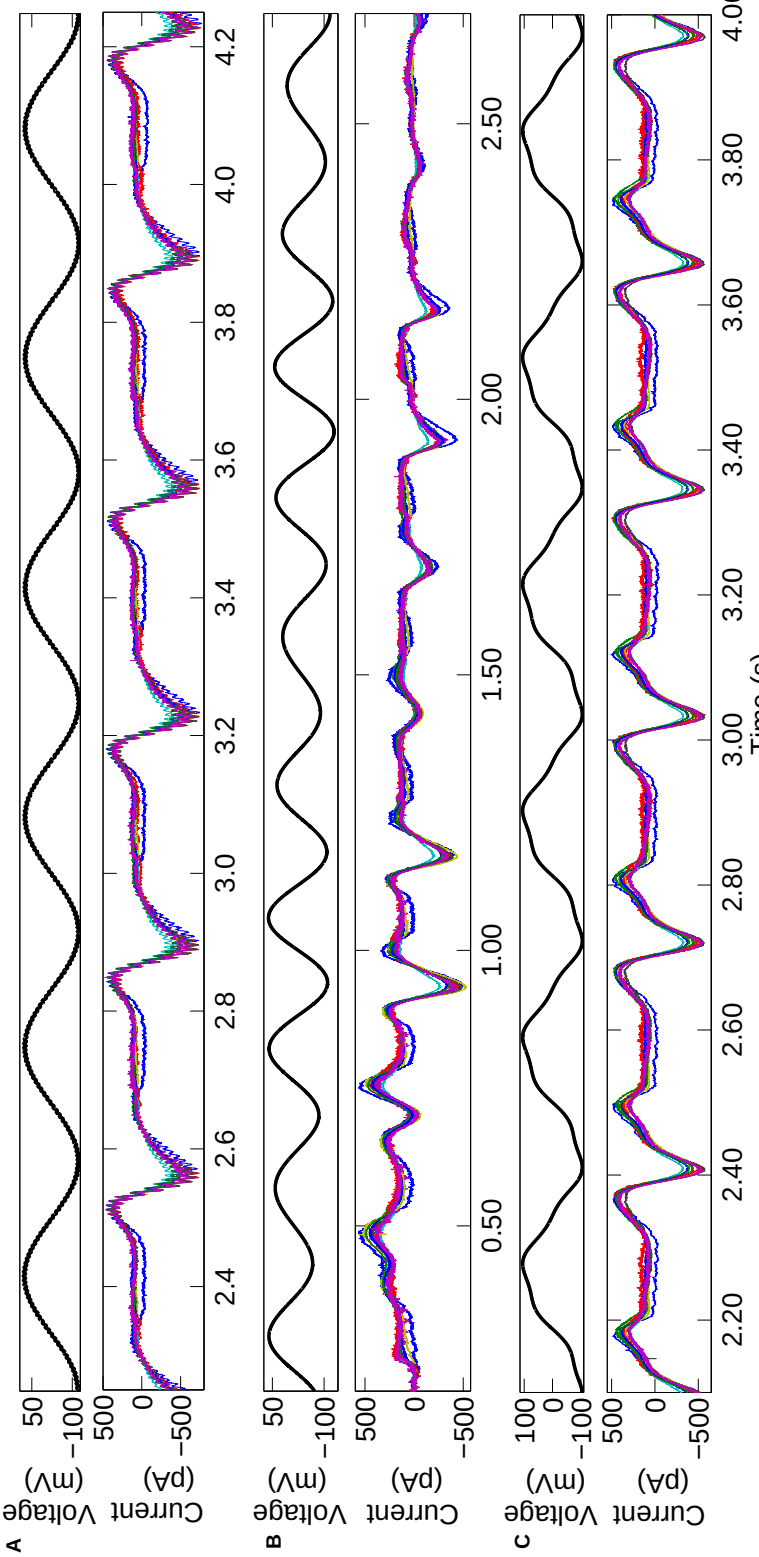


Figure 6.8: Section of the current traces recorded in 10 different cells when applying the *sine wave 2 protocol* in A and the *sine wave 3 protocol* in B and the *sine wave 4 protocol* in C in patch clamp experiments. Currents have been normalised to minimise the absolute difference between the each recording and a reference recording.

6.5.4 Additional Checks to Ensure Currents Recorded were hERG Channel Currents

Due to the non-standard nature of the sine wave protocols, as compared to the typical voltage-step protocols performed in patch clamp experiments, we performed a number of additional checks to be assured that the currents we were recording were those of hERG channel currents and that the protocols weren't inadvertently activating endogenous channels within the untransfected HEK 293 cells. We first checked that the current we recorded was blocked in the presence of a moderate dose of the hERG channel blocker dofetilide. Figure 6.9 shows the hERG channel current recorded when applying the *sine wave 1 protocol* in the presence of a range of different dofetilide concentrations. The maximum concentration tested is 0.3 μM which is of the order of 10 times greater than the estimated hERG channel IC_{50} value of dofetilide. A dose of this concentration would be expected to fully block hERG channel current.

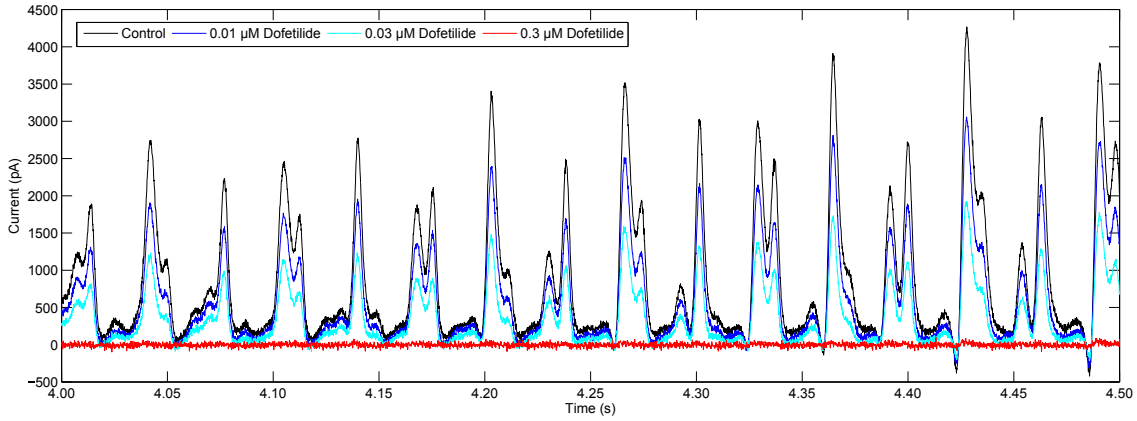


Figure 6.9: Current recorded when applying the *sine wave 1 protocol* in the presence of different concentrations of dofetilide.

In Figure 6.9 we see a concentration-dependent decrease in current amplitude with increasing dofetilide concentration. The recorded current is almost zero in the presence of 0.3 μM of dofetilide. These observations are expected and provide evidence that the current we are recording is predominantly that flowing through the hERG channel.

As an additional check that any contribution to the recorded current from endogenous channels expressed in HEK 293 cells is minimal, a 3.2 second version of the *sine wave 1 protocol* was performed in an untransfected HEK 293 cell; the current recorded is shown in Figure 6.10.

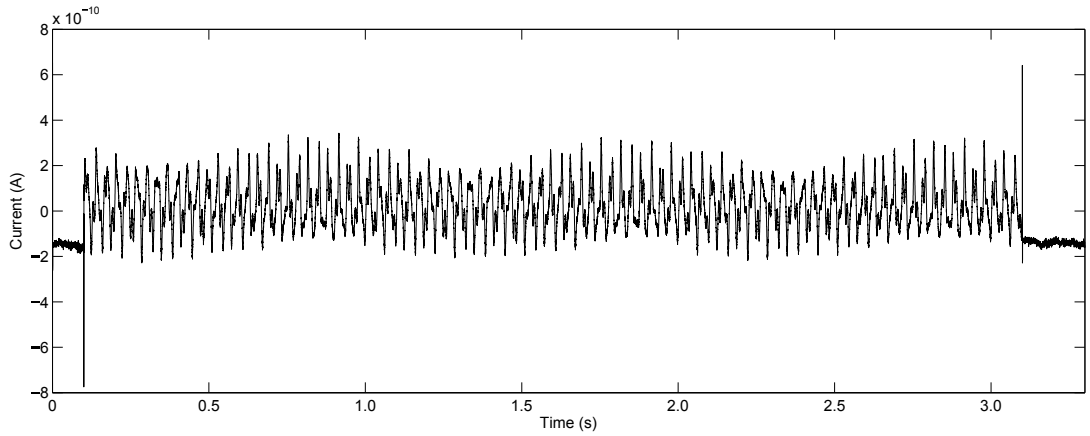


Figure 6.10: Current recorded in untransfected HEK 293 cell when applying a 3.2 second version of the *sine wave 1 protocol* to assess the current contribution from endogenous channels in HEK 293 cells. Recording provided by Teun de Boer, University of Utrecht. No P/N leak subtraction has been applied in this recording. A current of magnitude 200 pA is recorded during this recording, but as there is also evidence of a leak current, it may be expected that some of this current would be further reduced with P/N leak subtraction applied.

The level of current measured upon application of the *sine wave 1 protocol* in an untransfected HEK 293 cell can be seen to be relatively low. We also may expect further reduction of this current if P/N leak subtraction had been applied. Consequently, the expected contribution to the current recorded due to endogenous cells is expected to be reasonably small. As this trace has not been corrected with P/N leak subtraction, it provides an example to see the effects of leak current at the start and end of the recording (offset from zero) and apparent capacitive artefacts at the start and end of the sine wave portions of the protocol are also evident.

In Figure 6.11 we illustrate the advantage of performing P/N leak subtraction with these

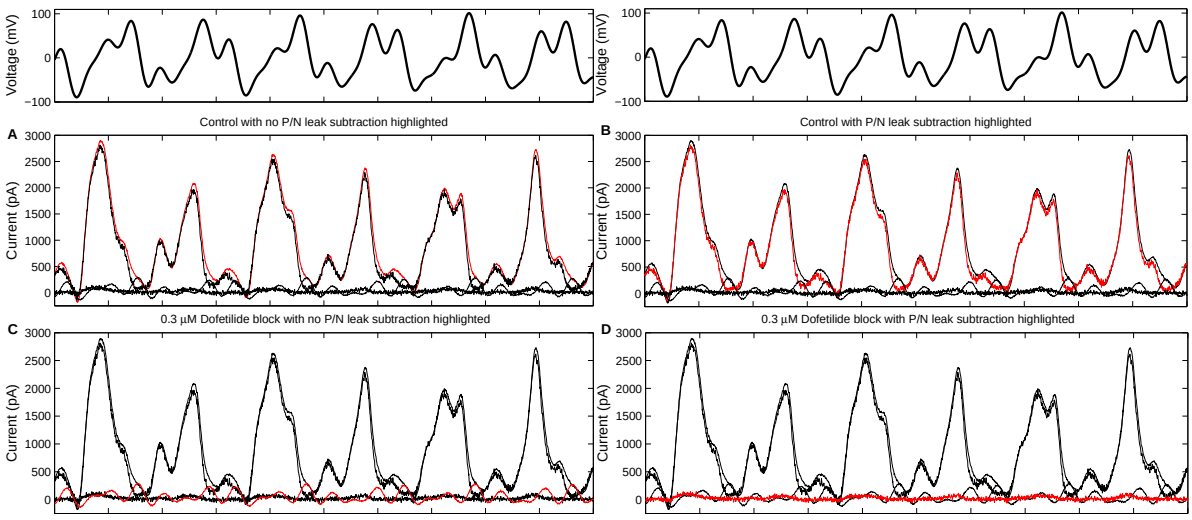


Figure 6.11: Raw recordings, P/N subtraction, and 0.3 μM dofetilide controls. Top row: input voltage trace from the *sine wave 1 protocol* (same on both sides for comparison with traces below). In A–D we show the current in response to this voltage protocol in four situations — each panel shows the same four traces, with a different one highlighted in red. In A we highlight the raw recording of whole current; in B the whole current recording after P/8 leak subtraction, in C the raw recording after the addition of dofetilide; and in D the dofetilide recording after P/8 leak subtraction. In C we still see a persistent level of current even at a concentration of dofetilide which would be expected to fully block hERG. In D however, we see the highlighted current is almost zero when applying P/8 leak subtraction.

non-standard voltage protocols.

In Figure 6.11 we show the result of the P/N correction and the dofetilide control. When dofetilide is added, a small current remains in the non-P/N leak subtracted trace, which must be a non-hERG current as 0.3 μM dofetilide would block hERG. With both the addition of dofetilide and P/N leak subtraction, the remaining current is almost zero. These comparisons provide us with further confidence that we are recording predominantly hERG current.

When comparing our input voltage protocol with that recorded in the results file output using the pClamp10 software, we observed a time shift of 0.1 ms in the applied voltage protocol relative to that recorded during the patch clamp experiments. This shift is consistent in all recordings and appears to be specific to the setup used. It is possibly attributed to a delay in digital-analog conversion in the amplifier mechanism. To account for this shift in the ex-

perimentally applied voltage, we apply a time shift of 0.1 ms to the input protocol for the comparisons in Section 6.6 and when using the experimental data for fitting in Chapter 7.

6.6 Quantitatively Comparing Existing hERG Channel Models

We next quantitatively assessed the ability of existing hERG channel models to replicate our experimental recordings. We present this comparison for the two individual sine wave protocols which we identified in Chapter 5 as allowing the most identifiable parameters to be recovered: the *sine wave 1 protocol* and *sine wave 3 protocol*. We took an average of the 40 experimental traces shown in Figure 6.7, scaled the conductance parameter of each of the existing hERG channel model to minimise the absolute value between the averaged experimental trace and each model simulation, and reported mean absolute difference in current between each simulated trace and the averaged experimental trace. We then repeated this for the 10 experimental traces obtained for the *sine wave 3 protocol*. The results are presented in Table 6.1, where we rank the existing models based on the error between the simulated trace and experimental trace for both protocols. It should be noted however that this ranking only describes how well these models replicate data in response to our protocols in a HEK hERG 1a expression system at room temperature, not I_{Kr} at physiological temperatures in a myocyte. However, such an approach may be useful as a first step of exploring the suitability of existing models.

In Figure 6.12 we show a selection of simulated traces as compared to the averaged experimental trace for both the *sine wave 1* and *sine wave 3 protocols*.

Table 6.1 Table detailing mean absolute differences (in pA) between each model simulation and the average of the 40 experimental traces of *sine wave 1 protocol* and 10 experimental traces of *sine wave 3 protocol* The 'Δ' measure is the mean (along the trace) absolute difference between the mean experimental trace and the model prediction (in units of pA), after model conductance is scaled to minimise this quantity. The Δ values were ranked for each sine wave, and then sum of ranks was used as an overall rank to order the table.

Model	Δ Sine wave 1 (pA)	Δ Sine wave 3 (pA)	Rank
Mazhari et al. (2001)	150.5	48.1	=1
Di Veroli et al. (2013) (Room Temp.)	127.8	89.0	=1
Fink et al. (2008)	246.3	101.9	3
Lu et al. (2001)	490.2	109.2	4
Wang et al. (1997)	244.6	122.3	5
Clancy and Rudy (2001)	465.9	124.7	6
Aslanidi et al. (2009)	548.7	120.0	=7
Di Veroli et al. (2013) (37°C)	553.3	117.6	=7
Oehmen et al. (2002)	535.6	125.4	9
Courtemanche et al. (1998)	587.7	105.7	10
Ramirez et al. (2000)	548.9	124.9	=11
O'Hara et al. (2011)	580.6	118.4	=11
Inada et al. (2009)	542.0	131.3	13
Priebe and Beuckelmann (1998)	604.9	107.2	14
Seemann et al. (2003)	551.6	125.5	15
Matsuoka et al. (2003)	545.7	138.0	=16
Hund and Rudy (2004)	625.5	119.3	=16
Ten Tusscher et al. (2004)	522.1	145.5	18
Winslow et al. (1999)	572.9	133.6	19
Zeng et al. (1995)	589.9	132.3	20
Kurata et al. (2002)	586.5	139.3	=21
Liu et al. (1996)	593.3	137.1	=21
Nygren et al. (1998)	656.0	130.2	23
Zhang et al. (2000)	560.0	156.7	24
Lindblad et al. (1996)	647.2	133.3	=25
Noble et al. (1998)	610.3	137.9	=25
Shannon et al. (2004)	636.2	143.3	27
Grandi et al. (2010)	668.2	142.1	28
Severi et al. (2012)	660.4	154.6	29
Fox et al. (2002)	712.8	157.5	30

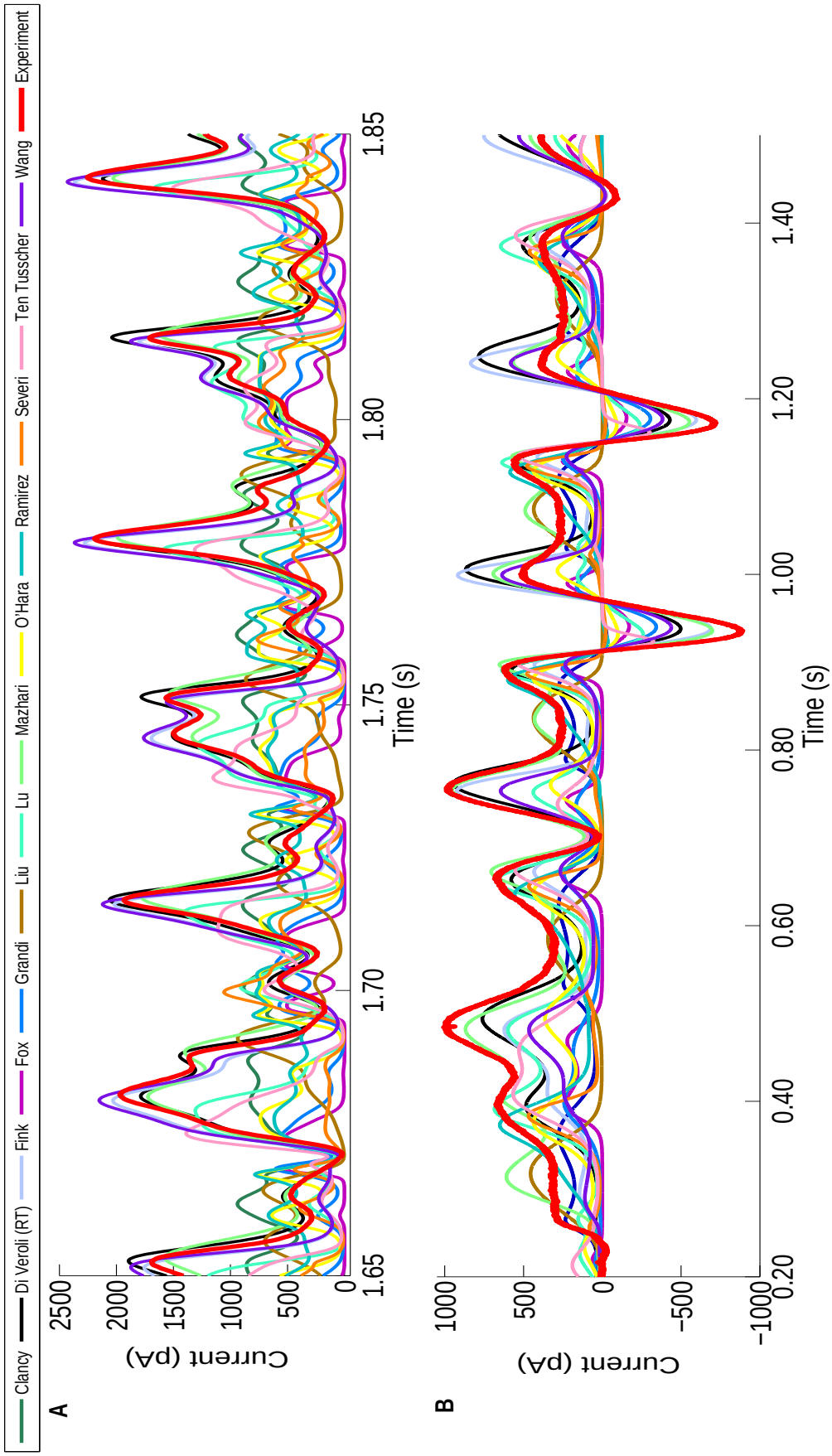


Figure 6.12: We show a selection of existing hERG channel model simulated current traces as compared to our averaged experimental trace from 40 cells for *sine wave 1 protocol* in A and 10 traces for *sine wave 3 protocol* in B. We see that a diverse range of behaviour is exhibited and that even the models ranked best in Table 6.1 are not able to fully describe the experimental observations from both sine wave protocols.

In Table 6.1 we see that the top 6 ranked models are all Markov models, illustrating that these are able to replicate our experimental data in response to the novel protocols better than the existing Hodgkin-Huxley style models. Also of note is that the range of errors for the *sine wave 1 protocol* is much greater than for the *sine wave 3 protocol*. This illustrates that the *sine wave 1 protocol* is able to discriminate better between models, as was the intention behind its design. In Figure 6.12 we see that while some models are able to replicate the experimental data better than others, none of the models are able to fully describe our experimental recordings. These observations demonstrate the necessity of careful model construction and validation of cell-specific mathematical models describing hERG channel kinetics, which we will describe in Chapter 7.

6.7 Discussion

By performing the experiments described in this chapter we have produced a large data set consisting of the *sine wave 1 protocol* recorded in 40 different cells and the full set of standard and non-standard protocols performed in 6 different cells. This provides a good basis for exploring the use of the protocols described in Chapter 5 for determining cell-specific models of hERG channel kinetics. By performing sequences of standard and non-standard protocols as described we have also generated additional data which will allow us to ensure that the behaviour of the cells is not altered by the use of the non-standard voltage protocols, and will enable us to validate our models.

As we were performing non-typical voltage protocols we had to ensure that currents measured were those of the hERG channel current in response to the protocols applied and not affected substantially by any additional factors. We have described a series of considerations and checks which were performed to attempt to verify this. A number of observations were made while performing these experiments which may help to improve the experimental design of the approach for future work.

We conclude this chapter by comparing quantitatively the model simulations of existing hERG channel models with our experimental recordings. We find that existing Markov models are consistently able to replicate better our experimental recordings. However, we see that none of the existing models are satisfactorily able to replicate both of the experimental traces presented. This motivates the research presented in Chapter 7 where we construct and validate cell-specific models of hERG channel kinetics — using a range of Markov model structures as model candidates.

6.8 Summary

In this chapter we have focused in detail on the methodology used to perform the voltage protocols described in Chapter 5 using the whole-cell patch clamp technique. In addition to describing the experimental setup we have also outlined the underlying electrical theory of the technique and described additional considerations and checks which were performed due to the non-standard nature of the sine wave voltage protocols to try to ensure good quality data recordings. We have described some observations made with regards to variability in experimental recordings which we wish to investigate further. We have also ranked existing hERG channel models on their ability to replicate our observed experimental data. We found that Markov model structures are able to better replicate the behaviour exhibited in our experiments but that none of the models were able to describe satisfactorily all experimental observations. In Chapter 5, we assessed the parameter identifiability of different model structures when fitted using data from different combinations of the novel voltage protocols. In Chapter 7 we use the experimental recordings described in this chapter to construct cell-specific models of hERG channel kinetics; with the aim of determining robust models able to accurately replicate a range of experimental recordings. Using the currents recorded in different cells we are also able to compare distributions of fitted parameters in the identifiable model structures to explore quantitatively variability in the experimental recordings described in this chapter.

7

Constructing Cell-Specific Models of hERG Channel Kinetics

In this chapter we use the experimental data described in Chapter 6 to construct cell-specific models of hERG channel kinetics. We parameterise a range of candidate model structures and consider the most appropriate model for describing our observed experimental data. We perform the model construction process using data from different combinations of the experimental protocols described in Chapter 5. We employ a Bayesian inference approach for model parameterisation to determine distributions of estimates for model parameters. This approach allows us to quantify the uncertainty in model parameters and also provides a means for investigating whether variability exists between cells. We outline the challenges we encountered and considerations taken when simultaneously selecting and parameterising models.

7.1 Introduction

In this chapter we use the experimental data described in Chapter 6 to construct cell-specific models describing hERG channel kinetics. In Chapter 5 we assessed the identifiability of parameters of existing hERG channel models when using simulated data from different voltage protocols for parameterisation. We demonstrated that many parameters of existing model structures are identifiable when using our sine wave voltage protocols, based on the criteria we defined using the Mean Absolute Percentage Error (MAPE). We used this analysis to conclude that our new protocols are theoretically appropriate for determining parameters of

many existing hERG channel models. However, the use of simulated data in this way represents the best possible scenario as there are no confounding factors to complicate the inverse parameter fitting problem. The synthetic data comes from a defined model structure and the true underlying parameter values are also known. In particular, this latter characteristic means that we know when an optimum solution to the parameter optimisation problem has been obtained and that the MCMC algorithm is behaving as anticipated.

In Chapter 5 we explored parameter distributions obtained when starting the MCMC algorithm at the true simulated parameter values. In order to use the Bayesian inference approach we described in Chapter 5, to construct models describing the experimental data we presented in Chapter 6, we must first explore automated techniques to find starting points for the MCMC chains, which are at, or very close to, globally optimal parameter set for each experimental data set. We begin this chapter with a discussion of the necessity of locating a globally optimal solution to the parameterisation problem. We then describe the methods we use to do this in Section 7.4. Using these additional methods, we can then go on to consider the extent to which we can recover model parameters from our experimental data to determine cell-specific models describing our experimental observations. We provide a summary of the experimental data sets that we have collected from HEK-cell line cells in Section 7.2 and detail how we use the data in this chapter.

7.2 Summary of Experimental Data Sets

As described in Chapter 6, we performed the *sine wave 1 protocol* on 40 different cells and the full set of standard voltage-step and novel sine wave protocols on 6 cells. Where we have recorded multiple experiments on the same cell, we can use subsets of the data for constructing a mathematical model, and we refer to these data as *training data* and then use the remaining data recorded on the same cell for validating the model, which we refer to as *validation data*. Figure 7.1 provides an illustration of the use of training and validation data sets recorded in the same cell, and serves as an illustration of how we present the results in

this chapter. We show both the model fit of a three state model (with one open, one closed and one inactivated state as shown in Figure 7.2) constructed using training data from the *sine wave 2 protocol* and the subsequent model predictions for the validation data sets. We describe the methods used to construct the model in Section 7.4.

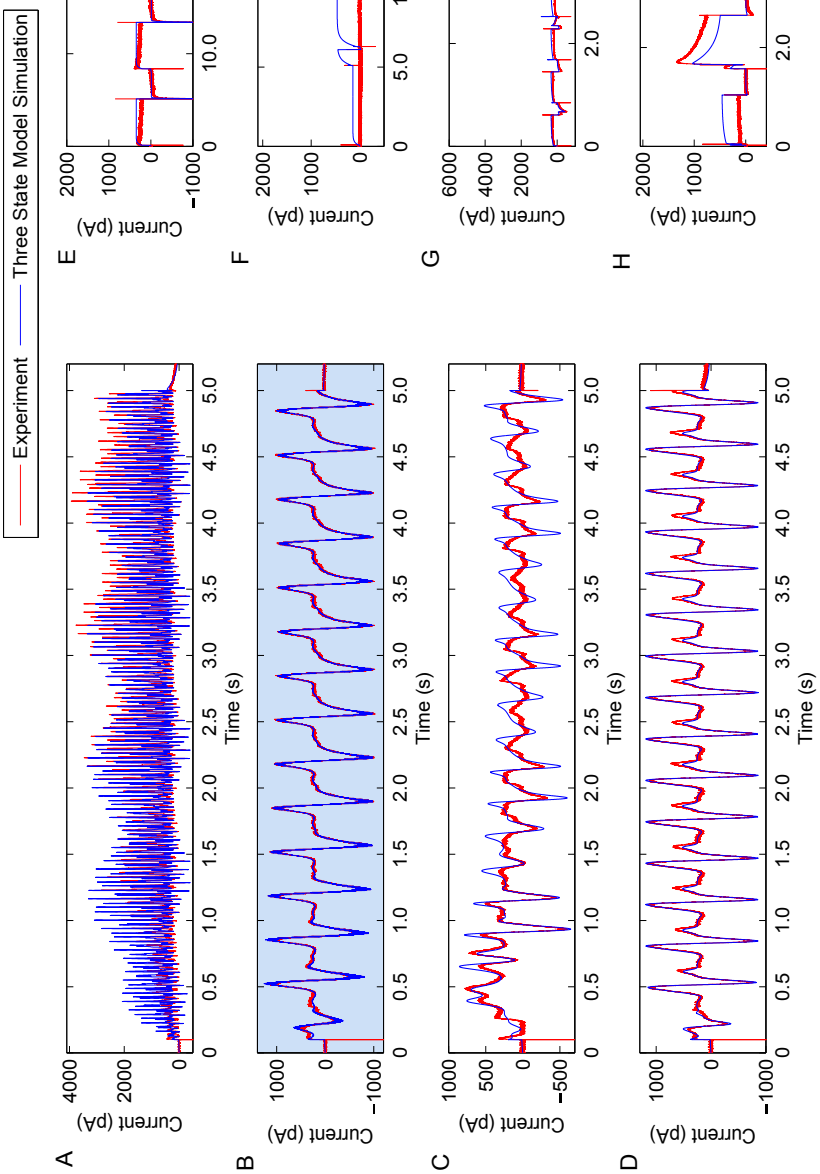


Figure 7.1: In this example, a three state model structure, consisting of one open state, one closed state and one inactivated state (as illustrated in Figure 7.2) was fitted to training data from the *sine wave 2 protocol*. The model simulation as compared to the training data is shown in the blue highlighted panel B. We validate the model using the experimental recordings of the other protocols recorded on the same cell by comparing experimental recordings with the model predictions. We show predictions of the *sine wave 1 protocol* in A, the *sine wave 3 protocol* in C, the *sine wave 4 protocol* in D, the deactivation protocol in E, the steady-state activation protocol in F, the inactivation protocol in G, and the activation kinetics protocol in H. Throughout this chapter the training data takes the form of different combinations of the sine wave protocols shown in A-D, while all of the remaining data not used in model construction is used as validation data.

In this chapter we consider the use of different combinations of the sine wave protocols for constructing models of hERG channel kinetics and validate these models using the data recorded in response to the standard voltage-step protocols from the same cell. As the validation set used in each case is the same, regardless of the choice of training data, this allows direct quantitative comparison of goodness of fit both when using models with different structures constructed using the same set of training data, and when using different sets of training data recorded from the same cell to construct the model. We do also however show a comparison when we construct a model using the traditional data and validate it based on its ability to predict the behaviour of the sine wave protocols.

We will demonstrate that using experimental data presents a much more difficult challenge than when applying the methods to simulated data and we reflect upon and suggest several possible reasons for these difficulties. As we will show, we have been able to determine parameter distributions for all model parameters for 24 individual cells when fitting to the five state model structure (shown in Figure 7.2) which is akin to that of the Wang et al. (1997) model. 18 of these model parameterisations satisfy the criteria we outline for being an optimal fit to the parameterisation problem (which we describe in Section 7.4). We are therefore able, for the first time, to quantify inter-cellular variability in hERG channel models for individual cell-line cells.

We go on to use our experimental data sets to attempt to select the most appropriate model structure to describe the experimental recordings from 3 of the cells for which we obtained both recordings of sine wave protocols and standard voltage-step protocols on the same cell. The candidate model structures we consider include existing Markov model structures and simpler sub-models of these models, as shown in Figure 7.2[1].

We explore further the validity of these individual models when used to predict cellular

[1]Note here that we do not include the Mazhari et al. (2001) model structure for model selection as we found that the behaviour of this model could be adequately replicated with a reparameterisation of the Wang et al. (1997) model structure when simulating our new voltage protocols (Figure 6.6).

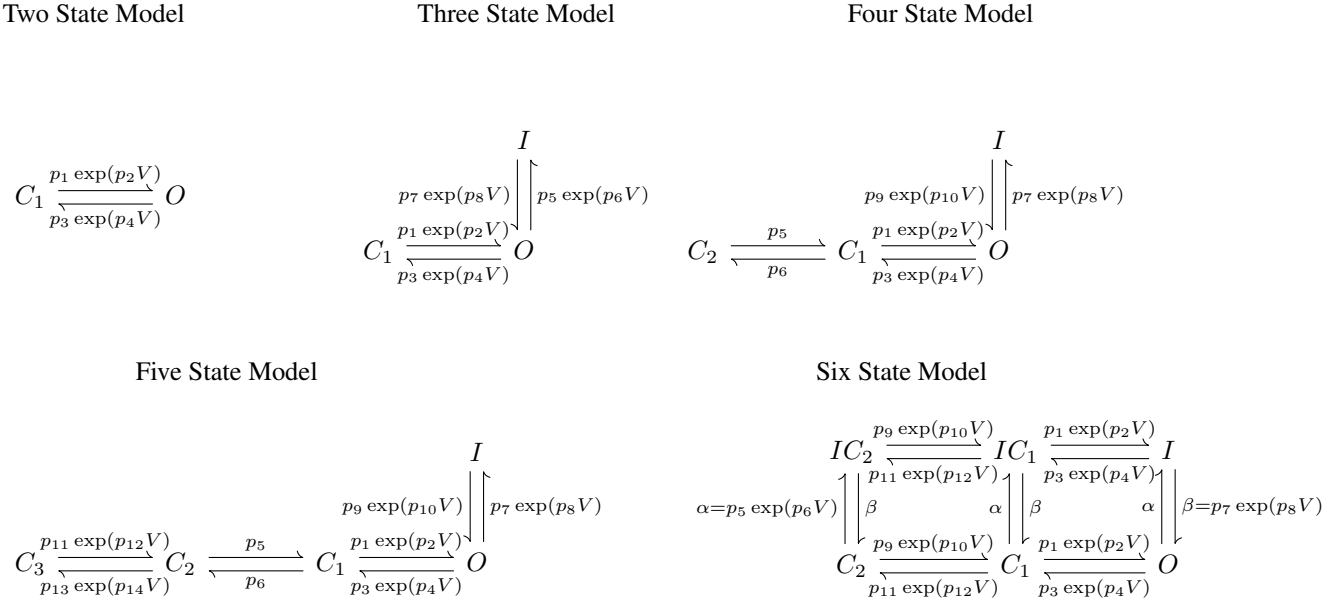


Figure 7.2: Candidate Markov model structures in the model selection process. Closed states are represented by C, inactivated states by I and open states by O. Model names and the p_i labels indicate the model structures and ordering of the parameters referred to throughout this chapter.

behaviour in response to alternative experimental protocols that were not used in the fitting procedure. We conclude with a discussion of the lessons learned, the difficulties arising and the next steps required to resolve the outstanding issues.

The data set that we have collected is extensive by comparison with other studies, and even the subset of the data we present here is large. We also must highlight that the criteria we specify for being satisfied with the parameterisation of a model is considerably more stringent than that normally considered; however, we believe this to be a necessity to ensure that models are accurate and maximally predictive. We now describe the considerations we deem necessary during mathematical model construction as well as outlining the ideal properties of a model.

7.3 Considerations when Performing Mathematical Model Construction

We identify possible explanations for a mismatch between experimental results and model simulations as being; a non-optimal model parameterisation; an inadequate model structure; insufficient information contained within experimental data for determining model parameters; and low quality experimental data. We now elaborate on each of these issues in turn.

7.3.1 Non-optimal Model Parameterisation

During model construction, we wish to select the most appropriate model structure and parameterise this structure to replicate best our observed experimental data. We want to select a model structure with the lowest complexity, for which all model parameters are identifiable, that provides a good fit to the training data, and is highly predictive of the validation data. As we are simultaneously parameterising and selecting kinetic model structures, it is important to identify an optimal fit to experimental data when assessing the appropriateness of a particular model structure. If a model parameterisation is not optimal then we cannot differentiate between an inappropriate model structure and a sub-optimal parameterisation of a suitable model structure. Consequently we have focused much effort on trying to identify the best possible fit of a model to experimental data.

To illustrate the importance of finding the best possible fit of a model to a set of data we can consider the simplified situation represented in Figure 7.3.

The simple example in Figure 7.3 demonstrates the possibility of identifying what may appear to be a good fit to experimental data but which is in fact just the solution of a local minimum of the negative log-likelihood surface. It may still be possible to obtain an MCMC distribution about such a local minimum as MCMC acts primarily to explore the local region of parameter space, and so it is possible that an MCMC chain may become trapped in a local minimum of the negative log-likelihood surface. A sub-optimal model parameterisation may

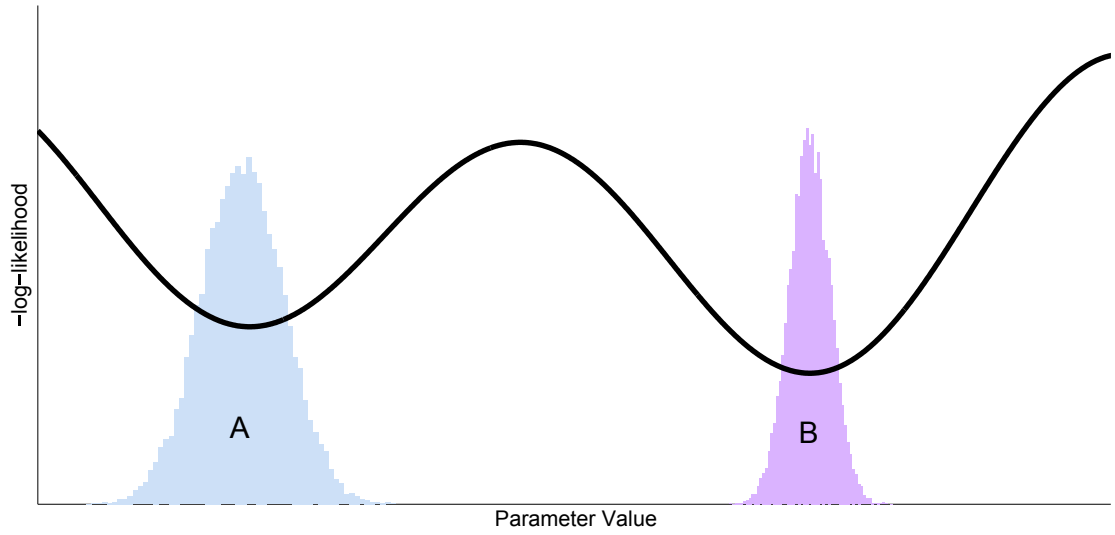


Figure 7.3: Diagram illustrating the possibility of finding multiple solutions to a model parameterisation problem. Here we consider the negative log-likelihood in a model with just one parameter. We see the possible existence of local and global minima of the negative log-likelihood curve across the parameter space. A sub-optimal fit to experimental data could be identified (as in A) and an MCMC distribution obtained about this parameter value, when it could be the case that there is a more optimal point (B) in parameter space which has not been identified by the search algorithm.

appear to allow a good fit of a model to experimental data, but the resulting model may not be an accurate representation of the system being modelled, as it is not parameterised with the parameter set with the global maximum likelihood (or minimum negative log-likelihood). We have to try to ensure that we have located the global optimal solution to the parameterisation problem (as in B in Figure 7.3) before we can start to make conclusions about the appropriateness of any given model structure.

7.3.2 Inadequate Model Structure

A model structure may not be appropriate for representing the underlying biological system, and so would not be able to demonstrate sufficient predictive ability even with an optimal parameterisation. We assess the suitability of a model structure in two ways; first in the traditional sense of identifying the model which yields the best fit to experimental data, and additionally by evaluating models based on their ability to predict validation data recorded in the same cell, as we describe in more detail in Section 7.4.1. This validation step is not

commonly undertaken in cardiac model development.

7.3.3 Suitability of Experimental Data

The results presented in Chapter 5 give us some confidence that the information content of our new protocols is sufficient to determine model parameters, as we showed that many parameters of existing hERG channel models are identifiable when using our new protocols with synthetic data. This is an aspect which we can also assess from the resulting distributions following Bayesian inference. A further consideration is whether (or not) the experimental data are of good quality and whether the data contain sufficient information to determine model parameters. We have tried to ensure that experimental recordings are dominated solely by the hERG channel current using the approaches described in Chapter 6. In addition, any data identified as having a large amount of leak current or unusual artefacts were discarded during recording.

However, despite the additional considerations outlined, it may still be the case that a model identified as being best from a selection of candidate models is not able adequately to describe the kinetics of the channel. We reason how we can identify when this may be the case; if when parameterising a model using data from different sets of training data performed on the same cell, if the parameter distributions obtained are distinct, with no overlap, this is an indication that the model is not able to describe robustly the behaviour of the cell. While the models constructed using the individual training data sets may be adequate representations of the behaviour of the channel when performing each particular experiment, neither model may be an accurate model of the channel. If either model was a sufficient model of the channel, then it would also be identified as an optimal parameter set for describing the behaviour of the channel when any of the experiments were performed. Finally, if a model is able to adequately describe the experimental data with which it was constructed, but is not able to satisfactorily predict the behaviour of validation protocols recorded on the same cell, it also cannot adequately describe the behaviour of the channel. Having outlined the motivation behind the methodology proposed for use in model construction, in Section 7.4 we describe

the technical details of the methods and in Appendix C.2 we provide some further details of the methods used.

7.4 Methods

We first provide an overview of the parameter fitting process we undertake for any given model structure in Figure 7.4.

For a given model structure, we begin our parameter optimisation using multiple initial random parameter sets to try to ensure that we explore different parts of parameter space during parameter fitting. This approach helps to identify situations where an optimisation algorithm simply finds, and gets trapped in, a local minimum of the negative log-likelihood surface, (see Figure 7.3). We explored the use of different optimisation algorithms for model parameterisation, including simplex search methods, simulated annealing and evolution strategy methods (as we describe in more detail in Section 7.4.2). We found that each algorithm was often able to identify the same optimal region of parameter space with the smaller model structures. However, we found that the methods were less consistent when considering larger model structures (with more parameters), and we had to evaluate the use of each method for different combinations of models and data. We also found that algorithm options were critical to performance in some cases, and careful consideration should be given to this. By employing different optimisation algorithms we hoped to eliminate further the potential for identifying non-optimal model parameterisations. We note that we refer to the search algorithms locating an optimal region of parameter space rather than a specific single parameter set. We have the option of using our computational resources either to either optimise fewer parameter guesses more thoroughly or a larger set of parameter guesses from different points in parameter space less thoroughly. As we apply MCMC following an initial global search, which adds an additional layer of optimisation, we choose to explore more initial points in parameter space less thoroughly and choose our search algorithm settings accordingly. We are therefore satisfied with identifying multiple starting points in the same region of param-

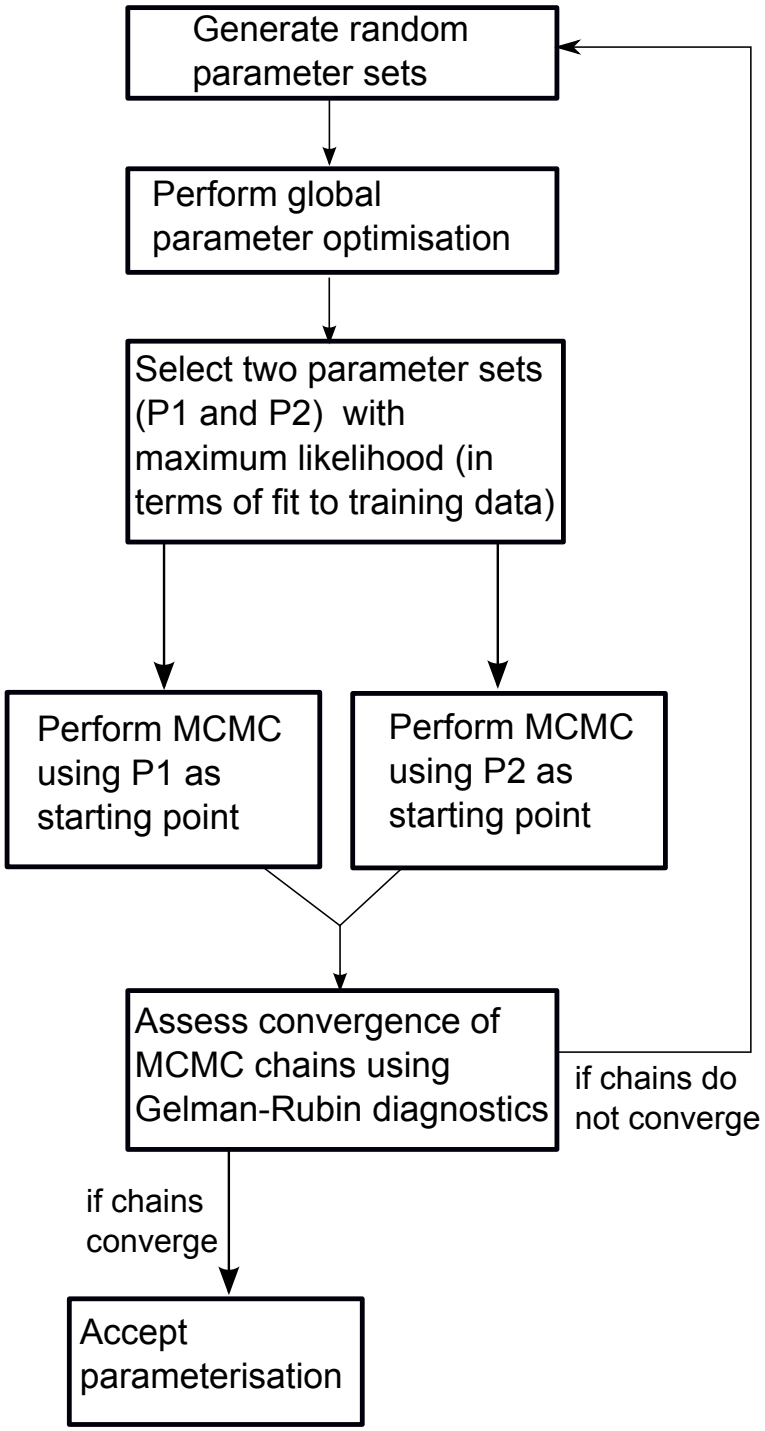


Figure 7.4: Outline of steps performed in parameter fitting process, indicating our criteria for accepting a model parameterisation as optimal. We describe each of the steps in the process in detail in the text.

eter space rather than exhaustively optimising a single parameter set.

Following a global search of the parameter space, we identify the two existing (distinct) parameter sets with the highest likelihoods. We then use these as starting points for MCMC chains. For smaller model structures, it may be the case that the optimisation methods identify the same parameter set every time the algorithm is run. In this case we run only one MCMC chain. For the models where multiple different parameter sets are obtained following a global search, it is important (for efficiency of the MCMC algorithm) that the two starting parameter sets have similar initial likelihood values.

After the individual MCMC chains have been deemed to have converged (assessed through simple visual inspection of the chains), we formally assess the convergence of the two chains. We do this using Gelman-Rubin diagnostics (Gelman and Rubin, 1992; Brooks and Gelman, 1998), as we describe in Section 7.4.3. Performing MCMC after the pre-MCMC optimisation routine is a further level of optimisation, which provides an additional opportunity for a non-optimal fit to be identified if i) the MCMC chains move away from the initial parameter set, or ii) a wide distribution of parameters gives the same likelihood. If two MCMC chains converge in this way then we accept the model parameterisation as being optimal; we then identify the parameter set in the parameter distribution obtained which has maximal likelihood and use these parameters for subsequent model simulations. We perform this process for each of the candidate model structures presented in Figure 7.2. We describe how we select between different candidate models in Section 7.4.1 and provide a more detailed description of the optimisation and MCMC algorithm used in Section 7.4.2 and Section 7.4.3.

7.4.1 Model Selection

A subset of the models included as candidate models as shown in Figure 7.2 have the same structures as models considered in Chapter 5. The criterion used for selecting the most appropriate model for describing the observed experimental data from the candidates presented

in Figure 7.2 is the *Bayesian information criterion* (BIC (Schwarz, 1978)). BIC is a measure of the goodness of fit of a model to the observed data and can be used to select between a finite number of candidate model structures. BIC penalises unnecessary complexity in a model, in the form of additional parameters, which do not lead to a notable improvement in the ability of the model to describe the observed data. BIC is defined as

$$BIC = -2\log(L) + p\log(n), \tag{7.1}$$

where L is the maximium likelihood value, p is the number of unknown parameters to be estimated in the model and n is the number of observed data points used to calculate the likelihood. Different models can be compared based on their BIC score. The model with the lowest BIC value is the model selected as being the simplest model that is able to best replicate the training data.

In addition, we calculate the negative log-likelihood of the predictions of the validation data for each fitted model structure. The validation data are recorded from the standard voltage-step protocols performed on the same cell as the sine wave training protocols. We can therefore directly compare different models quantitatively, regardless of the training data used to construct the model, as we described in Section 7.2.

7.4.2 Pre-MCMC Parameter Optimisation Algorithms

As discussed in Section 7.1, since MCMC primarily provides a local exploration of the likelihood surface, we need to perform a global search of the parameter space prior to initiating MCMC.

Simulated Annealing

The first optimisation method used is a combination of simulated annealing followed by simplex search on the best points identified using simulated annealing. We used the Matlab simulated annealing function in the Global Optimisation Toolbox (MatLab, 2012). This algorithm searches parameter space by taking steps sized according to a specified temperature

function. Higher temperatures correspond to larger steps and the temperature decreases during the search. A number of options can be adjusted in the simulated annealing algorithm (we include values for the options we set in the algorithm in Appendix C.2.1). We explored the use of different settings to try to identify combinations which allowed efficient and effective searching of the parameter space to determine optimal parameterisation. However this was found to be highly dependent on model and protocol combination and more work will be necessary to further refine this. We used the Boltzmann temperature function, which lowers the temperature rapidly initially before slowing down, allowing exploration with moderately large jumps for a longer period than with the default setting. The Matlab documentation for the simulated annealing function provides a clear summary of the different algorithm settings (MatLab, 2012).

Following simulated annealing initiated from random starting parameters sets, we performed a simplex search, using Matlab's `fminsearch` function (MatLab, 2012). The parameter sets identified as having the highest likelihood following simulated annealing were used as starting points for the simplex search. The simplex search was repeated until the exit flag for the optimisation algorithm indicated that termination was due to convergence of the solution. A tolerance of 0.05 on the change in likelihood value for termination of the algorithm was used. The tolerances were selected for termination of the algorithm within a reasonable time period which still yielded a parameter estimate with a good fit to experimental data, with further opportunity for optimisation when MCMC is performed.

We found that this combined method was very dependent on initial random parameter sets. For this reason the algorithm tolerances were selected so that the algorithm terminated sooner but had the capabilities of much higher-throughput. This meant that this method had to be performed with many different initial parameter sets for a best value to be identified. It was also found that when performing MCMC with the starting parameter sets identified using this method, for the larger models the MCMC chains would typically need to be run for longer before convergence of the chains, compared to using parameter sets identified using

the second method, which we now describe.

Covariance Matrix Adaptation - Evolution Strategy

The second algorithm used was a Matlab implementation of the Covariance Matrix Adaptation - Evolution Strategy (CMA-ES) (Hansen et al., 2003)[2]. This method searches parameter space using distributions of parameter values around the current parameter estimate on each iteration of the search. These distributions are updated via a covariance matrix so that the likelihood of the parameter set chosen in each step increases. This algorithm has been proposed as being suitable for optimisation problems without the need for extensive tweaking of algorithm settings. We therefore use the algorithm with all settings as default, changing only the step size parameter (we use a value of 0.15), which corresponds to the standard deviation of the initial parameter distributions.

This method was found to be much less dependent on the initial parameter guess. However, it still identified parameter sets with a range of different likelihoods when run from multiple different starting parameter sets. This indicated that for some protocol combinations for the higher dimensional parameter spaces considered (with the larger candidate models) this is a fairly complex optimisation problem. An ideal optimisation algorithm for a problem would identify the same region of parameter space as being optimal regardless of the initial parameter set. Further work on tuning the existing optimisation algorithms or consideration of additional methods is necessary to ensure that this ideal is realised and to reduce the computational time needed with the current approaches.

7.4.3 Markov Chain Monte Carlo

For the best starting points identified from pre-MCMC parameter optimisation, we then determine parameter distributions using MCMC. The same methods as described in Section 5.2.3 are used. In some cases, particularly for the models with fewer parameters, all of the

[2]A detailed description of this algorithm and source code can be found at https://www.lri.fr/ hansen/cmatutorial.pdf.

pre-MCMC parameter estimates from each algorithm identified the same parameter set. In this case, only one MCMC chain was run with this parameter set as starting point. In the cases where multiple parameter sets with similar likelihood values were identified for the larger models, two MCMC chains were run starting from the two best parameter sets and the convergence of the MCMC chain was assessed using the Gelman-Rubin diagnostics (Gelman and Rubin, 1992; Brooks and Gelman, 1998). To obtain the diagnostic we calculate the variance within each individual MCMC chain and the variance between different chains. An estimate of the variance of the parameter is then calculated as a weighted sum of the variance within the chain and between the chains. The diagnostic used to assess convergence is a ratio of the variance in the parameter and the variance within the chains. A diagnostic value close to 1 indicates convergence of the chains (Gelman and Rubin, 1992). The formulae used for calculating the variances and diagnostic described here are shown in Appendix C.2.2. We assessed convergence of the chains both using individual metrics on each parameter value and the multivariate version of the algorithm, which is thought to be a more strict measure. Diagnostics were calculated on chains of length 100,000 after discarding the burn-in. The length of the burn-in period is determined by inspection of each chain, and is dependent on the quality of the pre-MCMC parameter estimates. We accepted that chains have converged if the diagnostic value calculated using both measures is less than 1.1 (Gelman and Rubin, 1992; Brooks and Gelman, 1998). However, in practice, we obtained Gelman-Rubin convergence diagnostic values of less than 1.01 for many of our MCMC chains. Typically only a small improvement in the likelihood value is seen when performing MCMC, however, this step provides further opportunity for a sub-optimal parameters to be identified if the MCMC chain moves away from its initial starting point and towards a region of higher likelihood.

We used the methods outlined here to construct cell-specific models describing hERG channel kinetics observed in our experimental data and to begin to explore inter-cellular variability in our recordings. These results are presented in Section 7.5.

7.5 Results

In Chapter 6 we observed some variability between experimental recordings. We highlighted the example in Figure 6.5 where we observed that one cell exhibited the behaviour demonstrated by the Wang et al. (1997) hERG channel model and others that of the Mazhari et al. (2001) model. We therefore wanted to investigate further this observation and to explore quantitatively the underlying variability in the data. As the Mazhari et al. (2001) model was one of the highest ranked models in our preliminary assessment shown in Section 6.6, and as we noted that the behaviour of the Mazhari et al. (2001) model could be encapsulated in a reparameterisation of the Wang et al. (1997) model (as we demonstrated in Figure 6.6), we decided first to construct cell-specific models of hERG channel kinetics using the five state model structure shown in Figure 7.2 as in the Wang et al. (1997) model, which we discuss in Section 7.5.1.

7.5.1 Using Cell-Specific Models to Investigate Variability

As described in Section 7.2, we performed the *sine wave 1 protocol* on 40 different cells. We performed the model parameterisation process as described in Figure 7.4 on each of these sets of data. We accepted the resulting model parameterisation, after obtaining two convergent MCMC chains from different starting points for 18 of these cells. We obtained one convergent MCMC distribution for a further 6 cells, but did not conclude that these were optimal parameterisations as we did not have a second convergent chain to support this. We present the parameter estimates with maximum likelihood for these 18 cells in Figure 7.5 to assess the variability between cells. We also show a comparison with the parameter values in the Wang et al. (1997) model with the same structure.

In Figure 7.5 we can see that there are greater differences in estimates for some parameters than others. We also see that some parameters are much more distinct than those in the published Wang et al. (1997) model than others. We include the five state model structure with parameters and states numbered as we refer to them in this section in Figure 7.6 for

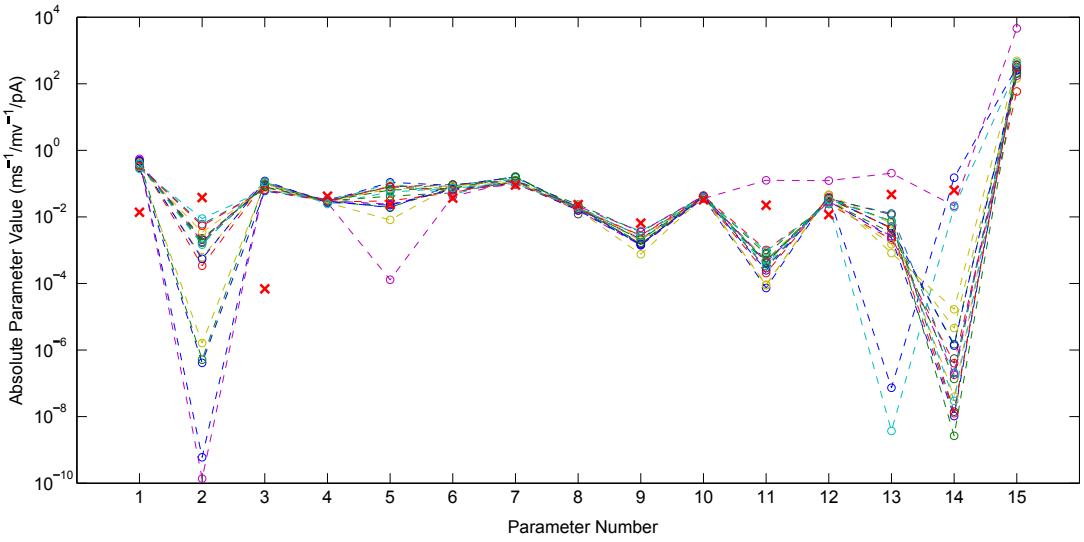


Figure 7.5: Plot of absolute values of parameter estimates with maximum likelihood from 18 different cells for the five state model structure estimated from current recordings in response to the *sine wave 1 protocol*. Estimates from each cell are connected so that parameter dependencies can be observed. Parameter numbering corresponds to that shown for the five state model structure, as recalled in Figure 7.6 below. Red crosses indicate the parameter values in the original Wang et al. (1997) model for comparison.

easy reference.

Five State Model

$$C_3 \underset{p_{13}\exp(p_{14}V)}{\overset{p_{11}\exp(p_{12}V)}{\rightleftharpoons}} C_2 \underset{p_6}{\overset{p_5}{\rightleftharpoons}} C_1 \underset{p_3\exp(p_4V)}{\overset{p_1\exp(p_2V)}{\rightleftharpoons}} O \qquad O \underset{p_7\exp(p_8V)}{\overset{p_9\exp(p_{10}V)}{\rightleftharpoons}} I$$

Figure 7.6: Five state Markov model structure with parameters (p_i) labelled as referred to in this section. p_{15} corresponds to the conductance parameter for this model.

We see in particular that our estimates for parameters 1-3 and 11-14 are most distinct from those in the Wang et al. (1997) model, in particular we see much less voltage-dependence on the transition between C_1 and O (parameter 2) and between C_2 and C_3 (parameter 14) than in the original model. Parameter 1, which is the parameter multiplying the exponential term in the activation transition rate between the closed C_1 state and the open state, is estimated fairly consistently in all the cells, with this being noticeably larger than in the Wang et al. (1997) model. Parameter estimates which vary most noticeably between cells are parame-

ters 2, 13 and 14. Parameters 13 and 14, which together define the deactivation transition rate between the C_2 and C_3 closed states appear to exhibit some parameter dependence — cells with a small estimate for parameter 13 have a larger estimate for parameter 14, and vice-versa, balancing out the overall difference in the transition rate. However, parameter 2, describing the voltage-dependence of the activation transition between the C_1 state and the open state appears to take a wide range of values. Parameters which we find have the smallest range of estimates from the 18 different cells are parameters 4 and 10, which correspond to the parameters describing voltage-dependence of the deactivation rate between the C_1 closed state and the open state and the transition between the inactivated and open state. These parameters are also similar to those in the Wang et al. (1997) model parameterisation.

From Figure 7.5 one trace stands out as being parameterised differently to the rest; having much lower values for parameter 5 (activation rate between the C_2 and C_1 states) and larger parameters describing the transition rates between the C_3 and C_2 states (described by parameters 11-14). We do note however, that this cell also has a larger conductance estimate (parameter 15) than the other cells. This may be an issue with larger parameters dominating the behaviour of the model and not allowing other parameters to be adequately determined. We also observed that this is the data set for which the highest values of the Gelman-Rubin diagnostics were calculated (indicating the worst convergence of the MCMC chains in our accepted set). This could indicate that this is not the optimal parameterisation and that we may need to make our criteria for MCMC convergence more stringent.

We next compared the parameter distributions obtained from the model parameterisation of the 18 cells. In Figure 7.7 we plot the parameter distributions for estimates of parameters 4 and 10, which are two parameters with fairly consistent values estimated across the different cells.

We see from the examples in Figure 7.7 that the parameter distribution estimates for the 18 cells all fall within an interval of less than 0.01 mV^{-1} for the estimates for both pa-

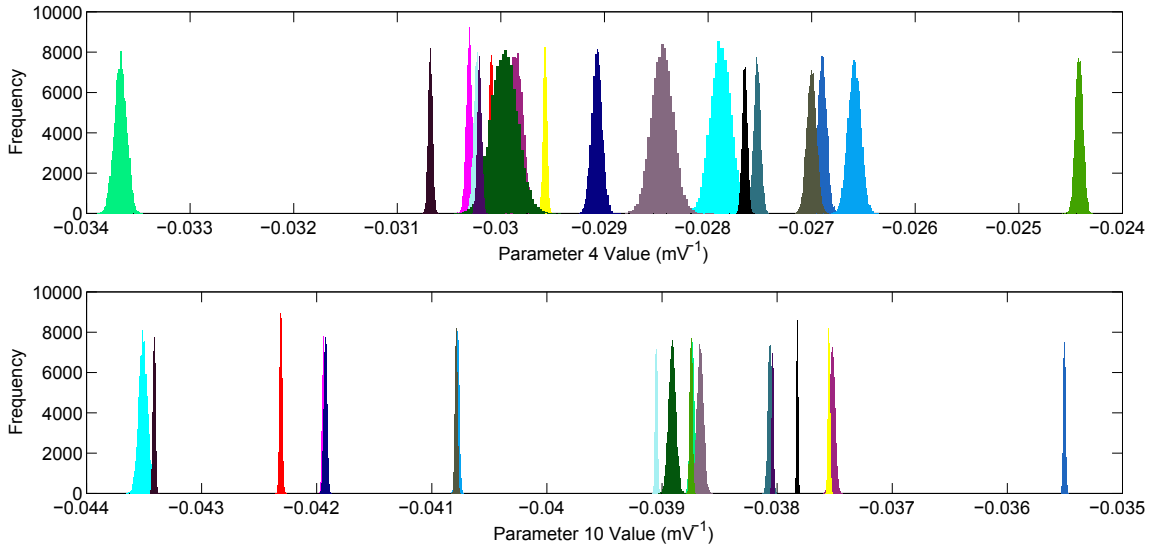


Figure 7.7: Histograms of parameter estimates for parameter numbers 4 and 10 from the five state model structure (parameter numberings correspond to those shown for the five state model in Figure 7.6). The histograms plotted are based on Markov chains of length 100,000 after discarding the burn-in period.

rameters. We also see that most of the histograms are clustered towards the centre of the intervals for both parameters, with just a couple of cells with distributions at the outer edges of the range. We see that while there is some overlapping of histograms, in many instances the distributions do not overlap, indicating there may be value of determining cell-specific models. Varying widths of the parameter distributions can be seen, which correspond to the uncertainty in the parameter estimates, in line with the amount of noise on the individual experimental recordings and any unidentifiability of the parameters with this training protocol. We should also note that the remaining 6 cells for which we have so far recovered just one MCMC chain for (and so have not yet accepted as optimal parameterisations according to our criteria) also lie within these intervals and follow the same patterns of parameter estimates.

In Figure 7.8, we show the distributions for parameters identified in Figure 7.5 as varying more between cells. The distributions are situated over a larger parameter range for parameter 5, which even more clearly demonstrates the advantage of cell-specific models.

However, in the case of parameter 2, the range is similar to the size of that for the parameters which appear to have more consistent values between cells, it is just that we see a number of cells with parameter distributions very close to zero, with the maximum likelihood being very small. This indicates very little voltage-dependence in the activation parameter in the transition between the closed state proximal to the open state and the open state for these cells.

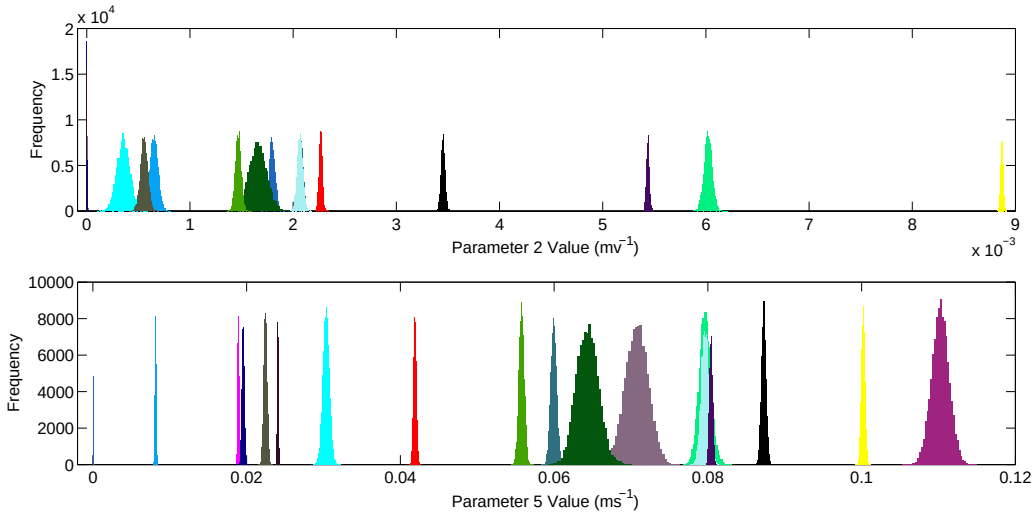


Figure 7.8: Histograms of parameter estimates for parameter numbers 2 and 5 for the five state model structure (parameter numberings correspond to those shown for the five state model in Figure 7.6). The histograms plotted are based on Markov chains of length 100,000 after discarding the burn-in period.

As noted in Section 7.2, we are not able to include the results from the full set of 40 sine wave traces collected in Chapter 6 as we were unable to obtain two converging MCMC distributions for all the parameter sets. The subset of 18 cells that we do include still constitutes a large data set for this investigation. However, we now reflect on possible explanations for the challenges encountered.

For the cells where only one MCMC chain was obtained, this was simply due to the search not encountering another parameter set with a comparable likelihood to the first parameter set. Starting an MCMC chain from a sub-optimal region of parameter space would make the

performance of the MCMC algorithm inefficient. Further search of parameter space using pre-MCMC optimisation is necessary to find a second chain to assess convergence and ultimately accept these parameterisations as optimal. We do however note that the parameter distributions for these remaining data sets with only one MCMC distribution are concordant with the trends observed from those presented here.

For the majority of cells for which we were unable to obtain any satisfactory parameter distributions, it was noted that the parameter sets returned from the pre-MCMC optimisation all contained a single pair of parameters which were typically much larger than the other parameters in the model (and in the models constructed which we accepted as being parameterised with an optimal solution). In these cases we found that there were identifiability issues with the models, with the large parameters being able to take a wide range of values without much change to the likelihood of the model fit. The large parameters would often get so large that numerical errors would begin in the solver and the MCMC algorithm would be terminated. The positions of these large parameters were consistently encountered either in the p_5 and p_{13} or p_{11} and p_{13} parameters (as labelled in Figure 7.6).

Taking an example of the simplest two state model, with one closed (C) and one open state (O) of the form shown in Equation (7.2), we demonstrate how possible identifiability issues can occur when large parameter values occur on the transitions between two states in the model.

$$C \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} O \tag{7.2}$$

where α and β are exponential functions of voltage (V) of the form $A\exp(BV)$ and $D\exp(EV)$ respectively. Recalling that $O = 1 - C$, the ODE describing this model is written as

$$\frac{\mathrm{d}C}{\mathrm{d}t} = -\alpha C + \beta(1 - C). \tag{7.3}$$

We can rewrite this as

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \frac{C_\infty - C}{\tau_c}, \tag{7.4}$$

where

$$C_\infty = \frac{\beta}{\alpha + \beta}, \tag{7.5}$$

and

$$\tau_c = \frac{1}{\alpha + \beta}. \tag{7.6}$$

We will call τ_v the timescale over which V varies (this is dictated by our protocol). If $\tau_v \gg \tau_c$ then, to first approximation, V is constant over the timescale τ_c and the solution to Equation (7.4) is given by

$$C(t) = C_\infty + (C(0) - C_\infty) \exp\left(\frac{-t}{\tau_c}\right). \tag{7.7}$$

Recalling that α and β are in the form $A\exp(BV)$ and $D\exp(EV)$ respectively, for this solution to be a good approximation we require

$$\tau_v \gg \tau_c = \frac{1}{\alpha + \beta} = \frac{1}{A\exp(BV) + D\exp(EV)}. \tag{7.8}$$

We now consider the extremes under which this must still hold — small τ_v and large τ_c. The smallest possible τ_v is simply the period of the shortest sine wave in the voltage protocol. The largest value of τ_c occurs when the denominator is at its minimum (V is at a minimum in the case where B is positive or maximum in the case where E is negative), which we can determine directly from our protocols. Then the above expression becomes

$$\tau_v \gg \frac{1}{A\exp(B\min(V)) + D\exp(E(\max(V)))}. \tag{7.9}$$

So, for any combination of A, B, D, E and τ_v for which equation (7.9) holds, then the solution rapidly decays to C_∞, and a quasi-equilibrium approximation of

$$C(t) \approx C_\infty = \frac{\beta}{\alpha + \beta} = \frac{D\exp(EV(t))}{A\exp(BV(t)) + D\exp(EV(t))} \tag{7.10}$$

will describe the solution well. In this situation, the dynamics of the transition will never be seen, as the states will quickly reach quasi-steady equilibrium. This suggests that the transition could be removed from the model, and a simpler Markov diagram would describe the data equally well. We will have unidentifiable parameters, as any choice of A, B, C and D that give the same C_∞ value will not change the model fit. The simplest way for this to

occur is if B and E are small (removing the effect of changing voltage), and A and D are large.

This just serves as an illustration for a simple model problem, and further work is required to try to apply the same reasoning to explain our observations of identifiability issues when large parameters are obtained on transitions which are not just between the same two states in the model (such as for the parameters p_5 and p_{11} and p_{11} and p_{13} parameters in the five state model structure shown in Figure 7.6). This example, however, provides some insight into what may be occurring within such models. This observation could be indicating that during model parameterisation, one of the states is effectively being removed from the model for these cells; this could be providing information about the model structure, possibly indicating that a simpler model structure is more appropriate.

We attempted to overcome this problem by restricting the search region for each parameter to the interval [0,1] (excluding the conductance). We did this to attempt to perform a more thorough search of this particular region (which contained the parameter values of the models which we had previously successfully fitted). We believed this would help us to determine whether the estimates identified using pre-MCMC optimisation were simply a result of the search algorithms getting stuck in local sub-optimal minima. However, we found that for the parameters which were estimated to be large when imposing the initial less strict prior, the parameter estimates were found to be at the upper boundary of the [0,1] interval, indicating that the parameterisation with maximal likelihood includes the large parameter values identified in the first instance.

When comparing the experimental recordings from cells for which we found large parameters, we see faster activation of the channel in the initial period of the protocol (with higher currents recorded more quickly) than in the training data where the fitted parameters were all within the [0,1] range. Incidentally, the two example cells we highlighted in Chapter 6 as one exhibiting behaviour more akin to that seen in the Wang et al. (1997) model and the other more like that of the Mazhari et al. (2001) model are examples of two cells where we

observed such differences; all the parameter estimates for the data more akin to that of the Wang et al. (1997) model were in the [0,1] range whereas we found larger parameters in the data which behaved more like the Mazhari et al. (2001) model.

Another observation is that two different types of behaviour are, for the most part, observed consistently on different days — with one type seen in all recordings made on any particular day. There were only two days which were an exception to this; with all but one of the cells being parameterised with values all within [0,1] on these days. These observations, in addition to recalling that the Mazhari et al. (2001) model describes dynamics at physiological temperature and the Wang et al. (1997) model describes kinetics at room temperature, could be an indication that this observed difference simply reflects temperature differences at the time of recording. This would also be concordant with the more rapid kinetics, and the correspondingly faster fitted transition rates, for the models which behave more akin to the Mazhari et al. (2001) model. Alternatively, there could be another more structure-based explanation for these observations, which may be worthy of further exploration. Further work will involve uncovering the challenges underlying the parameterisation of models to describe these additional data and then attempting to determine which of these explanations is more plausible. We will also record the temperature at the time of each future experiment.

The observations presented in this section indicate the potential of the use of a Bayesian approach for parameter estimation of cell-specific models for investigating variability between experimental recordings. This approach additionally allows uncertainty in parameter estimates to be propagated through to model predictions. Further work will see us add to these results to determine whether there are two distinct types of cells which can be differentiated by their estimated parameter values. We next attempt to select between candidate model structures shown in Figure 7.2 to determine the most appropriate model for describing hERG channel kinetics, as we describe in Section 7.5.2.

7.5.2 Determining and Validating Cell-Specific Models

We again follow the model parameterisation methodology outlined in Figure 7.4, this time repeating the process for each candidate model structure shown in Figure 7.2. We do this when using different combinations of sine wave protocols as the training data and we use the standard set of voltage-step protocols as the validation data.

We first show one example where we have used the data from the *sine wave 2 protocol* as our training data. We first examine the parameter distributions obtained from fitting both the three and four state model structures to data from the *sine wave 2 protocol*. In addition to those already outlined, another advantage of using a Bayesian inference approach for model parameterisation is the ability to compare the parameter distributions obtained from fitting different model structures to the same training data as this allows us to make further inferences about the appropriateness of a given structure. As an example, in Figure 7.9 we show the parameter distributions obtained from fitting the three and four state models when using the *sine wave 2 protocol* recordings as training data for one cell (we refer to this cell as cell 3 in Figures 7.11 and 7.12).

We see in Figure 7.9 that there is concordance between the parameter estimates for the three state and four state model structures when parameterised using the *sine wave 2 protocol*. This firstly provides further indication that this is the optimal parameterisation of the model structures for this data. It also provides an indication that the complexity of the three state model is almost sufficient to replicate this particular data. We can then assess the suitability of different model structures constructed using a subset of training data by assessing the predictions of validation data recorded in the same cell. In Figure 7.10 we show model predictions from both the three and four state models constructed using data from the *sine wave 2 protocol*.

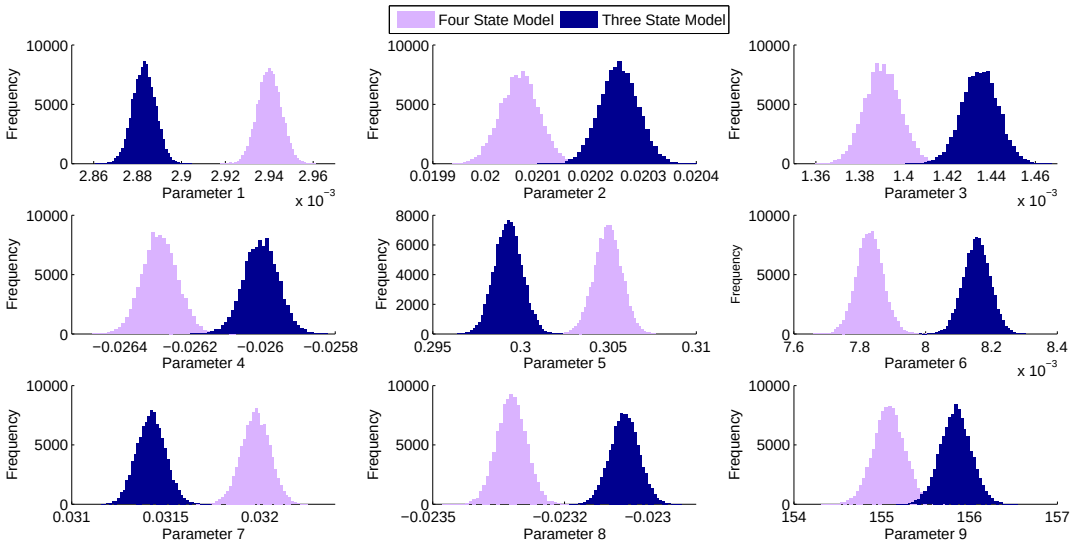


Figure 7.9: Comparison of parameter distributions obtained when fitting three and four state models using *sine wave 2 protocol* data from cell 3. Note that parameters (for both the three state and four state models) are numbered according to those in the three state model indicated in Figure 7.2. That is parameters 1-4 correspond to those on the transition between the closed and open state and parameters 5-8 correspond to those describing transitions between the open and inactivated states.

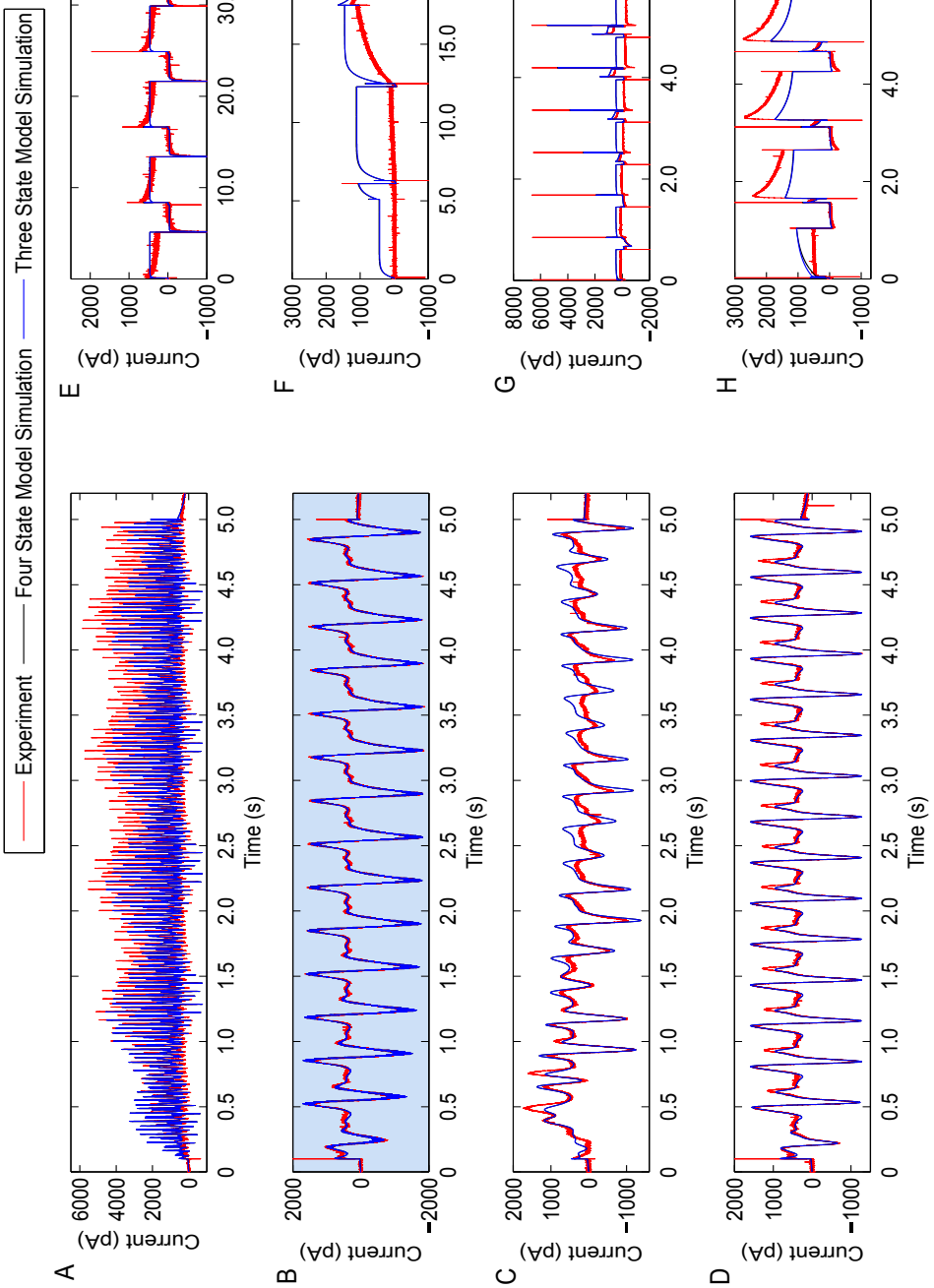


Figure 7.10: Comparison of model fit to experimental data and predictions from four state and three state model structures parameterised using training data from the *sine wave 2 protocol*. The model fits to this data are shown in panel B highlighted in blue. We show the model predictions of the *sine wave 1 protocol* in A, the *sine wave 3 protocol* in C, the *sine wave 4 protocol* in D, the deactivation protocol in E, the steady-state activation protocol in F, the inactivation protocol in G, and the activation kinetics protocol in H. It is apparent that the fits of each model to the experimental data from the *sine wave 2 protocol* are almost indistinguishable, as too are the model predictions.

We see from Figure 7.10 that the fits of the three and four state models to the experimental data from the *sine wave 2 protocol* are almost indistinguishable. This illustrates that the additional complexity represented in the four state model is somewhat unnecessary for encapsulating the behaviour of the training protocol. We see that the model predictions in Figure 7.10, while replicating some of the behaviour of the experimental data from this cell which were not used for fitting, is not able to fully replicate the observations.

As described in Section 7.4.1, we quantify the quality of the fit of a model to training data, penalising unnecessary model complexity using the BIC score. We also quantify the ability of a model to predict the validation data based on the likelihood of observing the experimental data given the parameter set identified during the fitting process. For the example shown in Figure 7.10, a marginally better fit is achieved when using the four state model (log-likelihood of -4.50$\times 10^5$ for the four state model and -4.51$\times 10^5$ for the three state model). However, the BIC score still indicates that the four state model is superior despite the added complexity (with a BIC score for the four state model of 9.00$\times 10^5$ and 9.02$\times 10^5$ for the three state model). We do however find that the three state model is more predictive than the four state model based on its ability to predict the behaviour of the standard voltage-step protocols in the validation set (with a log-likelihood of -7.87$\times 10^8$ for the four state model and -7.79$\times 10^8$ for the three state model). This also provides an illustration of how a better fit to a set of experimental data can often be achieved with a model of higher complexity, but that a better fit to training data does not necessarily translate into a more predictive model when assessed with validation data.

We performed the process outlined here using data recorded from 3 different cells. In each of these cells we performed the full set of sine wave and standard voltage-step protocols. We use data from each of the *sine wave protocols 1-3* (as well as considering the use of data from all three of these together) as the training data[3] when considering each of the candidate

[3]It should be noted that we disregard the *sine wave protocol 4* in this chapter for constructing models after the results presented in Chapter 5 indicated that this protocol did not provide as much information about model parameters as the other sine wave protocols; we instead use it as an additional validation protocol.

model structures shown in Figure 7.2. We summarise the quality of fits obtained in each case and in addition the ability to predict the behaviour observed in the standard voltage-step protocols in Figures 7.11 and 7.12 respectively.

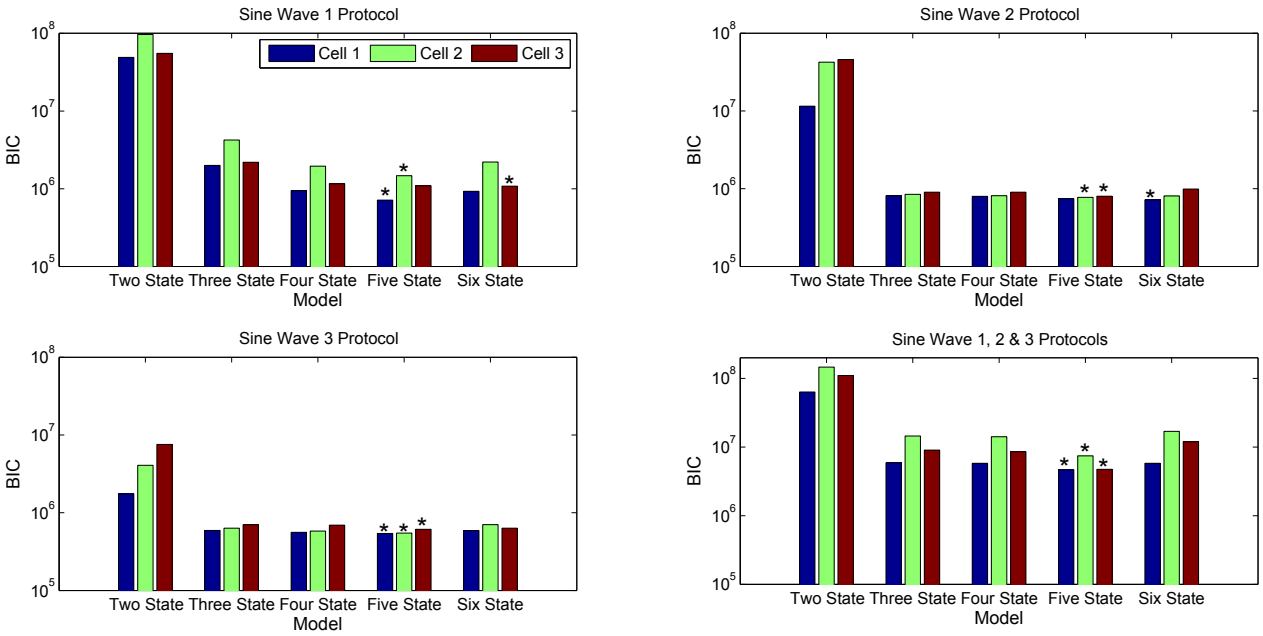


Figure 7.11: Calculated BIC values for 3 different cells based on the quality of fit of each candidate model structure shown in Figure 7.2 to data from different combinations of our sine wave protocols. The asterisk in each plot for each cell indicates the model with the lowest BIC score for that cell. This identifies the most appropriate model structure (from our candidate selection) for describing that particular set of experimental data based on the quality of model fit.

We see that in the majority of cases, the five state model is selected as being the model for which the best fit can be achieved for all of the cells when using data from each of the different voltage protocols. The five state model has more parameters than any of the other models, so we would expect that a better fit would be obtained. In only two cases an alternative model is selected; this being the six state model with the same structure as that proposed by Di Veroli et al. (2013). The two state model was consistently the model for which the poorest fit was obtained. We notice in particular, however, that when fitting models with each of the different protocol combinations, a reasonably good fit can be achieved for almost all of the model structures, excluding only the two state model.

Our experimental approach of performing multiple protocols on the same cell means that we can also validate the models constructed based on their ability to predict the behaviour of the standard voltage-step protocols after fitting the models to different combinations of the sine wave protocols performed on the same cell. The plot of the negative log-likelihood values calculated based on model predictions is shown in Figure 7.12.

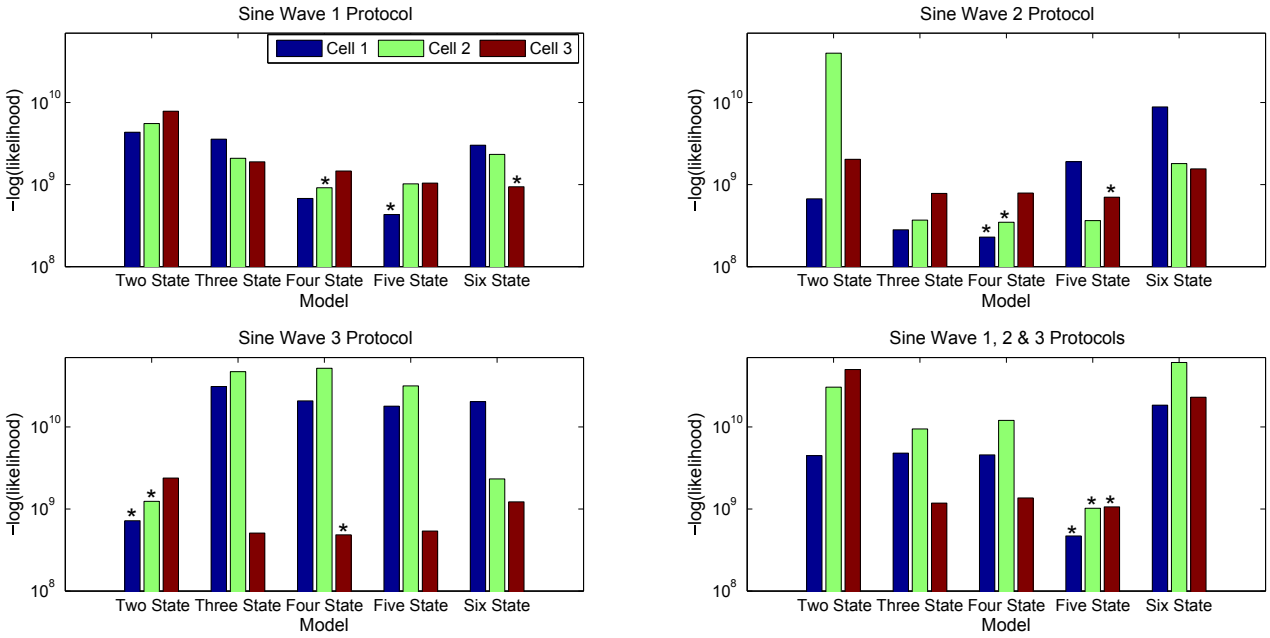


Figure 7.12: Calculated negative log-likelihood values values for 3 different cells based on the predictions from each model structure in Figure 7.2 of the behaviour of the standard voltage-step protocols when fitted using data from different combinations of sine wave protocols performed on the same cell. The asterisk in each plot for each cell indicates the model with the lowest negative log-likelihood, which indicates the best model for predicting the behaviour of the validation data.

Comparing Figures 7.11 and 7.12, we would make quite different conclusions about the appropriateness of different model structures for describing experimental behaviour when basing our decisions on the different BIC scores relative to the predictive power for validation data. When fitting the models to the combination of the *sine wave 1, 2 and 3 protocols*, the five state model structure is identified as being able to best replicate the data when using

the BIC score based on quality of model fit. This model is also selected as the best, as it was when the BIC was based on training data, for cell 1 fitting the model using data from the *sine wave 1 protocol* and cell 2 when using the *sine wave 2 protocol*. When fitting models for cell 3, using the *sine wave 1 protocol* as training data, the six state model was also identified as being most appropriate both for achieving a better quality of fit to the experimental data and for making the best predictions. However, we see that the model with the best fit to a given set of training data does not necessarily translate to be the most predictive model of the behaviour of the cell when the validation protocols are performed, and for the remainder of the cells and protocol combinations, a simpler model was found to have the lowest negative log-likelihood when considering validation than it did when the BIC score was calculated based on training data. This observation further indicates the necessity of considering not just the fit of a model to a subset of data when constructing a model, but also to thoroughly validate the model against independent data sets.

We saw in Figure 7.11 that a good fit could be achieved with almost all the model structures when parameterising the models using data from the *sine wave 3 protocol*. However, we see in Figure 7.12 that for cells 1 and 2, the best model predictions come from the two state model. These predictions are very poor, but so too are the rest of the predictions from the other models. This shows that achieving a very good fit of a model to training data does not mean that a model is appropriate for describing the behaviour of the system in general.

As a demonstration that the model structure to which the best fit to data can be achieved does not always consequently lead to the most predictive model, we can consider the example of cell 3 fitted to the *sine wave 3 protocol*. The model that fitted best to this set of experimental data was the five state model, with a BIC score of 6.12 $\times 10^5$ calculated. The three state model had a BIC score of 7.01 $\times 10^5$ when parameterised using this data. When validating each of these models by assessing the likelihood of the model parameters based on the ability to predict the behaviour of standard protocols performed on the same cell, we found that the three state model produced better predictions. The negative log-likelihood

values for predicting the behaviour of the standard voltage-step protocols being 5.39 $\times 10^8$ and 5.08 $\times 10^8$ for the five state and three state models respectively. The predictions from both the three and five state models are shown in Figure 7.13.

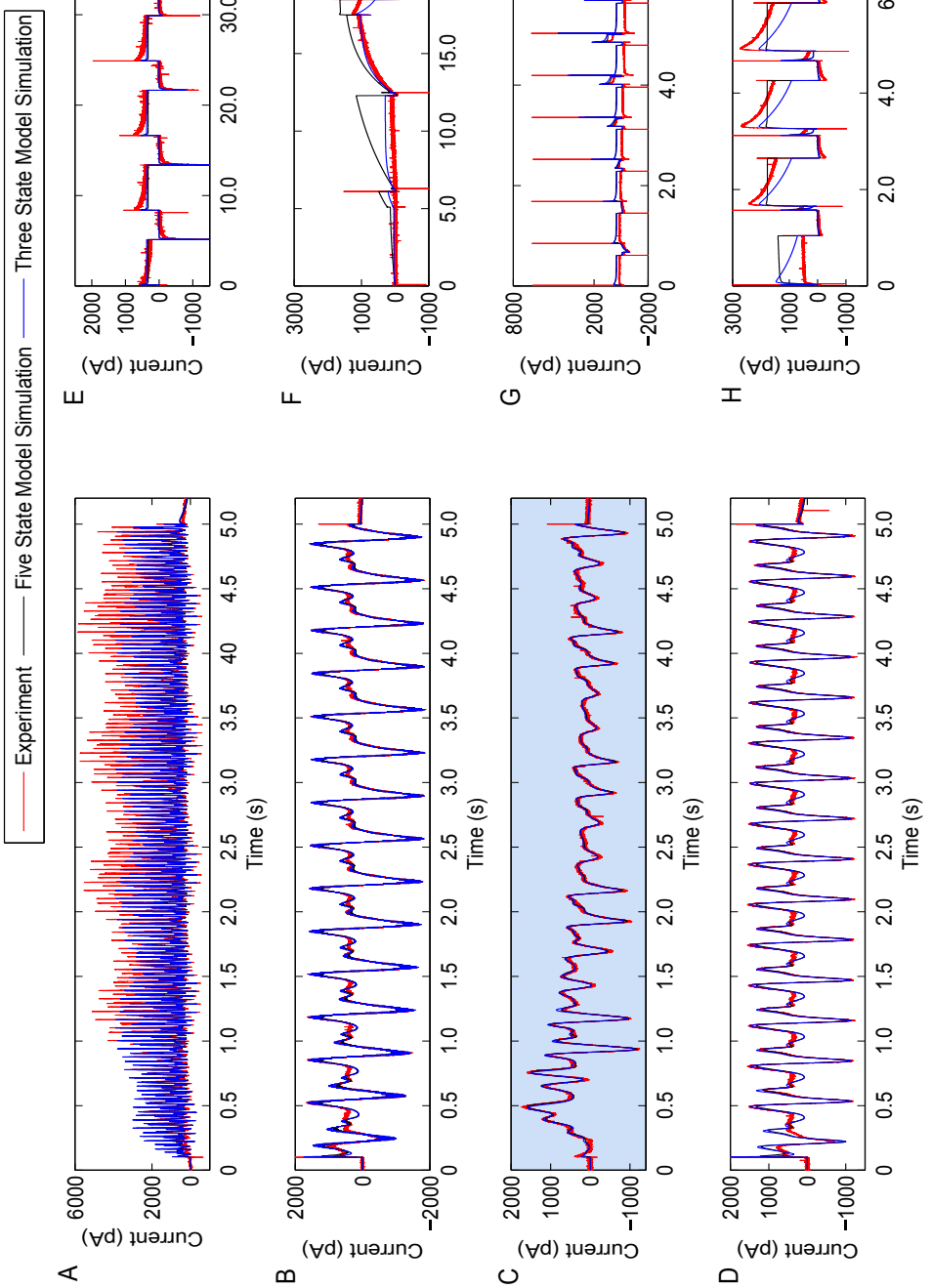


Figure 7.13: Comparison of predictions from five state and three state model structures parameterised using data from the *sine wave 3 protocol* (fits to this protocol highlighted in the blue panel C) for cell 3. While a better fit is achieved with the five state model, improved predictions are seen when using the three state model structure. We show the prediction of the *sine wave 1 protocol* in A, the *sine wave 2 protocol* in B, the *sine wave 4 protocol* in D, the deactivation protocol in E, the steady-state activation protocol in F, the inactivation protocol in G, and the activation kinetics protocol in H.

We can see from Figure 7.13 that both the five state and three state models are able to predict the behaviour of the *sine wave 1, 2 and 4 protocols* performed on the same cell well, with very similar predictions. However, the predictions for the standard voltage-step protocols are better when using the three state model than the five state model. It may be the case that the most appropriate parameterisation of the five state model has not been identified and an improved prediction may be possible. We note that the three state model, while it provides better predictions, is not able to capture all of the behaviour of the standard voltage-step protocols, and so we would still not deem this to be acceptable as a model for describing the kinetics of channels in this particular cell either.

Another interesting point to observe is that for cell 3, when fitting the models to the *sine wave 1 protocol*, the six state model was identified as being both the model to which the data was able to achieve a best fit, as well as the model which produced the best predictions. However, this was the only cell for which this was observed. We visualise the predictions in Figure 7.14, alongside those from the five state model.

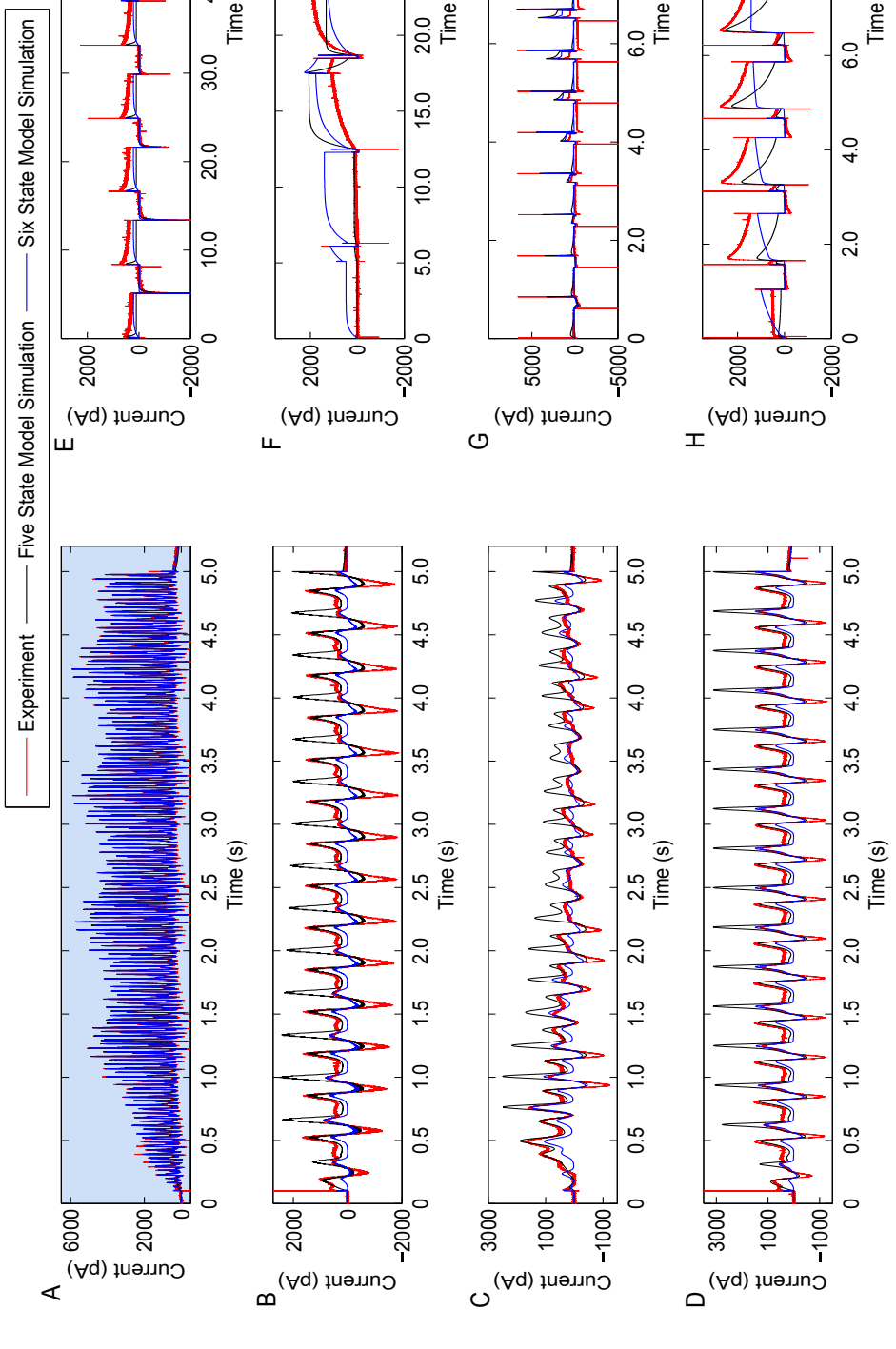


Figure 7.14: Comparison of predictions from five state and six state model structures parameterised to data from the *sine wave 1 protocol* (fits to this protocol highlighted in the blue panel A) for cell 3. A better fit and resulting predictions are seen when using the six state model structure. We show the prediction of the *sine wave 2 protocol* in B, the *sine wave 3 protocol* in C, the *sine wave 4 protocol* in D, the deactivation protocol in E, the steady-state activation protocol in F, the inactivation protocol in G, and the activation kinetics protocol in H.

We see in Figure 7.14 that while the six state model is able to predict the standard protocols reasonably well, neither it, nor the five state model, are able to predict the behaviour of the sine wave protocols adequately. We can compare the predictions presented in both Figures 7.13 and 7.14 and see that the predictions of the behaviour of the sine wave protocols were better when using the three state model (and fitting to the *sine wave 3 protocol*) than when using the five or six state models and fitting to the *sine wave 1 protocol*. In addition we should note that, as seen with simulated data, there is a very strong dependence between parameter 1 and the conductance parameter in the six state model structure when using the *sine wave 1 protocol* which means that the parameters cannot be adequately constrained.

If a model was to be identified as an appropriate model of the kinetics of the hERG channel then we would expect, not only concordance between the model structure identified as being the most appropriate when fitting to different experiments performed in the same cell, but also overlap of the parameter distributions when using training data from different experiments. This would then indicate that the models constructed were appropriate models of hERG channel kinetics and not just elaborate curve fitting to particular experiments. If we compare for example the parameter distributions obtained from cell 1 when fitting the five state model structure (identified as providing the best fit and best predictions in both cases) to just the *sine wave 1 protocol* and the combination of the *sine wave 1, 2 and 3 protocols*, as shown in Figure 7.15, we see a great discrepancy in the parameter distributions obtained.

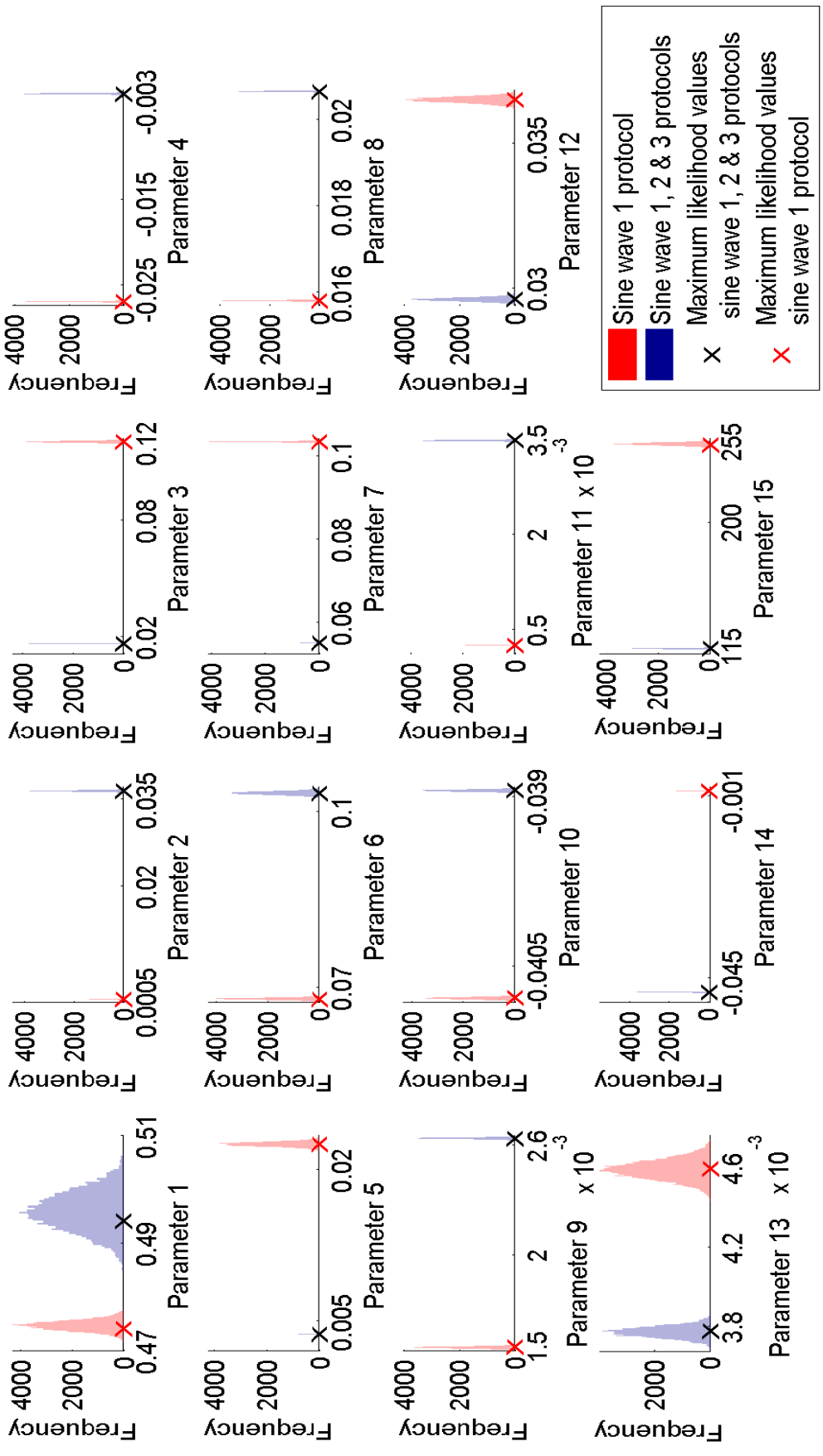


Figure 7.15: Comparison of parameter distributions obtained when training the five state model to cell 1 data from the *sine wave 1 protocol* or the combination of the *sine wave 1, 2 and 3 protocols*.

In Figure 7.15 we see no overlap between any of the parameter estimate distributions when the five state model is fitted to two different combinations of protocols recorded on cell 1. This could be an indication that non-optimal parameters have been identified, despite successfully performing the full model construction process as described in Section 7.4. Alternatively, it could be an indication that none of the existing hERG channel model structures are appropriate for describing the behaviour observed in experiment. The optimal parameter set when training the model to each experimental data set is sub-optimal when fitted to the other data set, indicating that if the problem is a sub-optimal parameterisation, this would mean that the parameterisations are sub-optimal in both cases. We also see in Figure 7.16 that the model predictions indicate that neither of these models are adequately able to describe the behaviour of the standard voltage-step protocols performed on the same cell, but this could be a consequence of either of these possibilities.

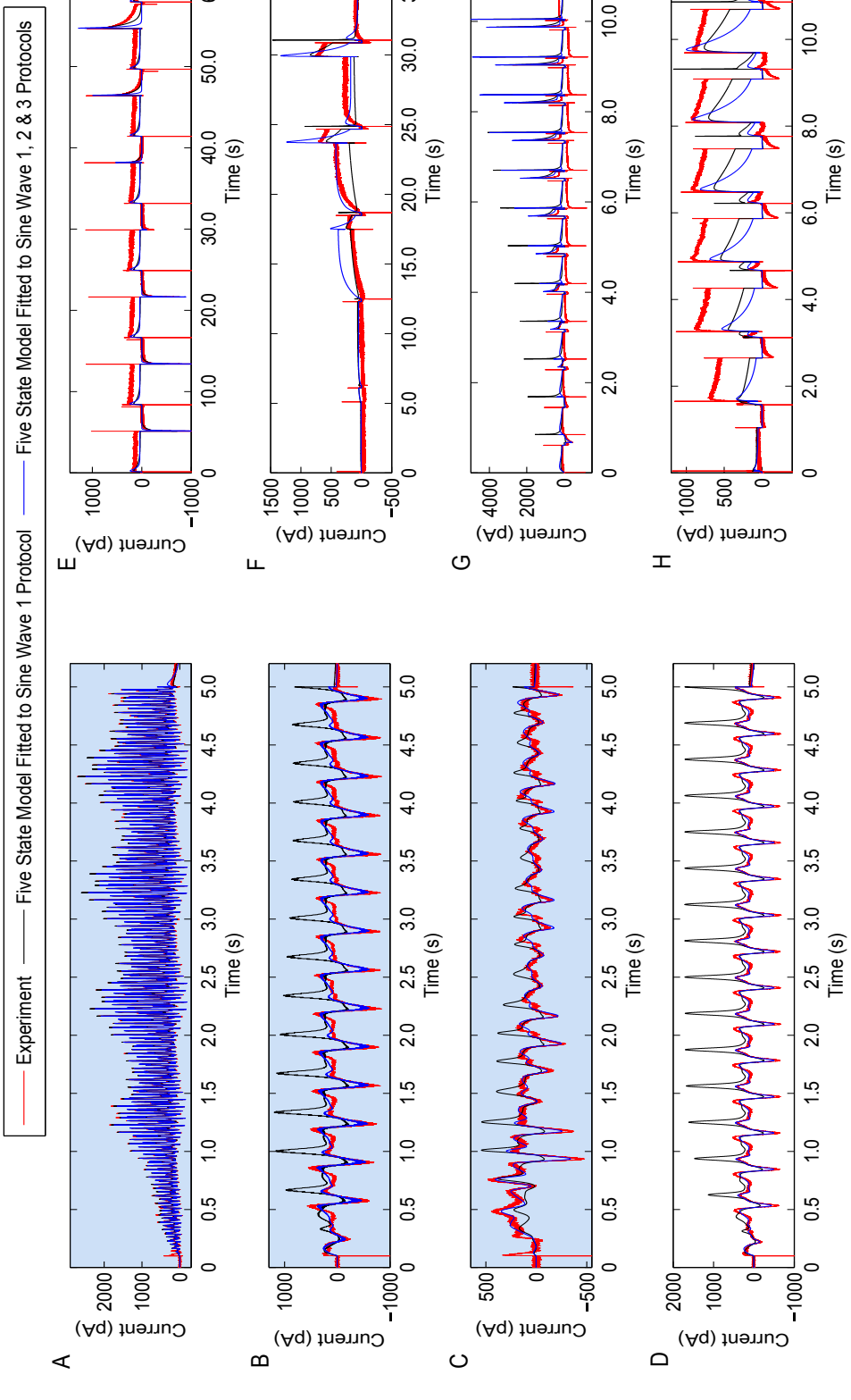


Figure 7.16: Comparison of predictions made when fitting the five state model to data from the *sine wave 1 protocol* and the combination of the *sine wave 1, 2 and 3 protocols* (traces fitted to highlighted in blue panels A, B and C) recorded in cell 1. We show the prediction of the *sine wave 1 protocol* in A, the *sine wave 3 protocol* in C, the *sine wave 4 protocol* in D, the deactivation protocol in E, the steady-state activation protocol in F, the inactivation protocol in G, and the activation kinetics protocol in H.

In Figure 7.16 we see that using the combination of the *sine wave 1, 2 and 3 protocols* for parameterisation leads to an improved prediction of the *sine wave 4 protocol*, as compared to when fitting with just the *sine wave 1 protocol* alone. However, worse predictions of the standard voltage-step protocols are seen, as is apparent from both visual inspection and from the calculated log-likelihood. It could be the case that in order to achieve a satisfactory fit to the additional two protocols the quality of the fit to the *sine wave 1 protocol* is compromised, which provides more information about the behaviour of the channel for predicting the standard voltage-step protocols. However if this were the case then we might expect that the prediction of the *sine wave 4 protocol* should also be better when fitting to just the one sine wave protocol. This could therefore provide indication that the fit to the combination of 3 sine wave protocols is not optimal, even though this has satisfied our criteria for being optimal as having 2 converging MCMC chains.

We encountered some of the same challenges outlined in Section 7.5.1 when parameterising certain model and protocol combinations. This provides some limitations to the results described in this section as we were not able to obtain two convergent MCMC chains in many cases (and so would not conclude that the optimal parameterisation has been identified), and so some of the results presented are based on the best parameter set identified by pre-MCMC optimisation. However, these challenges further illustrate the value of performing the thorough parameterisation process we outline, to ensure we construct maximally predictive models. We also see the value of performing MCMC, not just for obtaining model parameter distributions, but also in helping to identify model fits which are unsuitable or not optimal. In cases where a slight perturbation of the model parameters leads to a very large change in likelihood or indeed to numerical errors, this is an indication that this is not actually an appropriate parameterisation of the model and alerts us to the need to further refine our optimisation algorithms. We now reflect on some of the specific challenges we have encountered.

We found that large conductance values were fitted for many of the model structures and pro-

tocol combinations. We also saw again that pairs of parameters in the models were identified as being much larger than other model parameters identified by pre-MCMC optimisation. In particular, we saw that p_1 and p_6 were often much larger than others across the range of protocol combinations in the four state model shown in Figure 7.2. We also found instances of parameters p_5 and p_6 being parameterised with large values in the five state model. In all of these cases we again encountered issues with identifiability or low acceptance in the initial stages of MCMC leading to a very tight distribution about a point in parameter space. We consequently suspect that our resulting predictions were not as good as if we had been able to better constrain all of the parameters.

It is interesting to note that the Clancy and Rudy (2001) hERG channel model stands out from the other existing hERG channel models because it has large transition rates between the C_2 and C_1 states. In this model the p_5 and p_6 parameters take the values of 2.172 and 1.077 respectively. We also observe that this model is particularly poor at replicating our experimental data, an example of which we show in Figure 7.17 for the *sine wave 1 protocol*. This could perhaps be an indication of unidentifiability or a sub-optimal parameterisation in this model and of parameters possibly being able to be better constrained. However, we should also acknowledge that the Clancy and Rudy (2001) model describes hERG kinetics at physiological temperature and our data is recorded at room temperature, so differences may be attributed to temperature differences.

It could be the case that the identifiability issues encountered with the presence of large parameters in the models are suggesting that there is redundancy in the larger model structures, and that one transition (or a combination of transitions) immediately reaches a steady state, as we alluded to in Section 7.5.1. This could be indicating that a transition is not required within the model structure, a possible explanation which may be supported by our observations here that often simpler model structures are able to provide better predictions than the larger model structures which were identified as providing the best fits to the data. Alternatively it could be the case that we have not been able to identify the global optimal

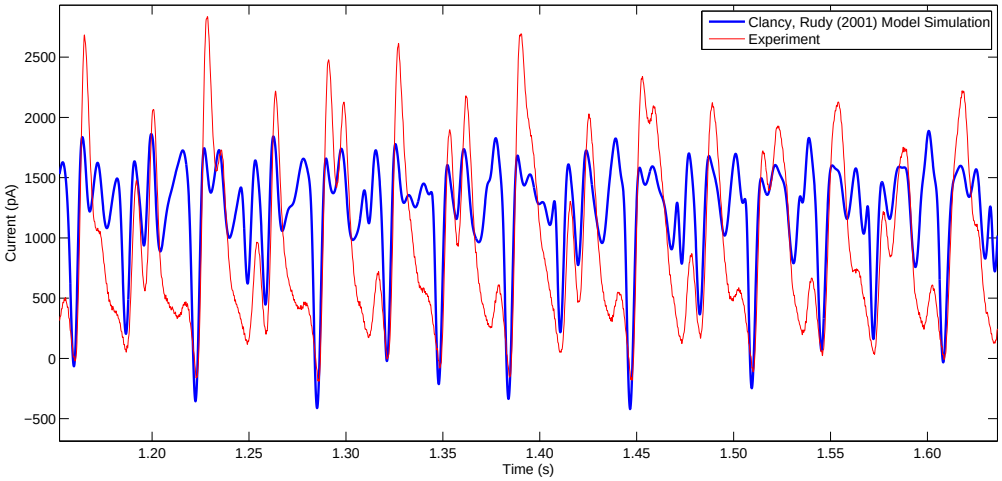


Figure 7.17: Comparison of the simulation of the *sine wave 1 protocol* using the Clancy and Rudy (2001) model simulation with recorded experimental data. The model used is parameterised with the published model parameters, and the conductance parameter has been scaled to minimise the absolute difference between the simulated and experimental trace to allow direct comparison.

solution for the larger models (where this behaviour is evident), and as we concentrate on optimising the optimiser performance for the particular protocol and model combinations, it should emerge if this is the case. A third possible explanation is that this is simply an indication that none of the models considered are adequate and that we need to expand our range of candidate structures.

While further work is required before these results can be used conclusively to infer the best model for describing the observed experimental data, we have highlighted a number of observations here which should be taken into account when determining kinetic models; we have clearly demonstrated the need for careful model parameterisation and adequate model validation. We have also learned much about the nature of the different protocols that we are using and their appropriateness for determining models of hERG channel kinetics. We saw consistently in each of the three different cells that the best fit was achieved in all cases when using the *sine wave 3 protocol* and the five state model structure, when basing this choice on the BIC score for quality of fit. However, when basing selection on the negative log-likelihood scores for the quality of model predictions of the standard voltage-step proto-

cols in the validation set from the same cell, we saw that the best model emerged as the four state model structure (parameterised via the *sine wave 2 protocol* in cells 1 and 2 and the *sine wave 3 protocol* for cell 3); this could not have been determined based upon considering the model fit to data alone.

Additionally, we should acknowledge that our assessment as to whether a model is able to adequately replicate our experimental data would be considered overly strict by many. For instance, if we plot the results of the predictions from the three state model shown in Figure 7.13 as it would normally be evaluated, through comparison of standard current-voltage relationships, then the results are as presented in Figure 7.18.

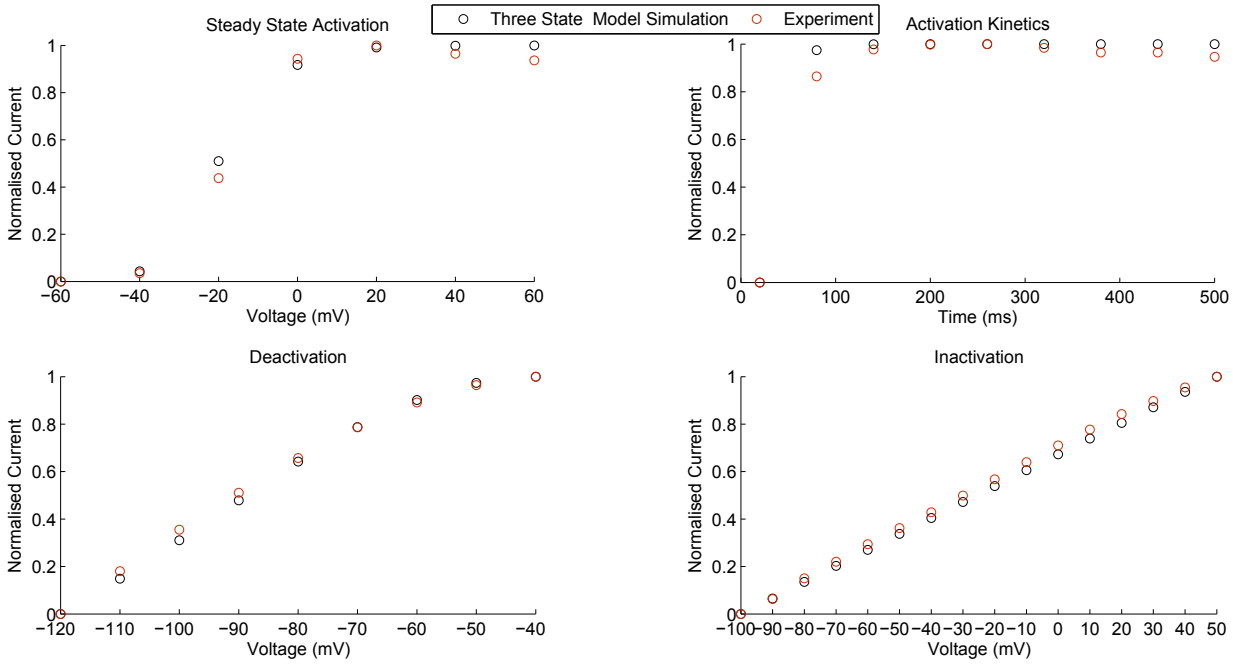


Figure 7.18: Current-voltage relationship curves based on maximum peak tail currents plotted for simulation and experimental data for the three state model structure describing the behaviour of cell 3 when fitted to the *sine wave 3 protocol*.

The results presented in Figure 7.18 would commonly be interpreted as demonstrating excellent agreement between model predictions and experiment. Based on Figure 7.18 alone, one may conclude once again that the proposed three state model (constructed from fitting

to just the 5.2 second *sine wave 3 protocol*) is able to adequately describe the kinetics of the hERG channel. However, as we saw in Figure 7.13 the model was not able to describe the full kinetic behaviour of the channel. This once again demonstrates not only the need for validation, but also the need for careful use of validation data — depending on the way the available data is used for validating models could lead to very different conclusions. Another point to note from this example is that the three state model, which is of much lower complexity than the majority of hERG channel models in use today, is able to replicate the behaviour of the observed data reasonably well, and compared with the data presented in Figure 7.18, would be considered an adequate model.

As an additional comparison we fitted the three state model structure to the experimental recordings in the set of four standard voltage-step protocols from the same cell (cell 3). The predictions of the sine wave protocols are shown in Figure 7.19.

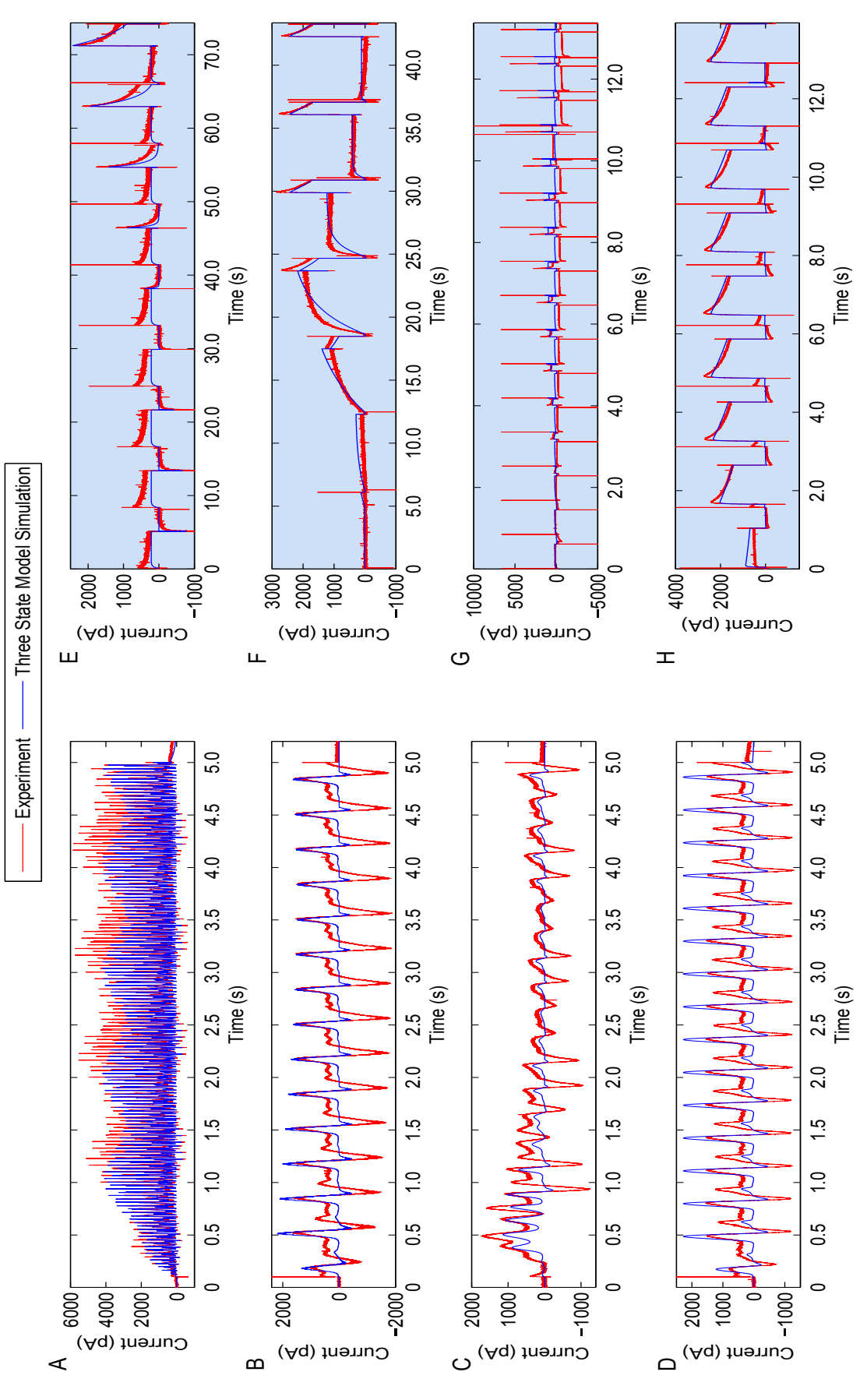


Figure 7.19: Comparison of predictions from the three state model structure parameterised using the set of data from the four standard voltage-step protocols (fits to these protocols are highlighted in the blue panels E-H) recorded in cell 3. We show the predictions of the *sine wave 1 protocol* in A, the *sine wave 2 protocol* in B, the *sine wave 3 protocol* in C and the *sine wave 4 protocol* in D.

We see much worse predictions when we fit the three state model to the standard protocols and predict the behaviour of the sine wave protocols than vice-versa (as compared with Figure 7.13). This may reflect additional information in the sine wave protocols for characterising the behaviour of the channel or the parameter set identified as optimal (consistently from multiple starting points) when fitting to the standard protocol data may not be the optimal fit. As part of further work we could consider parameterising different parts of the kinetic model using data from each of the different standard voltage-step protocols as would typically be done rather than parameterising the model based on using all the experimental data at once as we have done here. However, the results obtained at present further indicate the promise and applicability of using sine wave protocols for determining models of hERG channel kinetics.

7.6 Discussion

In this chapter we have illustrated an approach for simultaneous model selection and parameterisation to determine mathematical models of hERG channel kinetics. Our experimental approach of performing novel and standard voltage-step protocols on the same cell allows for thorough validation of the constructed models. The extensive experimental data set we have collected and the Bayesian approach for model parameterisation has allowed us to explore quantitatively the variability between cells. We see that many parameters are consistently estimated with similar values in different cells, and we identify parameters which vary on a cell-to-cell basis.

We have described the necessity of careful model construction, and in particular distinguishing between a poor parameter fit to a possible suitable model structure and a good fit to an unsuitable model structure. We have described the challenges encountered in practice when using experimental data, which were not seen when we explored the use of these protocols using simulated data with added noise as we described in Chapter 5. We have highlighted some interesting observations; notably how a model to which the data can be fitted does not

necessarily translate into being the most predictive model. While we consistently saw the same model structures identified using a BIC score based on the quality of the fit to training data, often simpler models had a greater likelihood based on their ability to predict the validation data. This further demonstrates the need for thorough validation and careful model construction.

We have also reflected on how the use of data for the purpose of validation may lead to very different conclusions about the appropriateness of a model and that there is value in looking at validation data in full detail rather than simply relying upon summary data, as we illustrated in Figure 7.18. From the present results it appears that none of the candidate model structures we have considered are appropriate for fully describing the full range of our experimental data. Further work will involve resolving the issues identified in this chapter to be able to determine this conclusively. We will then go on to consider additional possible model structures to determine a robust and accurate model of hERG channel kinetics. The parameter estimates obtained for the model structures considered so far will help to guide the choice of additional candidate models; for example our consistent observation of a more rapid activation transition than in existing models would be a first feature to try to encapsulate within a new model structure.

The results presented in this chapter demonstrate how a rigorous approach and careful consideration during the model construction process, coupled with experimental design which allows the opportunity for thorough model validation, can identify models that are good representations of the experimental data with which they were constructed but are not robust models of the physiological system they represent. The approach outlined here, and the experimental data amassed, provide a solid foundation and promise to enable an accurate and robust model of the kinetics of the hERG channel to be built. If we ultimately conclude that there is sufficient information for determining the kinetics of the channel in addition to determining the model parameters as we have shown using a 5.2 second voltage protocol, this will open up great opportunity for studying the kinetics of the hERG channel with a number

of different experimental interventions.

The results presented in this chapter also demonstrate the value of considering cell-specific models and the necessity, not only of performing validation, but performing it in such a way that does not lead to inappropriate conclusions being drawn. We have indicated that while the best model fits appear to be achieved with the more complex models, in the majority of cases, a simpler model is able to replicate better the validation data in the form of the behaviour of standard voltage-step protocols performed on the same cell.

We will build on the work presented here in future; refining aspects of the approach that have been identified as being necessary to reconsider or investigate. The data set we have collected and the methodology presented in this chapter provides us with the possibility of concluding this research by identifying an accurate and robust model of hERG channel kinetics.

7.7 Summary

In this chapter we have considered the problems of model selection and model parameterisation for constructing cell-specific models of hERG channel kinetics. We have also used such models to explore variability between the behaviour of different cells and we have highlighted some interesting observations worthy of further investigation. Our experimental approach allowed us both to construct and to validate models. We have discussed challenges encountered when attempting to construct models using experimental data, highlighting how we have attempted to avoid these and we have identified aspects of the methodology which need to be refined or further investigated as part of future work. In the results presented, we have demonstrated that a model which can best describe the training data used to construct it does not always translate to being the most predictive model when assessed using validation data from the same cell. We did not see concordance between parameters estimated using different subsets of data from the same cell, indicating either that none of the candidate

model structures considered are appropriate for describing hERG channel kinetics or that we have not yet found the best parameterisation of all of the model structures. As part of further work we will further investigate the observations and challenges discussed in this chapter in order to conclude which of the latter possibilities is correct and go on to determine the most appropriate model for describing hERG channel kinetics.

In Chapter 8 we provide an overview of the research presented in this thesis and describe the contributions and also limitations of the work as well as indicating possible avenues for extending this research.

8

Discussion and Future Research

We conclude with an overview of the research presented in this thesis, reflecting on how different aspects of the work contribute to our overall aims. We highlight both the key contributions and limitations of the research. We also identify a number of extensions and further applications of the work which will be pursued as future research.

8.1 Summary and Main Contributions

The research presented in this thesis is centered around the use of mathematical modelling for cardiac safety assessment of pharmaceutical compounds. We have described the potential of computational approaches for providing an inexpensive, and very high-throughput capability for assessing cardiac safety during drug development which also have the potential for reducing the use of animals needed for experimentation. We have placed the motivation behind this research in context, discussing intended changes to the existing regulatory guidelines for assessing the cardiac safety liability of pharmaceutical compounds: a core element of the proposal includes the use of *in silico* modelling for assessing the cardiac risk posed by new compounds. Consequently, increased adoption of computational approaches into routine pharmaceutical drug development is envisaged in the near future.

We have focused on two core concepts; the importance of careful model construction to ensure that mathematical models are accurate and able to replicate a range of experimental

behaviours, not just those from which they were constructed, and the necessity of thorough model validation. Careful consideration of these different aspects of model development will then hopefully lead to more predictive models, with rapid and successful uptake into pharmaceutical safety assessment. This could help alleviate the current high levels of attrition, reduce the use of animal experimentation, and facilitate the rapid progression of more safe drugs to market.

After an introduction outlining the aims and context of this thesis in Chapter 1, in Chapters 2 and 3 we described the biological and mathematical background of the research. In Chapter 4 we evaluated an *in silico* approach for predicting the results of a rabbit-based cardiac safety test; integrating a mathematical action potential model with high-throughput ion channel screening data collected in the earliest stages of drug development. The main strength of this study is the unprecedented number of compounds used for evaluation; with many compounds from multiple different ion channel screens included. The number of compounds used allows us to have greater confidence in the evaluation results and to gain a clearer understanding of the capabilities and limitations of the method.

We indicate that before adoption of computational approaches for use in routine pharmaceutical development, thorough validation must be performed. Validation has to extend beyond demonstrating the ability of a model to predict behaviour in response to a minimal number of reference compounds for it to be truly informative. We highlight how, with increasing numbers of compounds included for evaluation, the performance metric values calculated begin to converge rather than fluctuating as they do when only a handful compounds are considered (as is typical when validating such computational approaches). In addition, the thorough evaluation allows us to identify the best use of the evaluation data for the computational approach to demonstrate the highest level of predictive ability.

Following the large-scale evaluation presented in Chapter 4, this *in silico* screen is now in routine use in safety pharmacology at GSK. It is used prior to conducting the rabbit wedge

assay to prioritise compounds for screening in the rabbit wedge assay based on their expected effects on cardiac electrical activity and also to help assess the range of concentrations to be tested during experiment. The model predictions are also compared with the experimental results to help explain any unexpected experimental behaviour. The adoption of this *in silico* screen has led to a decrease in the number of rabbits used for experimentation, and while the computational approach demonstrated a good level of predictive ability, an enhanced ability would be desired before it could be considered as a complete replacement for an animal screen.

The results of this evaluation motivated the research presented in the rest of the thesis. In an attempt to enhance the predictive ability of the *in silico* assay we identified one of the main possible limitations of the work; the lack of kinetic representation of drug-ion channel interactions. We decided to address this limitation by focusing initially on the mathematical representation of the kinetics of the cardiac potassium channel hERG; a channel with which a diverse range of compounds interact and which plays a fundamental role in action potential repolarisation, which is implicated in many cardiac safety biomarkers. While our initial motivation was to create mathematical models describing drug interactions with the hERG channel, a review of the behaviour of existing hERG channel models demonstrated that these models can exhibit a wide range of different behaviours when simulated with the same voltage protocols. This led us to question whether the differences observed reflect biological variability (and that all the different model behaviours could be expected experimentally) or whether some of the differences were a consequence of the construction process of the mathematical models. We also demonstrated that the choice of hERG channel representation within an action potential model could have a profound effect on simulations of drug effects on cardiac electrical activity; further motivating the necessity of having an accurate representation of hERG channel kinetics.

We subsequently focused on the construction of accurate and robust mathematical models of hERG channel kinetics. Before embarking on this process we identified potential prob-

lems that could be encountered during model construction, which may consequently lead to inaccurate or non-robust models being proposed. Such problems included; not selecting the most representative model of the biological system, non-optimal model parameterisation, unidentifiable model parameters due to insufficient information contained within the experimental data or the excessive complexity of the model, *a priori* unidentifiability of the model formulation, or inadequate model validation. In Chapters 5 - 7, we described the full model construction process for determining models of hERG channel kinetics; beginning with the experimental design of novel experimental protocols, assessing the suitability of these protocols for determining parameters of existing models, performing the experiments and finally using the experimental data to select and parameterise mathematical models describing our experimental data, which we then go on to validate.

In Chapter 5 we described the methodology used to design new voltage protocols. We designed protocols specifically to aid with determining the most appropriate model structure and others to attempt to provide information for model parameterisation. We demonstrated how existing hERG channel models exhibit very different behaviour when simulated with these protocols; the dynamic nature of these protocols exposed even greater differences in the existing models than was apparent when simulating them with standard voltage-step protocols. The protocols proposed are very short compared to protocols typically performed for investigating hERG channel kinetics. If the protocols are able to more rapidly explore kinetics of the hERG channel, they offer great potential for being able to take multiple recordings from the same cell and to be able to collect data at a high-throughput. The ideas behind the design of each of the protocols are simple but may give more information for model selection and parameterisation.

In the second part of Chapter 5 we explored the identifiability of parameters of existing hERG channel models when using simulated data from different combinations of voltage protocols for model parameterisation. The use of a Bayesian inference approach for parameter estimation, allows us to quantify parameter uncertainty. We can then use this as a

guide for determining the information content within different voltage protocols for estimating each model parameter. This gives an indication of whether there is sufficient information contained within our new protocols for determining the parameters of different hERG channel model structures. The criteria used for defining identifiability are a subjective choice and may be considered strict, however, this analysis was used merely as a guide and we find that many of the parameters of existing models appear to be identifiable when using different combinations of the new protocols. We also observed that the more complex models have less-identifiable parameters, indicating that adding more and more equations to incorporate biological detail causes additional challenges for accurate model construction. This method for assessing the information content of different experimental protocols gives a good indication of whether there would be enough information within experimental data for determining a mathematical model before even performing the experiment, an important detail which is rarely considered, but could serve as a very important part of the experimental design process.

The conclusions of Chapter 5 encouraged us to pursue the use of our new protocols and we went on to perform the protocols in patch clamp experiments. In Chapter 6 we describe the experimental techniques used to record hERG channel currents in response to the protocols. We outlined some of the problems that could occur when performing these non-typical protocols and the considerations that were taken into account to attempt to prevent these problems from affecting the quality of the data. The observations and considerations noted in this chapter are important for being able to perform the sine wave protocols in patch clamp. Having the opportunity to undertake both the modelling and experimental aspects of this work allowed such considerations to be recognised and rectified after preliminary exploration of the use of these protocols and enabled greater familiarity and understanding of the data.

In Chapter 7 we used the experimental data collected and the results presented in Chapter 5 to attempt to construct cell-specific models of hERG channel kinetics. We emphasised

the necessity of careful consideration of both the model structure and parameterisation. We considered existing hERG channel Markov model structures as well as models nested within these as possible candidate models to try to identify the simplest possible model able to describe the observed experimental data. We described considerations taken to try to ensure that an optimal fit to the experimental data was obtained for each model structure, which we note as being important for being able to draw conclusions about the appropriateness of a given model structure. The Bayesian inference approach we used for model parameterisation provided a means for exploring parameter variability, and this can then be used to assess the suitability of a single model for describing the behaviour of a system as compared to considering models on a cell-specific basis. This approach also provides the possibility for parameter uncertainty to be propagated through to model predictions. A more careful approach to model construction as illustrated in this chapter will also hopefully lead to more informative and reliable model predictions.

We used the Bayesian information criterion as an indicator for model selection in Chapter 7. This only provides a guide as to the most appropriate model from a selection of models and the score does not give a direct indication as to how much better a given model is. Therefore a more sophisticated method could be used to guide model selection in future work. We were able to validate the models we constructed using data from protocols recorded on the same cell which were not used for parameterisation. This is a particular strength of this work, something very rarely seen when constructing mathematical models but we illustrate the great value of this approach in being able to identify models which may provide a good fit to the experimental data but which are not able to replicate the behaviour of the validation data as well.

We outlined a number of general potential reasons why models may not be as accurate or robust as possible, and we reflected on problems we encountered during the model selection and parameterisation process. The key contribution of the work presented in this chapter is not only illustrating the model construction and validation process, but, and possibly of more

value, highlighting the problems encountered and recognising the potential challenges when constructing mathematical models, which may lead to non-robust or inaccurate models.

We concluded Chapter 7 recognising that further work is necessary to conclusively determine an appropriate model of hERG channel kinetics, stating that existing hERG channel model structures appear inadequate for describing the experimental behaviour we observed in response to our new protocols. To determine a model which is able to replicate our observations we will have to reassess the suitability of our protocols for parameterising additional model structures and then perform the same process outlined in Chapter 7 for the new structures. The parameterisations we have already identified for existing model structures will help to guide the choice of which modifications to the existing structures to consider. We should however acknowledge that while our criteria for an adequate model representation may be considered rather strict, imposing more stringent assessment of model suitability in this way will hopefully be beneficial for determining more accurate and robust models. The careful consideration of the model construction process outlined in this chapter indicates that it is possible to determine a good fit to experimental data using many different possible model structures, but care must be taken to ensure that models constructed are accurate representations of the biological system they are modelling and not merely just good descriptions of the data used to construct them if they are to be used for predictive purposes.

The number of experimental recordings collected as part of this study provided the possibility of exploring cell-to-cell variability in response to the novel protocols. The results presented in this thesis based both on our qualitative observations as indicated in Chapter 6, and when considering the idea of variability more quantitative manner in Chapter 7, provide some interesting potential avenues for further exploration. In Section 8.2, we describe an array of extensions of the research presented in this thesis which could be pursued as part of further research.

8.2 Future Research

There are many potential avenues for future research, some of which have been outlined throughout the thesis. First and foremost, the methodology described in Chapter 7 will be applied to an extended range of candidate model structures, to identify the most appropriate model structure for describing our experimental data. The second step is to extend this methodology to determine models describing drug-ion channel interactions. After doing this we would then replace the existing model of hERG channel current in the Shannon et al. (2004) rabbit ventricular action potential model used for the evaluation presented in Chapter 4 with kinetic models, and re-evaluate the predicted effects of a selection of drugs on the rabbit electrophysiology to assess whether an improvement in predictive ability is demonstrated. We would also want to look beyond solely considering drug effects on the APD and explore the use of the hERG channel kinetic model within action potential models used for predicting additional arrhythmic biomarkers such as early and delayed after depolarisations.

If an improvement in predictive ability is seen when using kinetic models, the natural next step is to assess the feasibility of performing such protocols in high-throughput ion channel screens used within pharmaceutical drug development, to determine whether the data collected is of sufficient quality to be able to determine parameters for kinetic models. If the high-throughput screening data could be used to determine kinetic models, our next step would be to refine the methods presented in this thesis so that the model construction process could be performed automatically. This would involve increasing the computational efficiency of the current methods, as well as optimising parameter estimation for the particular protocols identified. With an automated model construction process, high-throughput ion channel screening data would feed directly into the mathematical model construction framework to determine kinetic models from the data to then be incorporated within electrophysiology models to make (hopefully more accurate) predictions of drug cardiac side effects. This would be an extremely valuable tool for use in an industrial setting.

The methodology described in this thesis, from the experimental design process through to the model construction, is also directly applicable for characterising the kinetics of different ion channels of interest, such as the CaV1.2 channel. We could then include the kinetic representation of multiple ion channels in the same way within action potential models and assess the impact on model predictions.

One of the key advantages of the voltage protocols presented in this thesis is that they can be performed in just 5.2 seconds rather than taking a number of minutes as with traditional voltage-step protocols. This opens us the possibility of studying a number of different factors which may affect ion channel kinetics on a single cell. Such experiments include measuring the effect of an increased number of a number of different drug concentrations on the same cell, and exploring temperature-dependence by obtaining recordings at different temperatures on the same cell. This is of particular interest due to the observations made in Chapter 7 regarding the potential temperature-dependence of the recordings. We have already begun work exploring temperature-dependence of hERG channel kinetics when applying our new protocols.

There are a number of specific aspects of the methods presented in this thesis which could be further refined. For instance, it would be of great value to compare the use of a more efficient voltage-step protocol, which is better tolerated by cells, than the protocol we designed (but did not present in this thesis). We would refine our criteria for maximum voltage levels in the design process to achieve this. Voltage-step protocols are more practical to implement on the full range of high-throughput screens so if it is determined that the sine wave protocols do not yield high enough quality data when performed on the automated platforms, a more efficient voltage-step protocol would be the next to consider. In addition, it would be of great value to further optimise the new protocols by also optimising the protocol length to maximise the information content for selecting and estimating parameters of models in the shortest possible time period. As sine wave protocols are periodic, simply reducing them to the minimum length before repetition would be the first natural step. If we were to redesign

the sine wave protocols then we would make a smoother transition from the period at holding potential to the start of the sine wave to minimise capacitive effects.

There would be value in including action potential waveforms as part of the validation data set. As our first aim is to improve the representation of hERG channel kinetics to consequently improve the prediction of drug effects on the cardiac action potential, then this should certainly form part of the validation set. Alternatively, simple ramps to mimic the shape of an action potential may suffice. As our data possibly indicates that there may have been some temperature differences in the recordings, it is also necessary to record the temperature at the time of recording of each experiment.

It would be interesting to further explore the observation of differences in behaviour between cells and to try to uncover potential underlying explanations. It would also be useful to consider designing a voltage protocol specifically for ease of parameter fitting to further increase the efficiency of the full model construction process and reduce the potential for optimisation terminating in a non-optimal parameterisation. In the MCMC results presented in this thesis we have assumed that the level of noise is constant throughout each current recording, and this assumption may be a limitation which could be addressed. We could in addition include the noise as a parameter in the model, the estimate for which is refined during fitting.

8.3 Concluding Remarks

To conclude, in this thesis we have highlighted the potential that computational approaches could have within pharmaceutical drug development for predicting the cardiac safety liability of new drugs. We have also emphasised the need for thorough model validation and careful model construction to ensure that mathematical models are accurate and robust if they are to be used for predictive purposes. We have evaluated a computational approach for predicting the effect of drugs on QT prolongation as measured in the rabbit left-ventricular wedge

cardiac safety assay; based on drug interactions with cardiac ion channels. We identified a possible limitation of this approach as being the lack of kinetic representation of drug-ion channel interactions. We subsequently focused specifically on the mathematical representation of the cardiac potassium channel hERG within biophysical cardiac electrophysiology models. We saw large differences between simulations of different existing hERG channel kinetic models when simulated with the same voltage protocol. We also saw great differences between model simulations of the effect of a drug on the cardiac action potential when using different hERG channel representations within cardiac action potential models. We identified possible reasons for the discrepancy of model simulations as being simply representative of experimental variability or perhaps due to limitations in the model construction process. To explore this we illustrated the execution of the full model construction process; designing new experiments, assessing the information content of the experiments for determining kinetic models with identifiable model parameters, performing the experiments and using the resulting experimental data to construct and validate cell-specific models describing hERG channel kinetics. We used these models to begin to explore variability in our experimental recordings. We have highlighted potential challenges encountered and aspects to be considered when constructing mathematical models which may result in obtaining inaccurate and non-robust models. We have also discussed specific problems that we encountered during the process. We conclude by identifying that it may be that none of the existing hERG channel model structures considered are appropriate for describing our experimental data. Further work is necessary to consider additional candidate model structures to determine an appropriate model structure using the methods described in this thesis.

It is hoped that the research presented in this thesis has raised some interesting questions about hERG channel kinetics that are worthy of further exploration, as well as providing an interesting perspective and thoughts on mathematical model construction in the field of computational cardiac electrophysiology and beyond. Furthermore, we hope that some of the issues and potential solutions discussed here will help build towards future improvement of predictive ability of computational approaches for cardiac safety assessment and the wider

uptake of these solutions into routine pharmaceutical drug development.

A

Appendix

A.1 Mathematical Models of hERG Channel Kinetics

This appendix details the model equations for each of the hERG channel models included in Table 3.1. The models included in this appendix either take the form of Markov model or Hodgkin-Huxley style equations. The hERG channel current in each model is calculated according to

$$I_{Kr} = g_{Kr} P_O (V - E_{Kr}), \tag{A.1}$$

where g_{Kr} is the maximum conductance of the I_{Kr} channel, V is the membrane potential and E_{Kr} is the reversal potential. In Markov models, P_O is the probability of being in the open state, as determined from solution of the model ODE equations. Model equations in Markov models are composed of one ODE for each conformational state represented in the model and describe each of the possible transitions between states.

In Hodgkin-Huxley style models, P_O is represented as the equations describing the product of the activation gate and the inactivation gate. In some models the inactivation gate is time-independent and in others it is not. In some Hodgkin-Huxley style models it may be assumed that there are fast and slow activation in certain proportions and so this assumption is incorporated into the current calculation accordingly.

In each model formulation, the reversal potential is calculated by

$$E_{Kr} = \frac{RT}{F} \ln \left(\frac{[K]_o}{[K]_i} \right),$$

where $[K]_o$ is the extracellular potassium concentration, $[K]_i$ is the intracellular potassium concentration, R is the ideal gas constant, T is the temperature in Kelvin and F is Faraday's constant. Throughout we assume a temperature of 295.15 Kelvin, corresponding to room temperature at which our experimental recordings were taken.

The model formulations for each of the hERG channel models considered in this thesis are detailed below. Markov models are simply represented by their Markov model structure with transition rates indicated, and Hodgkin-Huxley style model equations are given in full. In the general form of these models, X_{Kr} correspond to activation gates and R_{Kr} to inactivation gates. Parameters which are not dimensionless are expressed in units of milliseconds and millivolts as appropriate. In each model an additional parameter describes the conductance term g_{Kr}.

Each model is described in terms of the model parameters. The parameter numberings are concordant with the numbering in the thesis text for each model. Parameters are indicated by p_i and the parameters corresponding to the published model parameters are included below the model formulation. Where constants appear in the model equations and are not assigned a parameter number, this is to ensure that the model is *a priori* identifiable - without assuming these parameters are constant the model would be *a priori* unidentifiable. In these cases, the assumption of constant parameters is made for that particular model when assessing parameter identifiability as in Chapter 5. In other cases we express parameter values as a ratio of parameter values - this is to indicate how we have incorporated two parameter values into one, again to make the model *a priori* identifiable. Models are listed in alphabetical order.

Aslanidi et al. (2009) model

$\tau_{X_{Kr}} = \frac{p_1}{1+\exp(\frac{V}{p_2})} + p_3,$

$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_4)}{p_5})},$

$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty}-X_{Kr}}{\tau_{X_{Kr}}},$

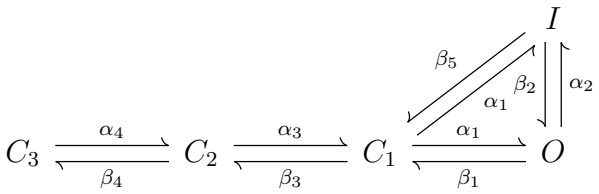
$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_6)}{p_7})},$

$g_{Kr} = p_8,$

$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V - E_{Kr}).$

p_1 = 900 p_2 = 5 p_3 = 100 p_4 = 0.085 p_5 = 12.25 p_6 = -5.4 p_7 = 20.4

Clancy and Rudy (2001) model



$\alpha_1 = p_1 \exp(p_2(V - 36))$ $\beta_1 = p_3 \exp(p_4 V)$

$\alpha_2 = p_7 \exp(p_8 V)$ $\beta_2 = p_9 \exp(p_{10}(V + 25))$

$\alpha_3 = p_5$ $\beta_3 = p_6$

$\alpha_4 = p_{11} \exp(p_{12}(V - 12))$ $\beta_4 = p_{13} \exp(p_{14} V)$

$\beta_5 = \frac{\beta_1\beta_2}{\alpha_2}$ (to satisfy microscopic reversibility condition)

$g_{Kr} = p_{15},$

$I_{Kr} = g_{Kr}P_O(V - E_{Kr}).$

p_1 = 0.0655 p_2 = 0.05547153 p_3 = 0.0029357 p_4 = -0.02158 p_5 = 2.172

p_6 = 1.077 p_7 = 0.6796 p_8 = 0.000942 p_9 = 0.4939 p_{10} = -0.02352

p_{11} = 0.0555 p_{12} = 0.05547153 p_{13} = 0.002357 p_{14} = -0.036588

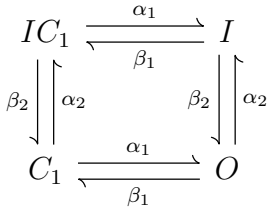
Courtemanche et al. (1998) model

$$\alpha_{X_{Kr}} = \frac{p_3(V+p_1)}{1-\exp(\frac{-(V+p_1)}{p_4})},$$
$$\beta_{X_{Kr}} = \frac{p_5(V+p_6)}{\exp(\frac{(V+p_6)}{p_7})-1},$$
$$\tau_{X_{Kr}} = \frac{1}{\alpha_{X_{Kr}}+\beta_{X_{Kr}}},$$
$$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_1)}{p_2})},$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty}-X_{Kr}}{\tau_{X_{Kr}}},$$
$$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_8)}{p_9})},$$
$$g_{Kr} = p_{10},$$
$$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V-E_{Kr}).$$

p_1 = 14.1 p_2 = 6.5 p_3 = 0.0003 p_4 = 5 p_5 = 0.000073898 p_6 = -3.3328

p_7 = 5.1237 p_8 = 15 p_9 = 22.4

Di Veroli et al. (2013) physiological temperature model



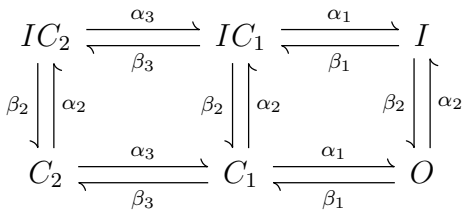
$$\alpha_1 = p_1 + p_2\exp(-(\tfrac{(V-p_3)}{p_4})^2) \qquad \beta_1 = p_5 + p_6\exp(-(\tfrac{(V-p_7)}{p_8})^2)$$
$$\alpha_2 = p_9 + p_{10}\exp(-(\tfrac{(V-p_{11})}{p_{12}})^2) \qquad \beta_2 = p_{13} + p_{14}\exp(-(\tfrac{(V-p_{15})}{p_{16}})^2)$$
$$g_{Kr} = p_{17},$$
$$I_{Kr} = g_{Kr}P_O(V-E_{Kr}).$$

p_1 = 0.000034 p_2 = 0.105 p_3 = 70 p_4 = 35.9 p_5 = 0.000000255 p_6 = 0.0427

p_7 = -145.3 p_8 = 62.7 p_9 = 0.127 p_{10} = 79 p_{11} = 470 p_{12} = 196

p_{13} = 0.00897 p_{14} = 63.9 p_{15} = -292 p_{16} = 102.2

Di Veroli et al. (2013) room temperature model



$\alpha_1 = p_1 \exp(p_2 V)$ $\beta_1 = p_3 \exp(p_4 V)$

$\alpha_2 = p_5 \exp(p_6 V)$ $\beta_2 = p_7 \exp(p_8 V)$

$\alpha_3 = p_9 \exp(p_{10} V)$ $\beta_3 = p_{11} \exp(p_{12} V)$

$g_{Kr} = p_{13}$,

$I_{Kr} = g_{Kr} P_O (V - E_{Kr})$.

p_1 = 0.00401 p_2 = 0.0263 p_3 = 0.000117 p_4 = -0.0401 p_5 = 0.0556 p_6 = 0.0199

p_7 = 0.0176 p_8 = -0.0274 p_9 = 0.025 p_{10} = 0.000124 p_{11} = 0.117 p_{12} = -0.0420

Fink et al. (2008) model

$$C_3 \underset{\beta_4}{\overset{\alpha_4}{\rightleftharpoons}} C_2 \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} C_1 \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} O \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} I$$

$\alpha_1 = \frac{T}{310} \exp(22.098 + \frac{310}{T}(0.0365V - 25.914))$

$\beta_1 = \frac{T}{310} \exp(7.313 + \frac{310}{T}(-0.0399V - 15.707))$

$\alpha_2 = \frac{T}{310} \exp(22.746 + \frac{310}{T}(-25.914))$

$\beta_2 = \frac{T}{310} \exp(13.193 + \frac{310}{T}(-15.707))$

$\alpha_3 = \frac{T}{310} (\frac{5.4}{4})^{0.4} \exp(30.016 + \frac{310}{T}(0.0223V - 30.888))$

$\beta_3 = \frac{T}{310} \exp(30.061 + \frac{310}{T}(-0.0312V - 33.243))$

$\alpha_4 = \frac{T}{310} \exp(24.335 + \frac{310}{T}(0.0112V - 25.914))$

$\beta_4 = \frac{T}{310} \exp(13.688 + \frac{310}{T}(-0.0603V - 15.707))$

Rewritten as:

$\alpha_1 = \frac{T}{310}\exp(p_2 V - p_1)$

$\beta_1 = \frac{T}{310}\exp(p_4 V - p_3)$

$\alpha_2 = \frac{T}{310}\exp(p_5)$

$\beta_2 = \frac{T}{310}\exp(p_6)$

$\alpha_3 = \frac{T}{310}(\frac{5.4}{4})^{0.4}\exp(p_8 V - p_7)$

$\beta_3 = \frac{T}{310}\exp(p_{10} V - p_9)$

$\alpha_4 = \frac{T}{310}\exp(p_{12} V - p_{11})$

$\beta_4 = \frac{T}{310}\exp(p_{14} V - p_{13})$

$g_{Kr} = p_{15},$

$I_{Kr} = g_{Kr} P_O (V - E_{Kr}).$

p_1 = 5.1198 p_2 = 0.0383 p_3 = 9.1843 p_4 = -0.0419 p_5 = -4.4718 p_6 = -3.3043

p_7 = 2.4261 p_8 = 0.0234 p_9 = 4.8546 p_{10} = -0.0328 p_{11} = 2.8828 p_{12} = 0.0118

p_{13} = 2.8093 p_{14} = -0.0633

Fox et al. (2002) model

$\tau_{X_{Kr}} = p_1 + \frac{1}{\exp(p_2 + p_3 V) + \exp(p_4 + p_5 V)}$

$X_{Kr\infty} = \frac{1}{1+\exp(p_6 + p_7 V)},$

$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty} - X_{Kr}}{\tau_{X_{Kr}}}.$

$R_{Kr} = \frac{1}{1 + p_8 \exp(p_9 (V+28))},$

$g_{Kr} = p_{10},$

$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}),$

p_1 = 43 p_2 = -5.495 p_3 = 0.1691 p_4 = -7.677 p_5 = -0.0128 p_6 = -2.182

p_7 = -0.1819 p_8 = 2.5 p_9 = 0.1

Grandi et al. (2010) model

$\tau_{X_{Kr}} = \frac{p_1}{1+\exp(\frac{-(p_2+V)}{p_3})} \frac{1}{1+\exp(\frac{(V+p_4)}{p_3})} + \frac{p_5}{1+\exp(\frac{(V+p_6)}{p_7})},$

$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_8)}{p_9})}$,
$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty}-X_{Kr}}{\tau_{X_{Kr}}}$.
$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_{10})}{p_{11}})}$,
$g_{Kr} = p_{12}$,
$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V - E_{Kr})$.

p_1 = 550×6 p_2 = 22 p_3 = 9 p_4 = 11 p_5 = 230 p_6 = 40 p_7 = 20 p_8 = 10
p_9 = 5 p_{10} = 74 p_{11} = 24

Hund and Rudy (2004) model

$\tau_{X_{Kr}} = \frac{1}{\frac{p_3(V+p_4)}{1-(\exp(-p_5(V+p_4)))}+\frac{p_6(V+p_7)}{\exp(p_8(V+p_7))-1}}$
$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_1)}{p_2})}$,
$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty}-X_{Kr}}{\tau_{X_{Kr}}}$.
$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_9)}{p_{10}})}$,
$g_{Kr} = p_{11}$,
$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V - E_{Kr})$,

p_1 = 10.085 p_2 = 4.25 p_3 = 0.0006 p_4 = -1.7384 p_5 = 0.136 p_6 = 0.0003
p_7 = 38.3608 p_8 = 0.1522 p_9 = 10 p_{10} = 15.4

Inada et al. (2009) model

$X_{Krfast_\infty} = \frac{1}{1+\exp(\frac{(V+p_1)}{-p_2})}$,
$\tau_{X_{Krfast}} = \frac{1}{p_3\exp(p_4V)+p_5\exp(-p_6V)}$,
$X_{Kr_{slow\infty}} = X_{Kr_{fast\infty}}$,
$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \frac{X_{Kr_{fast}\infty}-X_{Kr_{fast}}}{\tau_{X_{Kr_{fast}}}}$,
$\tau_{X_{Krslow}} = p_7 + p_8(\exp(\frac{-(V+p_9)^2}{p_{10}}))$,
$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \frac{X_{Kr_{slow}\infty}-X_{Kr_{slow}}}{\tau_{X_{Kr_{slow}}}}$,
$X_{Kr} = p_{19}X_{Krfast} + (1-p_{19})X_{Krslow}$,

$R_{Kr_\infty} = \frac{1}{1+\exp(\frac{(V+p_{11})}{p_{12}})}(1 - (p_{13}\exp(\frac{-V^2}{p_{14}})))$,

$\alpha_{R_{Kr}} = p_{15}\exp(-p_{16}V)$,

$\beta_{R_{Kr}} = p_{17}\exp(p_{18}V)$,

$\tau_{R_{Kr}} = \frac{1}{\alpha_{R_{Kr}}+\beta_{R_{Kr}}}$,

$\frac{\mathrm{d}R_{Kr}}{\mathrm{d}t} = \frac{R_{Kr\infty}-R_{Kr}}{\tau_{R_{Kr}}}$,

$g_{Kr} = p_{20}$,

$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V - E_{Kr})$.

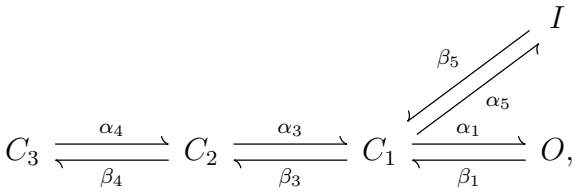
p_1 = 10.22 p_2 = 8.5 p_3 = 0.017 p_4 = 0.0398 p_5 = 0.000211 p_6 = 0.0510

p_7 = 335.81 p_8 = 906.73 p_9 = 10 p_{10} = 988.05 p_{11} = 4.9 p_{12} = 15.14

p_{13} = 0.3 p_{14} = 500 p_{15} = 0.09201 p_{16} = 0.0183 p_{17} = 0.6036 p_{18} = 0.00942

p_{19} = 0.9

Kiehn et al. (1999) model



Voltage (mV)	α_1	β_1	α_3	β_3	α_4	β_4	α_5	β_5
100	0.472	0.402	0.0997	0	0.0996	0	0.701	0.00167
80	0.577	0.370	0.0721	0	0.0721	0	0.596	0.00347
60	0.699	0.370	0.0497	0	0.0497	0	0.474	0.00597
40	0.588	0.323	0.0353	0.0171	0.0340	0	0.568	0.0106
-120	1.102	0.353	0.000245	0.0713	0.00162	0	0	0.111

$\mathrm{I}_{Kr} = g_{Kr}P_O(V - E_{Kr})$.

Kurata et al. (2002) model

$X_{Krfast_\infty} = \frac{1}{1+\exp(-\frac{(V+p_1)}{p_2})}$,

$\tau_{X_{Krfast}} = \frac{1}{p_3\exp(\frac{V}{p_4})+p_5\exp(\frac{-V}{p_6})}$,

$$X_{Kr_{slow}\infty} = X_{Kr_{fast}\infty},$$
$$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \frac{X_{Kr_{fast}\infty} - X_{Kr_{fast}}}{\tau_{X_{Kr_{fast}}}},$$
$$\tau_{X_{Krslow}} = \frac{1}{p_7 \exp(\frac{V}{p_8}) + p_9 \exp(\frac{-V}{p_{10}})},$$
$$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \frac{X_{Kr_{slow}\infty} - X_{Kr_{slow}}}{\tau_{X_{Kr_{slow}}}},$$
$$X_{Kr} = p_{17} X_{Krfast} + (1 - p_{17}) X_{Krslow},$$
$$R_{Kr\infty} = \frac{1}{1+\exp(\frac{(V+p_{11})}{p_{12}})},$$
$$\tau_{R_{Kr}} = \frac{1}{p_{13} \exp(\frac{-V}{p_{14}}) + p_{15} \exp(\frac{V}{p_{16}})},$$
$$\frac{\mathrm{d}R_{Kr}}{\mathrm{d}t} = \frac{R_{Kr\infty} - R_{Kr}}{\tau_{R_{Kr}}},$$
$$g_{Kr} = p_{18},$$
$$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$$

p_1 = 23.2 p_2 = 10.6 p_3 = $\frac{0.0372}{0.84655354}$ p_4 = 15.9 p_5 = $\frac{0.00096}{0.84655354}$ p_6 = 22.5
p_7 = $\frac{0.0042}{0.84655354}$ p_8 = 17 p_9 = $\frac{0.00015}{0.84655354}$ p_{10} = 21.6 p_{11} = 28.6 p_{12} = 17.1
p_{13} = 0.1 p_{14} = 54.645 p_{15} = 0.656 p_{16} = 106.157 p_{17} = 0.6

Lindblad et al. (1996) model

$$\alpha_{X_{Kr}} = p_1 \exp(\frac{V}{p_2}),$$
$$\beta_{X_{Kr}} = p_3 \exp(\frac{V}{-p_4}),$$
$$\tau_{X_{Kr}} = \frac{1}{\alpha_{X_{Kr}} + \beta_{X_{Kr}}}$$
$$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_5)}{p_6})},$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty} - X_{Kr}}{\tau_{X_{Kr}}}.$$
$$\alpha_{R_{Kr}} = p_7 \exp(\frac{V}{-p_8}),$$
$$\beta_{R_{Kr}} = p_9 \exp(\frac{V}{p_{10}}),$$
$$\tau_{R_{Kr}} = \frac{1}{\alpha_{R_{Kr}} + \beta_{R_{Kr}}}$$
$$R_{Kr\infty} = \frac{\alpha_{R_{Kr}}}{\alpha_{R_{Kr}} + \beta_{R_{Kr}}},$$
$$\frac{\mathrm{d}R_{Kr}}{\mathrm{d}t} = \frac{R_{Kr\infty} - R_{Kr}}{\tau_{R_{Kr}}}.$$
$$g_{Kr} = p_{11},$$
$$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$$

p_1=0.009 p_2 = 25.371 p_3 = 0.0013 p_4 = 13.026 p_5 = 5.1 p_6 = 7.4 p_7 = 0.1

p_8 = 54.645 p_9 = 0.656 p_{10} = 106.157

Liu et al. (1996) model

$$C_2 \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} C_1 \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} O \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} I$$

$\alpha_1 = p_1 \exp(p_2 V)$ $\beta_1 = p_3 \exp(p_4 V)$

$\alpha_2 = p_7 \exp(p_8 V)$ $\beta_2 = p_9 \exp(p_{10} V)$

$\alpha_3 = p_5$ $\beta_3 = p_6$

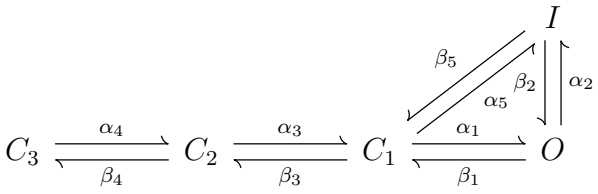
$g_{Kr} = p_{11}$,

$I_{Kr} = g_{Kr} P_O (V - E_{Kr})$.

p_1 = 0.0438 p_2 = 0.000583 p_3 = 0.0755 p_4 = -0.0315 p_5 = 0.00976 p_6 = 0.190

p_7 = 0.00446 p_8 = 0.0117 p_9 = 0.0121 p_{10} = -0.0324

Lu et al. (2001) model



$\alpha_1 = p_1 \exp(p_2 V)$ $\beta_1 = p_3 \exp(p_4 V)$

$\alpha_2 = p_7 \exp(p_8 V)$ $\beta_2 = p_9 \exp(p_{10} V)$

$\alpha_3 = p_5$ $\beta_3 = p_6$

$\alpha_4 = p_{11} \exp(p_{12} V)$ $\beta_4 = p_{13} \exp(p_{14} V)$

$\alpha_5 = p_{15} \exp(p_{16} V)$

$\beta_5 = p_{17} \exp(p_{18} V)$ (β_5 is calculated using the microscopic reversibility condition in practice)

$g_{Kr} = p_{19}$,

$I_{Kr} = g_{Kr} P_O (V - E_{Kr})$.

p_1 = 0.0578 $p_2 = \frac{11.2686}{T}$ p_3 = 0.000349 $p_4 = \frac{-12.3246}{T}$ p_5 = 0.1235 p_6 = 0.1911

$p_7 = 0.2533$ $p_8 = \frac{6.9085}{T}$ $p_9 = 0.0522$ $p_{10} = \frac{-9.5266}{T}$ $p_{11} = 0.1161$ $p_{12} = \frac{3.4699}{T}$

$p_{13} = 0.2442$ $p_{14} = \frac{-18.6146}{T}$ $p_{15} = 0.000052$ $p_{16} = \frac{17.6978}{T}$ $p_{17} = 0.0000000085$

$p_{18} = \frac{-21.3766}{T}$

Matsuoka et al. (2003) model

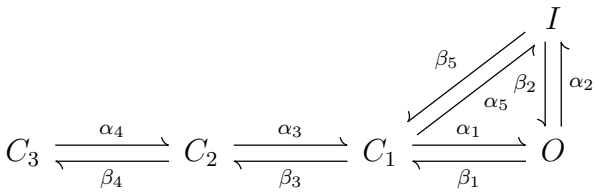
$$\alpha_{X_{Krfast}} = \frac{1}{p_1 \exp(\frac{-V}{p_2}) + p_3 \exp(\frac{-V}{p_4})},$$
$$\beta_{X_{Krfast}} = \frac{1}{p_5 \exp(\frac{V}{p6}) + p_7 \exp(\frac{V}{p_8})} + \frac{1}{p_9 \exp(\frac{V}{p_{10}})},$$
$$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \alpha_{X_{Kr_{fast}}}(1 - X_{Kr_{fast}}) - \beta_{X_{Kr_{fast}}}(X_{Kr_{fast}}),$$
$$\alpha_{X_{Krslow}} = \frac{1}{p_7 \exp(\frac{-V}{p_{11}}) + p_1 \exp(\frac{-V}{p_4})},$$
$$\beta_{X_{Krslow}} = \frac{1}{p_{12} \exp(\frac{V}{p_6}) + p_{13} \exp(\frac{V}{p_8})} + \frac{1}{p_{14} \exp(\frac{V}{p_{10}})},$$
$$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \alpha_{X_{Kr_{slow}}}(1 - X_{Kr_{slow}}) - \beta_{X_{Kr_{slow}}}(X_{Kr_{slow}}),$$
$$\alpha_{R_{Kr}} = \frac{1}{p_{15} \exp(\frac{V}{p_{16}}) + p_{17} \exp(\frac{V}{p_4})},$$
$$\beta_{R_{Kr}} = \frac{1}{p_{18} \exp(\frac{-V}{p_{16}}) + p_{19} \exp(-\frac{V}{p_{20}})},$$
$$\frac{\mathrm{d}R_{Kr}}{\mathrm{d}t} = \alpha_{R_{Kr}}(1 - R_{Kr}) - \beta_{R_{Kr}} R_{Kr},$$
$$X_{Kr} = p_{21} X_{Krfast} + (1 - p_{21}) X_{Krslow},$$
$$g_{Kr} = p_{22},$$
$$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$$

$p_1 = 20$ $p_2 = 11.5$ $p_3 = 5$ $p_4 = 300$ $p_5 = 160$ $p_6 = 28$ $p_7 = 200$ $p_8 = 1000$

$p_9 = 2500$ $p_{10} = 20$ $p_{11} = 13$ $p_{12} = 1600$ $p_{13} = 2000$ $p_{14} = 10000$ $p_{15} = 10$

$p_{16} = 17$ $p_{17} = 2.5$ $p_{18} = 0.35$ $p_{19} = 2$ $p_{20} = 150$ $p_{21} = 0.6$

Mazhari et al. (2001) model



$\alpha_1 = p_1 \exp(p_2 V)$ $\beta_1 = p_3 \exp(p_4 V)$

$\alpha_2 = p_7 \exp(p_8 V)$ $\beta_2 = p_9 \exp(p_{10} V)$

$\alpha_3 = p_5$ $\beta_3 = p_6$

$\alpha_4 = p_{11}\exp(p_{12}V)$ $\beta_4 = p_{13}\exp(p_{14}V)$

$\alpha_5 = p_{15}\exp(p_{16}V)$ $\beta_5 = \frac{\beta_1\beta_2\alpha_5}{\alpha_1\alpha_2}$ (to satisfy microscopic reversibility condition)

$g_{Kr} = p_{17}$,

$I_{Kr} = g_{Kr}P_O(V - E_{Kr})$.

p_1 = 0.0218 p_2 = 0.0262 p_3 = 0.0009 p_4 = -0.0269 p_5 = 0.0266 p_6 = 0.1348

p_7 = 0.0622 p_8 = 0.0120 p_9 = 0.0059 p_{10} = -0.0443 p_{11} = 0.0069 p_{12} = 0.0272

p_{13} = 0.0227 p_{14} = -0.0431 p_{15} = 0.0000129 p_{16} = 0.00000271

Noble et al. (1998) model

$\alpha_{X_{Krfast}} = \frac{p_1}{1+\exp(\frac{-(V-p_2)}{p_3})}$,

$\beta_{X_{Krfast}} = p_4\exp(\frac{-(V-20)}{p_5})$,

$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \alpha_{X_{Kr_{fast}}}(1 - X_{Kr_{fast}}) - \beta_{X_{Kr_{fast}}}(X_{Kr_{fast}})$,

$\alpha_{X_{Krslow}} = \frac{p_1}{1+\exp(\frac{-(V-p_2)}{p_3})}$,

$\beta_{X_{Krslow}} = p_6\exp(-((\frac{(V+p_7)}{p_8})^3))$,

$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \alpha_{X_{Kr_{slow}}}(1 - X_{Kr_{slow}}) - \beta_{X_{Kr_{slow}}}(X_{Kr_{slow}})$,

$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_9)}{p_{10}})}$,

$g_{Kr_{fast}} = p_{11}$,

$g_{Kr_{slow}} = p_{12}$,

$I_{Kr} = (g_{Kr_{fast}}X_{Kr_{fast}} + g_{Kr_{slow}}X_{Kr_{slow}})R_{Kr}(V - E_{Kr})$.

p_1 = 0.05 p_2 = 5 p_3 = 9 p_4 = 0.00005 p_5 = 15 p_6 = 0.0004 p_7 = 30

p_8 = 30 p_9 = 9 p_{10} = 22.4

Nygren et al. (1998) model

$X_{Kr_\infty} = \frac{1}{1+\exp(\frac{(V+p_1)}{-p_2})}$,

$\tau_{X_{Kr}} = p_3 + p_4\exp(-(\frac{(V+p_5)}{p_6})^2)$,

$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr_\infty} - X_{Kr}}{\tau_{X_{Kr}}}$,

$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_7)}{p_8})}$,
$g_{Kr} = p_9$,
$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr})$.

p_1 = 15 p_2 = 6 p_3 = 31.18 p_4 = 217.18 p_5 = 20.1376 p_6 = 22.1996 p_7 = 55
p_8 = 24

Oehmen et al. (2002) model

$$C_2 \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} C_1 \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} O \underset{\beta_2}{\overset{\alpha_2}{\rightleftharpoons}} I$$

$\alpha_1 = p_1 \exp(p_2(V + 10))$ $\beta_1 = p_3 \exp(p_4(V + 10))$
$\alpha_2 = p_7 \exp(p_8(V + 10))$ $\beta_2 = p_9 \exp(p_{10}(V + 10))$
$\alpha_3 = p_5$ $\beta_3 = p_6$
$g_{Kr} = p_{11}$,
$I_{Kr} = g_{Kr} P_O (V - E_{Kr})$.

p_1 = 0.0787 p_2 = 0.0378 p_3 = 0.0035 p_4 = -0.0252 p_5 = 0.0176 p_6 = 0.684
p_7 = 0.2977 p_8 = 0.0164 p_9 = 0.0862 p_{10} = -0.0454

O'Hara et al. (2011) model

$A_{X_{Krfast}} = \frac{1}{1+\exp(\frac{(V+p_1)}{p_2})}$,
$A_{X_{Krslow}} = 1 - A_{X_{Krfast}}$,
$\tau_{X_{Krfast}} = p_3 + \frac{1}{p_4 \exp(\frac{(V-31.66)}{p_5}) + p_6 \exp(\frac{-(V-47.78)}{p_7})}$,
$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_8)}{p_9})}$,
$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \frac{X_{Kr\infty} - X_{Kr_{fast}}}{\tau_{X_{Kr_{fast}}}}$,
$\tau_{X_{Krslow}} = p_{10} + \frac{1}{p_{11} \exp(\frac{(V-34.7)}{p_{12}}) + p_{13} \exp(\frac{-(V-29.74)}{p_{14}})}$,
$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \frac{X_{Kr\infty} - X_{Kr_{slow}}}{\tau_{X_{Kr_{slow}}}}$,
$X_{Kr} = A_{X_{Krfast}} \cdot X_{Kr_{fast}} + A_{X_{Krslow}} \cdot X_{Kr_{slow}}$,

$R_{Kr} = \frac{1}{(1+\exp(\frac{(V+p_{15})}{p_{16}}))\cdot(1+\exp(\frac{(V+p_{17})}{p_{18}}))},$

$g_{Kr} = p_{19},$

$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$

p_1 =54.81 p_2 = 38.21 p_3 = 12.98 p_4 = 0.3652 p_5 = 3.869 p_6 = 0.00004123

p_7 = 20.38 p_8 = 8.337 p_9 = 6.789 p_{10} = 1.865 p_{11} = 0.06629 p_{12} = 7.355

p_{13} = 0.00001128 p_{14} = 25.94 p_{15} = 55 p_{16} = 75 p_{17} = -10 p_{18} = 30

Priebe and Beuckelmann (1998) model

$\alpha_{X_{Kr}} = \frac{p_1 \exp(p_2(V+4.067))}{1+\exp(-p_3(V+4.067))},$

$\beta_{X_{Kr}} = \frac{p_4 \exp(p_5(V+65.66))}{1+\exp(p_6(V+65.66))},$

$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \alpha_{X_{Kr}}(1 - X_{Kr}) - \beta_{X_{Kr}} X_{Kr},$

$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_7)}{p_8})},$

$g_{Kr} = p_9,$

$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$

p_1 = 0.005 p_2 = 0.0005266 p_3 = 0.1262 p_4 = 0.016 p_5 = 0.0016 p_6 = 0.0783

p_7 = 26 p_8 = 23

Ramirez et al. (2000) model

$\alpha_{X_{Kr}} = \frac{p_1(V-p_2)}{1-\exp(\frac{-(V-p_2)}{p_3})},$

$\beta_{X_{Kr}} = \frac{p_4(V+p_5)}{\exp(\frac{(V+p_5)}{p_6})-1},$

$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_7)}{p_8})},$

$\tau_{X_{Kr}} = \frac{1}{\alpha_{Kr}+\beta_{Kr}},$

$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty}-X_{Kr}}{\tau_{X_{Kr}}},$

$R_{Kr} = p_9 + \frac{p_{10}}{1+\exp(\frac{V+p_{11}}{p_{12}})},$

$g_{Kr} = p_{13},$

$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$

p_1 = 0.04 p_2 = 248 p_3 = 28 p_4 = 0.028 p_5 = 163 p_6 = 21 p_7 = 7.654

p_8 = 5.377 p_9 = 0.07 p_{10} = 0.58 p_{11} = 15 p_{12} = 22.4

Seemann et al. (2003) model

Original model formulation:

$$\alpha_{X_{Kr}} = \frac{0.005\exp(0.00044761(V+2.23685))}{\exp(-0.10727(V+2.23685))},$$
$$\beta_{X_{Kr}} = \frac{0.016\exp(0.00232(V+55.811))}{\exp(0.066555(V+55.811))},$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \alpha_{X_{Kr}}(1-X_{Kr}) - \beta_{X_{Kr}}X_{Kr},$$
$$R_{Kr} = \frac{1}{1+\exp(0.023913(V+29.9))},$$

Rewritten as:

$$\alpha_{X_{Kr}} = p_1\exp(p_2(V+2.23685)),$$
$$\beta_{X_{Kr}} = p_3\exp(p_4(V+55.811)),$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \alpha_{X_{Kr}}(1-X_{Kr}) - \beta_{X_{Kr}}X_{Kr},$$
$$R_{Kr} = \frac{1}{1+\exp(p_6(V+p_5))},$$
$$g_{Kr} = p_7,$$
$$I_{Kr} = g_{Kr}R_{Kr}X_{Kr}(V-E_{Kr}).$$

p_1 = 0.005 p_2 = 0.10771761 p_3 = 0.016 p_4 = -0.064235 p_5 = 29.9 p_6 = 0.023913

Severi et al. (2012) model

$$X_{Kr_{fast\infty}} = \frac{1}{1+\exp(\frac{-(V+p_1)}{p_2})},$$
$$\tau_{X_{Krfast}} = \frac{1}{p_3\exp(\frac{V}{p_4})+\exp(\frac{-V}{p_5})},$$
$$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \frac{X_{Kr_{fast\infty}}-X_{Kr_{fast}}}{\tau_{X_{Krfast}}},$$
$$X_{Kr_{slow\infty}} = X_{Kr_{fast_\infty}},$$
$$\tau_{X_{Krslow}} = \frac{1}{p_6\exp(\frac{V}{p_7})+p_8\exp(\frac{-V}{p_9})},$$
$$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \frac{X_{Kr_{slow\infty}}-X_{Kr_{slow}}}{\tau_{X_{Krslow}}},$$

$$\tau_{R_{Kr}} = \frac{1}{p_{10}\exp(\frac{-V}{p_{11}})+p_{12}\exp(\frac{V}{p_{13}})},$$
$$R_{Kr_\infty} = \frac{1}{1+\exp(\frac{(V+p_{14})}{p_{15}})},$$
$$\frac{\mathrm{d}R_{Kr}}{\mathrm{d}t} = \frac{R_{Kr_\infty}-R_{Kr}}{\tau_{R_{Kr}}},$$
$$X_{Kr} = p_{16}X_{Krfast} + (1-p_{16})X_{Krslow},$$
$$g_{Kr} = p_{17},$$
$$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V-E_{Kr}).$$

p_1 = 14.8 p_2 = 8.5 p_3 = 0.03 p_4 = 10 p_5 = 12 $p_6 = \frac{0.0042}{0.84655}$ p_7 = 17
$p_8 = \frac{0.00015}{0.84655}$ p_9 = 21.6 p_{10} = 0.1 p_{11} = 54.645 p_{12} = 0.656 p_{13} = 106.157
p_{14} = 28.6 p_{15} = 17.1 p_{16} = 0.9

Shannon et al. (2004) model

$$\tau_{X_{Kr}} = \frac{1}{\frac{p_3(V+p_4)}{1-\exp(-p_5(V+p_4))}+\frac{p_6(V+p_7)}{\exp(p_8(V+p_7))-1}},$$
$$X_{Kr\infty} = \frac{1}{1+\exp(-\frac{(V+p_1)}{p_2})},$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty}-X_{Kr}}{\tau_{X_{Kr}}},$$
$$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_9)}{p_{10}})},$$
$$g_{Kr} = p_{11},$$
$$I_{Kr} = g_{Kr}X_{Kr}R_{Kr}(V-E_{Kr}).$$

p_1 = 50 p_2 = 7.5 p_3 = 0.00138 p_4 = 7 p_5 = 0.123 p_6 = 0.00061 p_7 = 10
p_8 = 0.145 p_9 = 33 p_{10} = 22.4

Ten Tusscher et al. (2004) model

$$\alpha_{X_{Kr_1}} = \frac{p_1}{1+\exp(\frac{-(p_2+V)}{p_3})},$$
$$\beta_{X_{Kr_1}} = \frac{1}{1+\exp(\frac{(V+p_4)}{p_5})},$$
$$\tau_{X_{Kr_1}} = \alpha_{X_{Kr_1}}\beta_{X_{Kr_1}},$$
$$X_{Kr_1\infty} = \frac{1}{1+\exp(\frac{-(p_6+V)}{p_7})},$$
$$\alpha_{X_{Kr_2}} = \frac{p_8}{1+\exp(\frac{-(p_9+V)}{p_{10}})},$$

$\beta_{X_{Kr_2}} = \frac{1}{1+\exp(\frac{(V-p_9)}{p_{10}})}$,
$\tau_{X_{Kr_2}} = \alpha_{X_{Kr_2}}\beta_{X_{Kr_2}}$,
$X_{Kr_{2\infty}} = \frac{1}{1+\exp(\frac{(V+p_{11})}{p_{12}})}$,
$\frac{\mathrm{d}X_{Kr_1}}{\mathrm{d}t} = \frac{X_{Kr_{1\infty}}-X_{Kr_1}}{\tau_{X_{Kr_1}}}$,
$\frac{\mathrm{d}X_{Kr_2}}{\mathrm{d}t} = \frac{X_{Kr_{2\infty}}-X_{Kr_2}}{\tau_{X_{Kr_2}}}$.
$g_{Kr} = p_{13}$,
$I_{Kr} = g_{Kr}X_{Kr_1}X_{Kr_2}(V - E_{Kr})$.

p_1 = 2700 p_2 = 45 p_3 = 10 p_4 = 30 p_5 = 11.5 p_6 = 26 p_7 = 7
p_8 = 3.36 p_9 = 60 p_{10} = 20 p_{11} = 88 p_{12} = 24

Wang et al. (1997) model

$$C_3 \underset{\beta_4}{\overset{\alpha_4}{\rightleftharpoons}} C_2 \underset{\beta_3}{\overset{\alpha_3}{\rightleftharpoons}} C_1 \underset{\beta_1}{\overset{\alpha_1}{\rightleftharpoons}} O \underset{\alpha_2}{\overset{\beta_2}{\rightleftharpoons}} I$$

$\alpha_1 = p_1 \exp(p_2 V)$ $\beta_1 = p_3 \exp(p_4 V)$
$\alpha_2 = p_7 \exp(p_8 V)$ $\beta_2 = p_9 \exp(p_{10} V)$
$\alpha_3 = p_5$ $\beta_3 = p_6$
$\alpha_4 = p_{11} \exp(p_{12} V)$ $\beta_4 = p_{13} \exp(p_{14} V)$
$g_{Kr} = p_{15}$,
$I_{Kr} = g_{Kr}P_O(V - E_{Kr})$.

p_1 = 0.013733 p_2 = 0.038198 p_3 = 0.0000689 p_4 = -0.04178 p_5 = 0.023761
p_6 = 0.036778 p_7 = 0.090821 p_8 = 0.023391 p_9 = 0.006497
p_{10} = -0.03268 p_{11} = 0.022348 $p_{12} = 0.01176$ p_{13} = 0.047002 p_{14} = -0.0631

Winslow et al. (1999) model

$$\alpha_{X_{Kr}} = 0.001 \times (\exp(-p_1 + p_2 V)),$$
$$\beta_{X_{Kr}} = 0.001 \times (\exp(-p_3 - p_4 V)),$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \alpha_{X_{Kr}}(1 - X_{Kr}) - \beta_{X_{Kr}} X_{Kr},$$
$$R_{Kr} = \frac{1}{1+p_5 \exp(p_6 V)},$$
$$g_{Kr} = p_7,$$
$$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$$

p_1 = 5.495 p_2 = 0.1691 p_3 = 7.677 p_4 = 0.0128 p_5 = 1.4945 p_6 = 0.0446

Zeng et al. (1995) model

$$X_{Kr\infty} = \frac{1}{1+\exp(\frac{-(V+p_1)}{p_2})},$$
$$\tau_{X_{Kr}} = \frac{1}{\left(\frac{p_3(V+p_4)}{1-\exp(-p_5(V+p_4))}\right) + \left(\frac{p_6(V+p_7)}{\exp(p_8(V+p_7))-1}\right)},$$
$$\frac{\mathrm{d}X_{Kr}}{\mathrm{d}t} = \frac{X_{Kr\infty} - X_{Kr}}{\tau_{X_{Kr}}},$$
$$R_{Kr} = \frac{1}{1+\exp(\frac{(V+p_9)}{p_{10}})},$$
$$g_{Kr} = p_{11},$$
$$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$$

p_1 = 21.5 p_2 = 7.5 p_3 = 0.00138 p_4 = 14.2 p_5 = 0.123 p_6 = 0.00061
p_7 = 38.9 p_8 = 0.145 p_9 = 9 p_{10} = 22.4

Zhang et al. (2000) model

$$X_{Krfast_\infty} = \frac{1}{1+\exp(\frac{-(V+p_1)}{p_2})},$$
$$\tau_{X_{Krfast}} = \frac{1}{p_3 \exp(\frac{(V-9)}{p_4}) + p_5 \exp(\frac{-(V-9)}{p_6})},$$
$$X_{Kr_{slow\infty}} = X_{Kr_{fast\infty}},$$
$$\frac{\mathrm{d}X_{Kr_{fast}}}{\mathrm{d}t} = \frac{X_{Kr_{fast_\infty}} - X_{Kr_{fast}}}{\tau_{X_{Kr_{fast}}}},$$
$$\tau_{X_{Krslow}} = \frac{1}{p_7 \exp(\frac{(V-9)}{p_8}) + p_9 \exp(\frac{-(V-9)}{p_{10}})},$$
$$\frac{\mathrm{d}X_{Kr_{slow}}}{\mathrm{d}t} = \frac{X_{Kr_{slow\infty}} - X_{Kr_{slow}}}{\tau_{X_{Kr_{slow}}}},$$
$$X_{Kr} = p_{14} X_{Krfast} + (1 - p_{14}) X_{Krslow},$$
$$R_{Kr_\infty} = \frac{1}{1+\exp(\frac{(V+p_{11})}{p_{12}})},$$

$\tau_{R_{Kr}} = p_{13},$

$\frac{\mathrm{d}R_{Kr}}{\mathrm{d}t} = \frac{R_{Kr\infty} - R_{Kr}}{\tau_{R_{Kr}}},$

$g_{Kr} = p_{15},$

$I_{Kr} = g_{Kr} X_{Kr} R_{Kr} (V - E_{Kr}).$

p_1 = 14.2 p_2 = 10.6 p_3 = 0.0372 p_4 = 15.9 p_5 = 0.00096 p_6 = 22.5 p_7 = 0.0042

p_8 = 17 p_9 = 0.00015 p_{10} = 21.6 p_{11} = 18.6 p_{12} = 10.1 p_{13} = 2 p_{14} = 0.6

B

Appendix

This appendix includes the performance metrics calculated when predicting QT prolongation and shortening when considering additional combinations of use of the available ion channel screening data to those presented in Chapter 4.

B.1 Further Combinations of Ion Channel Data Use

Table B.1 Prolongation metric values calculated (with 95% confidence intervals indicated) for evaluation of the assay's predictivity of QT prolongation in the rabbit ventricular wedge experiments when considering the different uses of the available ion channel data described which were not included in the main text. Entries are ordered alphabetically by the data type used in the simulations.

Case Number	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
S1	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value *(121 Compounds)*	**65.3** *(56.5-73.2)*	**0.35** *(0.19-0.51)*	**89.2** *(75.3-95.7)*	**54.8** *(44.1-65.0)*	**46.5** *(35.4-58.0)*	**92.0** *(81.2-96.8)*
S2	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(121 Compounds)*	**66.1** *(57.3-73.9)*	**0.36** *(0.20-0.52)*	**89.2** *(75.3-95.7)*	**56.0** *(45.3-66.1)*	**47.1** *(35.9-58.7)*	**92.2** *(81.5-96.9)*
S3	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value & Hill Coefficient *(121 Compounds)*	**75.2** *(66.8-82.0)*	**0.46** *(0.29-0.63)*	**73.0** *(57.0-84.6)*	**76.2** *(66.1-84.0)*	**57.4** *(43.3-70.5)*	**86.5** *(76.9-92.5)*
S4	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**71.4** *(60.5-80.3)*	**0.47** *(0.28-0.66)*	**100** *(86.7-100)*	**57.7** *(44.2-70.1)*	**53.2** *(39.2-66.7)*	**100** *(88.6-100)*
S5	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient, One-Dimensional Simulation *(77 Compounds)*	**70.1** *(59.2-79.2)*	**0.41** *(0.21-0.61)*	**84.0** *(65.3-93.6)*	**63.5** *(49.9-75.2)*	**52.5** *(37.5-67.1)*	**89.2** *(75.3-95.7)*
S6	QSAR Data, hERG Block Only *(372 Compounds)*	**54.6** *(49.5-59.6)*	**0.16** *(0.07-0.26)*	**94.4** *(89.8-97.1)*	**23.8** *(18.6-30.0)*	**48.9** *(43.4-54.4)*	**84.7** *(73.5-91.8)*

Table B.2 Shortening metric values calculated (with 95% confidence intervals indicated) for evaluation of the assay's predictivity for QT shortening in the rabbit ventricular wedge experiments considering the different uses of available ion channel data described which were not included in the main text. Entries are ordered alphabetically by the data type used in the simulations. Missing entries are due to the inability to calculate the metrics due to there being no compounds in the relevant categories for their computation.

Case Number	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
S1	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value *(121 Compounds)*	**74.4** *(65.9-81.3)*	**0.00** *(-0.30-0.30)*	**0.00** *(0.00-11.0)*	**100** *(95.9-100)*	-	**74.4** *(65.9-81.3)*
S2	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(121 Compounds)*	**74.4** *(65.9-81.3)*	**0.00** *(-0.30-0.30)*	**0.00** *(0.00-11.0)*	**100** *(95.9-100)*	-	**74.4** *(65.9-81.3)*
S3	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value & Hill Coefficient *(121 Compounds)*	**71.1** *(62.4-78.4)*	**0.22** *(0.00-0.44)*	**38.7** *(23.7-56.2)*	**82.2** *(73.1-88.8)*	**42.9** *(29.8-57.0)*	**79.6** *(70.3-86.5)*
S4	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(77 Compounds)*	**76.6** *(66.0-84.7)*	**0.00** *(-0.40-0.40)*	**0.00** *(0.00-17.6)*	**100** *(93.9-100)*	-	**76.6** *(66.0-84.7)*
S5	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient, One-Dimensional Simulation *(77 Compounds)*	**74.0** *(63.3-82.5)*	**0.27** *(0.00-0.55)*	**44.4** *(24.6-66.3)*	**83.1** *(71.5-90.5)*	**44.4** *(30.2-59.6)*	**83.1** *(71.5-90.5)*
S6	QSAR Data, hERG Block Only *(372 Compounds)*	**82.8** *(78.6-86.3)*	**0.00** *(-0.22-0.22)*	**0.00** *(0.00-5.66)*	**100** *(98.8-100)*	-	**82.8** *(78.6-86.3)*

Performance-Related Metrics Calculated For 44 Compounds With Data from Each Ion Channel Assay

Table B.3 Prolongation metric values calculated (with 95% confidence intervals indicated) for evaluation of the assay's predictivity of QT prolongation in the rabbit ventricular wedge experiments when considering the different uses of the available ion channel data described for a subset of 44 compounds for which there were ion channel data available from all three assays. Entries are ordered alphabetically by the data type used in the simulations.

Case Number	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
S1	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value *(44 Compounds)*	**72.7** *(58.2-83.7)*	**0.47** *(0.22-0.73)*	**92.9** *(68.5-98.7)*	**63.3** *(45.5-78.1)*	**54.2** *(35.1-72.1)*	**95.0** *(76.4-99.1)*
S2	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**72.7** *(58.2-83.7)*	**0.47** *(0.22-0.73)*	**92.9** *(68.5-98.7)*	**63.3** *(45.5-78.1)*	**54.2** *(35.1-72.1)*	**95.0** *(76.4-99.1)*
S3	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value *(44 Compounds)*	**81.8** *(68.0-90.5)*	**0.60** *(0.34-0.85)*	**78.6** *(52.4-92.4)*	**83.3** *(66.4-92.7)*	**68.8** *(44.4-85.8)*	**89.3** *(72.8-96.3)*
S4	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.50** *(0.22-0.77)*	**71.4** *(45.4-88.3)*	**80.0** *(62.7-90.5)*	**62.5** *(38.6-81.5)*	**85.7** *(68.5-94.3)*
S5	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value *(44 Compounds)*	**79.5** *(65.5-88.8)*	**0.55** *(0.29-0.81)*	**78.6** *(52.4-92.4)*	**80.0** *(62.7-90.5)*	**64.7** *(41.3-82.7)*	**88.9** *(71.9-96.1)*
S6	IonWorks/FLIPR Data, Multiple Channel Fitting For IC_{50} Value & Hill Coefficient Including KCNQ1 Interactions *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.50** *(0.22-0.77)*	**71.4** *(45.4-88.3)*	**80.0** *(62.7-90.5)*	**62.5** *(38.6-81.5)*	**85.7** *(68.5-94.3)*
S7	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.56** *(0.32-0.80)*	**100** *(78.5-100)*	**66.7** *(48.8-80.8)*	**58.3** (38.8-75.5)	**100** *(83.9-100)*

S8	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**75.0** *(60.6-85.4)*	**0.52** *(0.28-0.77)*	**100** *(78.5-100)*	**63.3** *(45.5-78.1)*	**56.0** *(37.1-73.3)*	**100** *(83.2-100)*
S9	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value *(44 Compounds)*	**70.5** *(55.8-81.8)*	**0.33** *(0.03-0.64)*	**57.1** *(32.6-78.6)*	**76.7** *(59.1-88.2)*	**53.3** *(30.1-75.2)*	**79.3** *(61.6-90.2)*
S10	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**75.0** *(60.6-85.4)*	**0.45** *(0.18-0.73)*	**71.4** *(45.4-88.3)*	**76.7** *(59.1-88.2)*	**58.8** *(36.0-78.4)*	**85.2** *(67.5-94.1)*
S11	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value, One-Dimensional Simulation *(44 Compounds)*	**68.2** *(53.4-80.0)*	**0.38** *(0.12-0.65)*	**85.7** *(60.1-96.0)*	**60.0** *(42.3-75.4)*	**50.0** *(31.4-68.6)*	**90.0** *(69.9-97.2)*
S12	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient, One-Dimensional Simulation *(44 Compounds)*	**63.6** *(48.9-76.2)*	**0.32** *(0.05-0.58)*	**85.7** *(60.1-96.0)*	**53.3** *(36.1-69.8)*	**46.2** *(28.8-64.5)*	**88.9** *(67.2-96.9)*
S13	QSAR Data, hERG Block Only *(44 Compounds)*	**50.0** *(35.8-64.2)*	**0.19** *(-0.05-0.43)*	**100** *(78.5-100)*	**26.7** *(14.2-44.4)*	**38.9** *(24.8-55.1)*	**100** *(67.6-100)*
S14	QSAR Data, Multiple Channel Block *(44 Compounds)*	**54.5** *(40.1-68.3)*	**0.24** *(0.00-0.49)*	**100** *(78.5-100)*	**33.3** *(19.2-51.2)*	**41.2** *(26.4-57.8)*	**100** *(72.2-100)*

Table B.4 Shortening metric values calculated (with 95% confidence intervals indicated) for evaluation of the assay's predictivity for QT shortening in the rabbit ventricular wedge experiments considering the different uses of available ion channel data described for a subset of 44 compounds for which there were data available from all three assays. Entries are ordered alphabetically by the data type used in the simulations. Missing entries are due to the inability to calculate the metrics due to there being no compounds in the relevant categories for their computation.

Case Number	Ion Channel Data Use Combination	Accuracy (%)	Kappa Value	Sensitivity (%)	Specificity (%)	Positive Predictive Value (%)	Negative Predictive Value (%)
S1	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.00** *(-0.54-0.54)*	**0.00** *(0.00-27.8)*	**100** *(89.8-100)*	**-**	**77.3** *(63.0-87.2)*
S2	IonWorks Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.00** *(-0.54-0.54)*	**0.00** *(0.00-27.8)*	**100** *(89.8-100)*	**-**	**77.3** *(63.0-87.2)*
S3	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value *(44 Compounds)*	**70.5** *(55.8-81.8)*	**0.13** *(-0.27-0.53)*	**30.0** *(10.8-60.3)*	**82.4** *(66.5-91.7)*	**33.3** *(15.6-57.6)*	**80.0** *(64.1-90.0)*
S4	IonWorks/FLIPR Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**72.7** *(58.2-83.7)*	**0.27** *(-0.08-0.62)*	**50.0** *(23.7-76.3)*	**79.4** *(63.2-89.7)*	**41.7** *(21.5-65.0)*	**84.4** *(68.2-93.1)*
S5	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value *(44 Compounds)*	**70.5** *(55.8-81.8)*	**0.13** *(-0.27-0.53)*	**30.0** *(10.8-60.3)*	**82.4** *(66.5-91.7)*	**33.3** *(15.9-56.9)*	**80.0** *(64.1-90.0)*
S6	IonWorks/FLIPR Data, Multiple Channel Block Including KCNQ1 Interactions, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**70.5** *(55.8-81.8)*	**0.19** *(-0.18-0.56)*	**40.0** *(16.8-68.7)*	**79.4** *(63.2-89.7)*	**36.4** *(17.7-60.3)*	**81.8** *(65.6-91.4)*
S7	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.00** *(-0.54-0.54)*	**0.00** *(0.00-27.8)*	**100** *(89.8-100)*	**-**	**77.3** *(63.0-87.2)*

S8	PatchXpress Data, hERG Block Only, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.00** *(-0.54-0.54)*	**0.00** *(0.00-27.8)*	**100** *(89.8-100)*	-	**77.3** *(63.0-87.2)*
S9	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value *(44 Compounds)*	**79.5** *(65.5-88.8)*	**0.23** *(-0.22-0.68)*	**20.0** *(5.67-51.0)*	**97.1** *(85.1-99.5)*	**66.7** *(41.7-84.8)*	**80.5** *(66.0-89.8)*
S10	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.18** *(-0.27-0.63)*	**20.0** *(5.67-51.0)*	**94.1** *(80.9-98.4)*	**50.0** *(28.5-71.5)*	**80.0** *(65.2-89.5)*
S11	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value, One-Dimensional Simulation *(44 Compounds)*	**72.7** *(58.2-83.7)*	**0.10** *(-0.34-0.53)*	**20.0** *(5.67-51.0)*	**88.2** *(73.4-95.3)*	**33.3** *(18.0-53.3)*	**78.9** *(63.7-88.9)*
S12	PatchXpress Data, Multiple Channel Block, Fitting For IC_{50} Value & Hill Coefficient, One-Dimensional Simulation *(44 Compounds)*	**75.0** *(60.6-85.4)*	**0.20** *(-0.20-0.61)*	**30.0** *(10.8-60.3)*	**88.2** *(73.4-95.3)*	**42.9** *(26.0-61.6)*	**81.1** *(65.8-90.5)*
S13	QSAR Data, hERG Block Only *(44 Compounds)*	**77.3** *(63.0-87.2)*	**0.00** *(-0.54-0.54)*	**0.00** *(0.00-27.8)*	**100** *(89.8-100)*	-	**77.3** *(63.0-87.2)*
S14	QSAR Data, Multiple Channel Block *(44 Compounds)*	**68.2** *(53.4-80.0)*	**-0.05** *(-0.51-0.40)*	**10.0** *(1.79-40.4)*	**85.3** *(69.9-93.6)*	**16.7** *(7.70-32.4)*	**76.3** *(60.8-87.0)*

C

Appendix

In this appendix we provide further details on some of the numerical aspects of the simulations performed in this thesis, as well as providing some more detail about the search algorithms and MCMC convergence diagnostics used.

C.1 Numerical Methods

C.1.1 CellML

Mathematical models within the CellML repository consist of separate components each describing a particular feature of the system being modelled. For example in the case of the Shannon et al. (2004) ventricular myocyte model described in Chapter 3 and used in Chapter 4, there is a component describing the cell, with the gates of different ionic currents represented within the model being a separate component. Each component contains variables and relationships between variables are expressed as mathematical equations written in MathML. The XML description of the mathematical model in CellML is converted into a programming language before it can be used to perform simulations. As we perform simulations in Chaste which is written in C++, the CellML version of the model is automatically converted to C++ using PyCml (Garny et al., 2008) within Chaste, as we describe in Appendix C.1.2.

C.1.2 Chaste

Chaste (Cancer, Heart and Soft Tissue Environment) is an open-source C++ library (Mirams et al., 2013) developed primarily within the Computational Biology Group in Oxford. Chaste provides a number of methods for performing single cell and tissue level simulations of cardiac electrophysiology. Part of the conversion process from the imported model from CellML to C++ code involves automatically ensuring that units in the CellML model are consistent with those represented within Chaste (milliseconds, millivolts, etc.). Following conversion within Chaste, the model can then be simulated using a choice of solvers. We use the PETSc PDE solvers (Balay et al., 2012) for the one-dimensional tissue simulations in Chapter 4 and CVODE integrated within Chaste for solving ODEs (Hindmarsh et al., 2005). CVODE was also used for performing the simulations described in Chapters 5 and 7 and so we describe this solver in more detail in Section C.1.3.

C.1.3 CVODE

The CVODE library is part of the SUNDIALS suite (Hindmarsh et al., 2005) and contains implementation of different methods for solving ODEs. We use an adaptive method based on backward differentiation formula. This method is particularly suitable for stiff differential equations, a property of both the action potential models (in particular around the upstroke of the action potential) and some hERG channel model simulations (where there are rapid changes in voltage in the experimental protocols). Backwards differentiation formulae are implicit and so require nonlinear equations to be solved at each time step. Newton's method is used to solve the resulting nonlinear system.

C.2 Parameter Optimisation

C.2.1 Simulated Annealing Options

Table C.1 Values for options set in simulated annealing function.

Option	Value
AnnealingFcn	`@annealingfast`
InitialTemperature	150
MaxFunEval	1000000
ReannealInterval	12000
StallIterLimit	10000
TemperatureFcn	`@temperatureboltz`
TolFun	1e-4

C.2.2 Gelman-Rubin Diagnostic

When using m MCMC chains of length n, the between-chain variance (B) and within-chain variance (W) are calculated by

$$W = \frac{1}{m}\sum_{j=1}^{m} {s_j}^2,$$

where

$${s_j}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\theta_{ij} - \bar{\theta_j})^2.$$

$$B = \frac{n}{m-1}\sum_{j=1}^{m}(\bar{\theta_j} - \bar{\bar{\theta}})^2,$$

where

$$\bar{\bar{\theta}} = \frac{1}{m}\sum_{j=1}^{m}\bar{\theta_j}.$$

The estimate of the variance of θ in the target distribution is calculated using

$$Var(\theta) = (1 - \frac{1}{n})W + \frac{1}{n}B.$$

The diagnostic is calculated according to

$$\hat{R} = \sqrt{\frac{Var(\theta)}{W}}.$$

For the multivariate version of the algorithm see Brooks and Gelman (1998).

Bibliography

Abbott, G. W., Sesti, F., Splawski, I., Buck, M. E., Lehmann, M. H., Timothy, K. W., Keating, M. T., and Goldstein, S. A. (1999). MiRP1 forms IKr potassium channels with HERG and is associated with cardiac arrhythmia. *Cell*, 97(2):175–187.

Abi-Gerges, N., Holkham, H., Jones, E., Pollard, C., Valentin, J.-P., and Robertson, G. (2011). hERG subunit composition determines differential drug sensitivity. *British Journal of Pharmacology*, 164(2b):419–432.

Aidley, D. J. (1996). *Ion channels: Molecules in action*. Cambridge University Press.

Akaike, H. (1974). A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716–723.

Andrieu, C. and Thoms, J. (2008). A tutorial on adaptive MCMC. *Statistics and Computing*, 18(4):343–373.

Armstrong, C. M. and Bezanilla, F. (1977). Inactivation of the sodium channel. II. Gating current experiments. *The Journal of General Physiology*, 70(5):567–590.

Asakura, K., Cha, C., Yamaoka, H., Horikawa, Y., Memida, H., Powell, T., Amano, A., and Noma, A. (2014). EAD and DAD mechanisms analyzed by developing a new human ventricular cell model. *Progress in Biophysics and Molecular Biology*, 116(1):11–24.

Aslanidi, O. V., Stewart, P., Boyett, M. R., and Zhang, H. (2009). Optimal velocity and safety of discontinuous conduction through the heterogeneous Purkinje-ventricular junction. *Biophysical Journal*, 97(1):20–39.

Balay, S., Brown, J., Buschelman, K., Eijkhout, V., Gropp, W., Kaushik, D., Knepley, M., McInnes, L., Smith, B., and Zhang, H. (2012). PETSc Users Manual Revision 3.3.

Bardenet, R. (2013). Monte Carlo methods. In *IN2P3 School of Statistics (SOS2012)*, volume 55, page 022002. EDP Sciences.

Beattie, K. A., Luscombe, C., Williams, G., Munoz-Muriedas, J., Gavaghan, D. J., Cui, Y., and Mirams, G. R. (2013). Evaluation of an in silico cardiac safety assay: using ion channel screening data to predict QT interval changes in the rabbit ventricular wedge. *Journal of Pharmacological and Toxicological Methods*, 68(1):88–96.

Beeler, G. W. and Reuter, H. (1977). Reconstruction of the action potential of ventricular myocardial fibres. *The Journal of Physiology*, 268(1):177–210.

Bett, G., Zhou, Q., and Rasmusson, R. (2011). Models of HERG gating. *Biophysical Journal*, 101(3):631–642.

Bezanilla, F. and Armstrong, C. M. (1977). Inactivation of the sodium channel. I. Sodium current experiments. *The Journal of General Physiology*, 70(5):549–566.

Bishop, M., Plank, G., Burton, R., Schneider, J., Gavaghan, D., Grau, V., and Kohl, P. (2010). Development of an anatomically detailed MRI-derived rabbit ventricular model and assessment of its impact on simulations of electrophysiological function. *American Journal of Physiology-Heart and Circulatory Physiology*, 298(2):H699–H718.

Bondarenko, V. E., Szigeti, G. P., Bett, G. C., Kim, S.-J., and Rasmusson, R. L. (2004).

Computer model of action potential of mouse ventricular myocytes. *American Journal of Physiology-Heart and Circulatory Physiology*, 287(3):H1378–H1403.

Bottino, D., Penland, R., Stamps, A., Traebert, M., Dumotier, B., Georgieva, A., Helmlinger, G., and Lett, G. (2006). Preclinical cardiac safety assessment of pharmaceutical compounds using an integrated systems-based computer model of the heart. *Progress in Biophysics and Molecular Biology*, 90(1-3):414–443.

Brennan, T., Fink, M., and Rodriguez, B. (2009). Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *European Journal of Pharmaceutical Sciences*, 36(1):62–77.

Brennan, T., Fink, M., Rodriguez, B., and Tarassenko, L. (2007). Modelling effects of sotalol on action potential morphology using a novel Markov model of the HERG channel. In *11th Mediterranean Conference on Medical and Biomedical Engineering and Computing 2007*, pages 50–53. Springer.

Britton, O. J., Bueno-Orovio, A., Van Ammel, K., Lu, H. R., Towart, R., Gallacher, D. J., and Rodriguez, B. (2013). Experimentally calibrated population of models predicts and explains intersubject variability in cardiac cellular electrophysiology. *Proceedings of the National Academy of Sciences*, 110(23):E2098–E2105.

Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455.

Burbidge, R., Trotter, M., Buxton, B., and Holden, S. (2001). Drug design by machine learning: Support vector machines for pharmaceutical data analysis. *Computers & Chemistry*, 26(1):5–14.

Carmeliet, E. and Vereecke, J. (2001). *Cardiac cellular electrophysiology*, volume 9. Springer.

Carro, J., Rodríguez, J. F., Laguna, P., and Pueyo, E. (2011). A human ventricular cell model for investigation of cardiac arrhythmias under hyperkalaemic conditions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1954):4205–4232.

Clancy, C. and Rudy, Y. (1995). Linking a genetic defect to its cellular phenotype in a cardiac arrhythmia. *Biology*, 9:295–305.

Clancy, C. and Rudy, Y. (2001). Cellular consequences of HERG mutations in the long QT syndrome: Precursors to sudden cardiac death. *Cardiovascular Research*, 50(2):301–313.

Clancy, C. E., Zhu, Z. I., and Rudy, Y. (2007). Pharmacogenetics and anti-arrhythmic drug therapy: a theoretical investigation. *American Journal of Physiology-Heart and Circulatory Physiology*, 292(1):H66–H75.

Clayton, R., Bernus, O., Cherry, E., Dierckx, H., Fenton, F., Mirabella, L., Panfilov, A., Sachse, F., Seemann, G., and Zhang, H. (2011). Models of cardiac tissue electrophysiology: progress, challenges and open questions. *Progress in Biophysics and Molecular Biology*, 104(1):22–48.

Cohen, J. (1968). Weighted Kappa: Nominal scale agreement provision for scaled disagreement or partial credit. *Psychological Bulletin*, 70(4):213.

Colquhoun, D., Dowsland, K. A., Beato, M., and Plested, A. J. (2004). How to impose microscopic reversibility in complex reaction mechanisms. *Biophysical Journal*, 86(6):3510–3518.

Cooper, J., Spiteri, R. J., and Mirams, G. R. (2014). Cellular cardiac electrophysiology modeling with chaste and cellml. *Frontiers in Physiology*, 5.

Courtemanche, M., Ramirez, R., and Nattel, S. (1998). Ionic mechanisms underlying human atrial action potential properties: Insights from a mathematical model. *American Journal of Physiology-Heart and Circulatory Physiology*, 275(1):H301–H321.

Darpo, B. (2010). The Thorough QT/QTc study 4 years after the implementation of the ICH E14 guidance. *British Journal of Pharmacology*, 159(1):49–57.

Davies, M. R., Mistry, H. B., Hussein, L., Pollard, C. E., Valentin, J.-P., Swinton, J., and Abi-Gerges, N. (2012). An in silico canine cardiac midmyocardial action potential duration model as a tool for early drug safety assessment. *American Journal of Physiology-Heart and Circulatory Physiology*, 302(7):H1466–H1480.

De Ponti, F. (2008). Pharmacological and regulatory aspects of QT prolongation. *Antitargets: Prediction and Prevention of Drug Side Effects*, pages 53–88.

Decker, K. F., Heijman, J., Silva, J. R., Hund, T. J., and Rudy, Y. (2009). Properties and ionic mechanisms of action potential adaptation, restitution, and accommodation in canine epicardium. *American Journal of Physiology-Heart and Circulatory Physiology*, 296(4):H1017–H1026.

Di Veroli, G. Y., Davies, M. R., Zhang, H., Abi-Gerges, N., and Boyett, M. R. (2013). High-throughput screening of drug-binding dynamics to herg improves early drug safety assessment. *American Journal of Physiology-Heart and Circulatory Physiology*, 304(1):H104–H117.

Di Veroli, G. Y., Davies, M. R., Zhang, H., Abi-Gerges, N., and Boyett, M. R. (2014). hERG inhibitors with similar potency but different binding kinetics do not pose the same proarrhythmic risk: Implications for drug safety assessment. *Journal of Cardiovascular Electrophysiology*, 25(2):197–207.

Elkins, R. C., Davies, M. R., Brough, S. J., Gavaghan, D. J., Cui, Y., Abi-Gerges, N., and Mirams, G. R. (2013). Variability in high-throughput ion-channel screening data and consequences for cardiac safety assessment. *Journal of Pharmacological and Toxicological Methods*, 68(1):112–122.

Eyring, H. (1935). The activated complex in chemical reactions. *The Journal of Chemical Physics*, 3(2):107–115.

Faber, G. M., Silva, J., Livshitz, L., and Rudy, Y. (2007). Kinetic properties of the cardiac L-Type Ca2+ channel and its role in myocyte electrophysiology: A theoretical investigation. *Biophysical Journal*, 92(5):1522–1543.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874.

Fermini, B. and Priest, B. (2008). *Ion channels*, volume 3. Springer Science & Business Media.

Ficker, E., Jarolimek, W., Kiehn, J., Baumann, A., and Brown, A. M. (1998). Molecular determinants of dofetilide block of HERG K+ channels. *Circulation Research*, 82(3):386–395.

Fink, M. and Noble, D. (2009). Markov models for ion channels: Versatility versus identifiability and speed. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1896):2161–2179.

Fink, M., Noble, D., Virag, L., Varro, A., and Giles, W. R. (2008). Contributions of HERG K+ current to repolarization of the human ventricular action potential. *Progress in Biophysics and Molecular Biology*, 96(1):357–376.

Fox, J. J., McHarg, J. L., and Gilmour Jr, R. F. (2002). Ionic mechanism of electrical alternans. *American Journal of Physiology-Heart and Circulatory Physiology*, 282(2):H516–H530.

Garny, A., Nickerson, D., Cooper, J., dos Santos, R., Miller, A., McKeever, S., Nielsen, P., and Hunter, P. (2008). CellML and associated tools and techniques. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1878):3017–3043.

Gelman, A., Roberts, G., and Gilks, W. (1996). Efficient Metropolis jumping rules. *Bayesian Statistics*, 5(599-608):42.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, pages 457–472.

Genschow, E., Spielmann, H., Scholz, G., Seiler, A., Brown, N., Piersma, A., Brady, M., Clemann, N., Huuskonen, H., Paillard, F., et al. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: Results of the definitive phase and evaluation of prediction models. European Centre for the Validation of Alternative Methods. *Alternatives to Laboratory Animals: ATLA*, 30(2):151.

Gima, K. and Rudy, Y. (2002). Ionic current basis of electrocardiographic waveforms. *Circulation Research*, 90(8):889–896.

Grandi, E., Pasqualini, F. S., and Bers, D. M. (2010). A novel computational model of the human ventricular action potential and Ca transient. *Journal of Molecular and Cellular Cardiology*, 48(1):112–121.

Grant, A. O. (2009). Cardiac ion channels. *Circulation: Arrhythmia and Electrophysiology*, 2(2):185–194.

Gurkiewicz, M. and Korngreen, A. (2007). A numerical approach to ion channel modelling using whole-cell voltage-clamp recordings and a genetic algorithm. *PLoS Computational Biology*, 3(8):e169.

Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*, pages 223–242.

Hamill, O. P., Marty, A., Neher, E., Sakmann, B., and Sigworth, F. (1981). Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv*, 391(2):85–100.

Hansen, N., Müller, S., and Koumoutsakos, P. (2003). Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11(1):1–18.

Hill, A. P., Perrin, M. J., Heide, J., Campbell, T. J., Mann, S. A., and Vandenberg, J. I. (2014). Kinetics of drug interaction with the Kv11.1 potassium channel. *Molecular Pharmacology*, 85(5):769–776.

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *Journal of Physiology (London)*, 40:4–7.

Hille, B. (1977). Local anesthetics: hydrophilic and hydrophobic pathways for the drug-receptor reaction. *The Journal of General Physiology*, 69(4):497–515.

Hille, B. et al. (2001). *Ion channels of excitable membranes*, volume 507. Sinauer Sunderland, MA.

Hindmarsh, A. C., Brown, P. N., Grant, K. E., Lee, S. L., Serban, R., Shumaker, D. E., and Woodward, C. S. (2005). Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software (TOMS)*, 31(3):363–396.

Hodgkin, A. and Huxley, A. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117(4):500.

Hund, T. J. and Rudy, Y. (2004). Rate dependence and regulation of action potential and calcium transient in a canine cardiac ventricular cell model. *Circulation*, 110(20):3168–3174.

Inada, S., Hancox, J., Zhang, H., and Boyett, M. (2009). One-dimensional mathematical model of the atrioventricular node including atrio-nodal, nodal, and nodal-his cells. *Biophysical Journal*, 97(8):2117–2127.

Inanobe, A., Kamiya, N., Murakami, S., Fukunishi, Y., Nakamura, H., and Kurachi, Y. (2008). In silico prediction of the chemical block of human ether-a-go-go-related gene (hERG) K+ current. *The Journal of Physiological Sciences: JPS*, 58(7):459–470.

Iyer, V., Mazhari, R., and Winslow, R. (2004). A computational model of the human left-ventricular epicardial myocyte. *Biophysical Journal*, 87(3):1507–1525.

Johnson, J., Mullins, F. M., and Bennett, P. B. (1999). Human ether-à-go-go–related gene K+ channel gating probed with extracellular Ca2+ evidence for two distinct voltage sensors. *The Journal of General Physiology*, 113(4):565–580.

Jones, E. M., Roti, E. C. R., Wang, J., Delfosse, S. A., and Robertson, G. A. (2004). Cardiac IKr channels minimally comprise hERG 1a and 1b subunits. *Journal of Biological Chemistry*, 279(43):44690–44694.

Kargol, A., Hosein-Sooklal, A., Constantin, L., and Przestalski, M. (2004). Application of oscillating potentials to the shaker potassium channel. *General Physiology and Biophysics*, 23:53–76.

Katz, A. M. (2010). *Physiology of the heart*. Lippincott Williams & Wilkins.

Keener, J. and Sneyd, J. (2010). *Mathematical physiology: I: Cellular physiology*, volume 1. Springer.

Kiehn, J., Lacerda, A., and Brown, A. (1999). Pathways of HERG inactivation. *American Journal of Physiology-Heart and Circulatory Physiology*, 277(1):H199–H210.

Kirsch, G. E., Trepakova, E. S., Brimecombe, J. C., Sidach, S. S., Erickson, H. D., Kochan, M. C., Shyjka, L. M., Lacerda, A. E., and Brown, A. M. (2004). Variability in the measurement of hERG potassium channel inhibition: effects of temperature and stimulus pattern. *Journal of Pharmacological and Toxicological Methods*, 50(2):93–101.

Kohl, P., Sachs, F., and Franz, M. R. (2011). *Cardiac mechano-electric coupling and arrhythmias*. Oxford University Press.

Kurata, Y., Hisatome, I., Imanishi, S., and Shibamoto, T. (2002). Dynamical description of sinoatrial node pacemaking: improved mathematical model for primary pacemaker cell. *American Journal of Physiology-Heart and Circulatory Physiology*, 283(5):H2074–H2101.

Lagarias, J. C., Reeds, J. A., Wright, M. H., and Wright, P. E. (1998). Convergence properties of the nelder–mead simplex method in low dimensions. *SIAM Journal on Optimization*, 9(1):112–147.

Laverty, H., Benson, C., Cartwright, E., Cross, M., Garland, C., Hammond, T., Holloway, C., McMahon, N., Milligan, J., Park, B., et al. (2011). How can we improve our understanding of cardiovascular safety liabilities to develop safer medicines? *British Journal of Pharmacology*, 163(4):675–693.

Li, L., Niederer, S., Idigo, W., Zhang, Y., Swietach, P., Casadei, B., and Smith, N. (2010). A mathematical model of the murine ventricular myocyte: a data-driven biophysically based approach applied to mice overexpressing the canine NCX isoform. *American Journal of Physiology-Heart and Circulatory Physiology*, 299(4):H1045–H1063.

Lindblad, D., Murphey, C., Clark, J., and Giles, W. (1996). A model of the action potential and underlying membrane currents in a rabbit atrial cell. *American Journal of Physiology*, 271:H1666–H1696.

Liu, S., Rasmusson, R., Campbell, D., Wang, S., and Strauss, H. (1996). Activation and inactivation kinetics of an E-4031-sensitive current from single ferret atrial myocytes. *Biophysical Journal*, 70(6):2704–2715.

Liu, T., Brown, B., Wu, Y., Antzelevitch, C., Kowey, P., and Yan, G. (2006). Blinded validation of the isolated arterially perfused rabbit ventricular wedge in preclinical assessment of drug-induced proarrhythmias. *Heart Rhythm*, 3(8):948–956.

Lloyd, C., Lawson, J., Hunter, P., and Nielsen, P. (2008). The CellML model repository. *Bioinformatics*, 24(18):2122–2123.

Lu, Y., Mahaut-Smith, M., Varghese, A., Huang, C., Kemp, P., and Vandenberg, J. (2001). Effects of premature stimulation on HERG K+ channels. *The Journal of Physiology*, 537(3):843–851.

Lundquist, A. L., Manderfield, L. J., Vanoye, C. G., Rogers, C. S., Donahue, B. S., Chang, P. A., Drinkwater, D. C., Murray, K. T., and George Jr, A. L. (2005). Expression of multiple KCNE genes in human heart may enable variable modulation of IKs. *Journal of Molecular and Cellular Cardiology*, 38(2):277–287.

Luo, C. and Rudy, Y. (1991). A model of the ventricular cardiac action potential. Depolarization, repolarization, and their interaction. *Circulation Research*, 68(6):1501–1526.

Luo, C. and Rudy, Y. (1994). A dynamic model of the cardiac ventricular action potential. I. simulations of ionic currents and concentration changes. *Circulation Research*, 74(6):1071–1096.

Mahajan, A., Shiferaw, Y., Sato, D., Baher, A., Olcese, R., Xie, L.-H., Yang, M.-J., Chen, P.-S., Restrepo, J. G., Karma, A., et al. (2008). A rabbit ventricular action potential model replicating cardiac dynamics at rapid heart rates. *Biophysical Journal*, 94(2):392–410.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics*, 11(2):431–441.

MatLab, M. (2012). The language of technical computing. *The MathWorks, Inc. http://www. mathworks. com*.

Matsuoka, S., Sarai, N., Kuratomi, S., Ono, K., and Noma, A. (2003). Role of individual ionic current systems in ventricular cells hypothesized by a model study. *The Japanese Journal of Physiology*, 53(2):105–123.

Mazhari, R., Greenstein, J., Winslow, R., Marbán, E., and Nuss, H. (2001). Molecular interactions between two Long-QT syndrome gene products, HERG and KCNE2, rationalized by in vitro and in silico analysis. *Circulation Research*, 89(1):33–38.

McAllister, R. E., Noble, D., and Tsien, R. (1975). Reconstruction of the electrical activity of cardiac Purkinje fibres. *The Journal of Physiology*, 251(1):1–59.

Milescu, L., Akk, G., and Sachs, F. (2005). Maximum likelihood estimation of ion channel kinetics from macroscopic currents. *Biophysical Journal*, 88(4):2494–2515.

Mirams, G., Cui, Y., Sher, A., Fink, M., Cooper, J., Heath, B., McMahon, N., Gavaghan, D., and Noble, D. (2011). Simulation of multiple ion channel block provides improved early prediction of compounds clinical torsadogenic risk. *Cardiovascular Research*, 91(1):53.

Mirams, G. R., Arthurs, C. J., Bernabeu, M. O., Bordas, R., Cooper, J., Corrias, A., Davit, Y., Dunn, S.-J., Fletcher, A. G., Harvey, D. G., Marsh, M. E., Osborne, J. M., Pathmanathan, P., Southern, J., Zemzemi, N., and Gavaghan, D. J. (2013). Chaste: An open source

C++ library for computational physiology and biology. *PLoS Computational Biology*, 9(3):e1002970.

Mirams, G. R., Davies, M. R., Brough, S. J., Bridgland-Taylor, M. H., Cui, Y., Gavaghan, D. J., and Abi-Gerges, N. (2014). Prediction of Thorough QT study results using action potential simulations based on ion channel screens. *Journal of Pharmacological and Toxicological Methods*.

Mirams, G. R., Davies, M. R., Cui, Y., Kohl, P., and Noble, D. (2012). Application of cardiac electrophysiology simulations to pro-arrhythmic safety testing. *British Journal of Pharmacology*, 167(5):932–945.

Mitcheson, J. S., Chen, J., and Sanguinetti, M. C. (2000). Trapping of a methanesulfonanilide by closure of the HERG potassium channel activation gate. *The Journal of General Physiology*, 115(3):229–240.

Moffatt, L. (2007). Estimation of ion channel kinetics from fluctuations of macroscopic currents. *Biophysical Journal*, 93(1):74–91.

Molecular-Devices (2008). *The axon guide, a guide to electrophysiology biophysics laboratory techniques*. Technical Report, Analytical Technologies, http://files. axon. com/downloads/manuals/Axon_Guide. pdf.

Molleman, A. (2003). *Patch clamping: An introductory guide to patch clamp electrophysiology*. John Wiley & Sons.

Moreno, J. D., Zhu, Z. I., Yang, P.-C., Bankston, J. R., Jeng, M.-T., Kang, C., Wang, L., Bayer, J. D., Christini, D. J., Trayanova, N. A., Ripplinger, C. M., Kass, R. S., and Clancy, C. E. (2011). A computational model to predict the effects of class I anti-arrhythmic drugs on ventricular rhythms. *Science Translational Medicine*, 3(98):98ra83–98ra83.

Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C., and Greyson, D. (2011). The cost of drug development: a systematic review. *Health Policy*, 100(1):4–17.

Näbauer, M. and Kääb, S. (1998). Potassium channel down-regulation in heart failure. *Cardiovascular Research*, 37(2):324–334.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, 7(4):308–313.

Nernst, W. (1888). Zur kinetik der in loesung befindlichen koerper.

Niederer, S., Fink, M., Noble, D., and Smith, N. (2009). A meta-analysis of cardiac electrophysiology computational models. *Experimental Physiology*, 94(5):486–495.

Noble, D. (1960). Cardiac action and pacemaker potentials based on the Hodgkin-Huxley equations. *Nature*, 188(4749):495–497.

Noble, D. (1962). A modification of the HodgkinHuxley equations applicable to Purkinje fibre action and pacemaker potentials. *The Journal of Physiology*, 160(2):317–352.

Noble, D., Garny, A., and Noble, P. J. (2012). How the Hodgkin–Huxley equations inspired the cardiac Physiome Project. *The Journal of Physiology*, 590(11):2613–2628.

Noble, D. and Noble, P. (2006). Late sodium current in the pathophysiology of cardiovascular disease: consequences of sodium–calcium overload. *Heart*, 92(suppl 4):iv1–iv5.

Noble, D., Noble, S., Bett, G., Earm, Y., Ho, W., and So, I. (1991). The role of sodium-calcium exchange during the cardiac action potential. *Annals of the New York Academy of Sciences*, 639(1):334–353.

Noble, D., Varghese, A., Kohl, P., and Noble, P. (1998). Improved guinea-pig ventricular cell model incorporating a diadic space, IKr and IKs, and length-and tension-dependent processes. *The Canadian Journal of Cardiology*, 14(1):123–134.

Nygren, A., Fiset, C., Firek, L., Clark, J., Lindblad, D., Clark, R., and Giles, W. (1998). Mathematical model of an adult human atrial cell the role of K+ currents in repolarization. *Circulation Research*, 82(1):63–81.

Obiol-Pardo, C., Gomis-Tena, J., Sanz, F., Saiz, J., and Pastor, M. (2011). A multiscale simulation system for the prediction of drug-induced cardiotoxicity. *Journal of Chemical Information and Modeling*, 51(2):483.

Oehmen, C., Giles, W., and Demir, S. (2002). Mathematical model of the rapidly activating delayed rectifier potassium current IKr in rabbit sinoatrial node. *Journal of Cardiovascular Electrophysiology*, 13(11):1131–1140.

O'Hara, T., Virág, L., Varró, A., and Rudy, Y. (2011). Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLoS Computational Biology*, 7(5):e1002061.

Pandit, S. V., Clark, R. B., Giles, W. R., and Demir, S. S. (2001). A mathematical model of action potential heterogeneity in adult rat left ventricular myocytes. *Biophysical Journal*, 81(6):3029–3051.

Perrin, M. J., Kuchel, P. W., Campbell, T. J., and Vandenberg, J. I. (2008a). Drug binding to the inactivated state is necessary but not sufficient for high-affinity binding to human ether-a-go-go-related gene channels. *Molecular Pharmacology*, 74(5):1443–1452.

Perrin, M. J., Subbiah, R. N., Vandenberg, J. I., and Hill, A. P. (2008b). Human ether-a-go-go related gene (hERG) K+ channels: Function and dysfunction. *Progress in Biophysics and Molecular Biology*, 98(2):137–148.

Picard, S., Goineau, S., Guillaume, P., Henry, J., Hanouz, J., and Rouet, R. (2011). Supplemental studies for cardiovascular risk assessment in safety pharmacology: A critical overview. *Cardiovascular Toxicology*, pages 1–23.

Piper, D., Varghese, A., Sanguinetti, M., and Tristani-Firouzi, M. (2003). Gating currents associated with intramembrane charge displacement in HERG potassium channels. *Proceedings of the National Academy of Sciences*, 100(18):10534.

Pollard, C., Abi Gerges, N., Bridgland-Taylor, M., Easter, A., Hammond, T., and Valentin, J. (2010). An introduction to QT interval prolongation and non-clinical approaches to assessing and reducing risk. *British Journal of Pharmacology*, 159(1):12–21.

Pourrier, M., Zicha, S., Ehrlich, J., Han, W., and Nattel, S. (2003). Canine ventricular KCNE2 expression resides predominantly in Purkinje fibers. *Circulation Research*, 93(3):189–191.

Priebe, L. and Beuckelmann, D. (1998). Simulation study of cellular electric properties in heart failure. *Circulation Research*, 82(11):1206–1223.

Puglisi, J. L. and Bers, D. M. (2001). LabHEART: an interactive computer model of rabbit ventricular myocyte ion channels and Ca transport. *American Journal of Physiology-Cell Physiology*, 281(6):C2049–C2060.

Ramirez, R. J., Nattel, S., and Courtemanche, M. (2000). Mathematical analysis of canine atrial action potentials: rate, regional factors, and electrical remodeling. *American Journal of Physiology-Heart and Circulatory Physiology*, 279(4):H1767–H1785.

Redfern, W., Carlsson, L., Davis, A., Lynch, W., MacKenzie, I., Palethorpe, S., Siegl, P., Strang, I., Sullivan, A., Wallis, R., et al. (2003). Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: Evidence for a provisional safety margin in drug development. *Cardiovascular Research*, 58(1):32.

Robert, C. and Casella, G. (2013). *Monte Carlo statistical methods*. Springer Science & Business Media.

Romero, L., Carbonell, B., Trenor, B., Rodríguez, B., Saiz, J., and Ferrero, J. M. (2011). Systematic characterization of the ionic basis of rabbit cellular electrophysiology using two ventricular models. *Progress in Biophysics and Molecular Biology*, 107(1):60–73.

Romero, L., Trenor, B., Yang, P.-C., Saiz, J., and Clancy, C. E. (2014). In silico screening of the impact of hERG channel kinetic abnormalities on channel block and susceptibility to acquired long QT syndrome. *Journal of Molecular and Cellular Cardiology*, 72:126–137.

Sager, P., Gintant, G., Turner, J., Pettit, S., and Stockbridge, N. (2014). Rechanneling the cardiac proarrhythmia safety paradigm: a meeting report from the cardiac safety research consortium. *American Heart Journal*, 167(3):292–300.

Saiz, J., Gomis-Tena, J., Monserrat, M., Ferrero, J. M., Cardona, K., and Chorro, J. (2011). Effects of the antiarrhythmic drug dofetilide on transmural dispersion of repolarization in ventriculum. A computer modeling study. *Biomedical Engineering, IEEE Transactions on*, 58(1):43–53.

Sakmann, B. and Neher, E. (2009). *Single-channel recording*. Springer Science & Business Media.

Sale, H., Wang, J., O'Hara, T. J., Tester, D. J., Phartiyal, P., He, J.-Q., Rudy, Y., Ackerman, M. J., and Robertson, G. A. (2008). Physiological properties of hERG 1a/1b heteromeric currents and a hERG 1b-specific mutation associated with Long-QT syndrome. *Circulation Research*, 103(7):e81–e95.

Sanguinetti, M. C. and Mitcheson, J. S. (2005). Predicting drug–hERG channel interactions that cause acquired long QT syndrome. *Trends in Pharmacological Sciences*, 26(3):119–124.

Sanguinetti, M. C. and Tristani-Firouzi, M. (2006). hERG potassium channels and cardiac arrhythmia. *Nature*, 440(7083):463–469.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Seemann, G., Sachse, F. B., Weiß, D., and Dössel, O. (2003). Quantitative reconstruction of cardiac electromechanics in human myocardium. *Journal of Cardiovascular Electrophysiology*, 14(s10):S219–S228.

Severi, S., Fantini, M., Charawi, L. A., and DiFrancesco, D. (2012). An updated computational model of rabbit sinoatrial action potential to investigate the mechanisms of heart rate modulation. *The Journal of Physiology*, 590(18):4483–4499.

Shah, R. R. (2013). Drug-induced QT interval prolongation: does ethnicity of the thorough QT study population matter? *British Journal of Clinical Pharmacology*, 75(2):347–358.

Shannon, T., Wang, F., Puglisi, J., Weber, C., and Bers, D. (2004). A mathematical treatment of integrated Ca dynamics within the ventricular myocyte. *Biophysical Journal*, 87(5):3351–3371.

Sher, A. A., Bond, A. M., Gavaghan, D. J., Harriman, K., Feldberg, S. W., Duffy, N. W., Guo, S.-X., and Zhang, J. (2004). Resistance, capacitance, and electrode kinetic effects in fourier-transformed large-amplitude sinusoidal voltammetry: Emergence of powerful and intuitively obvious tools for recognition of patterns of behavior. *Analytical Chemistry*, 76(21):6214–6228.

Shibasaki, T. (1987). Conductance and kinetics of delayed rectifier potassium channels in nodal cells of the rabbit heart. *The Journal of Physiology*, 387(1):227–250.

Sigg, D. C., Iaizzo, P. A., Xiao, Y.-F., and He, B. (2010). *Cardiac electrophysiology methods and models*. Springer.

Spector, P., Curran, M., Zou, A., Keating, M., and Sanguinetti, M. (1996). Fast inactivation causes rectification of the IKr channel. *The Journal of General Physiology*, 107(5):611–619.

Starmer, C., Grant, A., and Strauss, H. (1984). Mechanisms of use-dependent block of sodium channels in excitable membranes by local anesthetics. *Biophysical Journal*, 46(1):15–27.

Starmer, C., Ramasubba Reddy, M., Namasivayam, A., and Singh, M. (1994). Potassium channel blockade amplifies cardiac instability: Numerical studies of torsades de points. *Indian Journal of Physiology and Pharmacology*, 38:259–259.

Starmer, C., Romashko, D., Reddy, R., Zilberter, Y., Starobin, J., Grant, A., and Krinsky, V. (1995). Proarrhythmic response to potassium channel blockade numerical studies of polymorphic tachyarrhythmias. *Circulation*, 92(3):595–605.

Sundnes, J., Lines, G. T., Cai, X., Nielsen, B. F., Mardal, K.-A., and Tveito, A. (2007). *Computing the electrical activity in the heart*, volume 1. Springer.

Svetnik, V., Liaw, A., Tong, C., Culberson, J., Sheridan, R., and Feuston, B. (2003). Random forest: A classification and regression tool for compound classification and QSAR modeling. *Journal of Chemical Information and Computer Sciences*, 43(6):1947–1958.

Ten Tusscher, K., Noble, D., Noble, P., and Panfilov, A. (2004). A model for human ventricular tissue. *American Journal of Physiology-Heart and Circulatory Physiology*, 286(4):H1573–H1589.

ten Tusscher, K. H. and Panfilov, A. V. (2006). Alternans and spiral breakup in a human ventricular tissue model. *American Journal of Physiology-Heart and Circulatory Physiology*, 291(3):H1088–H1100.

Tolman, R. C. (1938). *The principles of statistical mechanics*. Courier Dover Publications.

Valentin, J. (2010). Reducing QT liability and proarrhythmic risk in drug discovery and development. *British Journal of Pharmacology*, 159(1):5–11.

Vandenberg, J., Perry, M., Perrin, M., Mann, S., Ke, Y., and Hill, A. (2012). hERG K+ channels: Structure, function, and clinical significance. *Physiological Reviews*, 92(3):1393–1478.

Vandenberg, J. I., Varghese, A., Lu, Y., Bursill, J. A., Mahaut-Smith, M. P., and Huang, C. L.-H. (2006). Temperature dependence of human ether-a-go-go-related gene K+ currents. *American Journal of Physiology-Cell Physiology*, 291(1):C165–C175.

Vecchietti, S., Rivolta, I., Severi, S., Napolitano, C., Priori, S. G., and Cavalcanti, S. (2006). Computer simulation of wild-type and mutant human cardiac Na+ current. *Medical and Biological Engineering and Computing*, 44(1-2):35–44.

Wang, L. J. and Sobie, E. A. (2008). Mathematical model of the neonatal mouse ventricular action potential. *American Journal of Physiology-Heart and Circulatory Physiology*, 294(6):H2565–H2575.

Wang, S., Liu, S., Morales, M., Strauss, H., and Rasmusson, R. (1997). A quantitative analysis of the activation and inactivation kinetics of HERG expressed in Xenopus oocytes. *The Journal of Physiology*, 502(1):45–60.

Wang, W., Xiao, F., Zeng, X., Yao, J., Yuchi, M., and Ding, J. (2012). Optimal estimation of ion-channel kinetics from macroscopic currents. *PLoS One*, 7(4):e35208.

Weerapura, M., Nattel, S., Chartier, D., Caballero, R., and Hébert, T. E. (2002). A comparison of currents carried by HERG, with and without coexpression of MiRP1, and the native rapid delayed rectifier current. Is MiRP1 the missing link? *The Journal of Physiology*, 540(1):15–27.

Wilson, E. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158):209–212.

Winslow, R. L., Rice, J., Jafri, S., Marban, E., and O'Rourke, B. (1999). Mechanisms of altered excitation-contraction coupling in canine tachycardia-induced heart failure, II Model studies. *Circulation Research*, 84(5):571–586.

Zaza, A. and Rosen, M. (2000). *An introduction to cardiac electrophysiology*. Taylor & Francis.

Zemzemi, N., Bernabeu, M. O., Saiz, J., Cooper, J., Pathmanathan, P., Mirams, G. R., Pitt-Francis, J., and Rodriguez, B. (2013). Computational assessment of drug-induced effects on the electrocardiogram: from ion channel to body surface potentials. *British Journal of Pharmacology*, 168(3):718–733.

Zeng, J., Laurita, K., Rosenbaum, D., and Rudy, Y. (1995). Two components of the delayed rectifier K+ current in ventricular myocytes of the guinea pig type: Theoretical formulation and their role in repolarization. *Circulation Research*, 77(1):140–152.

Zhang, H., Holden, A., Kodama, I., Honjo, H., Lei, M., Varghese, T., and Boyett, M. (2000). Mathematical models of action potentials in the periphery and center of the rabbit sinoatrial node. *American Journal of Physiology-Heart and Circulatory Physiology*, 279(1):H397–H421.

Zhou, Q., Zygmunt, A., Cordeiro, J., Siso-Nadal, F., Miller, R., Buzzard, G., and Fox, J. (2009). Identification of IKr kinetics and drug binding in native myocytes. *Annals of Biomedical Engineering*, 37(7):1294–1309.

Zhou, Z., Gong, Q., Ye, B., Fan, Z., Makielski, J., Robertson, G., and January, C. (1998). Properties of HERG channels stably expressed in HEK 293 cells studied at physiological temperature. *Biophysical Journal*, 74(1):230–241.

Zou, A., Curran, M. E., Keating, M. T., and Sanguinetti, M. C. (1997). Single HERG delayed rectifier K+ channels expressed in Xenopus oocytes. *American Journal of Physiology*, 272(3 Pt 2):H1309–H1314.