Computational modelling of ion channel kinetics under drug intervention on cardiac myocytes



Hui Jia Farm
Green Templeton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy

Trinity 2025

Supervisors:

Prof. David J. Gavaghan
Dr. Michael Clerx
Dr. Fergus Cooper
Dr. Chon Lok Lei

Industrial supervisors:

Dr. Liudmila Polonchuk
Dr. Ken Wang

This thesis started during Michaelmas 2020, the year when the global outbreak of the COVID-19 pandemic started. At the start of this D.Phil., I was affected by the travel restrictions and worked remotely in Malaysia and Singapore for the first half year of my D.Phil. I worked on two side projects on COVID-19 that turned into two publications, which I am the co-first author on both:

1. S. A. van der Vegt, L. Dai, I. Bouros, H. J. Farm, R. Creswell, O. Dimdore-Miles, I. Cazimoglu, S. Bajaj, L. Hopkins, D. Seiferth, F. Cooper, C. L. Lei, D. Gavaghan, and B. Lambert. Learning transmission dynamics modelling of COVID-19 using comomodels. *Mathematical Biosciences*, 349:108824, 2022. doi: 10.1016/j.mbs.2022.108824.

2. R. Creswell, D. Augustin, I. Bouros, H. J. Farm, S. Miao, A. Ahern, M. Robinson, A. Lemenuel-Diot, D. J. Gavaghan, B. C. Lambert, and R. N. Thompson. Heterogeneity in the onward transmission risk between local and imported cases affects practical estimates of the time-dependent reproduction number. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 380: 20210308, 2022. doi: 10.1098/rsta.2021.0308.

Acknowledgements

I would like to express my utmost gratitude to my supervisors for their support and guidance, drawing from your expertise and your experience in life. To Dave, for the many insights and advice when I was finding my way; to Chon, for the support you have given as someone who had recently gone through the same journey; to Michael and Fergus, for your guidance on presentation, both in words or in figures; to Ken and Liudmila, for helping me relate my work back to the biological context. I am especially grateful for Dave's advice to be patient with the outcome, to be brutally honest in presenting results, and to be confident. All lessons that extend far beyond research. I am also grateful for the insightful discussions with Gary. I deeply apprieciate the support, expertise and mentorship I received throughout this journey, which have allowed me to hone my skills and contributed to my holistic growth in areas I had not anticipated, and without which this DPhil journey would have been different and undoubtedly more challenging. To everyone in the research group, thank you for creating a conducive environment for research and discussions. In particular, Richard for the encouragement and help on several occasions.

To my family, thank you for being supportive of my decision in whatever ways you could, for believing in me and for the good food whenever I am back as your way of showing love.

I thank all the friends I have come to know, who brought balance to this journey and made it far more fulfilling than I ever imagined it to be. May, a friend from home, who was always there when I needed it, who respected my boundaries, and with whom I shared many memorable trips. The squash gang, including Yun (my squash buddy) and Hongshan (my lunch buddy), who gave me something to look forward to every week, both the squash sessions and the conversations afterwards. Lisha, Lucrezia, Mark and Ken, who have kept me in the sport and made it my go-to hobby, one that I will continue in the years to come. I am grateful and will forever cherish the connections we fostered together.

I would also like to express my appreciation to Green Templeton College for providing a welcoming environment and community that allowed me to relax and recharge, whether through squash, the wonderful meals prepared by the kitchen team, or the kindness of the porters. In particular, Richard for his listening ear and for always checking in on me during the final stages of the writing-up phase. I will greatly miss the late-night conversations we shared.

Abstract

The pumping of the human heart is driven by the rhythmic contraction of the heart's muscle cells, the cardiomyocytes. The synchronised contraction of these cardiomyocytes relies on the dynamic flow of ions across the membrane. These ionic flows change the voltage potential across the cell membrane, which forms the action potential. The current conducting through the human Ether-à-go-go-Related Gene (hERG) ion channel repolarises the membrane potential. The inhibition of this current prolongs the action potential and increases the proarrhythmic risk. Since the hERG channel is highly susceptible to blocking by drugs, the degree of this blockage is crucial in assessing the proarrhythmic risk of drugs. However, the drug-binding process and its inhibitory effect on the hERG current are complex and poorly understood. Many researchers have developed mathematical models to capture the channel behaviour under drug blockage and to provide proarrhythmic risk predictions of drugs, but these developments have not led to satisfactory prediction outcomes. The Li model of drug-binding within the hERG channel [Li et al., 2017b], developed by research teams in the FDA, incorporated drug-trapping and improved the proarrhythmic risk predictions. This formulation performs as a benchmark for risk prediction models under the CiPA paradigm. It is therefore essential to ensure that the predictions of the Li drug-binding formulation are robust and reliable.

We first perform a sensitivity analysis on the Li drug-binding formulation to investigate the contribution of each parameter to the proarrhythmic risk predicted by the model. We then change the underlying cell models of the Li drug-binding formulation and repeat the analysis. The lack of contribution of the governing parameters of this formulation to model prediction suggests that these parameters cannot be accurately elucidated from experimental data. This is investigated by an identifiability analysis. Utilising the identifiability analysis results, we simplify the formulation by removing the non-identifiable parameter and compare the risk categorisation of the simplified formulations to the Li drug-binding formulation. We show that the simplified formulation is able to provide the same risk classification as the Li drug-binding formulation for drugs from a reference dataset. We end by discussing the importance of this type of in-depth analysis and model evaluation in ensuring the credibility of model predictions.

Contents

1 Introduction **1**

1.1 Thesis outline . . . 3

2 Background **4**

2.1 Electrophysiology of the Human Heart . . . 4

2.1.1 Cellular electrophysiology and the action potential . . . 7

2.1.2 Contribution of ionic currents to AP . . . 8

2.2 Mathematical Modelling of Cellular Electrophysiology . . . 8

2.2.1 The cell membrane . . . 8

2.2.2 The ionic currents . . . 9

2.2.2.1 The driving term . . . 9

2.2.2.2 The open probability . . . 10

2.3 Experimental Measurements to Mathematical Models . . . 12

2.3.1 Experimental techniques . . . 12

2.3.2 Forward simulation . . . 14

2.3.3 Inverse problem . . . 15

2.3.4 Computational methods for the optimisation problem . . . 16

2.4 Pharmaceutical Cardiac Safety Testing . . . 18

2.4.1 Arrhythmia . . . 18

2.4.2 Safety guidelines . . . 19

2.4.3 The CiPA initiative . . . 19

2.4.4 Other efforts towards improving proarrhythmic risk prediction 20

2.5 Drug binding of the hERG channel . . . 21

2.5.1 The hERG channel . . . 21

2.5.2 Mathematical models of hERG channel gating . . . 23

2.5.3 Drug-binding mechanisms . . . 24

2.5.4 Drug-binding models . . . 25

2.5.5 Example of a drug-binding model . . . 30

2.6 Credibility assessment of models: parameter identifiability, sensitivity, and uncertainty . . . 32
2.6.1 Sources of uncertainty in cardiac modelling . . . 33
2.6.2 Model evaluation . . . 34
2.6.3 Parameter sensitivity and identifiability . . . 37
2.7 Summary . . . 39

3 Importance of modelling hERG binding in predicting drug-induced action potential prolongations 40
3.1 Introduction . . . 41
3.2 Methods . . . 45
3.2.1 hERG-channel component . . . 46
3.2.2 State-dependent drug-binding formulation of hERG . . . 46
3.2.3 Calibration of a simple conductance scaling drug-binding formulation . . . 47
3.2.4 AP model . . . 49
3.2.5 Synthetic drugs . . . 49
3.2.6 Protocols . . . 50
3.2.7 Sensitivity analysis . . . 51
3.2.8 Measuring the difference between the drug-binding formulations 52
3.2.9 Simulations . . . 53
3.3 Results . . . 53
3.3.1 CS drug-binding formulation replicates SD formulation at steady state for a trapped drug example . . . 53
3.3.2 Trapping properties are apparent in the transient phase . . . 57
3.3.3 AP prolongation is dependent on the protocol used to stimulate the SD formulation . . . 60
3.3.4 Drugs were approximated to have shorter steady state APD_{90}s with the AP-CS model . . . 62
3.4 Discussion . . . 65

4 Impact of uncertainty in action potential models on *in silico* drug effect predictions 68
4.1 Introduction . . . 69
4.1.1 Naming convention . . . 70
4.1.2 Summary of the SD formulation and *drug-trapping* . . . 70
4.1.3 Extrapolation to new models . . . 71

4.1.4 Sensitivity analysis . . . 71
4.1.5 Summary . . . 73
4.2 Components of an action potential model . . . 73
4.2.1 AP models . . . 74
4.2.2 hERG-channel component of the AP model . . . 75
4.2.3 Drug-binding formulations to the hERG channel . . . 78
4.3 Methods . . . 79
4.3.1 Calibration of models . . . 79
4.3.2 Synthetic drugs . . . 81
4.3.3 Parameter space exploration and definition of similarity range 84
4.3.4 Measuring the difference between the drug-binding formulations 87
4.3.5 Simulations . . . 89
4.4 Results . . . 90
4.4.1 The relative I_{Kr} contribution in each AP model affects APD – dofetilide and verapamil . . . 90
4.4.2 The binding and unbinding rates remain the primary factors contributing to the differences between the SD and the CS drug-binding formulation . . . 94
4.4.3 Changes in the open state probability reduce the effect of drug binding . . . 95
4.5 Discussion . . . 97
4.6 Conclusion . . . 101

5 Identifiability of the trapping tendency parameter and its role within the state-dependent drug-binding formulation 103
5.1 Introduction . . . 103
5.1.1 Summary of the SD formulation . . . 105
5.1.2 Uncertainty quantification of the SD formulation . . . 105
5.1.3 Research question . . . 106
5.1.4 Summary . . . 106
5.2 Simulation of the SD formulation . . . 107
5.3 Identifiability analysis of the SD drug-binding formulation . . . 108
5.3.1 Generation of synthetic data . . . 108
5.3.2 Error profile . . . 110
5.3.3 Minimisation of RMSD . . . 111

5.3.4 Practical identifiability of the SD formulation for three example drugs . . . 113
5.3.5 Practical identifiability of the SD formulation for the parameter space . . . 116
5.4 Adjustments to the SD formulation . . . 120
5.4.1 Investigation of the range of parameter values in the SD drug-binding formulation – trapping rate K_t . . . 120
5.4.2 Investigation of the equations of the SD drug-binding formulation – membrane potential V . . . 123
5.5 Alternative drug-binding formulations without the trapping component 128
5.5.1 Introduction to the alternative drug-binding formulations . . . 128
5.5.2 Practical identifiability of the alternative drug-binding formulations . . . 132
5.5.3 Inference performance of all drug-binding formulations on experimental data . . . 139
5.5.4 Comparisons to the SD formulation . . . 150
5.5.5 Proarrhythmic risk categorisation of drugs using different drug-binding formulations . . . 153
5.5.6 Inclusion of trapping behaviour into the objective function . . . 158
5.6 Discussion . . . 161
5.6.1 Reasoning for the alternative drug-binding formulations . . . 162
5.6.2 Implications and future work . . . 164
5.7 Conclusion . . . 164

6 Conclusion and discussion **166**
6.1 Future work . . . 168
6.2 Concluding Remarks . . . 171

A Software availability and details **173**

B Data inconsistencies **174**

Bibliography **177**

Chapter 1

Introduction

The heart, as part of the circulatory system, pumps blood throughout the body to transport oxygen and nutrients and collect metabolic byproducts and carbon dioxide. The blood flow is facilitated by the contraction of the atria and ventricles. A cardiac cycle is initiated by the generation of electrical impulses by the pacemaker cells at the sinoatrial node, then conducted to the atria, which triggers their contraction. These signals then travel to the atrioventricular node before conducting to the Purkinje fibres and ventricles to generate the contraction of the ventricles. These electrical signals are conducted through the heart's muscle cell, known as the *cardiomyocyte*. The changes in the voltage potential of its membrane throughout a conduction cycle are known as the *action potential* (AP).

The AP is a result of the dynamic flow of ions, mainly potassium, sodium, and calcium ions, into and out of the cell via ion channels or transporters on the cell membrane. The ion channels are studied through *patch-clamp* experiments, performed on cells overexpressing the ion channels. A computational model is then developed based on experimental observations to capture the channel behaviour. Due to the lack of a complete understanding of the cardiomyocytes and the ion channels, there are many diverse mathematical models developed.

Disruption to the electrical activity of cardiac cells can lead to disruption in the heart rhythm, known as *arrhythmia*, and potentially bring about severe consequences. While this may be treated with drugs, they can have side effects and worsen the condition. Therefore, newly developed drugs must undergo a series of safety screening tests to ensure it is safe for consumption. This can become very costly in terms of time and monetary resources [Wouters et al., 2020, Sertkaya et al., 2024]. Existing guidelines rely heavily on an assay of a single ion channel and the costly Thorough QT study, each of which has low specificity [Stockbridge et al., 2012], and this has slowed and in some cases prevented potential drugs from further development [Sager

et al., 2014, Colatsky et al., 2016]. Several institutions, including the U.S. Food and Drug Administration (FDA), the Health and Environment Sciences Institute (HESI), and Japan National Institute of Health Science (NIHS), have come together under the Comprehensive *in vitro* Proarrhythmic Assay (CiPA) initiative to improve the assessment of pro-arrhythmic risk of drugs [Sager et al., 2014, Colatsky et al., 2016, Strauss et al., 2019, Li et al., 2020, Strauss et al., 2021]. The recommended approach of CiPA focuses on four streams, which are *in silico* models, *in vitro* data from ion channel assays, stem cell-derived myocytes, and *in vivo* evaluation, to provide a more holistic risk assessment.

This shifts the focus for the development of computational models, from fundamental research in describing and quantifying electrophysiology observations, to its use in drug safety testing, where the understanding of the underlying mechanism becomes important in developing robust and reliable models. Therefore, there is a need to reassess the model development process to allow for the functional transition of computational models to contribute to reliable drug-safety testing.

The human Ether-à-go-go-Related Gene (hERG) ion channel plays a critical role in current drug-safety guidelines [ICH E14/S7B Q&A, 2022]. This is due to its function in returning the membrane potential to its resting state and its high susceptibility to drug blockage. However, the underlying mechanism of how drugs bind and block the hERG current is poorly understood [Vandenberg et al., 2017]. This has led to the development of significantly different models in attempts to capture the drug-binding behaviour [Brennan et al., 2009, Lei et al., 2024]. One particularly important and influential model is the Li drug-binding formulation [Li et al., 2017b]. This risk prediction model and its associated metric scores were developed by research teams in the FDA [Li et al., 2016, 2017b, Dutta et al., 2017, Chang et al., 2017], and serve as a reference for the arrhythmic risk prediction under the CiPA paradigm [Li et al., 2020]. Given its role within the CiPA initiative and drug safety assessment, it is important to assess the predictivity of the Li drug-binding formulation.

The overall aim of this thesis is to perform a thorough sensitivity analysis and identifiability analysis on the Li drug-binding formulation. Our work assesses the robustness and reliability of the formulation. This is especially important when the model is used in drug-safety prediction.

In this thesis, we first compare the Li drug-binding formulation with a simpler drug-binding model to determine the reliability of the Li formulation out of its defined context. We find that the Li formulation does not perform better than the simple formulation in the prediction of the AP. We then embed both drug-binding

formulations into different cell models and repeat the comparison, but obtain the same results. This leads to the hypothesis that the Li drug-binding formulation is potentially not capturing the drug-binding mechanism accurately, or the parameters are not identifiable from the experimental data. Through the identifiability analysis of the Li drug-binding formulation, we simplify the formulation in several ways and compare them to the Li drug-binding formulation. Using the same experimental data that was used by Li et al. [2017b] in deriving their original model, we are able to obtain the same risk categorisation for a range of reference drugs using the simplified model. We conclude by discussing the implications of our findings.

1.1 Thesis outline

Chapter 2 introduces the background knowledge required for the thesis. This includes the necessary background on cell physiology, the mathematical models, and the statistical inference used in the thesis.

Chapter 3 presents the sensitivity analysis of the Li drug-binding formulation. We took a simpler model as a baseline to investigate the influence of the Li drug-binding formulation on AP prediction.

Chapter 4 extends the comparison between the Li drug-binding formulation and the simpler model to various ion-channel level and cell-level models. The Li drug-binding formulation is embedded into several different AP models, and its impact is again compared to the simpler model in each of the settings.

In Chapter 5, we investigate in more detail the defining equations of the Li drug-binding formulation, exploring the identifiability of the parameters. We then explore several simplifications of the Li drug-binding formulation and assess the performance of each model in categorising the proarrhythmic risk of a reference set of drugs.

In Chapter 6, we conclude that a simplified model provides the same risk categorisation as the Li drug-binding formulation. We discuss future research and open challenges.

Chapter 2

Background

Overview:

This chapter provides the background information necessary to understand the thesis. In this thesis, we are interested in using mathematical and computational models of the heart in the prediction of possible cardiotoxic side effects of potential new drugs, and so in categorising the safety level of those drugs. We will first describe the physiology of the heart that is relevant to the modelling of such adverse effects. Next, we will describe the mathematical models that reflect the biological phenomena, and hence allow the prediction of the side effects of drugs. Thirdly, we will focus on the development of the mathematical models from experimental data. Since these models are intended to be used in a clinical setting, we need robust and reliable predictions from the models. We will therefore finally discuss the concepts and methods required in assessing the reliability of a model.

2.1 Electrophysiology of the Human Heart

The heart pumps blood throughout the body, providing oxygen to all parts of the body and taking carbon dioxide back to the lungs for excretion. The blood is transported throughout the body by the pressure generated by the contraction of the *atria* and *ventricles*. A schematic of the human heart is shown in Figure 2.1. Deoxygenated blood is pumped by the right side of the heart towards the lungs, while the left atrium and left ventricle pump oxygenated blood from the lungs to the rest of the body. Since the atria pump blood only to the ventricles while the ventricles pump the blood to either the lungs or the rest of the body, the muscular walls of the ventricles are thicker than the muscular walls of the atria. The muscular wall of the heart is formed by three layers of tissue, where the inner layer is the *endocardium*, the middle

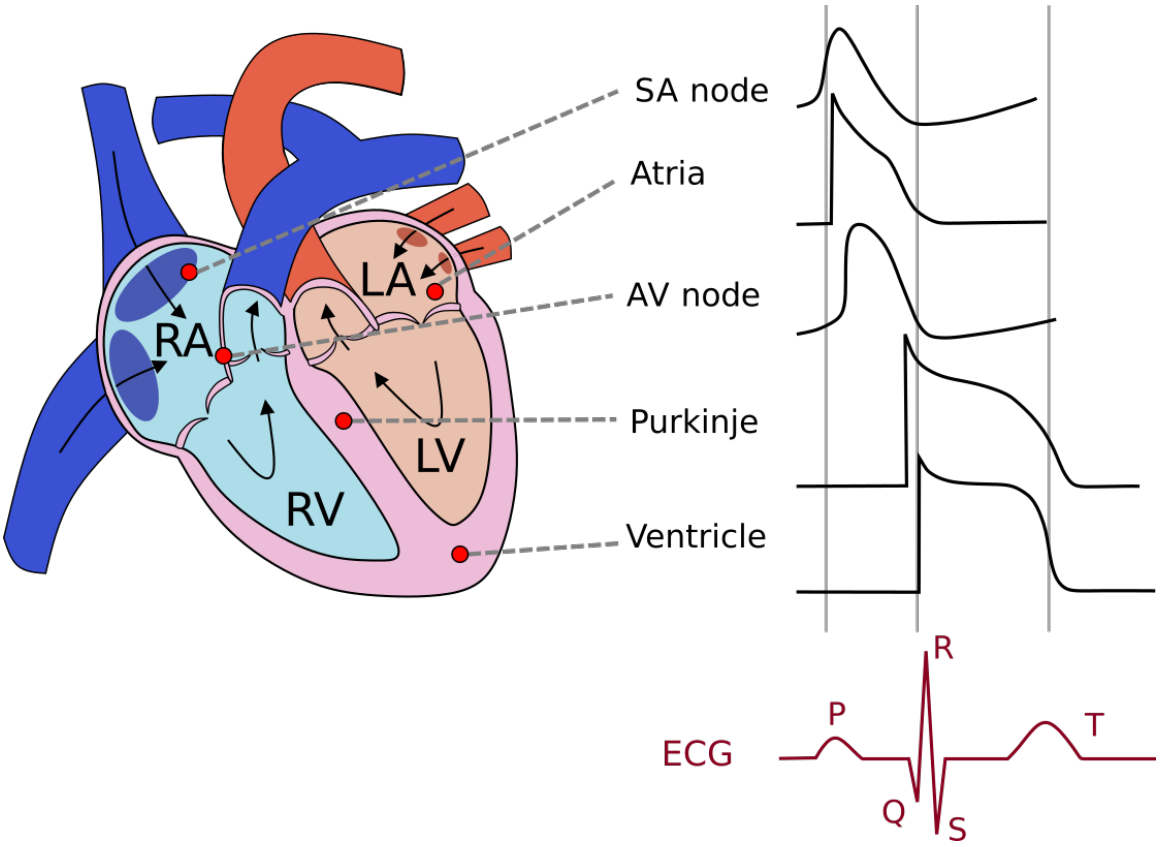


Figure 2.1: **A schematic of a human heart.** The figure shows the 4 chambers of the heart: the left atrium (LA), the right atrium (RA), the left ventricle (LV), and the right ventricle (RV). The electrical impulse that stimulates heart contraction is initiated at the sinoatrial (SA) node. The electrical impulse then travels to the atrioventricular (AV) node, the Purkinje fibres and finally to the ventricles. Based on [Hume and Grant, 2015]

layer is the *myocardium* and the outer layer is the *epicardium*. The muscle cells of the heart are known as *cardiomyocytes*.

The contractions of the atria and ventricles are performed in a controlled sequence at each heartbeat, known as the *cardiac cycle*. The cycle starts from the blood filling the ventricles by relaxation of the ventricles and contraction of the atria. The ventricles then contract, forcing blood out of the heart to the rest of the body and the lungs. This controlled sequence of muscle contraction of the heart can be visualised by the electrocardiogram (ECG) (Figure 2.1). The electrical signals within the cardiomyocytes trigger the contraction of the cardiomyocytes which propagates the signal to the adjacent muscle cells, generating a coordinated sequence of muscle contraction. The ECG starts with a P wave when the *sinoatrial* node generates electrical impulses to be propagated throughout the atria and causes the atria to depolarise and con-

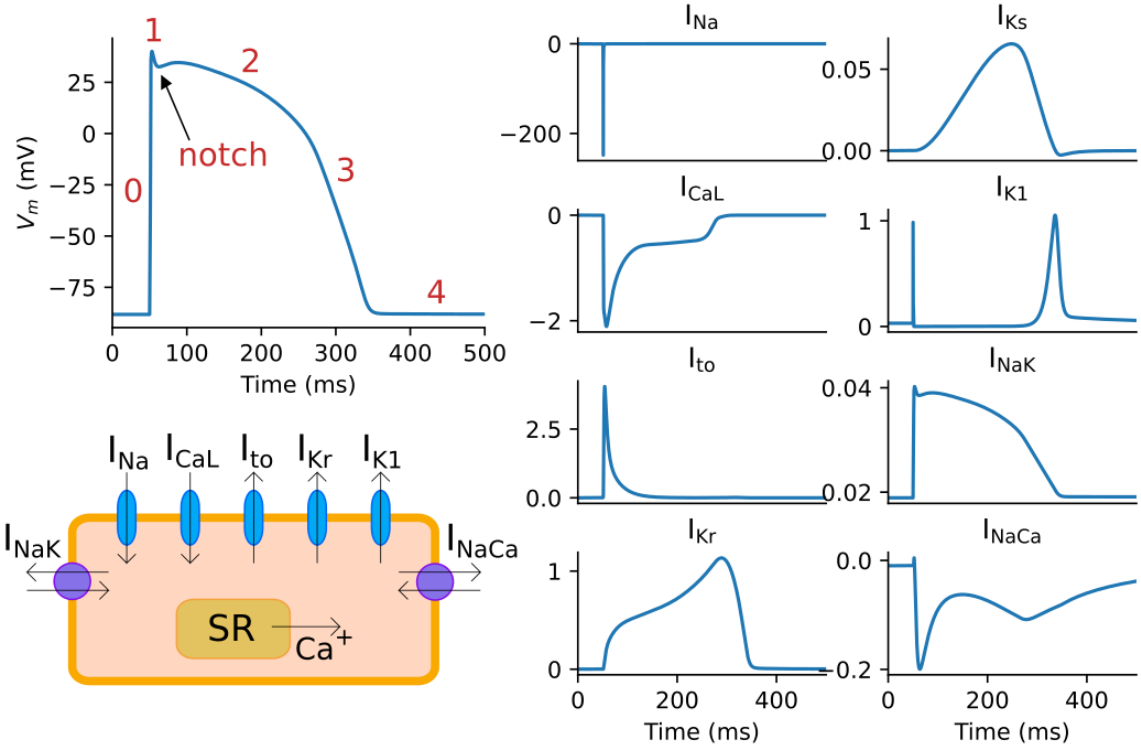


Figure 2.2: **Action potential and its ionic currents simulated with the O'Hara model [O'Hara et al., 2011] for epicardial ventricular myocytes, together with a schematic of a cardiomyocyte.** The numbers in red show the phases of the epicardial AP. The eight panels on the right show the major currents that play a role in the AP. The bottom left schematic shows a simplified cardiomyocyte model, showing some ion channels, pumps and exchangers. The ion channels are shown in blue, while pumps and exchangers are shown in purple. The sarcoplasmic reticulum (SR) compartment contains the Ca^{2+} storage for calcium release in the myocyte.

tract. The electrical signals then travel over the atria to the *atrioventricular* node; this is represented by the PQ segment of the ECG (see Figure 2.1). Subsequently, the waves travel to the Bundles of His in the interventricular septum and then to the Purkinje fibres, causing the ventricles to depolarise. This is represented by the QRS complex. This is followed by the contraction of ventricles, represented by the ST segment. Finally, the T wave indicates the repolarisation and relaxation of the ventricles.

2.1.1 Cellular electrophysiology and the action potential

Cardiomyocytes in different parts of the heart, such as the atria, ventricles and the Purkinje fibres, contract in synchronisation to form a heartbeat. This synchronisation relies on the propagation of the electrical signals across the cardiomyocytes. The *action potential* (AP) captures the sequential changes in voltage across the membrane of the cardiomyocyte during a contraction cycle.

The cardiomyocytes are bound by a lipid bilayer membrane, which is electrically insulated and prevents ions from flowing freely across the membrane [Hille, 2001]. The ions move across the membrane via *transmembrane proteins*. There are several mechanisms for ions to be transported across the cell membrane: ions flow through the membrane passively via the *ion channels* down the electro-chemical gradient; ions are transported actively against the concentration gradient and electric potential via *ion pumps* or *ion exchangers*, depending on the source of energy, to maintain the homeostasis of ion concentrations [Hille, 2001]. The ion pumps use energy from adenosine triphosphate (ATP), while the ion exchangers utilise the energy from passive transport to push other ions across the membrane. Reviews of ion transport processes are provided in Ravens et al. [2004] and Gadsby [2009].

The change in the membrane potential is caused by ions moving across the membrane of the cardiomyocyte. We take the AP of epicardium ventricular cells to exemplify the flow of ions across the cell membrane during an AP because we will focus on the ventricular AP in this thesis. Figure 2.2 (top left) shows a schematic of an AP of the outer layer (*epicardial*) ventricular myocytes. When the cardiomyocytes receive an electrical stimulus at the resting state, sodium ions flow into the cell, creating the fast sodium current I_{Na}. The membrane voltage increases and the membrane potential *depolarises*. This is known as the *depolarisation* phase (Phase 0). At the initial *repolarisation* phase (Phase 1), the membrane voltage decreases slightly from the efflux of potassium ions from the cell. This flow of potassium ions generates the transient outward potassium current I_{to}, which creates the notch in the AP. This is prominent in epicardial ventricular myocytes, but not in myocardial or endocardial ventricular myocytes. Next comes the plateau phase (Phase 2) where the influx of calcium ions, forming the L-type calcium current I_{CaL}, balances off the efflux of potassium ions, which mainly consists of the rapid delayed rectifier current I_{Kr} and the slow delayed rectifier current I_{Ks}. The entry of calcium ions into the cardiomyocyte causes calcium ions to be released from an internal reservoir, initiating the process of *calcium-induced-calcium-release*, which will then lead to the contraction of the cardiomyocyte. The sustained flow of potassium ions out of the cell and reduced influx

of calcium ions bring the cell to the final repolarisation phase (Phase 3). Towards the end of Phase 3, the inward rectifier current I_{K1} brings the membrane potential to its resting potential and stabilises the membrane voltage in Phase 4. In Phase 4, the membrane potential stays at its resting potential. A schematic of a simplified cell model with components involved in the ionic currents and the contraction of a cell is shown in the bottom left of Figure 2.2, along with the simulated membrane voltage and the major ionic currents.

2.1.2 Contribution of ionic currents to AP

The opening and closing of the ion channels are timed precisely to allow selected ions to transverse the membrane and generate the AP [Hille, 1978, Grant, 2009]. The selective permeability of ion channels allows only specific ions to pass through their pore, though other ions may still traverse the membrane at a significantly slower rate. The opening and closing of the pores of the channels, that is, the *gating* of the ion channels, are determined by the conformational changes of the channels. The *gating* of the ion channels can be driven by different forces, for example *voltage-gated*, *ligand-gated*, *light-gated*, and *mechano-sensitive gating*. Voltage-gated channels change its conformation in response to changes in the membrane potential.

The gating process may also be altered by drug blockers, inhibiting the ion flux across the membrane [Hille, 2001]. This will be described in detail in Section 2.5. This drug-induced inhibition may have significant consequences on the resulting AP and cause unwanted changes to the cardiac rhythm.

2.2 Mathematical Modelling of Cellular Electrophysiology

In this section, we will show how we can model the interactions between the ion currents and the membrane potential to develop an AP model.

2.2.1 The cell membrane

The electrically-insulated membrane effectively acts as a capacitor, with an electric field across the membrane. The capacitance of the membrane C_m, the storage of electric charge, is Q_m/V_m, where Q_m is the charge needed to generate the membrane potential of V_m between the internal and external environment of the membrane. V_m here refers to the potential difference between the intracellular potential and the

extracellular potential, following the usual convention. The ion channels act as a *conductor* to discharge the capacitor. The cell membrane and the ion channels thus complete an electric circuit.

Given that current is the rate of change of charge dQ/dt, the capacitive current is $C_m dV_m/dt$, assuming the C_m is constant. The equation of the electric circuit is therefore

$$C_m \frac{\mathrm{d}V_m}{\mathrm{d}t} + I_{\text{ion}} = 0. \tag{2.1}$$

where I_{ion} is the contribution of ionic currents to the action potential. I_{ion} can be described as the sum of the individual ionic currents I_j for some ion j. Therefore, the relationship between the membrane potential and the ionic currents can be written as

$$\frac{\mathrm{d}V_m}{\mathrm{d}t} = -\frac{1}{C_m} \sum_j I_j. \tag{2.2}$$

An external stimulus, e.g. from neighbouring cells, is included as

$$\frac{\mathrm{d}V_m}{\mathrm{d}t} = -\frac{1}{C_m} \left(\sum_j I_j + I_{\text{stim}} \right). \tag{2.3}$$

2.2.2 The ionic currents

The ionic currents flowing through the membrane can be expressed as a *driving term*, describing the electrochemical field that drives the ionic flow, multiplied by the conductivity of the channel

$$I_j = g_j \cdot \delta_j, \tag{2.4}$$

where g_j is the conductance (reciprocal of the resistance) and δ_j is the driving term.

2.2.2.1 The driving term

Ionic movement through ion channels is driven by the electric field around the membrane and the concentration gradients of the ions. The ions will flow until an equilibrium is reached.

Assuming the ionic current is Ohmic, where the current passing through a conductor (the ion channels) is directly proportional to the voltage, the ionic current is given by

$$I_j = g_j \cdot (V_m - E_{\text{S}}), \tag{2.5}$$

where E_S is the Nernst potential for ion S. The term $\delta_{\text{Ohm,S}} = V_m - E_S$ is the driving term in Equation 2.4. An alternative formulation of the driving term is the *Goldman-Hodgkin-Katz* (GHK) equation [Goldman, 1943, Hodgkin and Katz, 1949].

The Ohmic driving term and the GHK driving term have been used among many ionic current models. A list of commonly used driving terms is given in Agrawal et al. [2023].

2.2.2.2 The open probability

Although ion channels open and close stochastically, but due to the large number of ion channels of a specific type within the membrane, the macroscopic current can be approximated by the ensemble average. The macroscopic current can then be modelled by the proportion of channels in the open state, which is termed the *open state occupancy*. This means that the ionic conductance g_j is usually not constant. In this thesis, we are particularly interested in voltage-gated channels, where the gating mechanism of the channel is dependent on V_m. We can model the conductivity as the probability of open states and rewrite the general ionic current in Equation 2.4 as

$$I_j = \bar{g}_j \cdot O \cdot \delta_j, \tag{2.6}$$

where $\bar{g}_j$ is the maximal conductance for a channel of type j and $O \in [0, 1]$ is the *open probability* or the *open state occupancy*, i.e., the fraction of channels being in the open or conducting state. For voltage-gated channels, $O(V_m, t)$ is a function of the membrane voltage V_m and t.

The open probability describes the *kinetics* or *dynamics* of the voltage-gated channel, capturing the gating mechanism of the channels in terms of how and when the selected ions can pass through its pores. There are two common formulations of the open probability of ionic channels, the *Hodgkin-Huxley* formulation and the *Markov* formulation.

The Hodgkin-Huxley formulation [Hodgkin and Huxley, 1952] models the channel as having '*gates*' that allow (and stop) ion movement across the channel. There can be several layers and types of 'gates', for example, three activation gates and one inactivation gate for a sodium channel. It is assumed that each gate functions independently. They are expressed as the probability of opening, thus taking any values between 0 and 1. As an example, the opening probability of a channel can be

$$O(V_m) = \prod_k p_k = p_1(V_m, t) \cdot p_2(V_m, t), \tag{2.7}$$

where p_k is the open probability of gate k. Given a gate with opening probability p, the closing probability is $(1-p)$. The gating mechanism can be viewed as

$$(1-p) \underset{\beta(V_m)}{\overset{\alpha(V_m)}{\rightleftharpoons}} p, \tag{2.8}$$

with α and β as the opening and closing rates, respectively. The rate of change of the open probability p is

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \alpha(V_m)(1-p) - \beta(V_m)p. \tag{2.9}$$

The rate of change of the open probability $\dfrac{\mathrm{d}p}{\mathrm{d}t}$ can be rewritten in terms of steady state (p_∞) and time constant (τ_p)

$$\frac{\mathrm{d}p}{\mathrm{d}t} = \frac{p_\infty(V_m) - p}{\tau_p(V_m)} \tag{2.10}$$

where the steady state value $p_\infty = \alpha/(\alpha+\beta)$ and the time constant or rate at which the steady state is achieved $\tau_p = 1/(\alpha+\beta)$.

The Markov formulation views the channel as being in different states and models the kinetic states of the ion channels, for example, open, closed, and inactivated states. The Markov formulation allows more flexibility in the possibilities of states and the transitions between states, eliminating theassumption of independent gating in the Hodgkin-Huxley formulation.

A simple example of a Markov formulation that differentiates from a Hodgkin-Huxley formulation is a three-state channel model, where there is an open state O, an inactivated state I, and a closed state C:

$$C \underset{\beta(V_m)}{\overset{\alpha(V_m)}{\rightleftharpoons}} O \underset{\zeta(V_m)}{\overset{\gamma(V_m)}{\rightleftharpoons}} I, \tag{2.11}$$

where $\alpha(V_m)$, $\beta(V_m)$, $\gamma(V_m)$ and $\zeta(V_m)$ are the *transition rates* from closed state to open state, open state to closed state, open state to inactivated state and inactivated state to open state respectively. The system of differential equations describing this three-state channel model is

$$\frac{\mathrm{d}C}{\mathrm{d}t} = \beta(V_m)\cdot O - \alpha(V_m)\cdot C, \tag{2.12}$$

$$\frac{\mathrm{d}O}{\mathrm{d}t} = \alpha(V_m)\cdot C + \zeta(V_m)\cdot I - [\beta(V_m) + \gamma(V_m)]\cdot O, \tag{2.13}$$

$$\frac{\mathrm{d}I}{\mathrm{d}t} = \gamma(V_m)\cdot O - \zeta(V_m)\cdot I. \tag{2.14}$$

Note that the sum of the probabilities of all states equals one, $C + O + I = 1$. In this three-state Markov model, channel inactivation depends on the activation of the channel. Therefore, the assumption of independent gating in a Hodgkin-Huxley formulation no longer holds. A concern of using a Markov model is that it has to satisfy the thermodynamic properties of an equilibrium system, which is termed *microscopic reversibility*, i.e., the transition rate between any two states via any route must be the same.

For more details on the difference between the Hodgkin-Huxley formulation and the Markov formulation, see Carbonell-Pascual et al. [2016] and Fink and Noble [2009].

2.3 Experimental Measurements to Mathematical Models

We next describe the experimental techniques used to make observations and the mathematical techniques needed to translate the observations into the mathematical models, which allow us to study the impact of drugs on cardiac electrophysiology.

2.3.1 Experimental techniques

The electrical activity of an excitable cell can be recorded using *whole-cell patch-clamp experiments*. The basic idea of these experiments is to access the interior of a cell and connect it with the extracellular environment to allow measurements of the potential difference between the intracellular and extracellular solutions. To do so, the first step is to form a high-resistance seal, termed *gigaseal*, between the tip of the pipette and the cell membrane. This high-resistance seal is essential in minimising current leakage and improving the signal-to-noise ratio of the recording. This gigaseal is achieved by the micropipette applying suction to the cell membrane. Once a gigaseal is formed, electrical access to the cell's interior is made available by applying an abrupt pulse of suction to rupture the small patch of membrane under the pipette tip.

The *automated patch-clamp* was designed to gain a higher throughput of experimental recordings. In the automated patch-clamp experiments, a planar substrate containing a micron-sized aperture is used instead of a conventional glass micropipette to establish a recording. A single cell from a suspension is attracted to the aperture, often through the application of negative pressure. Once the cell is in place, a stronger pulse of suction is applied to form a high-resistance seal, followed by another suction

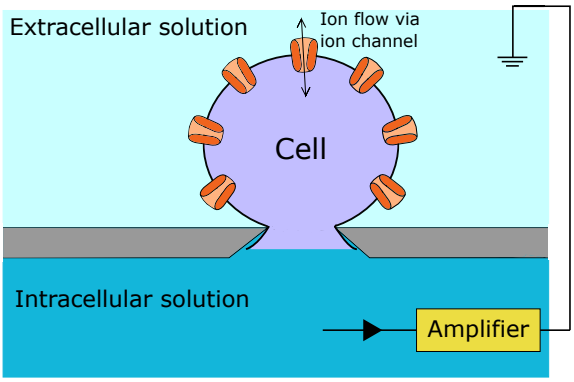


Figure 2.3: **A schematic of the automated patch-clamp experiment.** The figure shows how the current generated from ion flow via ion channels is measured.

to rupture the enclosed membrane patch to set up the whole-cell configuration. Figure 2.3 shows a schematic of the setup of an automated patch-clamp experiment. Comparisons between the manual patch-clamp and the automated patch-clamp are available in Franz et al. [2017], Li et al. [2017a].

There are two different recording modes in patch-clamp experiments: current-clamp and voltage-clamp. The current clamp measures the membrane potential, and is used to study the change in membrane potential over time, particularly in response to stimuli. On the other hand, the voltage clamp measures the ionic currents, and is used to study the properties of the ion channels. The ionic currents are measured while the membrane potential is held at a fixed, constant value (the *command potential* intended by the experimenter). To do so, the amplifier constantly measures the membrane potential of the cell and compares it to the command potential. When the measured potential and the command potential differ, the amplifier injects a *compensatory current* to ensure the membrane potential is held at the command potential. This compensatory current gives the recorded signal. A predefined sequence of command potentials applied over time to the ion channels is termed the *voltage-clamp protocol*.

The current measured in the patch-clamp experiments mentioned above is the total transmembrane current, without differentiating between the types of currents flowing through the membrane. The study of a specific type of ion channel can be performed on stably transfected cell lines overexpressing the gene for the channel of interest. Widely used cell lines are the Chinese Hamster Ovary (CHO) and the Human Embryonic Kidney (HEK) cell lines [Hill and Stephens, 2021, Gamper et al., 2005, Li et al., 2017a]. When combining these ionic currents into an overall action

potential model (see Equation 2.2), each of the currents will come from a different experimental system.

2.3.2 Forward simulation

From Equation 2.3, $C_m \mathrm{d}V_m/\mathrm{d}t + I_{\text{ion}} + I_{\text{stim}} = 0$, I_{stim} is given, as it is the stimulus current provided to the cell in experiments. We want to simulate V_m and I_{ion}. This is done by solving the equations for V_m and I_{ion}.

There is usually no analytical solution for V_m and I_{ion}, and the equations must be solved using numerical methods. For ease of reference, let the governing equations for V_m and I_{ion} be a system of functions $\mathbf{f}$ with variables $\mathbf{u}$. Note that the governing equations of I_{ion} are also made up of differential equations (Section 2.2). In general, we wish to solve equations of the form

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = \mathbf{f}(\mathbf{u}). \tag{2.15}$$

The simplest method is the forward Euler method, where the value $\mathbf{u}(t + \Delta t)$ at the next time step $t + \Delta t$ is estimated from the value $\mathbf{u}(t)$ at the present time step t, starting from the initial condition $\mathbf{u}(0)$.

$$\mathbf{u}(t + \Delta t) = \mathbf{u}(t) + \mathbf{f}(\mathbf{u})\Delta t. \tag{2.16}$$

The systems of ODEs we are interested in are usually stiff, i.e., some variables fluctuate on vastly different time scales [Cooper et al., 2015]. As can be seen in the constituent currents of an action potential in Figure 2.2, the sodium current I_{Na} has fast kinetics while the potassium current I_{Ks} has slower kinetics. Smaller step sizes are required to resolve the faster kinetics, but it will be inefficient for the slower kinetics. These problems are better solved by implicit numerical methods as they generally have a larger stability region. The *backward differentiation method* (an implicit multistep method), combined with Newton's method, is used to solve the systems of ODEs in this thesis. This is implemented via the software package CVODE [Hindmarsh et al., 2005] using a Python interface Myokit [Clerx et al., 2016], which is developed for cardiac electrophysiology systems. Throughout the thesis, the absolute tolerance and the relative tolerance are set to 10^{-7} and 10^{-8}, respectively, unless stated otherwise. A summary of the computational details for all coding implementations of this thesis is provided in Appendix A.

2.3.3 Inverse problem

The mathematical models described in Section 2.2 can be simulated with a range of parameter values, resulting in many possible model outputs. What we wish to find are the specific parameter values where the corresponding simulated model output "best matches" the experimental observation in a well-defined sense. This problem of searching for the optimal parameter values that best describe the observed data is usually termed the *inverse problem*, *model calibration*, or *model training*.

We first formulate the problem. We have a mathematical model $f(\boldsymbol{\theta})$ with model parameters $\boldsymbol{\theta}$ that describes the time-varying potential across the cell membrane of the cardiomyocyte (Section 2.2). We perform patch-clamp experiments on ion channels and obtain the experimental data Y in the form of a current trace (Section 2.3.1). In an ideal case, the mathematical model would match the observed data exactly, giving $Y = f(\boldsymbol{\theta})$. However, experimental data are noisy, so we assume the data Y_i can be approximated as

$$Y_i \approx y_i = f_i(\boldsymbol{\theta}) + \epsilon_i, \tag{2.17}$$

where i indexes the data point (and typically indexes time in this thesis), ϵ_i refers to the noise at each data point, and y_i is the simulated model output, i.e., the simulated voltage or current. In this thesis, we assume the noise follows a normal distribution with zero mean.

One way to solve the inverse problem is by the *ordinary least squares* method. The aim is to minimise the difference between the observed data Y and model simulation $f(\cdot)$. The optimal parameter vector $\hat{\boldsymbol{\theta}}$ can be found by searching for parameter values $\boldsymbol{\theta}$ within a parameter space $\boldsymbol{\Theta}$ that minimises the noise

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min} \sum_i \epsilon_i^2 = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min} \sum_i \left(Y_i - f_i\right)^2 \tag{2.18}$$

Alternatively, we can define the objective function as a likelihood function

$$\hat{\boldsymbol{\theta}} = \underset{\theta\in\Theta}{\arg\max} \prod_i \mathcal{L}_i\left(\boldsymbol{\theta}|Y_i\right), \tag{2.19}$$

where $\mathcal{L}$ is the likelihood of the parameter $\boldsymbol{\theta}$ given the observed data. With a likelihood function, we want to maximise the likelihood as a function of the parameters $\boldsymbol{\theta}$ given that we have observed a particular data set Y. This is the *maximum likelihood estimate*. In formulating the objective function, we assume the noise is independent and identically distributed (i.i.d.). If the noise is further assumed to be Gaussian

$\mathcal{N}(0, \sigma^2)$ with zero mean and σ standard deviation, the likelihood function is given by the product of the probability density function evaluated at each data point, i.e.,

$$\mathcal{L}(\boldsymbol{\theta}|Y) = p(Y|\boldsymbol{\theta}) = \prod_{i=1}^{N} \mathcal{N}(Y_i|f_i, \sigma^2) \tag{2.20}$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(Y_i - f_i)^2}{2\sigma^2}\right) \tag{2.21}$$

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{\sum_{i=1}^{N}(Y_i - f_i)^2}{2\sigma^2}\right). \tag{2.22}$$

When the noise is assumed to be Gaussian with zero mean and the standard deviation is assumed to be known, the maximum likelihood estimation and the ordinary least squares are equivalent, yielding the same optimised parameter $\hat{\boldsymbol{\theta}}$. These problems, formulated as minimisation and maximisation problems with objective functions as *cost function* and *likelihood function*, respectively, are commonly known as *optimisation* problems.

2.3.4 Computational methods for the optimisation problem

Suppose we have an inverse problem with an objective function $g(\boldsymbol{\theta})$ to be optimised for an optimal parameter $\hat{\boldsymbol{\theta}}$. This problem is commonly solved by taking the derivative of the objective function to be zero, i.e., $\nabla g(\boldsymbol{\theta}) = 0$ and solving the resultant equations. However, there is usually no closed-form solution to the equations given the objective function g of our problem. Again, we need numerical methods to search for the optimum.

As with all methods of optimisation, the challenge is to find the global optimum in the possible presence of multiple local optima. A global optimum is where $\hat{\boldsymbol{\theta}}$ gives the optimum of the objective function for all parameter values in the parameter space, whereas a local optimum satisfies the objective function only around the optimised parameter, that is

$$g(\hat{\boldsymbol{\theta}}) \leq g(\boldsymbol{\theta}) \text{ for all } \boldsymbol{\theta} \in \boldsymbol{\Theta} \text{ near } \hat{\boldsymbol{\theta}}. \tag{2.23}$$

The main approaches to solving the optimisation problem numerically are *gradient-based* methods and *derivative-free* methods. One of the gradient-based methods is the *gradient descent* method. The optimum is achieved by moving along or against the gradient (for maximisation and minimisation, respectively) of the objective function

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - \eta \nabla g(\boldsymbol{\theta}_n), \tag{2.24}$$

where η is the learning rate, a hyperparameter that controls the step size taken in the direction of the gradient. The method is initiated from a random point in the parameter space and repeated until the improvement in the objective value between iterations is below a certain predetermined threshold, i.e. when the gradient gets sufficiently close to zero. The derivatives provide the direction of the local optima, whereas the second derivatives indicate the curvature around the point, providing information on the rate of change of the slope. Newton's method utilises this to adapt the step size for a faster convergence rate, assuming the objective function is twice differentiable, giving

$$\boldsymbol{\theta}_{n+1} = \boldsymbol{\theta}_n - [H_g(\boldsymbol{\theta}_n)]^{-1} \nabla g(\boldsymbol{\theta}_n), \tag{2.25}$$

where H_g is the Hessian matrix of the objective function g. The Newton's method is computationally more expensive than the gradient descent method. There are many other gradient-based methods developed to improve the convergence rate and avoid evaluations of the Hessian matrix. For a more detailed explanation on gradient-based methods, see e.g. [Stein, 2024, Boyd and Vandenberghe, 2014, Martins and Ning, 2022].

Gradient-based methods assume the optimisation problems are continuous and smooth, at least around the optimum. However, the derivatives of our objective function do not always exist, nor can they be accurately approximated. There is a range of derivative-free methods to solve the optimisation problem without using the gradients, such as genetic algorithms, generalised pattern search, and many more. Methods that are more suitable for our problem are *evolutionary algorithms*, which are inspired by the principles of natural selection. These are widely used in cardiac modelling. The algorithm is initiated with a population of points. All these points are updated at each iteration, which allows a global search for the optimum. At each iteration, a subset of the current population that has better objective values is chosen as the parent set. At the next iteration, parameter vectors within the population with the best objective values are retained, while changes, chosen from a set of operators, are made to the others. These operators include crossover, where two parent vectors are combined, and mutation, where a small change is made to some elements of the parameter vector. There are great flexibilities on the choice of operators, leading to many variations of the evolutionary algorithm. Evolutionary algorithms are heuristic algorithms, which means they do not have a stopping criterion that guarantees optimality of the algorithm, i.e., it is not expected for the algorithms to achieve the global optimum, but only provides solutions near the optimum. For

a more detailed explanation on derivative-free methods and evolutionary algorithms, see e.g. Audet and Hare [2017], Conn et al. [2009], Martins and Ning [2022], Simon [2013].

The evolutionary algorithm that has proved to be most useful for our problems is the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) by Hansen et al. [2003]. Rather than performing the operators on the population of points at each iteration, the CMA-ES generates offspring by sampling from the search distribution. Moreover, this search distribution is updated by approximating the objective function space with a multivariate Gaussian distribution. This algorithm is implemented using PINTS, a Python package developed within our group to provide easy use of many optimisation and sampling methods [Clerx et al., 2019].

2.4 Pharmaceutical Cardiac Safety Testing

While cardiac models have been used to gain insights into the mechanism or working of the cardiomyocytes, cardiac models are also now increasingly used in clinical and drug safety testing, specifically predicting the side effects of drugs on the heart [Davies et al., 2016, Mirams, 2022, Bifulco et al., 2021, Corrado et al., 2021, Hill et al., 2016], shifting the focus from qualitative to quantitative predictions. This motivates the work presented in this thesis.

2.4.1 Arrhythmia

Irregularities in heart rate are called *arrhythmias*. Arrhythmias may indicate a disruption to the healthy functioning of the heart and, if severe, may lead to sudden cardiac death [Zipes and Wellens, 1998]. Arrhythmias can be caused by drugs that affect ion channel function, whether intended or unintended.

For patients, anti-arrhythmic drugs can regulate the function of ion channels, restore the action potential, and restore the heart rhythm to its regular rate. However, they may also be pro-arrhythmic [Nattel and Carlsson, 2006]. Drugs with usage unrelated to the heart or cardiac cells can also cause serious and harmful reactions (*adverse drug effects*) on cardiac cells [Nattel and Carlsson, 2006]. For example, some antipsychotic drugs can cause *Torsade de Pointes* (TdP), a type of arrhythmia beginning from the ventricles [Glassman and Bigger, 2001]. TdP is considered to be initiated by focal activities arising from early afterdepolarisations (EADs), where the membrane potential depolarises before it fully repolarises [El-Sherif et al., 1996]. When these EADs occur across cells in the ventricular subendocardium with a spatial

dispersion of repolarisation, it creates a highly vulnerable environment that could trigger the onset of TdP [Dunnink et al., 2016]. Once the initial irregular beats are triggered, the continuation of the TdP is induced by the reentrant excitation [El-Sherif et al., 1997, Rivaud et al., 2021] or continuous focal activities [Vandersickel et al., 2016].

At the cellular level, the prolongation of AP creates opportunities for EADs. When the prolonged AP duration occurs in a sufficient proportion of the ventricular myocardium, it leads to delayed ventricular repolarisation and is reflected as QT prolongation. Therefore, the prolongation of AP duration and the prolongation of the QT interval are used as predictors for a drug to cause TdP. This is reflected in the drug safety guidelines through the Thorough QT/QTc study [ICH E14, 2005].

2.4.2 Safety guidelines

It is essential to have drug safety testing within the drug development pipeline to predict potential risks before conducting human trials. To prevent drugs that may induce life-threatening cardiac arrhythmias from being in the market for human use, the International Conference on Harmonisation (ICH) introduced the S7B guidelines [ICH S7B, 2005] for nonclinical assessment and E14 guidelines for clinical assessments [ICH E14, 2005]. The ICH S7B guidelines [ICH S7B, 2005] suggest *in vivo* QT study and *in vitro* screening for blockage in the *human Ether-à-go-go-Related Gene* (hERG) channel (the alpha subunit of the potassium ion channel for I_{Kr}), which are currently standard procedures for testing of new drugs. This potassium ion current I_{Kr} of the hERG channel is responsible for the repolarisation of the action potential (Section 2.1) and thus the blocking of this channel can delay the repolarisation, prolong the action potential [Jurkiewicz and Sanguinetti, 1993] and increase the QT interval [Curran et al., 1995].

2.4.3 The CiPA initiative

However, the high false positive rate of this drug safety testing approach has prevented the further development of potential drugs [Sager et al., 2014, Colatsky et al., 2016]. To address this issue, that is to improve the assessment of a drug's proarrhythmia risk in the early drug development process, the Comprehensive *in vitro* Proarrhythmia Assay (CiPA) initiative has been proposed by a committee consisting of the U.S. Food and Drug Administration (FDA), the Health and Environment Sciences Institute (HESI), the Japan National Institute of Health Science (NIHS), and others

[Sager et al., 2014]. The CiPA committee identified four pillars: *in vitro* assessment on multiple ionic currents, *in silico* risk prediction, *in vivo* ECG biomarkers, and the use of human stem cell-derived cardiomyocytes [Sager et al., 2014, Fermini et al., 2015]. Towards this aim, the CiPA committee compiled a dataset of drugs [Colatsky et al., 2016], in each case providing their risk categories. This dataset comprises 28 drug compounds, categorised into high, intermediate, and very low proarrhythmic risk. These risk categories and the drug's risk are determined by the opinion of experienced professionals and published data. This dataset enables the training, testing, and validation of *in silico* models in development. Later, an *in silico* electrophysiological model, coupled with a risk metric, was developed and validated, showing good prediction performance [Strauss et al., 2019]. This will be described in detail in Section 2.5.5. The CiPA committee further published a list of general principles to evaluate and validate risk prediction models [Li et al., 2020].

More recently, Questions and Answers (Q&As) to the S7B and E14 guidelines [ICH E14/S7B Q&A, 2022] have been published to provide updates to the safety guidelines. Even though the guidelines are still heavily focused on the inhibition of the hERG and *in vivo* QT studies, proarrhythmic risk prediction models, including *in silico* models, can be included in the integrated risk assessment.

The main focus of this thesis is on these *in silico* risk prediction models. Updates on the work by CiPA regarding the other pillars are not discussed here. Interested readers can look in Colatsky et al. [2016], Strauss et al. [2019], Li et al. [2020] for the progress of the CiPA initiative under the other pillars.

2.4.4 Other efforts towards improving proarrhythmic risk prediction

Outside of the CiPA committee, lots of effort has been put towards improving computational models of cardiac electrophysiology, and proarrhythmic risk prediction [Trayanova et al., 2024, Moreno and Silva, 2023, Mirams, 2022, Niederer et al., 2019]. Various research has been performed across different aspects, including computational model of ion channels, myocytes [Krogh-Madsen et al., 2017, Tomek et al., 2019] and the heart [Sahli-Costabal et al., 2020], the proarrhythmic risk metric and the design of experiments for information-rich data. The advancement in ion channel models is not limited to the hERG channel. It involves other ion channels such as the sodium channel [Moreno et al., 2016]. Clark et al. [2022] designed a voltage-clamp protocol to identify ion channel-specific kinetics and provides alternative perspective to the proarrhythmic risk prediction.

In addition to the use of APD_{90} as a proarrhythmic risk predictor, other metrics have been developed. This includes the qNet (the net charge during an action potential) [Dutta et al., 2017], the morphology of the action potential [Lee et al., 2023], the electromechanical window [Passini et al., 2019], the axis of arrhythmia [Heitmann et al., 2025] and others. The population of models approach [Passini et al., 2017] incorporates the variability of the population as a range of parameter sets and predicts the variability in drug response of the population.

2.5 Drug binding of the hERG channel

Drug effects on the hERG channel remain as the main predictor of the arrhythmic risk of drugs in the drug safety guidelines. Although it is recognised that drugs may affect multiple ion channels, which may compensate for the effect on single ion channels [Martin et al., 2004, Mirams et al., 2011], the focus of much research is still the hERG channel as it is highly susceptible to drug compounds [Vandenberg et al., 2012].

2.5.1 The hERG channel

The $K_V11.1$ channel is a voltage-gated ion channel with permeability for potassium ions that conducts the rapid delayed rectifier current I_{Kr}. The pore-forming subunit of the channel is encoded by the human Ether-à-go-go Related Gene (hERG) [Sanguinetti et al., 1995]. Throughout this thesis, we will refer to the $K_V11.1$ channel as the hERG channel.

The transmembrane protein of the hERG channel consists of four identical alpha subunits, each containing six transmembrane segments (S1-S6) [Wang and MacKinnon, 2017]. Within each subunit, the S1-S4 segments form the voltage sensor domain, while the S5-S6 segments contribute to the central pore domain. The complete channel has four voltage sensor domain (one from each subunit) surrounding a single, central pore domain formed by the S5-S6 segments of all four subunits. Figure 2.4A shows the structure of an open hERG channel by Wang and MacKinnon [2017]. Figure 2.4B shows a schematic of the hERG channel. The 'central cavity' between all subunits is the central ion conduction pathway. There are two isoforms of the hERG channel, i.e., hERG 1a and hERG 1b isoforms. It is suggested that the hERG channels in human tissue are mainly 1a/1b heterotetramers [Jones et al., 2004]. The hERG 1a isoform and hERG1a/1b heterotetramers have different gating kinetics and drug

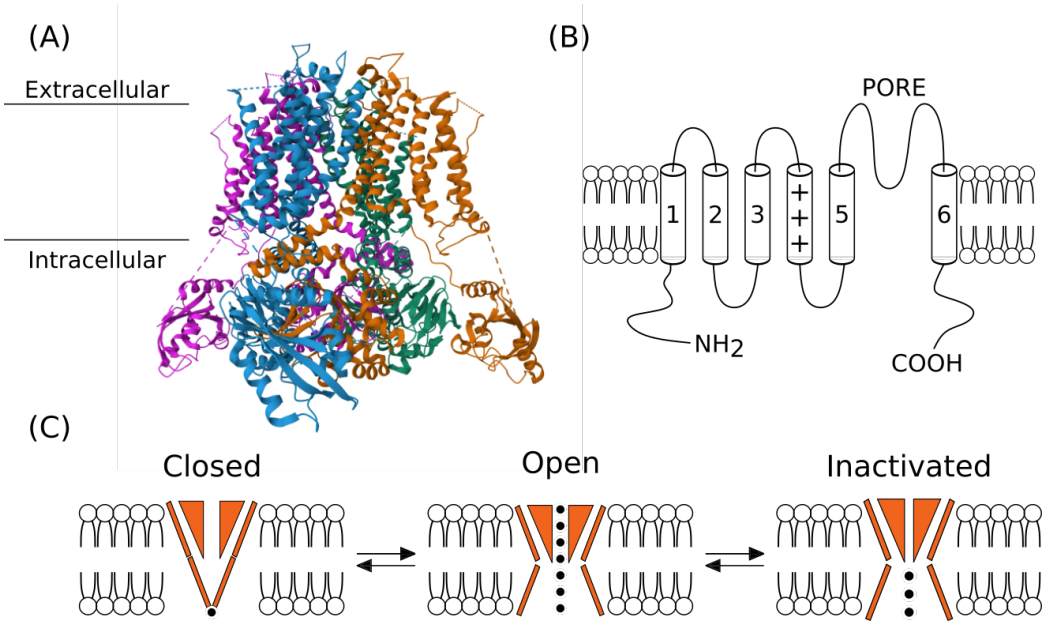


Figure 2.4: **The hERG channel and its gating mechanism.** **(A)** A cryo-EM structure of an open hERG channel with each subunit coloured differently. Created from the PDB file 5VA1 [Wang and MacKinnon, 2017]. **(B)** A schematic of the hERG channel lining the cell membrane. Based on Zünkler [2006]. **(C)** A schematic of the gating mechanism of the hERG channel. The black dots represent potassium ions. Based on Vandenberg et al. [2012].

sensitivity. The description of the hERG channel in this section will be based on the hERG1a isoform.

There are three conformational states of the hERG channel controlling its conduction pathway, i.e., open, closed, and inactivated states. A schematic of the hERG channel's conformation state is shown in Figure 2.4C. The opening and inactivation of the channel are voltage-dependent, with the inactivation process occurring much faster than the activation process. This plays a role in sustaining the plateau phase of the ventricular AP. Early in the AP during the depolarisation phase, hERG channels transition from the closed to the open state slowly, preventing a large efflux of potassium ions. The activation is followed by a rapid inactivation process, that occurs much faster than the activation. Therefore, at the plateau phase, most hERG channels are in the inactivated state, allowing a low potassium efflux. This contributes to the sustained plateau duration. At the repolarisation phase, hERG channels recover rapidly from inactivation as the membrane potential becomes more negative. The outward K^+ current actively drives the membrane potential back towards its resting state.

For a more detailed description of the hERG channel, see e.g. Vandenberg et al. [2012, 2017], Mitcheson et al. [2000a].

2.5.2 Mathematical models of hERG channel gating

With increasing understanding of the gating mechanism of the hERG channel and advancement in computational power, the mathematical models of the hERG channel have evolved accordingly (see Supplementary of Beattie et al. [2018] for a list of hERG channel models). The earliest model for the potassium current in cardiomyocytes is by Noble and Tsien [1969], but the different components of the potassium current were not identified. The simplest Hodgkin-Huxley formulation of the I_{Kr} consists of an activation gate and an inactivation gate, which is used in the ten Tusscher cell model [ten Tusscher et al., 2004, ten Tusscher and Panfilov, 2006]. The hERG channel of the O'Hara-Rudy cell model [O'Hara et al., 2011] has a structure that is modified from the classic Hodgkin-Huxley formulation to have a fast and a slow component of the activation gate, and also introduced a fast inactivation gate.

One of the earliest Markov formulations of the hERG channel was proposed by Wang et al. [1997a]. The authors claimed that three closed states were required to reproduce the kinetics of the channel, which is not possible using the Hodgkin-Huxley formulation. In the Wang et al. [1997a] model, hERG channels move through three closed states, before entering an open state, and with the possibility to enter an

inactivated state from the open state. Kiehn et al. [1999] further extended the Wang et al. [1997a] model to include a transition between the inactivated state and the "pre-open" closed state. It was speculated that the channels enter the inactivated states directly from the closed state, then rapidly recover from the inactivated state to produce bursts of opening of the channel because no current was detected before the recovery from inactivation. The direct transition from closed to inactivated state was also required to capture the transient peak during depolarisation. However, a comparison between competing hERG channel models suggested that the closed-inactivated transition is not required to reproduce the experimental data [Bett et al., 2011]. There are many other hERG channel models not described here. The various hERG channel models can be broadly categorised into seven model structures (see Supplementary of Beattie et al. [2018]). Given the diversity of the model structures describing the gating mechanism of the hERG channel, each parameterised to have different gating kinetics, each model will produce different predictions. Mangold et al. [2021] shows a study identifying the structure of the ion channel, reflecting the difficulty of the problem. It will be further discussed in Section 2.6 on how the mathematical models are assessed for their credibility.

2.5.3 Drug-binding mechanisms

The hERG channel is highly susceptible to inhibition by drugs, and as mentioned above, I_{Kr} is responsible for the repolarisation of the action potential and thus inhibition of I_{Kr} will prolong the action potential duration and increase the potential for arrhythmia. The main binding site of the hERG channel lies within its inner cavity, with the activation gate covering the entry [Mitcheson et al., 2000a]. Therefore, the hERG channels have to open for drug compounds to enter the binding site [Kiehn et al., 1996]. It is suggested that the hydrophobic pouches at the side of the inner cavity of the hERG channel [Wang and MacKinnon, 2017] increase the size of the pore cavity and allows the accommodation of larger-sized drugs, potentially explaining the high susceptibility of the hERG channel to so many drug compounds [Vandenberg et al., 2017]. Additionally, the residues Y652 and F656 on the S6 helices facing the inner cavity interact with a wide range of drug compounds, leading to the high susceptibility of the hERG channel [Mitcheson et al., 2000a].

The modulated receptor hypothesis states that the association and dissociation rate of drugs within the channel are regulated by the states of the channel, i.e., each state has its own characteristic rate independent of other rates [Hondeghem, 1987]. The *association rate* refers to the rate at which the drug compound binds to the

channel and forms a complex, while the *dissociation rate* refers to the rate at which the drug compound unbinds from the channel and the drug/channel complex separates. The binding rate of drugs when the channel is inactivated may not be the same as when it is open. Early results show that dofetilide (a hERG channel blocker) has reduced affinity during strong depolarisation, when channels are inactivated [Kiehn et al., 1996, Snyders and Chaudhary, 1996]. Later, similar behaviour is shown in other drugs, where the affinity of drugs, including dofetilide, is altered when the inactivation of the channels are impaired through mutations [Ficker et al., 1998, Wang et al., 1997b, Lees-Miller et al., 2000], indicating that the inactivated state plays a role in the binding of drugs. The inactivated state is indeed the preferential conformation state of the channel for the binding of some drugs [Perrin et al., 2008]. One hypothesis for such a preference is that the access to the binding site is favourable at the inactivated state [Chen et al., 2002].

In addition to the state-dependence of the drug-binding of the hERG channel, drugs can also be *trapped* within the hERG channel [Carmeliet, 1993, Kamiya et al., 2006, Mitcheson et al., 2000b, Stork et al., 2007, Windisch et al., 2011], indicated by a slow recovery from block. The dissociation of drugs from the channel can occur in two phases: a rapid phase, where drugs unbind from the open channel, and a slow phase, where drugs remain within the channel after the closure of the activation gate during repolarisation. This slow phase can be seen as *drug-trapping*. The other possible mechanism of drug-trapping is the foot-in-a-door mechanism, where drugs prevent the closure of the activation gate and delay the deactivation process [González et al., 2002, Armstrong, 1971]. The trapping of drugs within the channels may accumulate when drug compounds bound to the channel remain within the channel, and more drug compounds bind to the channels at repetitive activation. This is also termed *use-dependent block* to indicate the dependence of the effect of block on the rate of successive activations [Vandenberg et al., 2012].

2.5.4 Drug-binding models

As discussed in Section 2.4, drugs undergo safety testing to assess their potential to inhibit the hERG channel because of its association with proarrhythmic risk. We can utilise the modelling techniques discussed in Section 2.2 to capture the drug-binding mechanism into mathematical models and predict the proarrhythmic risk of drugs.

Many have used ligand-based approaches, such as quantitative structure-activity relationship (QSAR) models and statistical classification models, to identify structural features that correlate with drug block [Kalyaanamoorthy and Barakat, 2018]

and identify drugs that will inhibit I_{Kr}. A challenge for this approach is the lack of a complete understanding of the binding mechanism [Kalyaanamoorthy and Barakat, 2018]. The recent publication of the structure of the drug-channel complex [Asai et al., 2021, Miyashita et al., 2024] may provide insights into the binding modes of the hERG channel. Moreover, these approaches focus on the binding ability of drug compounds to the channel, and not their translated effect to QT prolongation or arrhythmic risk.

The action potential models described in Section 2.2 can be used to address this gap [Mirams et al., 2011]. Various modifications can be made to the ionic current component of the AP model (Section 2.5.2) to incorporate the effect of drug block and replicate some arrhythmic risk markers, such as the action potential duration, the TRIaD indicator [Hondeghem, 2006], and the axis of arrhythmia [Heitmann et al., 2023]. The simplest approach to incorporate a drug-block into an AP model is by reducing the ionic conductance of I_{Kr} (scaling the magnitude of the current). We refer to this approach as the *conductance scaling* formulation. An example of the use of the conductance scaling formulation is Davies et al. [2011]. This can be expressed as a scaling factor in the ionic current equation (Equation 2.6)

$$I_j = \bar{g}_j \cdot b \cdot O \cdot \delta_j, \tag{2.26}$$

where $b \in [0, 1]$ is the scaling factor. The intensity of the drug block can be derived using principles of mass-action kinetics

$$\mathrm{R} + n\mathrm{L} \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} \mathrm{RL}_n, \tag{2.27}$$

where R is the channel receptor (binding site of the channel), L is the ligand (drug molecules), n is the number of drug molecules, and k_{on} and k_{off} are association and dissociation rates, respectively. It is assumed that the drug molecules and the channel receptor are uniformly distributed and constantly available. The proportion of drug-free channels (no drugs bound to the channel) $b = [\mathrm{R}]/([\mathrm{R}] + [\mathrm{RL}_n])$, where square brackets denote concentrations, follows the equation

$$\frac{\mathrm{d}b}{\mathrm{d}t} = -k_{\text{on}} b [\mathrm{L}]^n + k_{\text{off}}(1 - b). \tag{2.28}$$

We then have the steady state b_∞ when $\dfrac{\mathrm{d}b}{\mathrm{d}t} = 0$ as

$$b_\infty = \frac{[\mathrm{IC}_{50}]^n}{[\mathrm{L}]^n + [\mathrm{IC}_{50}]^n}, \tag{2.29}$$

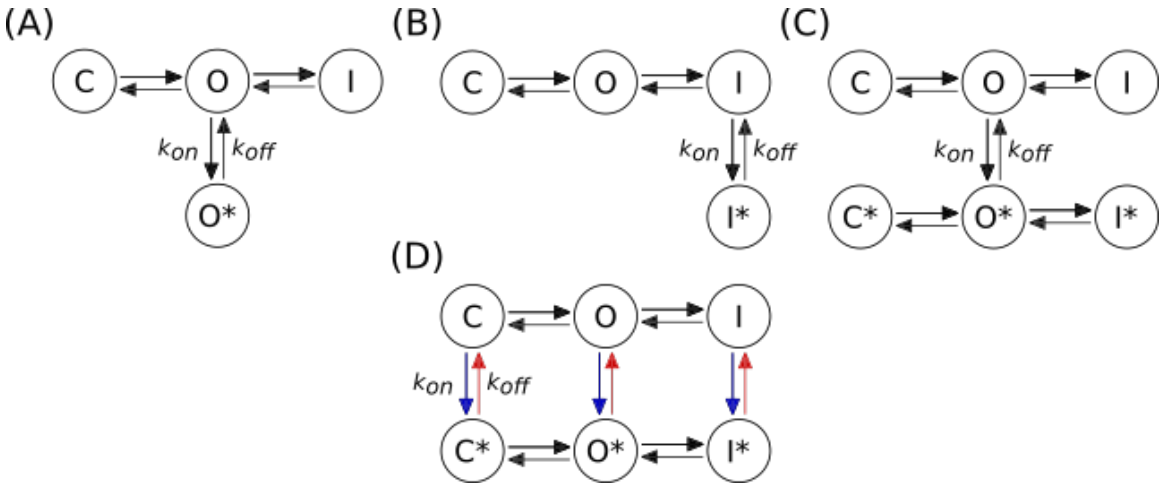


Figure 2.5: **Visualisation of drug-binding mechanism into Markov structures.** A simple 'closed-open-inactivated' formulation is used as example for the ion-channel formulation to demonstrate the various aspects of the drug-binding mechanism. The asterisks indicate drug-bound states. Figure adapted from [Mirams, 2022].

where $[\text{IC}_{50}]^n = k_{\text{off}}/k_{\text{on}}$ with $[\text{IC}_{50}]$ being the concentration when the inhibitory effect is 50%. This equation yields the *Hill equation*, with n being the *Hill coefficient* [Hill, 1910]. This equation gives the scaling factor or response of a drug to the current given a drug concentration. The IC_{50} and Hill coefficient n may be estimated from patch-clamp experiments by measuring the percentage of current inhibition from the control condition for several concentrations and fitting to the Hill equation. The estimated Hill coefficient n is usually a non-integer, even though it should be an integer based on the derivation, as it counts the number of drug molecules that bind to the channel. However, the uncertainty around the value is large. Some may assume n to be 1 to minimise the uncertainty.

The binding mechanism of drugs to the hERG channel is more complex than simply scaling the ionic conductance. Some aspects of the drug-binding mechanism, such as state-dependency and drug trapping (discussed in Section 2.5.3), are not considered in the conductance scaling formulation. We can use extensions to the Markov structure of an ion channel (Equation 2.11) to visualise the modelling of these complex drug-binding processes. We know that drug compounds can bind to the open state of a channel. In the Markov structure, we can add an additional state, for example a drug-bound state (O*), that transitions from the open state (O) (Figure 2.5A). The transition rate into the drug-bound state is termed the association rate k_{on}, and the transition rate back to the open state (*unbinding rate*) is termed the dissociation rate k_{off}.

Following the modulated receptor hypothesis, the state-dependency of the drug-binding can be modelled by having a different k_{on} for each state of the channel. A stronger assumption to the modulated receptor hypothesis is the *guarded receptor hypothesis*, where drugs bind only to a single state of the channel and the association rates k_{on} to the other states are zero [Starmer et al., 1984]. Examples of this are shown in Figures 2.5A and B, where drugs only bind to either an open state or an inactivated state, respectively. The guarded receptor hypothesis has been shown to replicate the use-dependency of drug block [Starmer et al., 1984]. This aligns well with the relation between the foot-in-a-door drug-trapping mechanism and the use-dependent block. This formulation can be viewed as the drug-bound state being unable to transition to other conformational states, which translates to the prevention of closure of the activation gate of the foot-in-a-door drug-trapping. These hypotheses introduce parameters and options to various model structures, thus increasing the difficulty in identifying the various rates and distinguishing between the binding modes. A potential approach is to carefully design voltage-clamp protocols to infer the ion channel kinetics and the state preference of drug binding [Jennings et al., 2024, Lee et al., 2019, Gomis-Tena et al., 2020, Escobar et al., 2022]. Some examples of incorporating state dependence in the drug-binding models are Jennings et al. [2024], Lee et al. [2016] As discussed in Section 2.5.1, the conformational states of the hERG channel are voltage-dependent, which allows the design of voltage-clamp protocols to identify the state preference of drug binding. However, this dependency coincides with the voltage-dependency of drug binding, making it difficult to differentiate between the two. Therefore, the state dependency and voltage dependency are frequently used to mean the same thing.

Another aspect of the drug-binding mechanism to take into consideration is drug-trapping. Unlike the foot-in-a-door drug-trapping, drug compounds that fit into the inner cavity of the channel leave the movement of the activation gate unaffected, therefore allowing conformational changes of the channel even in the presence of drugs. This can be translated to the Markov structure as transitions between the drug-bound states, and keeping the transition rates the same as those of the drug-free channel (Figure 2.5C). While drug-trapping of smaller molecules does not alter the conformational changes of the channel, larger compounds can prevent certain conformational changes of the channel. Furthermore, drugs can have an impact on the channel gating (see e.g., Balser et al. [1996], Huang et al. [2021], Zhou et al. [2011]). In Markov structures, this is reflected as different transition rates between

the conformational changes in drug-free channels and drug-bound channels. Some examples include Di Veroli et al. [2014], Gomis-Tena et al. [2020].

The conductance scaling formulation can also be visualised with a Markov diagram, as having a mirror image to the hERG channel model structure, as it does not interfere with the conformational changes of the ion channel when the drug is bound, with the k_{on} and k_{off} rates transitioning from the states of the channel to their corresponding drug-bound states (Figure 2.5D).

In addition to the modification to the Markov structure to reproduce the different aspects of the drug-binding mechanism, there are other factors affecting the transition rates, such as temperature and pH values. Temperature affects the channel behaviour [Zhou et al., 1998, Vandenberg et al., 2006, Lei et al., 2019a] and the binding kinetics (k_{on} and k_{off}) [Kirsch et al., 2004, Yao et al., 2005, Windley et al., 2016, Lo and Kuo, 2019]. These effects can be incorporated by changing the transition rates among the channel states [Li et al., 2016, Lei et al., 2019a] and the binding kinetics based on temperature. The pH value has also been shown to change the channel behaviour and its binding kinetics [Grant et al., 1980, Du et al., 2011, Wang et al., 2016], which can also be accommodated by adjusting the binding kinetics accordingly.

Finally, a note on the interaction of multiple drug compounds on a single channel. In settings of more than two types of drug compounds, the inhibition effect may be enhanced or reduced. An example of reduced inhibition effect is shown in Clarkson and Hondeghem [1985]. Due to competitive displacement between the drugs at the channel, it is unlikely for the drug-drug interaction to be modelled by the *simple additivity* model, where the combined effect is a summation of percentage block with a limit of 100%. The other models are the *Bliss independence* model and the *Loewe additivity* model. The Bliss independence model [Bliss, 1939] assumes the drug-binding process is independent and no competition for binding sites. This is implemented by the summation of percentage blocks and removing the double-counted interaction between the drugs. Similar to the Bliss independence model, the Loewe additivity model also assumes the drugs do not interact, but compete for the same binding site. These drug-drug interaction models are discussed in detail in [Wiśniowska and Polak, 2016]. These models require further verification with experimental observations.

Similar to the mathematical model of the hERG channel, due to the difficulty in identifying the model structure that captures the drug-binding mechanism, there are many proposed models in the literature, with a combination of the formulations described above. Lei et al. [2024] compared 13 models of the drug-binding mechanism and showed large uncertainty in model output, leading to different risk predictions for

some of the drugs. We will discuss in Section 2.6 methods to evaluate the uncertainty and assess the credibility of the model outputs. For more details on the modelling of drug action, see e.g. Mirams [2022], Brennan et al. [2009], Mirams et al. [2012].

2.5.5 Example of a drug-binding model

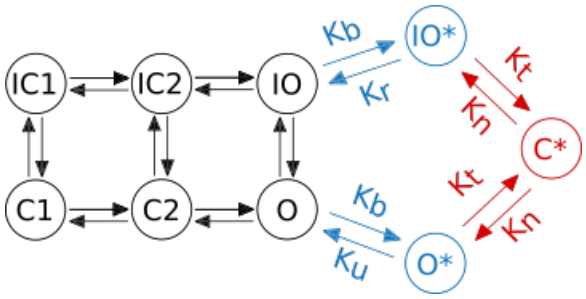


Figure 2.6: **The Markov structure of the SD drug-binding formulation (blue and red) [Li et al., 2017b] extended from the Li hERG-channel model in black [Li et al., 2016].** The blue and red parts refer to the binding and the trapping components, respectively.

Here we show an example of how all of the approaches (Section 2.5.4) can come together to give a formulation for drug-binding, and also how variability in these drug-binding models can arise. The drug-binding formulation by Li et al. [2017b] is taken as an example and is the main drug-binding formulation used throughout this thesis, and so we explain the model below in some detail. We refer to this drug-binding formulation as the *state-dependent (SD) drug-binding formulation*. Figure 2.6 shows a Markov diagram of the SD formulation in blue and red. The SD formulation is extended from the Li hERG-channel model [Li et al., 2016] (black parts of Figure 2.6), a choice made by the authors. Other structures of the hERG-channel model are possible.

The Li hERG-channel model is formulated with two inactivated-closed states (IC1 and IC2), two closed states (C1 and C2), an inactivated-open state (IO), and an open state (O). The SD formulation incorporates state-dependency and drug-trapping, visualised as the blue and red parts of Figure 2.6, respectively. The state-dependency of drug-binding arises from the binding of drugs to the open and the inactivated-open states, visualised by the extension from the hERG-channel model (black parts of Figure 2.6) to drug-bound states (indicated by asterisks) only from the O and the IO states. This binding component is indicated by the blue parts of Figure 2.6. Once the channel enters these two states, closure of the channel is allowed, reflecting

drug-trapping, where conformational changes of drug-bound channels are possible. At the closure of drug-bound channels, the channel enters the trapped state (C*). This trapping component is indicated by the red parts of Figure 2.6.

The transition rates among the drug-bound states are the binding rate (K_b), the unbinding rate (K_u), the trapping rate (K_t), the untrapping rate (K_n) and the microscopic reversibility rate (K_r). The transition rates are defined as follows

$$K_b = K_u \times K_{\max} \times \frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{2.30}$$

$$K_n = K_t \times \frac{1}{1 + e^{-(V - V_{\mathrm{half-trap}})/6.789}}, \tag{2.31}$$

$$K_t = 3.5 \times 10^{-5}, \tag{2.32}$$

where $K_{\max}$ is the maximum drug response on the binding rate, D is the drug concentration, n is the Hill coefficient, EC_{50} is the concentration at which half maximal binding rate is achieved, and V is the membrane voltage. The SD drug-binding formulation can be viewed as having a binding component and a trapping component. The binding component includes the binding rate K_b, the unbinding rate K_u, and the microscopic reversibility rate K_r, shown in blue in Figure 2.6. The trapping component includes the trapping rate K_t and the untrapping rate K_n, shown in red in Figure 2.6.

Drugs bind to the channel at a binding rate K_b that is dependent on drug concentration, following a sigmoidal function, as described in Equation 2.30. Note that EC_{50} defined in the SD formulation is different from IC_{50} of the Hill equation in Equation 2.29. The IC_{50} in the Hill equation refers to the inhibitory effect on the current, while EC_{50} in the SD formulation refers to the inducing effect on the binding rate. The drug-bound channel (in the O* and IO* states) can then close, trapping the drug compound in it, at a trapping rate that is assumed by Li et al. [2017b] to be a constant. This trapping rate is the same for any drug compound. The value of the trapping rate $K_t = 3.5 \times 10^{-5}$ (Equation 2.32) is chosen by Li et al. [2017b], out of five candidates: 1×10^{-6}, 2×10^{-5}, 3.5×10^{-5}, 5×10^{-5} and 1×10^{-4}. The criteria for choosing the optimal trapping rate K_t are determined by its fit to the experimental data and the separation of drugs based on their trapping tendency, indicated by the steady-state ratio between the O* and C* states at a resting potential of $-80\,\mathrm{mV}$ [Li et al., 2017b]. Drug compounds then exit the trapped state (C*) at an untrapping rate K_n that is voltage-dependent and follows Equation 5.2. The term $\frac{1}{1+e^{-(V-V_{\mathrm{half-trap}})/6.789}}$ is taken from the steady-state activation function for I_{Kr} in the O'Hara-Rudy model [O'Hara et al., 2011], but with $V_{\mathrm{half-trap}}$ replacing a constant

-8.337 in the original equation. The untrapping rate K_n can also be viewed as the opening or activation of the channel when drugs are bound within it, potentially explaining the choice of the activation gate for the untrapping rate. To differentiate between trapped and nontrapped drugs, trapped drugs have a slower transition rate to get out of the trapped state than nontrapped drugs, thus staying in the trapped state longer. Finally, this formulation has to satisfy microscopic reversibility, and that is chosen by Li et al. [2017b] to be enforced through the unbinding rate from IO* to IO (K_r).

As demonstrated by the SD formulation, there are numerous choices available for the structure of the drug-binding formulation and the definition of the transition rates, resulting in a variety of mathematical formulations that describe drug binding to ion channels. However, since the SD formulation is so influential, its detailed study will be the primary focus of this thesis.

2.6 Credibility assessment of models: parameter identifiability, sensitivity, and uncertainty

Considering models of the action potential of cardiomyocytes, the ionic currents of the hERG channel, and the drug-binding to the hERG channel across the three levels, we can develop a mathematical model to describe the effect of drugs on the action potential and further predict the risk of drugs in causing arrhythmia. However, due to the difficulty in capturing the physiological detail at every level, there are many diverse models describing the physiology at each level. This leads to large variability in the model outputs. Specific to our interest is the reliability of the prediction of proarrhythmic risk. It may have serious consequences if the predictions are unreliable or not robust to changes.

While we are able to develop a mathematical model that describes the action potential of cardiomyocytes with details up to the action of drugs on ion channels, all such models are still a simplified representation of the biological system. For example, at the action potential level, the action potential is approximated by a number of (not all) constituent current components. At the ion channel level, the channel behaviour is approximated by the ensemble average, giving the macroscopic current, and thus simplifies the stochasticity of the channel to a deterministic model. At the drug-binding level, most drug-binding processes are inferred based on experimental outputs without direct information. Even for parameters that are measurable, the

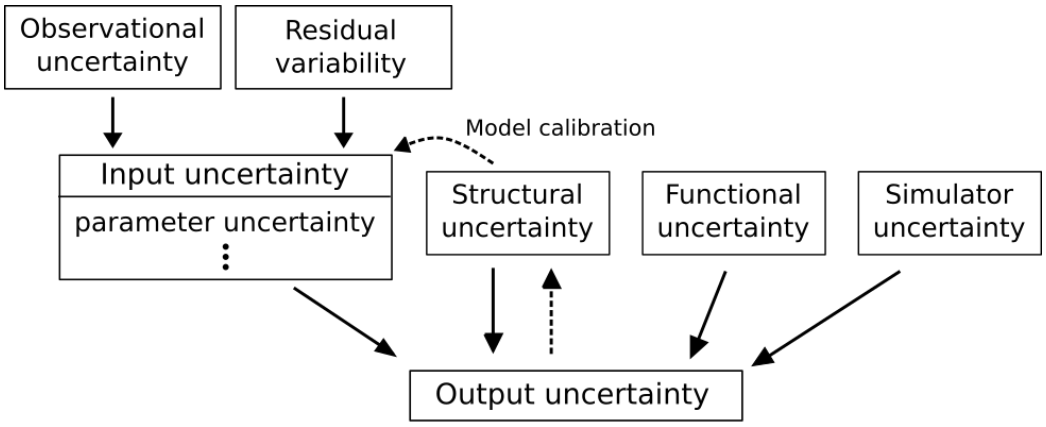


Figure 2.7: **Sources of uncertainty and their influence on the output uncertainty.** The dotted arrows show the reverse influence of output uncertainty on structural uncertainty and input uncertainty through model calibration. Figure adapted from [Mirams et al., 2016].

measurements are uncertain and vary randomly. For example, the concentration-effect curve, used to infer IC_{50}s and the Hill coefficients, has substantial variability [Elkins et al., 2013]. We therefore have uncertainty in both the model equations and the parameter values.

2.6.1 Sources of uncertainty in cardiac modelling

The variability and uncertainty in models of cardiac physiology are identified and classified by Mirams et al. [2016]. Here, we adapt Mirams et al.'s approach slightly to narrow down the focus on action potential models and their sublevel models. This variability and uncertainty can be broadly classified into *observational uncertainty* and *residual variability*. Observational uncertainty arises from experimental error and the lack of knowledge or information, such as those mentioned above, while residual variability refers to the variability between cells (*extrinsic variability*) and the temporal variability within a single cell (*intrinsic variability*). For example, the extrinsic variability in cell size leads to variability in cell capacitance, and intrinsic variability from the stochasticity of the channel opening, which may propagate to the beat-to-beat differences in action potential duration [Zaniboni et al., 2000]. Both the observation uncertainty and residual variability contribute to uncertainties in the model input, mainly the *parameter uncertainty* in our case, and the model equations, which is termed *structural uncertainty* or *model discrepancy*. In practice, the models are used outside of the conditions of calibration data, and the uncertainty surrounding

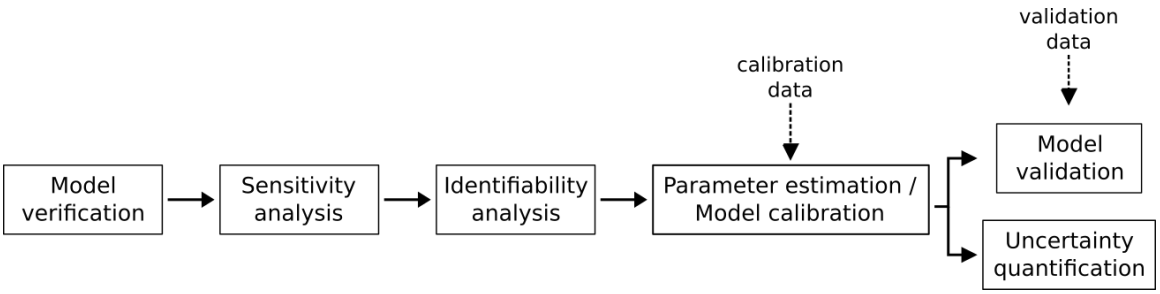


Figure 2.8: **A workflow for model evaluation.** The model evaluation process is iterative and may be repeated at any step. Figure adapted from [Braakman et al., 2022].

this is the *functional uncertainty*. Finally, the numerical approximations in solving or simulating the model contribute to the *simulator uncertainty* [Pathmanathan et al., 2012]. All these uncertainties in the model input and the model structure will propagate through the model simulations and lead to uncertainties in the model output, which may affect the reliability of model predictions.

2.6.2 Model evaluation

The complexity and variability of the biological system make it impossible to develop a model that represents the biological phenomena exactly. We usually address this problem by approximating and simplifying certain aspects of the biological system; thus, the mathematical model does not fully recapitulate the biological system. To ensure that the predictions we make are reliable and robust, the models we design have to be evaluated and assessed for their credibility, specifically whether the question of interest is addressed, and if the predictions are trustworthy. The question of interest may refer to the action potential or proarrhythmic risk for our problem. Given that the cardiac electrophysiology models we are interested in are potentially used in clinical settings, it is important for the models to be assessed for their credibility. An evaluation framework is suggested to establish model credibility for mechanistic systems models in the context of quantitative systems pharmacology (QSP) [Braakman et al., 2022], but also adaptable to cardiac models. The proposed framework integrates sensitivity analysis and identifiability analysis into the *verification, validation, and uncertainty quantification* (VVUQ) framework that has been suggested to evaluate the credibility of cardiac models [Pathmanathan and Gray, 2013].

Following the model evaluation phase proposed in Braakman et al. [2022], the first step to evaluate a mathematical model is *model verification*, where the implementation and numerical simulation of the model are assessed to depict the mathemati-

cal/conceptual description of the model accurately. This includes code verification, which ensures the model implementation is free of bugs, and calculation verification which tests whether the approximation of the numerical solution is sufficiently accurate to the true solution of the model equations [Pathmanathan and Gray, 2014]. In cardiac modelling, human error from encoding mathematical models of tens of equations and an order of magnitude more parameters can be minimised by using models encoded in a universal language [Noble et al., 2012] and shared through a public repository. Many cardiac models are now encoded with the CellML markup language [Garny et al., 2008] and uploaded to the CellML repository [Yu et al., 2011] for easy shareability. For calculation verification, the errors from numerical approximation are well-studied and addressed in the numerical solver CVODE [Hindmarsh et al., 2005] (as discussed in Section 2.3), which in this work is wrapped using the Myokit software package [Clerx et al., 2016] for easy implementation in cardiac modelling settings.

After model verification, the proposed next step is to perform *sensitivity analysis* on the model to determine the sensitivity of model output to perturbations in the parameter values. Sensitivity analysis can provide a direction for designing future experiments to achieve a more accurate estimation of influential parameters and reduce the number of free parameters either by setting a fixed value for or removing non-influential parameters. Sensitivity analysis is the main topic of Chapters 3 and 4, and will be discussed in more detail in Section 2.6.3.

Parameter sensitivity is closely related to the next step of the model evaluation, i.e., the identifiability analysis. A parameter is *identifiable* when, for a given observational data, there is a unique parameterisation of the model. [Wieland et al., 2021, Preston et al., 2025]. A parameter is likely to be non-identifiable if the model output is insensitive to changes in that parameter. The identifiability of the model parameters is an important factor when considering the reliability of model predictions, especially when the prediction outcome is dependent on the parameter value [Whittaker et al., 2020, Heinrich et al., 2025a, Preston et al., 2025]. Parameter identifiability is also essential for parameter estimation. Estimated parameter values of nonidentifiable parameters may be random [Raue et al., 2013] and do not reflect the model behaviour it is intended to capture. Identifiability analysis can also be used in experimental design to optimise data collection and provide insights for possible model reduction [Bellu et al., 2007]. Parameter identifiability is the focus of Chapter 5, it will be discussed in more detail in Section 2.6.3.

In the model evaluation framework, after establishing parameter sensitivity and parameter identifiability, model calibration or parameter estimation is performed, as

we have a theoretical guarantee that the parameters can be estimated. These can be performed with methods mentioned in Section 2.3 to find the optimal parameter values that best capture experimental observations.

After parameter estimation, the models can now reproduce the experimental data used for calibration, and at least in a sense represent the biological system within the space of the calibration dataset. The next step will be to validate (the validation process of VVUQ) the model, i.e., to determine when the model is representative of the biological system in terms of the intended use of the model [Pathmanathan and Gray, 2013]. This allows us to rule out overfitting and evaluate the predictive accuracy of the model. The intended use of the model is also referred to as *context of use*. The validation is usually performed by comparing predictions of the calibrated model with data that is not used in the calibration process, but is still relevant to the context of use. However, cardiac models are usually developed to reproduce the action potential , rather than having a specific context of use. Therefore, it is usually validated indirectly by relying on other evidence, such as supporting evidence on the model assumptions and formulations, and confidence from calibration results against experimental data [Pathmanathan and Gray, 2018]. For example, the predictive capability of ion channel models is evaluated by data collected from experiments with multiple voltage-clamp protocols [Patten-Elliott et al., 2025].

The final step of the workflow is *uncertainty quantification*, to quantify the uncertainties of the model prediction. In this process, the uncertainties to the model inputs (discussed in Section 2.6.1) are characterised as probability distributions and propagated to the model outputs [Pathmanathan and Gray, 2013]. One approach would be to use Monte Carlo techniques, where samples are drawn from the probability distribution of the parameters (which can be approximated by a normal distribution) and simulations are performed to obtain a distribution of the model outputs. An *emulator*, a statistical approximation of the model outputs given the model inputs, can be used when the number of simulations using the Monte Carlo technique becomes impractical to perform. An example of uncertainty quantification for cardiac electrophysiology models is [Pathmanathan et al., 2015] on sodium channels. On the other hand, structural uncertainty is often difficult to address and has gained much attention recently. Lei et al. [2020] demonstrates the difficulty in addressing model discrepancy and presents an approach to address the uncertainty by adding a discrepancy term in the calibration process, but notes that the additional parameters will lead to identifiability issues. Shuttleworth et al. [2023], Lei et al. [2024] showed how

structural uncertainty in ion-channel models and drug-binding models, respectively, can affect model predictions.

This completes all steps of the model evaluation framework [Braakman et al., 2022], adapted to our problem. The model evaluation process is iterative and could be repeated several times, possibly not in full sequence, before arriving at the finalised model. For example, model improvement or reduction performed based on the sensitivity or identifiability analyses results would sometimes require a repeat of model verification.

2.6.3 Parameter sensitivity and identifiability

Parameter identifiability determines whether the parameter can be estimated from experimental data. When a parameter is unidentifiable, multiple or a range of parameter values can reproduce the same model output. There are commonly two notions of identifiability: structural identifiability and practical identifiability. Following the definition used in Wieland et al. [2021], structural identifiability refers to the property of a model as having a unique parametrisation for any given model output when there is an infinite amount of data with zero noise. Structural identifiability is a property that solely depends on the model structure itself and determines whether a parameter can be estimated at all. There are various approaches to assess structural identifiability, such as the symmetries approach, which utilises Lie group theory [Yates et al., 2009, Massonis and Villaverde, 2020], the power series expansion approach [Pohjanpalo, 1978], and differential algebra [Ljung and Glad, 1994, Pia Saccomani et al., 2003, Margaria et al., 2001], which are available as software packages [Bellu et al., 2007, Ligon et al., 2017, Meshkat et al., 2014]. These methods are reviewed in Raue et al. [2014], Chis et al. [2011].

While there is a clear definition for structural identifiability, there is no consensus on the definition of practical identifiability [Whittaker et al., 2020, Heinrich et al., 2025a]. In a broad sense, practical identifiability considers the identifiability of the model under practical circumstances, where the data are limited and noisy [Wieland et al., 2021, Heinrich et al., 2025a]. In Heinrich et al. [2025a], there are two branches of practical identifiability: one uses sensitivity-based methods and the other uses likelihood-based methods. However, Krivorotko et al. [2020], Miao et al. [2011] take structural identifiability, sensitivity, and practical identifiability to be the three groups of identifiability. Regardless of the categorisation of the parameter sensitivity within the identifiability framework, we take parameter sensitivity to refer to the extent to which changes in the parameter values affect the model output, following the

definition in Heinrich et al. [2025a], Hamby [1994]. This is as opposed to practical identifiability, which refers to the ability to infer parameters accurately as defined by a particular measure of spread, e.g., a confidence interval. Following these definitions, parameter sensitivity can be viewed as part of the 'forward problem', while practical identifiability is part of the 'inverse problem'.

The aim of a sensitivity analysis is usually to rank the parameters by the amount of influence each has on the model output. The most common method of sensitivity analysis is computing the sensitivity matrix, a matrix giving the change in the model output given changes in the model input, then analysed by methods such as the rank or linearity of the matrix [Heinrich et al., 2025a, Hamby, 1994, Krivorotko et al., 2020, Miao et al., 2011] (examples in cardiac models [Fink and Noble, 2009, Sher et al., 2013]). Sometimes it is not straightforward to compute the sensitivity matrix as it is made up of derivatives. A simple approach to sensitivity analysis is by using factorial design [Box et al., 1978, Hamby, 1994]. The factorial analysis is performed by taking a number of samples of each parameter of interest, then simulating the model output for all parameter combinations. These methods do not take into account the interaction between parameters. Random sampling-based methods can address this problem by generating a large set of model outputs from randomly sampled model inputs, then computing the correlations between the model inputs and model outputs [Hamby, 1994]. Correlation coefficients such as the partial correlation coefficient and Pearson correlation coefficient are suitable for linear correlation between model inputs and outputs, while rank transform-based measures such as the Spearman rank correlation coefficient and partial rank correlation coefficient are better suited for nonlinear but monotonic relationships between model inputs and outputs [Marino et al., 2008]. When the correlation relationship is nonlinear and non-monotonic, variance-based methods may be more suitable. Variance-based methods suggest segmenting the distribution of model inputs and determining their contribution to the variance of the model outputs. Some of the common methods are the Sobol' indices, the Fourier amplitude sensitivity test (FAST), and the extended FAST [Saltelli et al., 1999, Marino et al., 2008]. Examples of sensitivity analysis performed on cardiac electrophysiology models are Sher et al. [2013], Huberts et al. [2013b,a], Sarkar et al. [2012].

Parameters that have minimal influence on the model output will likely not be identifiable. Practical identifiability analysis assesses whether a parameter can be estimated with sufficient accuracy when the calibration data is noisy and contains missing data points. Practical identifiability analysis can provide insights into whether the model output contains sufficient information to estimate the true parameter value

[Heinrich et al., 2025a]. Methods to analyse practical identifiability of parameters are mainly profile likelihood [Raue et al., 2009, Kreutz et al., 2013, Wieland et al., 2021] and MCMC sampling, where the posterior distribution is used to inform practical identifiability (see e.g. Hines et al. [2014], Gallaher et al. [2018]). The profile likelihood is usually performed with a single parameter at a time. If a parameter is identifiable, the profile likelihood will have a maximum value close to the true parameter value. One way to improve practical identifiability is by optimally designing experiments to allow for better estimation of the parameters or calibration of the model [Fink and Noble, 2009, Beattie et al., 2018, Lei et al., 2019b], and hence reduce the uncertainty in the parameter values. Alternatively, practical identifiability can be improved by reducing model complexity based on prior knowledge [Tivay et al., 2020], sensitivity analysis [Smith, 2013], or profile likelihood [Pironet et al., 2019, Maiwald et al., 2016].

For more details on identifiability, see Wieland et al. [2021], Heinrich et al. [2025b], Preston et al. [2025], Heinrich et al. [2025a].

2.7 Summary

We have now described how to translate the biological phenomena of the cardiomyocytes, ion channels, and drug-binding to mathematical models, explained the increasing importance of these models due to their use in drug safety testing, and how to assess the credibility of the model predictions. This thesis focuses on evaluating a specific model of the drug-binding mechanism. In the next chapter, we will perform a sensitivity analysis on the model and determine the influence of each parameter on its prediction of the risk of drugs.

Chapter 3

Importance of modelling hERG binding in predicting drug-induced action potential prolongations

This chapter is adapted from Farm et al. [2023], of which I am the first author of. The paper is titled "Importance of modelling hERG binding in predicting drug-induced action potential prolongations for drug safety assessment", available at https://doi.org/10.3389/fphar.2023.1110555. The associated GitHub repository is at https://github.com/FarmHJ/importance-of-binding-mechanism. The code is archived on Zenodo at https://doi.org/10.5281/zenodo.7689046. I led the research of this paper by designing the methodologies, and performing all simulations and analysis of this study.

Overview:

This chapter presents an exploration of drug-binding to the human Ether-à-go-go-Related Gene (hERG) channel and the resulting action potential (AP) prolongations. Inhibition of I_{Kr} increases the proarrhythmic risk of drugs. Drug-trapping is hypothesised to further increase this risk. The state-dependent (SD) drug-binding formulation incorporates drug-trapping and is shown to improve the risk categorisation of drugs. In this chapter, we perform sensitivity analysis on the SD formulation, particularly to investigate the influence of drug-trapping on the AP duration (APD). We use the conductance-scaling (CS) drug-binding formulation, which does not include drug-trapping, as a reference to investigate the sensitivity of APD to drug-trapping. We show that drug-trapping does not influence the difference in APD between the two drug-binding formulations.

3.1 Introduction

In Chapter 2, we introduced the crucial role of I_{Kr} in coordinating the repolarisation of the membrane potential back to its resting state for the subsequent cycle of action potential (AP). Reduction of I_{Kr} delays the repolarisation process and lengthens the AP, thus potentially disrupting the AP when membrane potentials do not return to their resting state before the next pulse. I_{Kr} inhibition is associated with increased risk of arrhythmias, including Torsades de Pointes [Curran et al., 1995, Heist and Ruskin, 2010]. Moreover, the hERG channel is highly susceptible to block or functional inhibition by a variety of drugs [Mitcheson, 2008, Thomas et al., 2004, Vandenberg et al., 2012], making it a key marker for the proarrhythmic risk assessment of drugs [ICH S7B, 2005].

Drugs mainly bind to the inner cavity of the hERG channel, where its main binding site lies [Mitcheson, 2008]. When the channel closes, some drugs can remain bound, *trapped* within the central cavity, so that unbinding can only happen on a subsequent re-opening [Mitcheson et al., 2000b, Stork et al., 2007]. This indicates that the space within the cavity is sufficiently large to contain the drug compounds [Yellen, 1998, Mitcheson, 2008]. For example, dofetilide has been shown to remain bound when the hERG channel closes [Kamiya et al., 2006], while verapamil can still unbind from the hERG channel after repolarisation [Zhang et al., 1999]. Simulation studies have suggested that such *trapped drugs* may be more prone to cause arrhythmia [Di Veroli et al., 2014, Pearlstein et al., 2016], and have higher arrhythmic risk [Li et al., 2017b].

Trapped drugs remain within the channel during the intervals between successive channel-opening pulses and maintain the inhibition of currents till the subsequent pulse. The channels do not recover from drug blockage during these intervals. This feature is utilised to investigate drug-trapping experimentally, distinguishing them by the rate of recovery from drug block between pulses [Stork et al., 2007, Windisch et al., 2011]. Specifically, the voltage-clamp protocol, the *modified Milnes' protocol* [Li et al., 2017b], modified from Milnes' protocol [Milnes et al., 2010], is designed to have a long inter-pulse interval for nontrapped drugs to unbind from the channel and have the current amplitude restored on the next opening of the channel, while trapped drugs stay bound within the channel and the current remains diminished upon re-opening of the channel.

As described in Chapter 2, there are many ways to incorporate drug inhibition effects into ion channel models (see Chapter 2 for more details). The most straightforward way of doing this is to assume a certain fraction of a cell's channels are blocked,

and to scale the current's *maximum conductance* (or permeability) variable accordingly [see e.g., Davies et al., 2011]. Alternatively, drug effects can be modelled by changing channel transition rates [Tsujimae et al., 2007] or by adding new states to a multi-state channel model [Di Veroli et al., 2012] to reduce the average time spent in open channels. This strategy of adding states has also been used to study trapping [Di Veroli et al., 2014, Li et al., 2017b].

In this chapter, we will focus on a particular drug-binding formulation for hERG-current inhibition, the state-dependent (SD) formulation [Li et al., 2017b] introduced in Chapter 2. To reiterate, the drug-binding process described by the SD formulation can be viewed as a two-step process, the *binding* and the *trapping* (blue and red parts in Figure 3.1A, respectively). The transition rates and channel states coloured in blue in Figure 3.1A are referred to as the *binding module*, while those in red are referred to as the *trapping module*. A drug compound first binds to the ion channel (entering the blue module of the SD formulation in Figure 3.1A), then gets trapped within the channel when the channel closes (entering the red module in Figure 3.1A). The other drug-binding formulation we will use as a reference is the *conductance scaling* (CS) formulation. This formulation scales the maximal conductance of the ion channel to mimic the inhibition of the current. Both models are described in detail in Section 3.2.

The inhibition effect of drugs within the SD formulation is modelled by the transition rates among the drug-bound states (the O*, IO*, and C* states in Figure 3.1A). This is illustrated in Figure 3.1B, where we show the experimental data (orange) and model's response (blue) to the modified Milnes' protocol (top panel) when parameterised for dofetilide, a trapped drug (middle panel), and verapamil, a nontrapped drug (lower panel) for hERG [Li et al., 2017b]. Only signals at 0 mV are shown for both the experimental data and the model's response in Figure 3.1B. The drug concentrations of dofetilide and verapamil were 30 nM and 1 µM, respectively, giving 75–90 % steady state block of hERG. In both cases, I_{Kr} rapidly activates during the high-potential pulses at 0 mV, then diminishes as the drug binds to and blocks the open channels. In the trapped case of dofetilide, the drug stays bound during the interpulse interval, so that the current at the start of each subsequent pulse maintains at the diminished magnitude of the previous pulse, and this process continues until the drug-binding saturates. In the nontrapped case of verapamil, the drug dissociates during the interpulse interval and recovers to unbound states, so that each successive pulse shows a similar I_{Kr} response. This mechanism is illustrated further in Figure 3.1C, which shows the state-occupancy during a single pulse of the modified Milnes' protocol using the SD formulation. With high trapping tendency (dofetilide),

most channels stay in drug-bound (blues) states during the interpulse interval, while with low trapping tendency (verapamil), the channel rapidly returns to a closed or inactivated state. Figure 3.1D shows the response and state occupancies when a predetermined elongated AP signal is applied. In this case, both drugs remain bound between pulses of APs. Since the trapping tendency of the drugs is not strongly reflected within the state-occupancy under an AP-clamp, it raises the question of whether we need to include these complex mechanisms when simulating their effect at the AP level.

This relates to the idea of parameter sensitivity, where we are concerned with the extent to which changes in the parameter values affect the model output, following the definition in Heinrich et al. [2025a], Hamby [1994]. The sensitivity analysis we perform in this chapter investigates the importance or influence of the binding and trapping modules of the SD formulation in prolonging the AP. We first describe the main elements of the sensitivity analysis, covering the parameters of interest, the data-generating model, and the model output or prediction. The parameters of interest are those of the SD drug-binding formulation, i.e., the parameters describing the transition rates among the drug-bound states (the O*, IO*, and C* states in Figure 3.1A) (more details in Section 3.2). The data-generating model refers to the mathematical model that can simulate/generate the model output or prediction. The SD drug-binding formulation was developed to improve the prediction of proarrhythmic risk of drugs; thus, the model output we are interested in is the Torsades Metric Score [Dutta et al., 2017]. The Torsades Metric Score is the change in the net charge across the cell membrane during an action potential, computed at multiples of the C_{max} concentration. C_{max} is the maximal therapeutic concentration, estimated from the measured plasma concentration data of a given drug, and is measured from blood samples [Byers and Sarver, 2009]. Therefore, not all parameter combinations within the SD formulation correspond to an actual drug where C_{max} is known. To mitigate this, we take the action potential duration at 90% repolarisation (APD_{90}) as a replacement to the Torsades Metric Score because it can be computed for any parameter combination of the SD formulation and has been used as a surrogate for QT interval prolongation, a predictor for Torsades de Pointes risk [Redfern et al., 2003, Hammond et al., 2001]. The data-generating model for the sensitivity analysis is thus the AP model used by Li et al. [2017b] to simulate the drug inhibition effect of the SD drug-binding formulation on the action potential. The AP model used is the ORd model [O'Hara et al., 2011] with a hERG-channel component developed by the same authors as the SD formulation [Li et al., 2016].

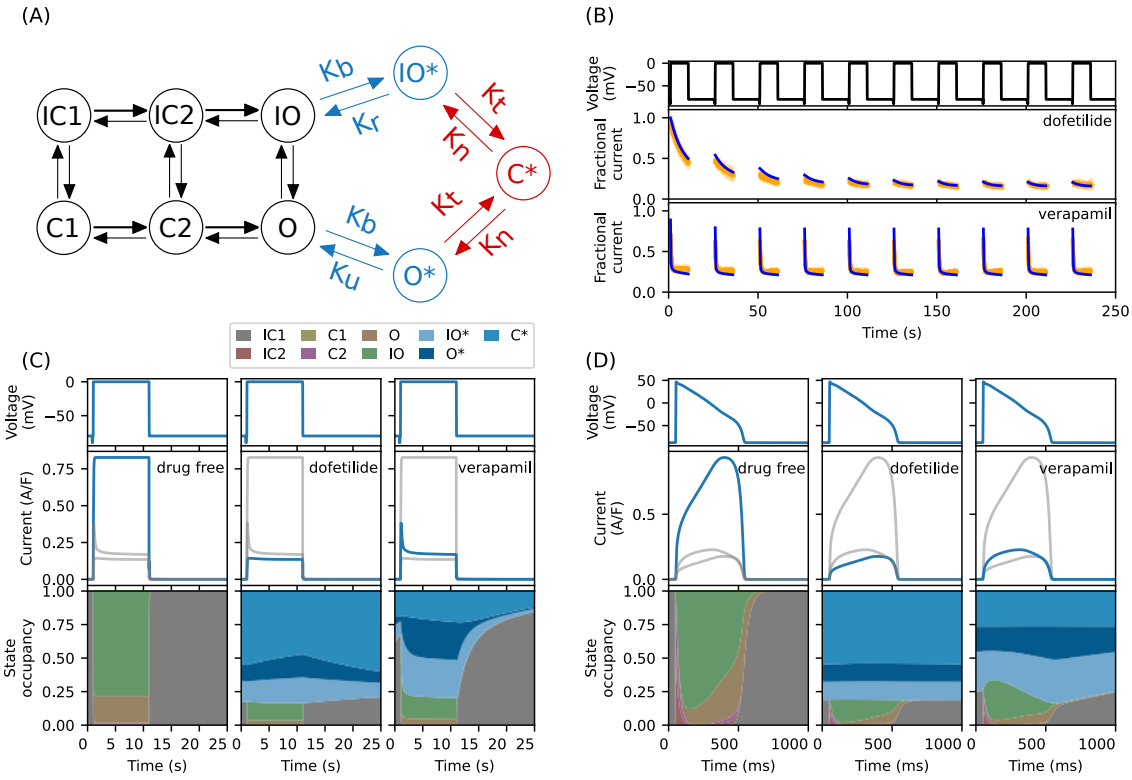


Figure 3.1: **Introduction of SD drug-binding formulation. (A)** The six-state hERG-channel component by Li et al. [2016] (in black) and its extension with the SD drug-binding formulation to model drug block and trapping by Li et al. [2017b] (blue and red parts). **(B)** Fractional block predictions for 10 paces of the modified Milnes' protocol (top) for the SD formulation with a trapped drug, dofetilide (middle), and a nontrapped drug, verapamil (lower). The SD formulation's simulation (blue) is overlaid on the experimental data (orange) from Li et al. [2017b]. **(C)** Steady-state response (after 1000 paces) for the SD formulation during a single step of the modified Milnes' protocol, for a drug-free (left), trapped (middle), and nontrapped configuration (right). The lower row shows the state occupancy: the fraction of channels in any one state at a given time. **(D)** Like (C), but using a predetermined elongated AP signal (1 s) instead of a rectangular pulse (25 s). For comparison, the grey lines in each column of (C) and (D) show the data from the other columns. Note that the labelling of the IC and C states, adapted from the original model [Li et al., 2017b], does not always correspond to their respective physical states.

Our objective for the sensitivity analysis is to investigate the parameters of the SD drug-binding formulation, specifically on how the binding module and the trapping module (blue and red parts of Figure 3.1A) prolong the AP. Our interest lies in whether the SD formulation is required to capture the AP prolongation as opposed to a simpler formulation, such as the conductance-scaling (CS) drug-binding formulation. Therefore, instead of comparing the change in APD_{90} with respect to changes in the parameter values, as is performed in a conventional sensitivity analysis, we take the difference in APD_{90} (ΔAPD_{90}) between the SD formulation and the CS formulation, treating the CS formulation as a reference point. This is to investigate the additional AP prolongation caused by the SD formulation that cannot be replicated by the CS formulation, and thus justifies the need for the SD formulation.

As mentioned in Chapter 2, the most common method of sensitivity analysis is computing the sensitivity matrix, a matrix giving the change in the model output given changes in the model input, which is then analysed by methods such as the rank or linearity of the matrix [Heinrich et al., 2025a, Hamby, 1994, Krivorotko et al., 2020, Miao et al., 2011] (examples in [Fink and Noble, 2009, Sher et al., 2013]). In this chapter, we perform a simple sensitivity analysis using factorial design [Box et al., 1978, Hamby, 1994]. The factorial analysis is performed by taking a number of samples of each parameter of interest, then simulating the model output for all parameter combinations.

We first consider two SD formulation parameterisations representing *synthetic drugs* with properties similar to dofetilide (high trapping tendency) and verapamil (low trapping tendency). We then consider three of the parameters of the SD drug-binding formulation, and explore the parameter space using parameter combinations drawn from the samples of each parameter of interest. For each parameter combination, we calibrate the drug-inhibition effect for the CS formulation accordingly. We then simulate the APD_{90} under the drug-inhibition effect by the SD formulation and the equivalent CS formulation. Finally, the APD_{90} predicted by the AP model with the SD formulation is compared to that of the CS formulation.

3.2 Methods

We first describe in detail the drug-binding formulations used in this study, then the calibration procedure to make them comparable. Next, we describe the hERG-channel component of the AP model, the cell-level AP model, synthetic drugs, and protocols used to compare the drug-binding formulations. Finally, we explain the details of the

sensitivity analysis performed on the state-dependent (SD) drug-binding formulation and the metrics used to quantify the difference in the APs.

As an overview, we take the SD formulation, use it as a reference to calibrate the ionic conductance of the conductance-scaling (CS) drug-binding formulation, then input these formulations into the hERG-channel component of the AP model chosen by [Li et al., 2017b, 2016] for AP comparison.

3.2.1 hERG-channel component

The hERG-channel component used in this study is by Li et al. [2016], with its structure shown in Figure 3.1A in black. The hERG-channel model is referred to as a *component* to distinguish it from the cell-level AP model. This hERG-channel structure and its transition rates are used in the development of the SD formulation [Li et al., 2017b]. It is a six-state Markov model with two inactivated closed (IC) states, two closed (C) states, an inactivated open (IO) state, and an open (O) state. Scaling the ionic conductance of the hERG-channel component, which is equivalent to inhibiting the hERG current, gives the CS drug-binding formulation. Extending the hERG-channel component with drug-bound states of open bound (O*) state, inactivated open bound (IO*) state, and closed bound (C*) state gives the SD formulation, as shown in Figure 3.1A (blue and red parts).

3.2.2 State-dependent drug-binding formulation of hERG

As described in Chapter 2, and repeated here for ease of reference, the SD formulation by Li et al. [2017b] captures the binding mechanism of drugs to the hERG channel. We refer to the drug-binding component as *formulation* to differentiate from the cell-level AP model and the hERG-channel component. The drug-binding formulation is described by the drug-bound states of O* state, IO* state, and C* state, and their respective transition rates, as shown in Figure 3.1A (in blue and red).

In this drug-binding formulation, compounds can bind to the O or IO states of the hERG channel and transition to the O* or IO* states, respectively, with a binding rate (K_b) given by

$$K_b = K_u \times E_{\max}(D), \tag{3.1}$$

$$= K_u \times K_{\max} \times \frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{3.2}$$

where K_u is the unbinding rate and $E_{\max}$ is a sigmoid function describing a drug's response at concentration D. The $E_{\max}$ function is scaled by $K_{\max}$, the maximum

response of a drug when it is at its saturating concentration. Finally, n is the Hill coefficient and EC_{50} is the drug concentration when the binding rate is at 50% of its maximum. The compound can be *trapped* in the C* state with a trapping rate (K_t) that is fixed at $K_t = 3.5 \times 10^{-5}\,\mathrm{ms}^{-1}$ [Li et al., 2017b]. The rate at which the drugs get *untrapped* (K_n) is defined as

$$K_n = K_t \times X(V), \tag{3.3}$$

$$= K_t \times \frac{1}{1 + e^{-(V - V_{\mathrm{half-trap}})/6.789}}, \tag{3.4}$$

where $V_{\mathrm{half-trap}}$ is the membrane voltage when K_n is half the trapping rate K_t. According to Li et al. [2017b], the function $X(V)$ is adapted from the hERG channel's steady state activation of the O'Hara et al. [2011] model, which is $x_{r,\infty} = 1/\left(1 + e^{-(V+8.337)/6.789}\right)$. A trapped drug can be modelled with either a high trapping rate (K_t) or a low untrapping rate (K_n), and in this case, $V_{\mathrm{half-trap}}$ is the only parameter that controls the ratio of trapping to untrapping rate. Finally, the transition rate from the IO* state to the IO state, K_r, is defined to maintain the microscopic reversibility of the system.

3.2.3 Calibration of a simple conductance scaling drug-binding formulation

To reveal how the SD formulation alters I_{Kr} and the AP, we compared the SD formulation with a simple instantaneous full-state drug-binding formulation, the CS formulation. To ensure that the SD and the CS formulations are comparable, we use simulated data from the SD formulation to calculate the appropriate scaling of the CS formulation to ensure that the peak I_{Kr} of both models is equal, as detailed below.

The CS formulation is equivalent to modelling the current inhibition by scaling the ionic conductance, thereby reducing the overall flow of ions across the channel. This formulation assumes that the half maximal inhibitory concentration (IC_{50}) of a drug on the hERG channel does not change with the experimental protocols. We note that the IC_{50} of a drug can be protocol-dependent as demonstrated in the later sections, and we acknowledge the differences in the results (Section 3.3.3). However, an investigation into such a dependence is not the aim of this study. Therefore, for the majority of this study, we use the modified Milnes' protocol [Milnes et al., 2010, see Section 3.2.6] to simulate the models as in Li et al. [2017b].

The ionic conductance of the hERG channel was scaled to capture the drug response, such that the peak current of the CS formulation under the modified Milnes'

protocol matched that of the SD formulation. A schematic of the fitting procedure is shown in Figure 3.2. We assumed that the SD formulation captures the experimental data presented in Li et al. [2017b], which we used as a reference point. For each *synthetic drug* (see Section 3.2.5), I_{Kr} was simulated at a range of drug concentrations using the SD formulation stimulated by the modified Milnes' protocol. The peaks of these I_{Kr} were normalised and then fitted to the Hill curve. The Hill curve parameters —the Hill coefficient and IC_{50}—were determined by minimising the mean squared error between the peak currents and the Hill curve. The ionic conductance of the CS formulation was then scaled with the drug response output obtained from the Hill curve function. The peak hERG-current of the CS formulation will then have the same peak current as the SD formulation for a given drug and concentration, as required.

The optimisations of the Hill curves to the peak of hERG currents were performed with the covariance matrix adaptation-evolution strategy (CMA-ES) algorithm [Hansen et al., 2003] in the PINTS software [Clerx et al., 2019].

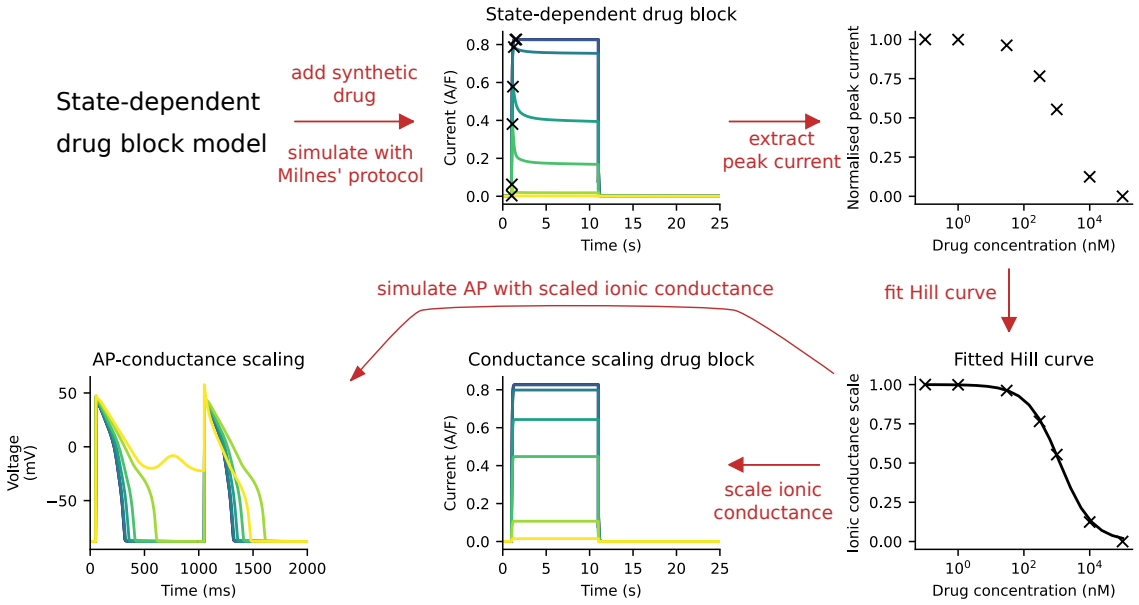


Figure 3.2: **A schematic of the calibration process for the CS formulation to the SD formulation.** First, simulations are run with the SD formulation using the modified Milnes' protocol at various drug concentrations. The peaks of the hERG currents are extracted and normalised. Subsequently, a Hill curve is fitted to the normalised data points. This Hill curve is then used to scale the ionic conductance of the CS formulation. Finally, the simulated APs of both models are compared.

Table 3.1: Parameter values of the SD formulation for all synthetic drugs, taken from Li et al. [2017b].

Drug	$K_{\max}$	K_u (ms^{-1})	EC_{50}^{n} (nM)	n	$V_{\mathrm{half-trap}}$ (mV)
Dofetilide	1.00×10^{8}	1.79×10^{-5}	5.483×10^{8}	0.9999	−1.147
Verapamil	4.646×10^{4}	7.927×10^{-4}	9.184×10^{6}	1.043	−100
Bepridil	3.735×10^{7}	1.765×10^{-4}	1.00×10^{9}	0.9365	−54.93
Terfenadine	9.884×10^{3}	8.18×10^{-5}	4.138×10^{4}	0.65	−77.49
Cisapride	9.997	4.161×10^{-4}	4.206×10^{1}	0.9728	−199.5
Ranolazine	5.584×10^{1}	1.929×10^{-2}	1.472×10^{5}	0.95	−94.87
Quinidine	5.770×10^{3}	1.00×10^{-2}	1.00×10^{6}	0.8311	−64.87
Sotalol	2.403×10^{3}	1.985×10^{-2}	9.619×10^{6}	0.7516	−55
Chlorpromazine	2.060×10^{5}	3.866×10^{-2}	5.677×10^{7}	0.8871	−14.57
Ondansetron	3.354×10^{4}	2.325×10^{-2}	9.950×10^{6}	0.8874	−82.11
Diltiazem	2.51×10^{2}	2.816×10^{-1}	1.00×10^{6}	0.9485	−90.89
Mexiletine	9.996	9.967×10^{-2}	2.308×10^{6}	1.304	−86.26

3.2.4 AP model

To study the effect of the hERG-channel component at the level of APs, the AP model by Dutta et al. [2017] was used. The model was based on the one used in the development of the SD formulation Li et al. [2017b]. To allow better quantification of each individual current on the AP, the ionic conductances of five currents were scaled [Dutta et al., 2017], including I_{Kr}, the slow rectifier potassium current (I_{Ks}), the inwardly rectifying potassium current (I_{K1}), the L-type calcium current (I_{CaL}), and the late sodium current (I_{NaL}). The AP model, when its drug-binding formulation is the SD formulation, is referred to as the AP-SD model; the AP model with its drug-binding formulation replaced with the CS formulation is referred to as the AP-CS model.

3.2.5 Synthetic drugs

The binding dynamics on the hERG channel of all twelve drugs in Li et al. [2017b] were taken as the synthetic drugs in this study. The parameters of the 12 synthetic drugs are provided in Table 3.1. These 12 drugs are referred to as *synthetic drugs*.

Here, dofetilide and verapamil were chosen as examples to compare trapping and nontrapping properties of the hERG channel, respectively. They have the highest and lowest $V_{\mathrm{half-trap}}$, excluding cisapride. The parameter combination of cisapride

cannot fully inhibit the hERG current at high concentrations, as it is limited by the low binding rate K_{max}, therefore cisapride is not taken as a nontrapped drug example. Verapamil, with the second lowest $V_{\text{half-trap}}$, is chosen instead. Although verapamil is a multi-ion channel blocker, we considered only its effect on the hERG channel, so any multi-channel effects on the AP were deliberately neglected, reflecting the focus of this study being only on the hERG-channel binding effects on the AP. The synthetic drug with parameters describing dofetilide is referred to as *example drug T*, while the synthetic drug with parameters describing verapamil is referred to as *example drug N*.

3.2.6 Protocols

Four voltage protocols are used in this work (Figure 3.8A). Firstly, we have the modified Milnes' protocol, modified by Li et al. [2017b] from that of Milnes et al. [2010]. This modified Milnes' protocol aims to differentiate between trapped and nontrapped drugs. The modified Milnes' protocol is repeated with a depolarisation step to 0 mV from the holding potential of −80 mV. The 0 mV step is held for 10 s before repolarising back to −80 mV for 15 s in-between pulses. While the depolarisation step allows the binding of drug compounds to the channel, the 15 s holding potential in-between pulses allows nontrapped drugs to unbind from the channel, thus allowing the current to recover to a drug-free condition. Experimental data obtained with this protocol were used to fit the parameters of the SD formulation [Li et al., 2017b]. Therefore, this modified Milnes' protocol is used as a reference for the comparison of the SD and CS drug-binding formulation.

The *Pneg80*, the *P0*, and the *P40* protocols from Gomis-Tena et al. [2020] are used to assess the dependency of the SD and the CS formulation comparison on the calibration protocol. These protocols were designed to maximise the channel's state occupancy in a certain state. When the hERG-channel component is stimulated by the Pneg80, P0, and P40 protocols, the channel will most likely be in the closed, open, and inactivated states, respectively, as discussed in Gomis-Tena et al. [2020]. The Pneg80 protocol is held at a holding potential of −80 mV, then depolarised to 20 mV for 0.5 s before a short pulse of −50 mV for 0.2 s. The time period of the protocol is 5.4 s. The P0 and P40 protocols are both held at −80 mV holding potential before depolarising to 0 mV and 40 mV respectively for 5 s. After that, a short pulse of −60 mV was applied for 0.2 s. Both these protocols have a pulse length of 5.2 s.

3.2.7 Sensitivity analysis

The behaviour of the SD drug-binding formulation is governed by its transition rates, which are defined by the five drug-dependent parameters — $K_{\max}$, K_u, $V_{\text{half-trap}}$, n, and EC_{50}^{n}. To understand the role of each drug-related parameter of the SD formulation, we performed a sensitivity analysis on the APD_{90} difference between the SD and the CS formulation.

First, we simplified the SD formulation by normalising the drug concentration D input to the EC_{50} value. The normalised drug concentration is therefore defined as

$$\tilde{D} = \frac{D}{\text{EC}_{50}}. \tag{3.5}$$

This reduces the number of parameters to four.

Second, we observed that the Hill coefficient n affects only the range of drug concentration where AP prolongation and early after depolarisation (EAD)-like behaviour are observed. To confirm this, we performed a sensitivity analysis on the Hill coefficient n for all synthetic drugs. We varied the value of n for each synthetic drug and monitored the APD_{90} differences between the two drug-binding formulations. The Hill coefficient n was sampled from a range of the minimum and maximum of n of all synthetic drugs. The APD_{90} differences were quantified by the root mean square difference (RMSD) (see Section 3.2.8). Figure 3.3A shows the distribution of the RMSD for each synthetic drug. Figure 3.3B shows the histogram of the difference in RMSD between the synthetic drugs (as reference) and those with their parameter n changed. The histogram has a mean of 4.313± 4.832 ms and a maximum of 25 ms(Figure 3.3). The SD formulation can thus be further simplified to three parameters, namely the $V_{\text{half-trap}}$, $K_{\max}$, and K_u. We then performed a detailed sensitivity analysis on this three-dimensional space.

In the sensitivity analysis, we repeated the comparison procedure as shown in Figure 3.2 for a range of parameter values within the three-dimensional space of the simplified SD formulation. The three parameters of the simplified SD formulation describe a hypothetical drug which we termed a *virtual drug*. Sweeping through the parameter space effectively performed a virtual drug screening for all possible drug-binding kinetics of hERG, within the SD formulation.

The two models, the AP-SD model and the AP-CS model, were considered to be similar if the RMSD of the APD_{90} values are small. This similarity threshold value is taken to be 30 ms. In the sensitivity analysis throughout the parameter space, a surface plane, where the RMSD values between the two models are within

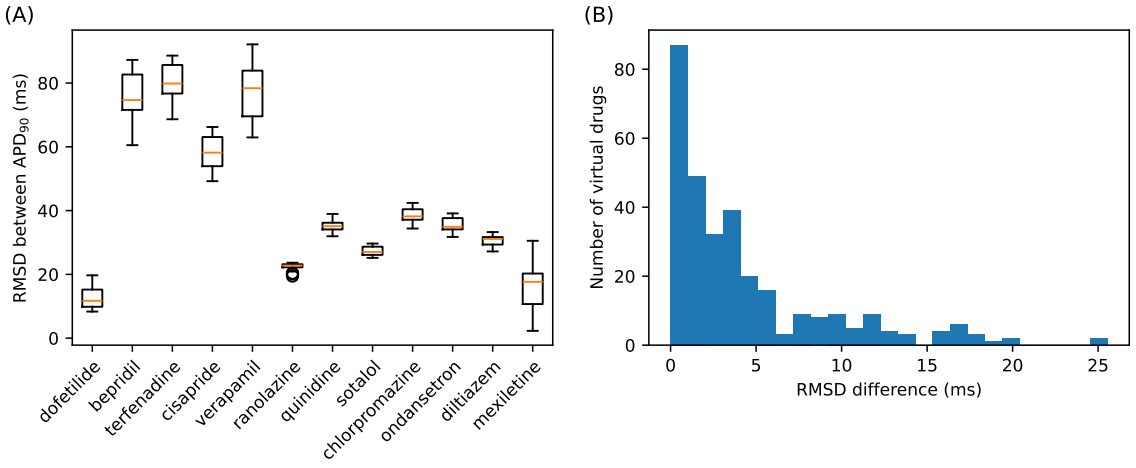


Figure 3.3: **RMSD is minimally affected by the Hill coefficient n** (A) The RMSD of each synthetic drug with the Hill coefficient n varied within the minimum and maximum of all synthetic drugs' Hill coefficients. (B) A histogram of the difference in RMSD with each synthetic drug's RMSD for all the synthetic drugs.

the threshold value, is formed. The parameter space around this surface plane was sampled more densely to obtain higher resolution.

3.2.8 Measuring the difference between the drug-binding formulations

In this study, a drug's effect on a cell, or more specifically on the hERG current, is determined by the prolongation of the AP, which we quantify by using the APD at 90% repolarisation (APD_{90}) and the *qNet* score. The APD_{90} acts as a surrogate to the QT interval, and is the time taken for the membrane potential to achieve 90% of repolarisation from depolarisation. The qNet score, defined by Dutta et al. [2017], is the sum of charge of a chosen list of currents — I_{CaL}, I_{NaL}, I_{Kr}, I_{Ks}, I_{K1} and the transient outward potassium current (I_{to}), which are the major currents in an AP. The qNet score is computed from an AP that is paced at 0.5 Hz with a stimulus duration of 0.5 ms and a time step resolution of 0.01 ms until steady state after 1000 paces. The qNet score is defined as

$$\text{qNet} = \int_t (I_{\text{CaL}} + I_{\text{NaL}} + I_{\text{Kr}} + I_{\text{Ks}} + I_{\text{K1}} + I_{\text{to}})\, dt, \tag{3.6}$$

approximated using the trapezium rule. High values of qNet are associated with a low risk of arrhythmia.

The difference between the drug-binding formulations will be measured by the signed root mean square difference (RMSD) in their APD_{90}s and qNet scores. The

signed RMSD is computed by the RMSD and the mean difference (MD). The RMSD is defined as

$$\text{RMSD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(SD_i - CS_i)^2}, \tag{3.7}$$

where N is the total number of drug concentrations tested, excluding points where behaviours similar to early after-depolarisation (EAD) are observed. We say an EAD-like event has occurred when the membrane potential does not fully repolarise before the next pulse of current stimulus. SD and CS can be APD_{90} or qNet scores from the SD and the CS drug-binding formulation, respectively. The MD, on the other hand, is taken as the average difference between the SD and the CS formulation, and is defined as

$$\text{MD} = \frac{1}{N}\sum_{i=1}^{N}(SD_i - CS_i). \tag{3.8}$$

A positive MD indicates larger APD_{90} or qNet score for the SD drug-binding formulation than the CS formulation. The signed RMSD is then defined as

$$\text{RMSD}_{\text{signed}} = \text{RMSD} \times \frac{\text{MD}}{|\text{MD}|}, \tag{3.9}$$

to encompass the magnitude and the directionality of the difference between the SD and the CS formulation. The signed RMSD between an AP-SD model and an AP-CS model will be represented using the symbol Δ for convenience, giving ΔAPD_{90} and ΔqNet for the signed RMSD of APD_{90} and qNet scores, respectively.

3.2.9 Simulations

All simulations, including voltage-clamp and current-clamp experiments, were run with Python 3.8 using Myokit 1.33 [Clerx et al., 2016] with the CVODE solver [Hindmarsh et al., 2005]. The absolute tolerance and relative tolerance were set to 10^{-7} and 10^{-8}, respectively. A summary of the computational details for all coding implementations of this thesis is provided in Appendix A.

3.3 Results

3.3.1 CS drug-binding formulation replicates SD formulation at steady state for a trapped drug example

Figures 3.1B and C show that the SD formulation can capture the trapping properties of a drug, as observed in the lack of recovery from block and the state occupancy of

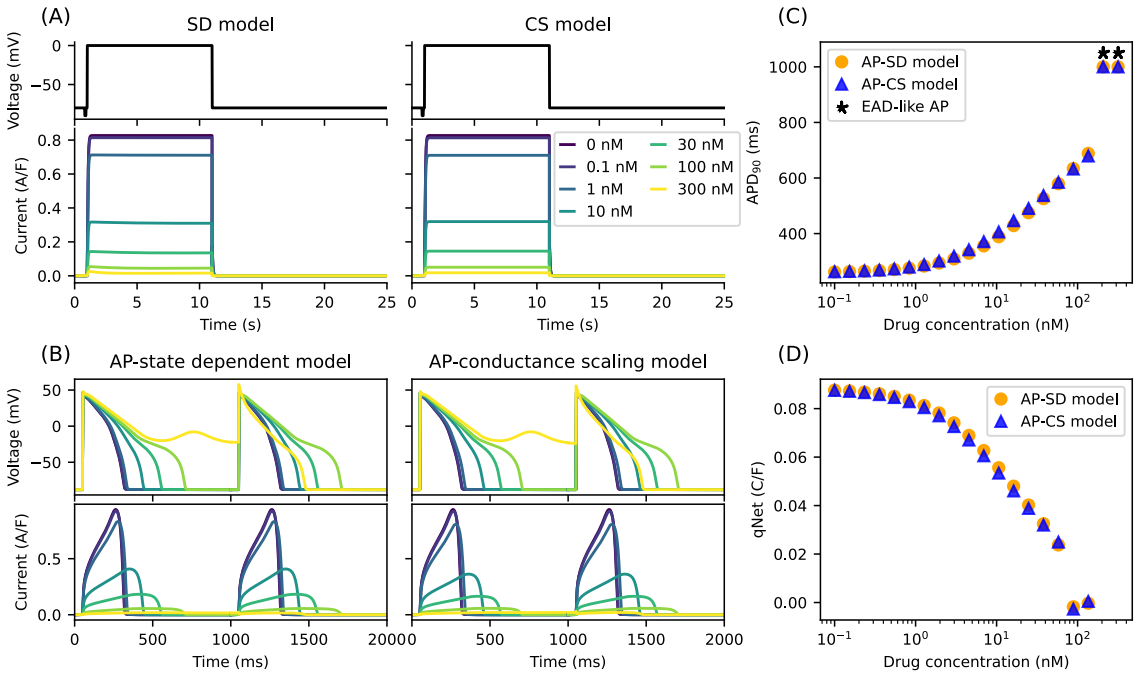


Figure 3.4: **The AP-CS model can provide a reasonable approximation of the AP-SD model for the example drug T. (A)** The matching of the peak hERG currents simulated by both drug-binding formulations under the modified Milnes' protocol (top row) at various drug concentrations of the example drug T. **(B)** The AP (first row) and its hERG current (second row) for the AP-SD model and the AP-CS model at steady state. **(C)** The APD_{90}s of the AP-SD model and the AP-CS model under the effect of the example drug T. APs that show EAD-like behaviour are indicated with asterisks. **(D)** The qNet scores of the AP-SD model and the AP-CS model under the effect of the example drug T. qNet values of APs with EAD-like behaviours are not shown.

drug-bound states. However, it is not clear whether a simpler AP-CS model could replicate the AP prolongation by the AP-SD model that has state dependency and a trapping mechanism included. We systematically compared the APs of the two models for the example drug T and the example drug N using the methods described in Section 3.2.3 and Figure 3.2.

Figure 3.4A shows the I_{Kr} of both the SD and the CS formulations under the modified Milnes' protocol with the example drug T present throughout. By design, the peaks of the I_{Kr} of the two formulations are the same at any drug concentration of the example drug T, which gives the same dose-response effect. The APs of the two models (the AP-SD model and the AP-CS model) under the effect of the example drug T are given in Figure 3.4B, showing two pulses of the AP at steady state, together with their corresponding I_{Kr}. As the drug concentration increases, the amplitude of the I_{Kr} decreases, and the APD increases. The two models show similar AP behaviours. A slight difference can be observed at 10 nM, where I_{Kr} of the AP-SD model has a slightly bigger amplitude and shorter period. The differences give rise to a slightly shorter APD because I_{Kr} aids in the repolarisation of APs. Moreover, both models show EAD-like behaviours at the same drug concentrations (>300 nM) of the example drug T.

In Figure 3.4C, APD_{90}s of the AP-SD model and the AP-CS model for the example drug T were calculated. APs that show EAD-like behaviour are indicated with asterisks. APD_{90}s of both models are similar, consistent with the APs shown in Figure 3.4B. The qNet scores, a proarrhythmic risk marker, of the example drug T shown in Figure 3.4D are similar for both models. The CS formulation can replicate the SD formulation for the example drug T at steady state.

We further compared the APs of the models for the example drug N (a nontrapped drug); the nontrapping phenotype of the drug is captured by the small value of $V_{\mathrm{half-trap}}$. Figure 3.5A shows the I_{Kr} of the two drug-binding formulations under the effect of the example drug N. The I_{Kr} simulated by the SD formulation shows dips after the initial increase at drug concentrations higher than or equal to ∼300 nM. Since the peak current is used to define the inhibition level of a drug, which is commonly done for IC_{50} calculations, the total amount of I_{Kr} (area under the curve) simulated by the CS formulation is higher than the SD formulation.

Figures 3.5B and C compare the APs and their corresponding I_{Kr} of the AP-SD model and the AP-CS model for the example drug N. The I_{Kr} and the APs are significantly different between the two models. At high drug concentrations, the AP-SD model shows longer APDs and lower I_{Kr} amplitudes than the AP-CS model. The

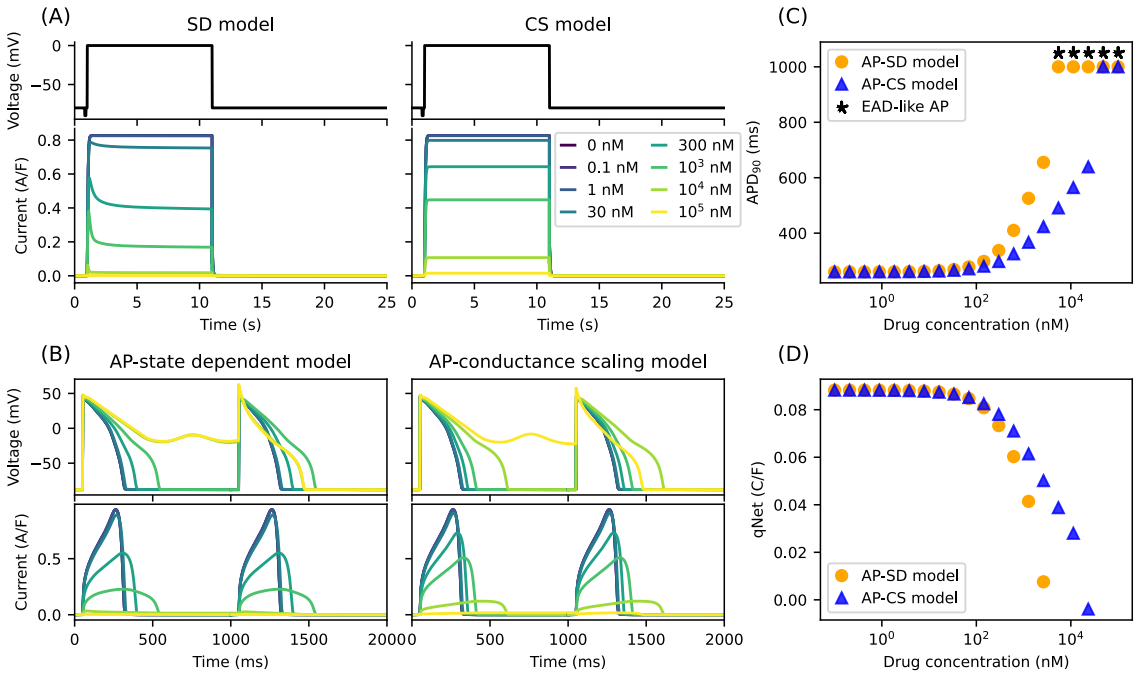


Figure 3.5: **The AP-CS model is less successful at replicating the model behaviour of the AP-SD model for the example drug N. (A)** I_{Kr} simulated by the SD and the CS formulations for the example drug N. The top row shows the modified Milnes' protocol used to stimulate the hERG-channel components. **(B)** The AP (first row) and its I_{Kr} (second row) for the AP-SD model and the AP-CS model at steady state. **(C)** The APD_{90}s of the AP-SD model and the AP-CS model under the effect of the example drug N. APs that show EAD-like behaviour are indicated with asterisks. **(D)** The qNet scores of the AP-SD model and the AP-CS model under the effect of the example drug N. qNet scores of APs with EAD-like behaviours are not shown.

AP of the AP-SD model is prolonged more than the AP-CS model and starts to display EAD-like behaviour at a lower drug concentration than the AP-CS model. The example drug N, for which the untrapping rate K_n is higher and the binding rate ($K_{\max} \times K_u$) is lower than the example drug T, generated shorter APD_{90}s with the AP-CS model. Similarly, the qNet scores of the example drug N (Figure 3.5D) are different between the two AP models. Example drug T and example drug N are taken as examples to show the difference in APD_{90} and qNet for drugs with varying trapping tendency.

Figure 3.6 shows the APD_{90}s and qNet scores of the AP-SD and the AP-CS model for the other synthetic drugs (Table 3.1). Most drugs show longer APD_{90} and lower qNet scores in the AP-SD model than in the AP-CS model, also starting to display EAD-like behaviour at lower drug concentrations. Cisapride and mexiletine are the two exceptions where the AP prolongations saturate at high drug concentrations, similarly with the qNet scores. This is due to the low $K_{\max}$ of these drugs, both at around 10, whereas the others range from 55.84 to 10^8. Low $K_{\max}$ limits the binding rate (Equation 3.2) and restricts the amount of drugs that could bind to the channel. With restricted drug-binding at the channel and limited inhibition of the current, the AP cannot prolong further and thus have a saturated AP prolongation.

3.3.2 Trapping properties are apparent in the transient phase

The main difference between the two drug-binding formulations is the inclusion of the trapping mechanism. The feature of this mechanism is the accumulation of the drug compound within the channel due to 'trapping'. Figures 3.7A–D show the progression of the APs and their I_{Kr} for the example drug T and the example drug N in the two AP models. In Figure 3.7A, the I_{Kr} decreases gradually with time for the AP-SD model with the example drug T, and the AP prolongation increases with pulses. In contrast, the I_{Kr} and the APs for the example drug N in the same AP-SD model are relatively stable with a cycle of two pulses, as shown in Figure 3.7B. The AP-SD model displays a much shorter transient phase with the example drug N than with the example drug T, which is due to the differences in the trapping phenotype of these synthetic drugs. Figures 3.7C and D compare the AP-CS model with the example drug T and the example drug N, respectively, and show that the APs and the I_{Kr} do not change much between cycles of two pulses.

The APD_{90}s from Figures 3.7A-D are given in Figures 3.7E and F for the example drug T and the example drug N, respectively. For the example drug T (Figure 3.7E), the AP-SD model showed a longer transient phase, then stabilised to give the same

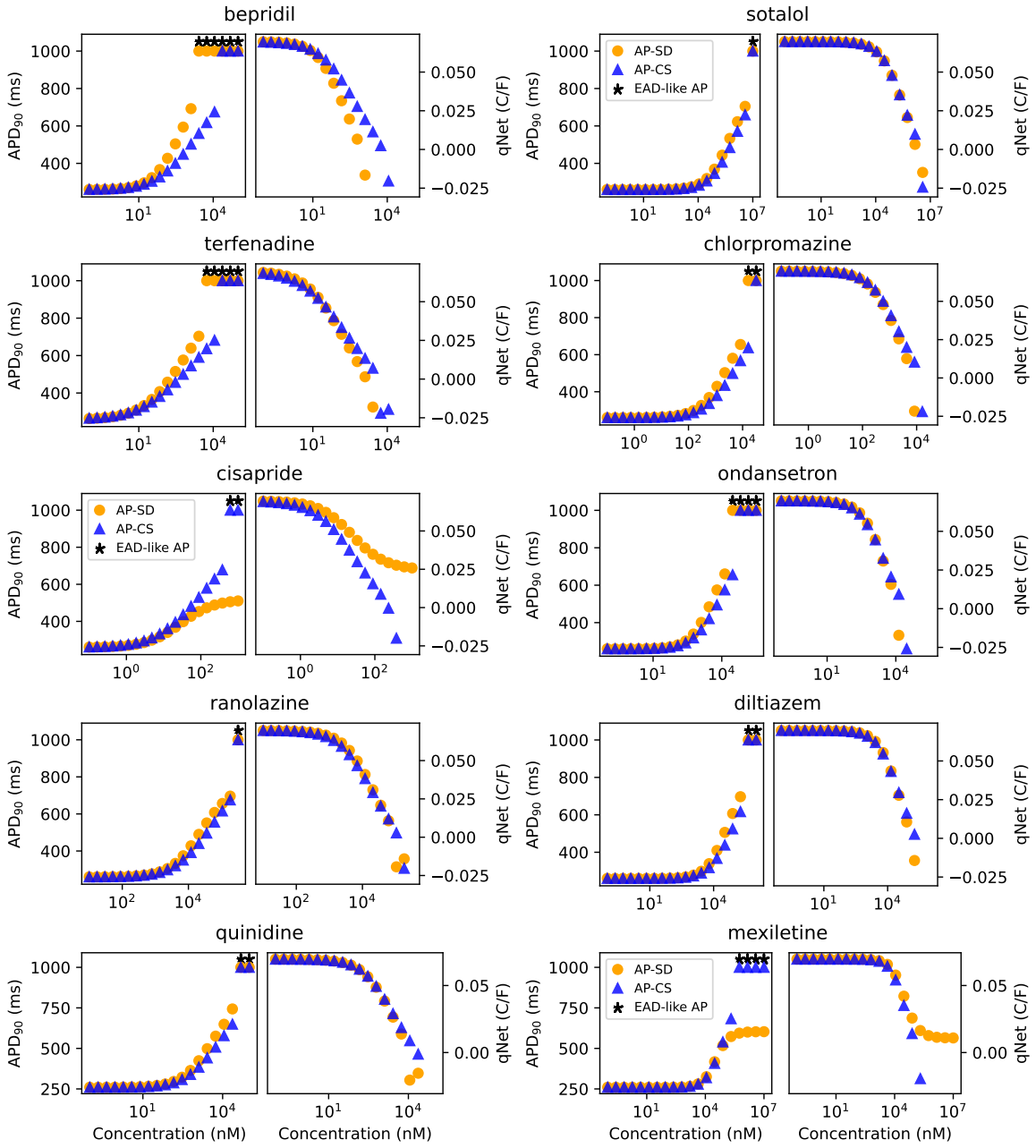


Figure 3.6: **The AP-SD model has higher APD_{90}s than the AP-CS model for most of the synthetic drugs, except for cisapride and mexiletine.** Each panel shows the APD_{90}s (left) and qNet scores (right) of the AP-SD model and the AP-CS model under the effect of a synthetic drug. APs that show EAD-like behaviour are indicated with asterisks. qNet scores of APs with EAD-like behaviours are not shown.

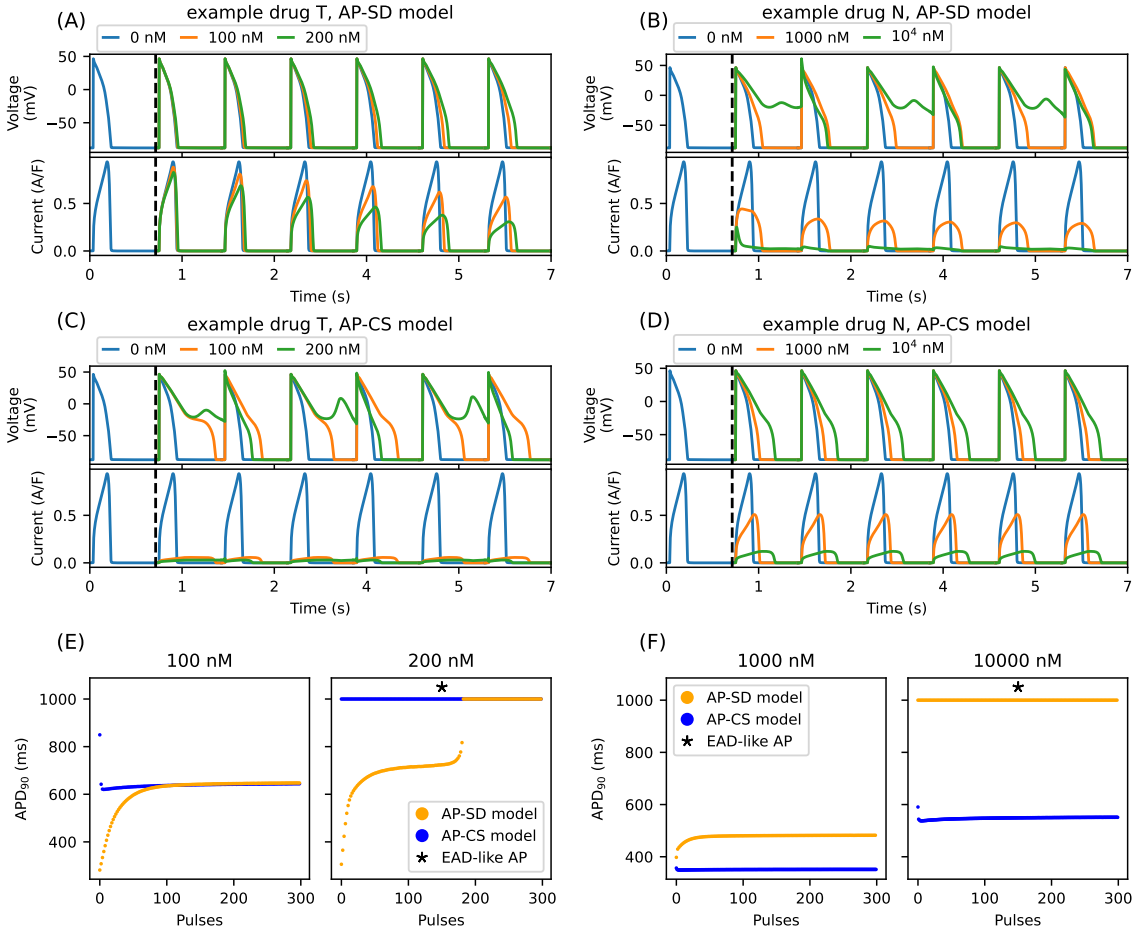


Figure 3.7: **The accumulation of drug compounds at the channel due to them being trapped is well-observed in the transient phase. The Example drug T shows a longer transient phase than the example drug N in the SD formulation.** The first six pulses of AP (top row) and its I_{Kr} (bottom row) for the AP-SD model after the addition of **(A)** the example drug T and **(B)** the example drug N. The vertical black dashed lines indicate the addition of the drug into the system after 1000 pulses of pacing. The first six pulses of the AP (top row) and the I_{Kr} (bottom row) for the AP-CS model after the addition of **(C)** the example drug T and **(D)** the example drug N. **(E)** APD_{90}s of both the SD and the CS formulations at 100 nM and 200 nM of the example drug T. **(F)** APD_{90}s of both the formulations at 1000 nM and 10^4 nM of the example drug N.

APD_{90} as the AP-CS model. In contrast, the two models with the example drug N (Figure 3.7F) have shorter transient phases and achieve steady state at different APD_{90}s.

3.3.3 AP prolongation is dependent on the protocol used to stimulate the SD formulation

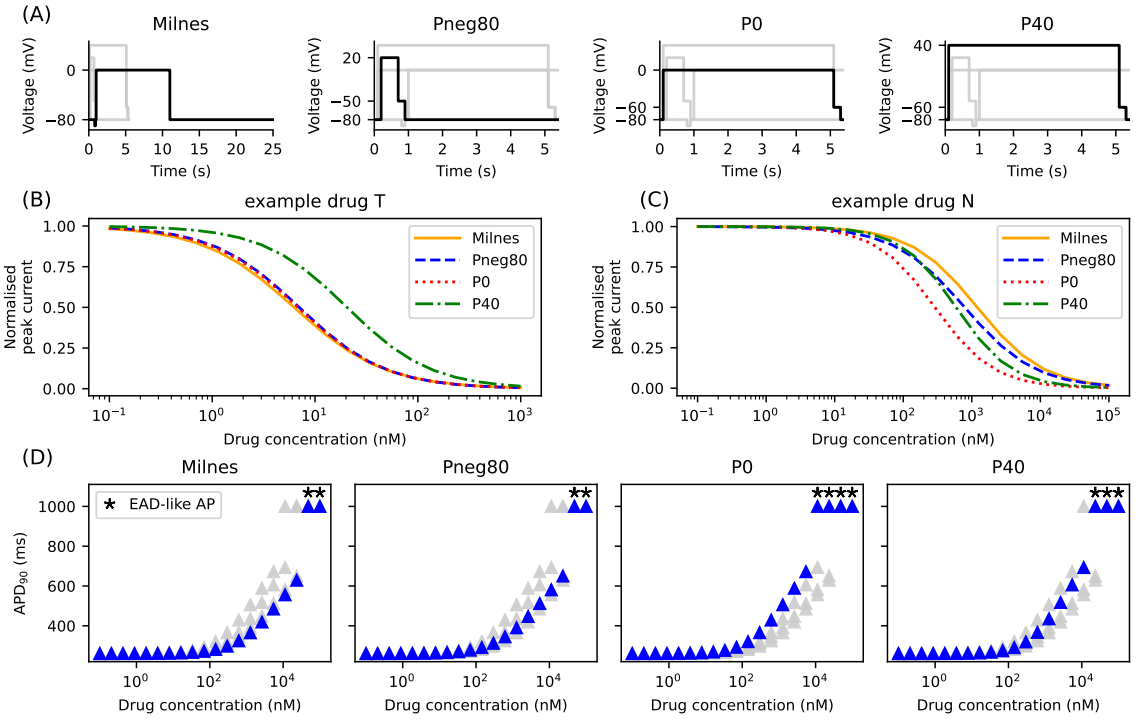


Figure 3.8: **The Hill curves for the example drug T and the example drug N are protocol dependent, thus affecting the resultant APD prolongation.** **(A)** The four protocols used to calibrate the hERG ionic-conductance for the CS formulation from the SD formulation. Note the differences in the time axes for visualisation purposes. The Hill curves based on the peak I_{Kr} simulated using the SD formulation under the protocols given in panel A for **(B)** the example drug T and **(C)** the example drug N. **(D)** APD_{90}s of the AP-CS model under the effect of the example drug N with its hERG conductance scaled based on the Hill curves shown in panel C.

In all model comparisons shown in Figures 3.4 and 3.5, we have assumed that the IC_{50} of a drug is independent of the protocol used to measure its response. Thus, the IC_{50} can be used to characterise a drug's effect [Mirams et al., 2011]. However, the assumption does not always hold [Yao et al., 2005, Kirsch et al., 2004]. Here, we demonstrate the change in the APD_{90} when different protocols are used to generate

the I_{Kr} for drug characterisation. Figure 3.8A shows the effect of using the modified Milnes', the *Pneg80*, the *P0*, and the *P40* protocols to stimulate the SD formulation. The normalised peak of the I_{Kr} simulated for both the example drug T and the example drug N are shown in Figure 3.8B and C, respectively. The example drug T (Figure 3.8B) yields similar dose-response curves for all of the protocols, except for the P40 protocol; while the example drug N (Figure 3.8C) displayed different dose-response curves for all protocols. Using the Hill curves in Figure 3.8C to characterise the drug effect for the CS formulation (see Section 3.2.3 and Figure 3.2), the APs of the AP-CS model were simulated and quantified. Figure 3.8D shows the APD_{90}s of the simulated APs with the example drug N. While using the Pneg80 protocol gave similar APD_{90}s to using the modified Milnes' protocol, the use of P0 and P40 protocols generate higher APD_{90} values, causing EAD-like behaviours to start appearing at lower concentrations of the example drug N.

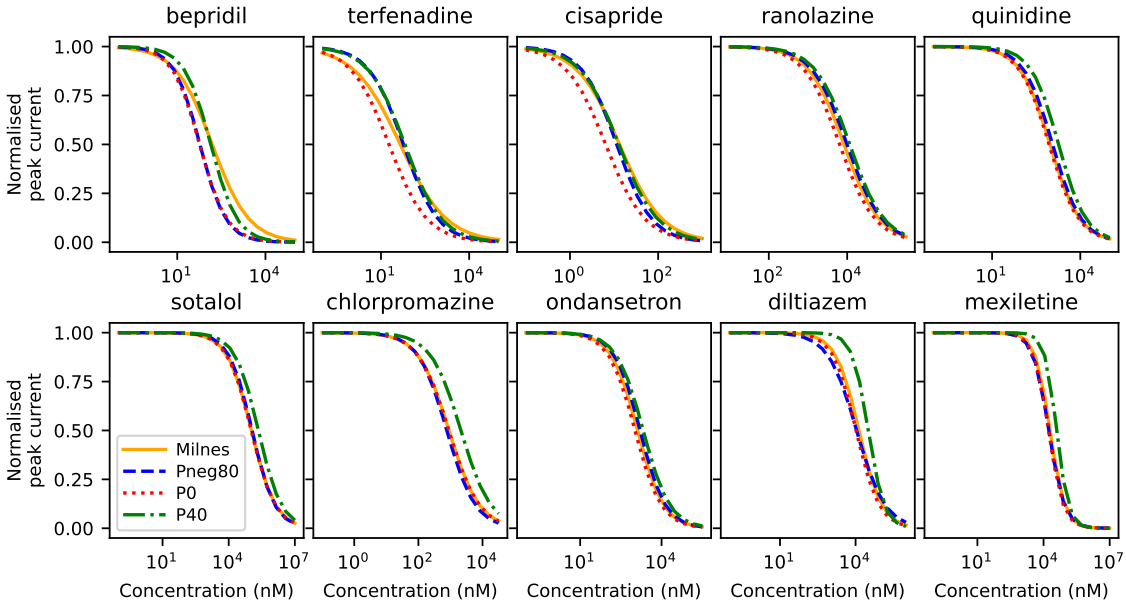


Figure 3.9: **The Hill curves of synthetic drugs from the SD formulation stimulated by the four protocols: the modified Milnes', the Pneg80, the P0, and the P40 protocols.**

Figure 3.9 shows the Hill curves for all protocols with the remaining synthetic drugs (Table 3.1). None of the drugs has the same Hill curves for all protocols, with some having a larger difference than the others.

3.3.4 Drugs were approximated to have shorter steady state APD_{90}s with the AP-CS model

A sensitivity analysis is performed on the simplified SD drug-binding formulation. The SD formulation is simplified to three drug-related parameters—$V_{\text{half-trap}}$, $K_{\max}$, and K_u—without loss of generality (Section 3.2.7). Given a virtual drug from the parameter space, the procedure shown in Figure 3.2 is repeated (Section 3.2.3). The differences between the AP-SD model and the AP-CS model are measured by the signed RMSD of the APD_{90}s (ΔAPD_{90}) (Equation 3.9).

Figure 3.10A shows the ΔAPD_{90} for various combinations of the three parameters. Each point represents a virtual drug, while its color indicates the APD_{90} difference between the AP-SD model and the AP-CS model when the virtual drug is added. The majority of the virtual drugs resulted in a positive ΔAPD_{90}, indicating that the AP-SD model generated APs that had longer durations than the AP-CS model. It is also shown that the ΔAPD_{90} shifts from positive to negative as $K_{\max}$ (related to the binding rate) decreases. The lower binding rate leads to less amount of drug compounds binding to the channel, which causes lesser hERG current inhibition and less AP prolongation. This would eventually have the AP-SD model show shorter APD_{90} than the AP-CS model and thus have a negative ΔAPD_{90}.

The parameter space where the two models give similar APD_{90}s (RMSD < 30 ms) is limited, as shown in Figure 3.10B in grey (see Section 3.2.7). Figures 3.10C and D show different viewing angles of the top panels. For these parameter combinations (Figure 3.10B in grey), the AP-SD model and the AP-CS model are interchangeable, as assessed by their APD. The triangles show the synthetic drugs (Section 3.2.5) within the parameter space. The parameter values for all 12 synthetic drugs are taken from Li et al. [2017b]. Of all synthetic drugs, dofetilide, ranolazine, sotalol, and mexiletine lie within the similarity threshold, and thus are said to be unaffected by the choice of drug-binding formulations. Cisapride, on the other hand, showed higher APD_{90} in the AP-CS model. The remaining synthetic drugs all caused higher APD_{90} values with the AP-SD model: bepridil, terfenadine, verapamil, quinidine, chlorpromazine, ondansetron, and diltiazem.

The results show that $V_{\text{half-trap}}$ plays a small role in affecting the differences between the AP-SD model and the AP-CS model. This is observed from similar ΔAPD_{90}s across the $V_{\text{half-trap}}$ axis. Indeed, Figure 3.11 shows similar ΔAPD_{90}s for three chosen $V_{\text{half-trap}}$ values of the $K_{\max}$-K_u plane. $K_{\max}$ and K_u, related to the binding and unbinding rates, respectively, are the major driving forces for the changes in

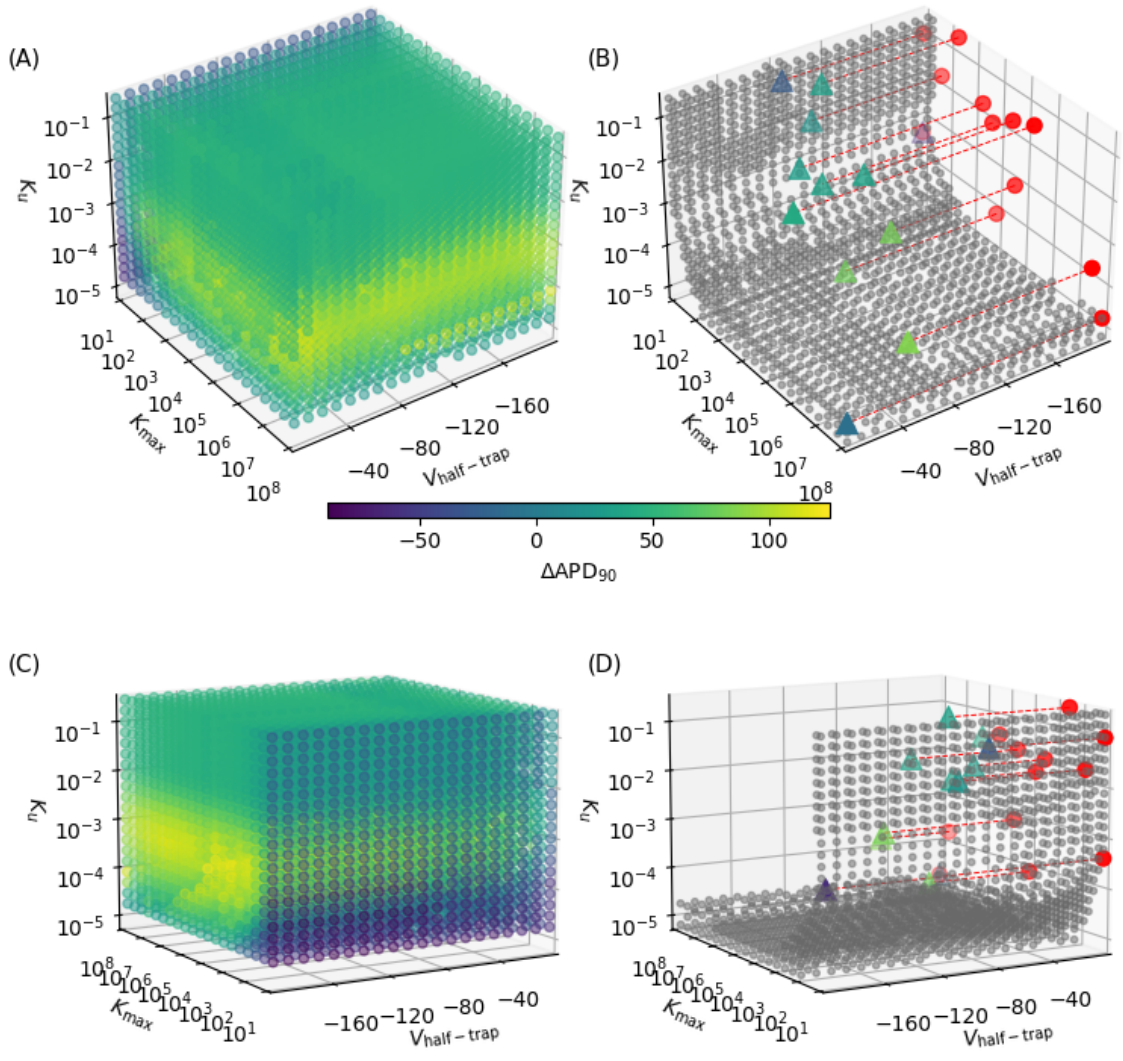


Figure 3.10: **The AP-SD model is most likely to return higher APD$_{90}$s than the AP-CS model. (A)** The APD$_{90}$ differences for combinations of $V_{\text{half-trap}}$, K_{max} and K_u parameters. The color of the markers indicates the ΔAPD$_{90}$ of each virtual drug in the parameter space. **(B)** The grey circles are parameter combinations (virtual drugs) where the ΔAPD$_{90}$ is between -30 ms and 30 ms. The triangles are the synthetic drugs taken from Li et al. [2017b], color coded with their ΔAPD$_{90}$ value. These triangles are projected to the K_{max} - K_u plane as red circles for better visualisation. Some data points are missing due to numerical issues with these parameter combinations. **(C, D)** Different viewing angles of (A) and (B), respectively.

the ΔAPD_{90}s. Furthermore, at the lower boundary of K_{max}, K_u does not affect the ΔAPD_{90} much; for large K_{max}, the two models differ the most around $K_u \approx 5 \times 10^{-4}$.

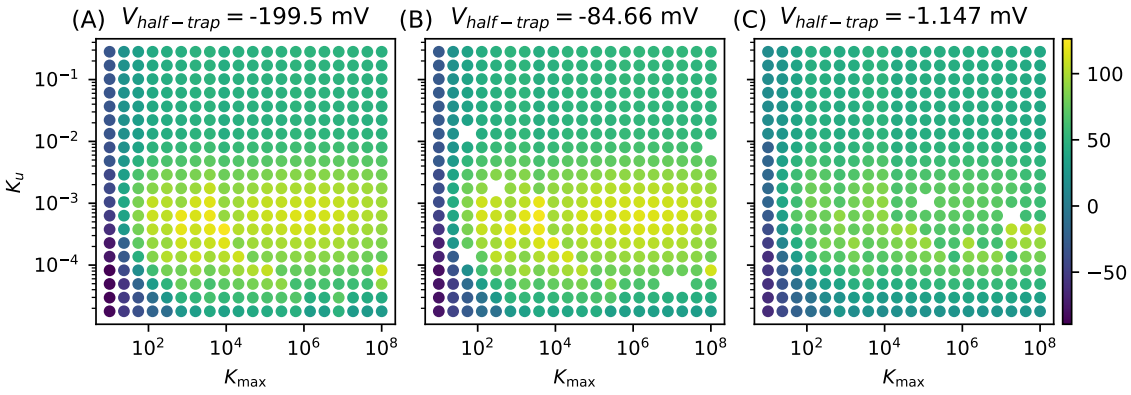


Figure 3.11: **ΔAPD_{90}s are similar for different values of $V_{half\text{-}trap}$. (A-C)** ΔAPD_{90}s for combinations of K_{max} and K_u at three $V_{half-trap}$ values. The colour of the markers indicates the ΔAPD_{90} of each virtual drug in the parameter space. Some data points are missing due to numerical issues with these parameter combinations.

However, this does not imply that $V_{half-trap}$ has no effect on the APs. Figure 3.12 shows the APD_{90}s for the same set of $V_{half-trap}$ values as in Figure 3.11, with a randomly chosen combination of K_{max} and K_u. The RMSD calculated for the three virtual drugs in Figure 3.12 are 57.15 ms, 58.88 ms, and 30.29 ms, indicating that the RMSDs are similar (difference between the RMSDs are < 30 ms) but the APD_{90}s are different.

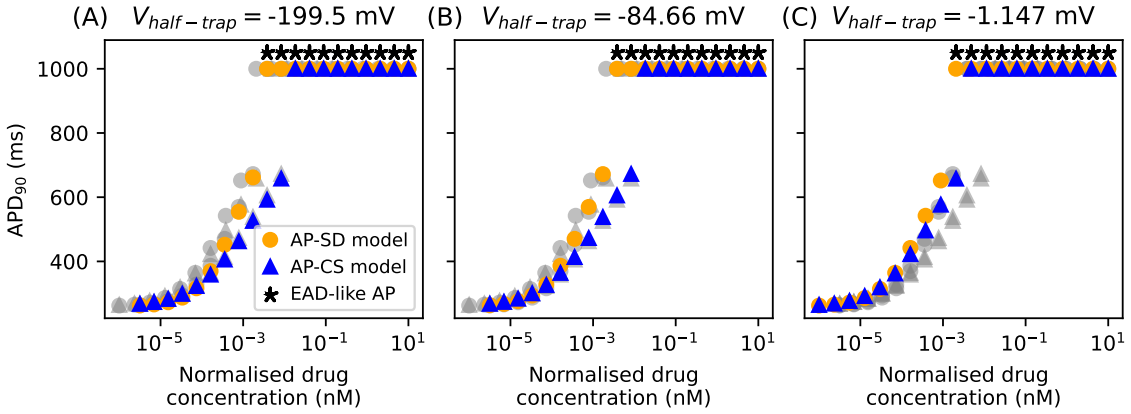


Figure 3.12: **The APD_{90}s of the AP-SD model and the AP-CS model are different despite showing similar ΔAPD_{90}. (A-C)** APD_{90}s for one randomly chosen combination of K_{max} and K_u, but for three distinct $V_{half-trap}$ values. The grey markers are repeats of the other panels for a better visual comparison.

3.4 Discussion

This study has compared the AP simulated by the AP-CS model and the AP-SD model. The conductance-scaling (CS) drug-binding formulation reduces the ionic conductance to approximate the effect of I_{Kr} reduction induced by drugs binding to the hERG channel, while the state-dependent (SD) drug-binding formulation includes the state dependency and the trapping mechanism to replicate the drug's effect. Our comparison intended to explore whether the more complex SD formulation and the simpler CS formulation give the same prediction of hERG inhibitor-induced APD change. A sensitivity analysis of the simplified SD formulation, with the $V_{\mathrm{half-trap}}$, K_{max} and K_u parameters, showed that there is only a small region of the parameter space where the two formulations produce similar steady state APD_{90}s (RMSD of the steady state APD_{90}s $< 30\,\mathrm{ms}$). The results suggested that the majority of the virtual drugs (parameter combinations) induce lower steady state APD_{90}s with the AP-CS model than with the AP-SD model, leading to possible lower arrhythmic risk prediction.

AP simulations of the AP-SD model and the AP-CS model with the example drug T (a trapped drug) show similar AP prolongations and qNet scores at steady state (Figure 3.4). The small difference in APD_{90} between the models contradict our expectations, as it was hypothesised that the AP-SD model with a trapping mechanism would predict trapped drugs like dofetilide to cause more AP prolongation [Pearlstein et al., 2016]. It was believed that the accumulation of the drug at the channel ('trapped') could inhibit the current more as compared to no trapping, thus prolonging the APD further. On the other hand, for the example drug N (a nontrapped drug), the AP-CS model predicted less AP prolongation than the AP-SD model (Figure 3.5), conforming to expectations. The large difference in the qNet scores for the example drug N, suggests that we have to be wary when categorising the drug, as we might be around the categorisation boundary and the choice of the drug-binding formulation could sway the outcome. However, for drugs like dofetilide, it does not matter which drug-binding formulation is used for AP prolongation prediction.

The effects of drug trapping in the hERG channel become apparent during the transient phase as it takes more pulses to achieve steady state (Figure 3.7). Trapped drugs accumulate at the channel, without recovering from the block during the interpulse interval, and inhibit the current pulse by pulse until steady state, creating a longer transient phase; nontrapped drugs unbind and rebind to the channel in between pulses of protocol, seemingly achieving steady state. However, in Figure 3.7,

the AP-SD model and the AP-CS model both predict similar steady state APDs for the example drug T, while both models predict different steady state APDs for the example drug N, suggesting that trapping is likely not a key driver for the differences in APDs.

To further explore the observation, our sensitivity analysis on the drug-related parameters of the SD formulation demonstrated that the AP-SD model and the AP-CS model predict significantly different AP prolongation for the majority of the virtual drugs. Therefore, modelling drug effect by simply reducing the ionic conductance is most likely insufficient to capture the dynamics of drug-binding, especially at the AP level. However, we also observed only a small change in the ΔAPD_{90} between the two models when $V_{\text{half-trap}}$ varies, which implies that no extra dynamics can be observed from the addition of the $V_{\text{half-trap}}$ parameter or the *trapping component*, consistent with all our previous observations. Furthermore, the parameters controlling the binding and unbinding rates—K_{max} and K_u—determine the difference in the steady state AP prolongation predictions between the two models (Figure 3.10). We note that it does not imply $V_{\text{half-trap}}$ has no influence on the APD_{90}s; although the difference in the APD_{90}s are small, the APD_{90}s vary with $V_{\text{half-trap}}$ as shown in Figures 3.11 and 3.12. The difference in AP prolongation for the majority of the parameter space implies that one cannot assume that with the same baseline AP model, switching between the SD and the CS drug-binding formulation can achieve the same prediction results as the outcome of the model training. It is not possible to determine the best model in this study as it depends on the type of input data the model is trained on. Since it is common practice to acquire IC_{50} data, it is more pragmatic to use the CS formulation. Without re-evaluation of a drug on the SD formulation, it cannot be assumed that the SD formulation can replace the CS formulation, or vice versa. It is important that the drug-binding formulation is used within the context of use on which the model is trained and validated (including the type of input data).

One of the limitations of this study is that we assumed that the SD formulation can replicate the experimental data perfectly and thus fitted the CS formulation to the Hill curve generated from the SD formulation. Moreover, the six-state Markov model and the Dutta et al. [2017] model were used as the hERG-channel component and the AP model, respectively, but changing the choice of the models and their components could affect the behaviour of the APs and hence the differences in APs between the SD and the CS formulation [Clayton et al., 2020, Lei et al., 2020]. This is further investigated in Chapter 4. Furthermore, in the sensitivity analysis, the ranges

of the parameter values were determined by the minimum and maximum of each of the parameters of the synthetic drugs, and the trapping rate of the SD formulation was fixed in the same manner as it was done in the original model [Li et al., 2017b], limiting the exploration of the model dynamics. Finally, we emphasise that we focus on the drug effect on the hERG channel only, and specifically on the effect of the inclusion of the state-dependent drug-binding of the hERG channel. Investigation of the effect of other ion channels is out of the scope of this study. Therefore, this study is not an integrative and comprehensive assessment of the change in AP or proarrhythmic risk of the drug compounds, particularly for a multi-channel blocker such as verapamil.

In this study, we have shown that the ΔAPD_{90} between the AP-SD model and the AP-CS model for the example drug T and the example drug N are not consistent with their trapping phenotypes. They are differentiable only in the transient phase. The ratio of trapping to untrapping rate, defined in the SD formulation by $V_{\text{half}-\text{trap}}$, does not affect the ΔAPD_{90} between the two models. The binding rates and unbinding rates, on the other hand, are the main determinants of the APD_{90} differences. This study demonstrates the importance of modelling drug-binding and highlights the need for improved understanding of drug trapping, which can have implications for the uses in drug safety assessment.

Chapter 4

Impact of uncertainty in action potential models on *in silico* drug effect predictions

We plan to submit this chapter to a journal for publication. The associated GitHub repository is at https://github.com/FarmHJ/binding-mechanism-extend. *I led the research of this paper by designing the methodologies, and performing all simulations and analysis of this study.*

Overview:

In Chapter 3, we compared the state-dependent (SD) drug-binding formulation by Li et al. [2017b] and the conductance-scaling (CS) drug-binding formulation in prolonging the action potential (AP). We did this in the context of embedding this drug-binding formulation in a hERG-channel component of a single AP model. However, a plethora of AP models are available in the literature which each have very different combinations of ionic currents and other components, and very different associated parameters. In this chapter, we therefore investigate the effect of embedding both the SD and the CS formulations in three further widely used AP models. We show that the AP prolongation predicted using the SD formulation is not replicable in all cases by the CS formulation. Drug-trapping still does not influence the AP prolongation, similar to the results we have shown in Chapter 3.

4.1 Introduction

In Chapter 3, we investigated the effect of the state-dependent (SD) drug-binding formulation, introduced by Li et al. [2017b], on action potential (AP) prolongation, by comparing it with a simpler drug-binding formulation, the conductance-scaling (CS) formulation. We showed that the AP prolongation predicted by the SD formulation cannot be replicated by the AP prolongation of the CS formulation. In particular, we showed that the parameter, $V_{\text{half-trap}}$, does not influence the difference in APD_{90} between the SD formulation and the CS formulation. This study was performed by embedding the drug-binding formulations into the hERG-channel component within the AP model. The AP model chosen by Li et al. [2017b] was the ORd model [O'Hara et al., 2011], with the addition of their own in-house hERG-channel component [Li et al., 2016].

As described in Chapter 2, the changes in membrane potential during an AP are controlled by the flow of ionic currents across the cell membrane. These ionic currents include the hERG current I_{Kr}, the slow delayed rectifier outward potassium current I_{Ks}, the inward rectifier potassium current I_{K1}, the L-type calcium current I_{CaL}, the late sodium current I_{NaL}, and others. AP models are built by combining computational models of the ionic currents, some of which are inherited from other AP models, and calibrated against experimental data and biomarkers chosen by the authors [Rudy and Silva, 2006, ten Tusscher et al., 2006, Niederer et al., 2009, Amuzescu et al., 2021] (see O'Hara et al. [2011], Grandi et al. [2010], ten Tusscher et al. [2004], Tomek et al. [2019] for examples of various AP models). The variation between AP models arises from the large variation in the experimental data, some of which may arise from the use of voltage-clamp protocols specifically designed to extract certain ion channel or drug-binding properties [Gomis-Tena et al., 2020]. Additionally, the variety of biomarkers, which include but are not limited to AP morphology, calcium transient, and current inhibition [Tomek et al., 2019], contributes to the variability between AP models. The diverse choices and lack of standardisation in the calibration procedure lead to large variability between AP models. The differences between AP models are reviewed in Niederer et al. [2009], Noble et al. [2012], Fink et al. [2011]. The variability between AP models also extends to the components within the AP models, such as the hERG channel [Bett et al., 2011, Vandenberg et al., 2012, Fink and Noble, 2009] (see the Appendix of Beattie et al. [2018] for a list of published hERG-channel components).

4.1.1 Naming convention

There are three scales of computational model discussed in this chapter (Figure 4.2). At the highest level, we have the action potential (AP) *models*, describing the change in membrane potential during a cardiomyocyte contraction cycle. The computational representation of an ion channel is referred to as a *component*, to differentiate it from a *model*, which we use to refer to the computational model at the AP level. In this thesis, we are focusing on the computational models of the inhibition effect induced by drugs on the human Ether-à-go-go-Related Gene (hERG) ion channel. To separate this sub-ion channel scale from the other scales (which are the AP level and the ion channel level), the computational models describing the drug-inhibition effect are referred to as the *drug-binding formulations*.

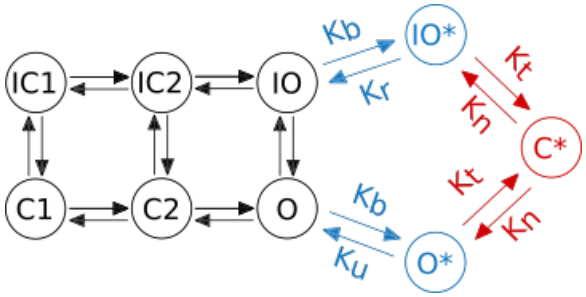


Figure 4.1: **The structure of the SD drug-binding formulation (blue and red) [Li et al., 2017b] attached to the Li hERG-channel component in black [Li et al., 2016].** The blue and red parts refer to the binding and the trapping components, respectively.

4.1.2 Summary of the SD formulation and *drug-trapping*

As in Chapter 3, the key focus of this chapter is again the SD drug-binding formulation of Li et al. [2017b] (the blue and red parts of Figure 4.1), which incorporates a *trapping module* (the red part of Figure 4.1). This drug-binding formulation, specifically the trapping module, was described in detail in Section 2.5.5. In summary, the drug-binding process described by the SD formulation can be viewed as a two-step process, *binding* and *trapping* (blue and red parts in Figure 4.1, respectively). A drug compound first binds to the ion channel (entering the blue module of the SD formulation in Figure 4.1), then gets trapped within the channel when the channel closes (entering the red module in Figure 4.1). Each drug has a different tendency to stay trapped within the trapped state (C*), and this is controlled by the trapping

tendency parameter $V_{\text{half}-\text{trap}}$, which controls the rate at which the drug-bound channel opens and allows drug compounds to unbind from the channel. Drugs that can get trapped within the ion channel are hypothesised to prolong the AP more than drugs that cannot be trapped [Di Veroli et al., 2014], and have higher arrhythmic risk [Li et al., 2017b].

4.1.3 Extrapolation to new models

Given the variability within the AP models and their hERG-channel components, and the sensitivity of the predicted AP to the hERG-channel component within the AP model, the lack of influence of the parameters of the SD formulation within the ORd model and the Li hERG-channel component [Li et al., 2016] to the difference in AP duration observed in Chapter 3, cannot be translated to consideration of other AP models or hERG-channel components. The variability between AP models may result in a high degree of variability in AP prediction when the models are reused or modified to new situations [Niederer et al., 2009, Whittaker et al., 2020]. In this chapter, we therefore further investigate the sensitivity of the SD drug-binding formulation (blue and red parts of Figure 4.1) when it is embedded within different AP models with differing hERG-channel components. The AP models chosen are the Grandi model [Grandi et al., 2010], the ten Tusscher model [ten Tusscher and Panfilov, 2006], and the Tomek model [Tomek et al., 2019, 2020], whereas the hERG-channel component chosen is the Lei component [Lei et al., 2019b]. The hERG-channel component and the drug-binding formulation within the chosen AP models are replaced with the Li component or the Lei component and the SD formulation or the CS formulation, respectively. We then perform a sensitivity analysis on the SD formulation when it is embedded in the various modified AP models, similar to the procedure described in Chapter 3. A figure summary of this chapter is shown in Figure 4.2.

4.1.4 Sensitivity analysis

We first reiterate the main elements of the sensitivity analysis performed in Chapter 3, which include the parameter of interest, the data-generating model, and the model output. The parameters of interest are the parameters K_{max}, K_u, and $V_{\text{half}-\text{trap}}$, referring to the binding rate, the unbinding rate, and the trapping rate, respectively. Due to the restriction of calculating the Torsades Metric Score (TMS) for all parameter combinations within the SD formulation, we choose the model output to be the

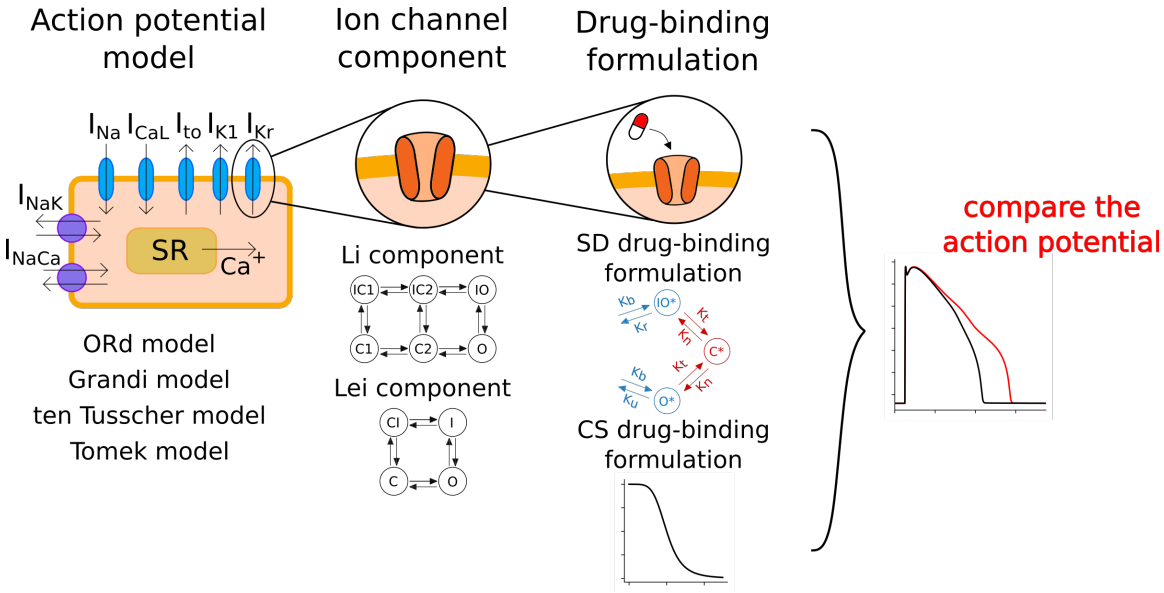


Figure 4.2: **Schematic of all the scales used in this chapter.** There are four AP models: the ORd model, the Grandi model, the ten Tusscher model, and the Tomek model. There are two ion-channel components: the Li component and the Lei component. Only a subset of the combinations (of the AP models and the ion-channel components) is used in the comparison between the SD and the CS drug-binding formulations. For a given AP model and ion-channel component combination, the SD and the CS drug-binding formulations are attached to it, and their action potentials are compared.

action potential duration at 90% repolarisation (APD_{90}). The data-generating model is therefore the AP model.

In a similar fashion to Chapter 3, we want to investigate whether drug-trapping influences the proarrhythmic risk of drugs. The conductance-scaling (CS) drug-binding formulation, which does not have drug-trapping, is used as a reference point. The sensitivity analysis is performed using a factorial design, similar to Chapter 3. This is performed by comparing the model outputs of parameter combinations of the parameter space in a grid-like fashion. For each parameter combination, the Hill equation for the CS formulation is calibrated accordingly, following the procedure described in Chapter 3. Finally, the APD_{90} predicted by the SD formulation and the equivalent CS formulation is compared.

4.1.5 Summary

In this chapter, we first introduce the computational models of all the scales used in this chapter: the AP models (the ORd, the Grandi, the ten-Tusscher, and the Tomek models), the hERG-channel components (the Li and the Lei components), and the drug-binding formulations (the SD and the CS formulations) (Section 4.2). We then perform the sensitivity analysis detailed above on two parameter combinations with different trapping tendency (different $V_{\text{half-trap}}$ values) for all the modified AP models (Section 4.4.1). We further extend the sensitivity analysis to parameter combinations of the parameter space of the SD formulation, using the factorial design (Section 4.4.2 & 4.4.3).

4.2 Components of an action potential model

We first describe the AP models and the components of AP models that are of interest in this study. We establish a naming convention to avoid confusion when referring to the AP models and their components. As described in the Overview, there are three scales of the AP model to consider: the AP *model*, the hERG-channel *component*, and the sub-channel scale of the drug-binding *formulation*. Each part is described in Sections 4.2.1, 4.2.2 and 4.2.3, respectively. A modified AP model will be named after these three scales, with a hyphen in between: *AP model - hERG-channel component - drug-binding formulation*. For example, the ORd-CiPAv1.0 model [Li et al., 2017b, Dutta et al., 2017] is referred to as the ORd-Li-SD model for its AP model chosen to be the ORd model [O'Hara et al., 2011], the hERG-channel component as the Li component developed by Li et al. [2016], and the drug-binding formulation as the SD

formulation [Li et al., 2017b]. When referring to the modified AP model (the AP model and hERG-channel component pair) for any drug-binding formulation, such a model combination can be referred to as *AP model - hERG-channel component* without specifying the drug-binding formulation. For example, the Grandi model with its hERG-channel component replaced with the Li component is referred to as the Grandi-Li model. When investigating the dependence on the hERG-channel component, the hERG channel component of the ORd-Li model is replaced with the Lei component (Section 4.2.2), retaining the optimised ionic conductance value for I_{Ks}, I_{K1}, I_{CaL}, and I_{NaL} performed by Dutta et al. [2017]. This model will be referred to as the ORd-Lei model when the drug-binding formulation is not specified. When referring to the SD formulation with any AP model and any hERG-channel component, it will be referred to as the AP-IKr-SD model.

4.2.1 AP models

We first consider the variability between the four AP models, the ORd-Li model itself [Dutta et al., 2017] (the optimised AP model used in the development of the SD drug-binding formulation) and three further widely used AP models — the Grandi model [Grandi et al., 2010], the ten Tusscher model [ten Tusscher and Panfilov, 2006] and the Tomek model [Tomek et al., 2019, 2020].

Figure 4.3 shows the major transmembrane currents of the AP models that contribute to the AP, separated into positive and negative currents. The APs are paced at 1 Hz with an amplitude of -80 A/F for 1000 paces to achieve steady state. The relative magnitude and behaviour of currents that make up the AP vary across the AP models. For example, the chloride currents (I_{ClCa} and $I_{\mathrm{Cl,B}}$) are included only in the Grandi model and the Tomek model, whereas I_{NaL} is included only in the ORd-Li model and the Tomek model. I_{K1} (green), on the other hand, has positive and negative values in the AP of the ORd-Li model, but remains positive throughout in other AP models. At the plateau phase, the main contributors of outward currents differ significantly across all four models, with the ORd-Li model mainly dominated by I_{Kr}, the Grandi model by $I_{\mathrm{Cl,B}}$, the ten Tusscher model by I_{Kb} and I_{Ks}, and the Tomek model by several currents including I_{Kr}, I_{Kb} and I_{to}. The behaviour of the hERG current (I_{Kr}), which is the interest of this study, differs between the AP models too. In the ORd-Li model, the magnitude of the hERG current increases at the start of the plateau phase and remains large till the repolarisation phase. Even though the hERG current in the Tomek model increases in magnitude at the start of the plateau phase, similar to the ORd-Li model, its relative contribution is lower than the ORd-Li

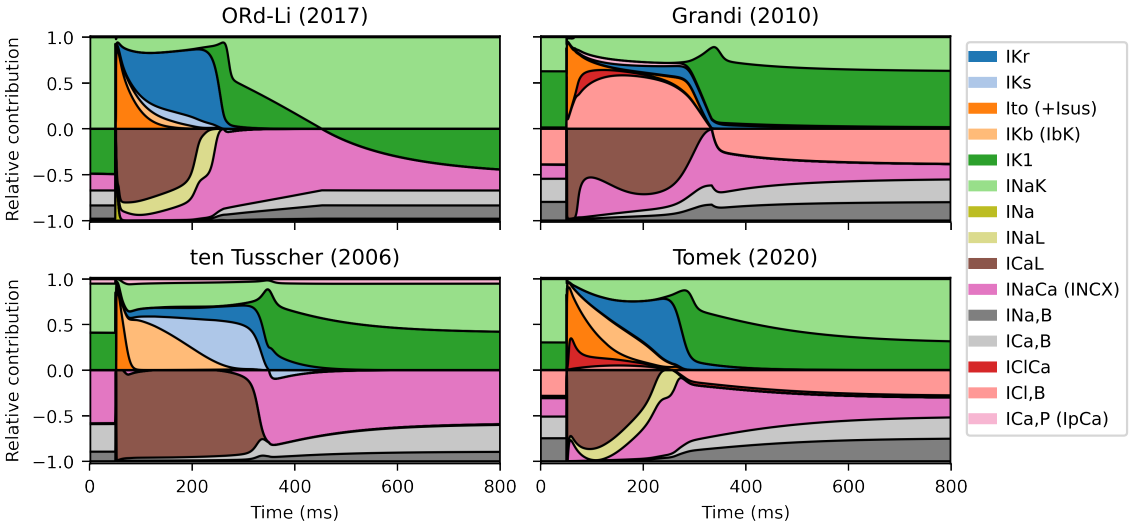


Figure 4.3: **Relative current contributions making up each AP model are different.** The ionic currents of each AP model are grouped into positive and negative currents and normalised. The relative current contribution of each current to the overall AP is shown for the ORd-Li model [Dutta et al., 2017], the Grandi model, the ten Tusscher model, and the Tomek model. The APs are paced at 1 Hz with an amplitude of −80 A/F till steady state. The idea for the plot is originally by Dr. Michael Clerx from the University of Nottingham.

model. The Grandi model and the ten Tusscher model, on the other hand, are less reliant on the hERG current in their repolarisation phase. These differences are not limited to the relative magnitude, as shown in Figure 4.3, but also in the absolute magnitude of the hERG current, as shown in Figure 4.5 dashed lines. These show the different reliance of the AP on the hERG current for these four AP models. Thus, the inhibition of the hERG current with the SD formulation or the CS formulation will affect the AP differently across all four models.

4.2.2 hERG-channel component of the AP model

Figure 4.4A and B in black show the Markov structure of two hERG-channel components, the Li component [Li et al., 2016] and the Lei component [Lei et al., 2019b], respectively, used in this study. Both structures, despite being different, can capture the dynamics of the hERG channel. Figure 4.4C shows the modified Milnes' protocol in the top panel and the hERG current of the Li component under the protocol in the middle panel. The bottom panel shows the state occupancy of the Li component, corresponding to the top panels. When the channel is conducting current, the Li component is in either open (O) or inactivated open (IO) states, indicated by yel-

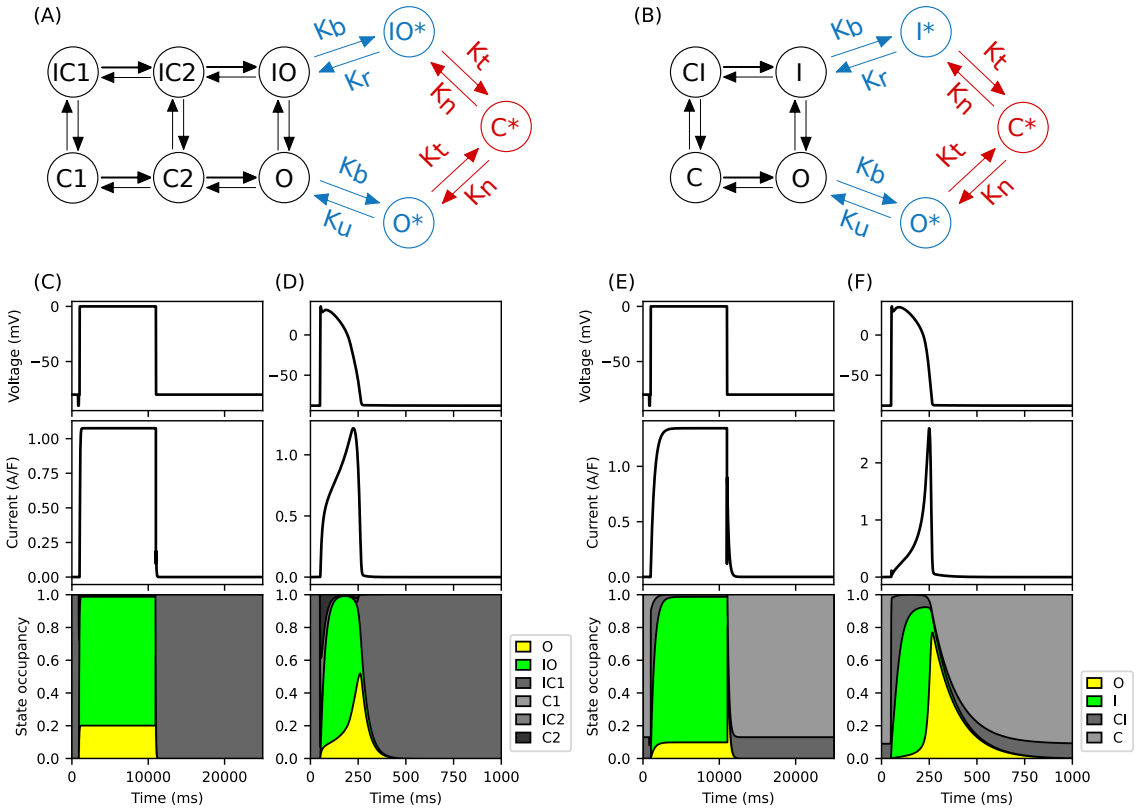


Figure 4.4: **hERG channel components used in the drug-binding formulation comparison.** **(A)** Structure of the Li component (black) [Li et al., 2016] with the SD formulation (blue and red) [Li et al., 2017b]. **(B)** Structure of the Lei component (black) [Lei et al., 2019b] with the SD formulation (blue and red). **(C)** Under the modified Milnes' protocol (top), the current signal (middle) and state occupancy (bottom) of the Li component. **(D)** (top) AP of the ORd-Li model. The current signal (middle) and state occupancy (bottom) of the Li component when it is under the AP-clamp protocol (shown in the top panel). **(E)** Under the modified Milnes' protocol (top), the current signal (middle) and state occupancy (bottom) of the Lei component. **(F)** (top) AP of the ORd-Lei model. The current signal (middle) and state occupancy (bottom) of the Lei component when it is under the AP-clamp protocol (shown in the top panel). Note that at the resting state, the Li component stays mainly in the IC1 state when it should be in a closed state. The labelling of the IC and C states of the Li component, adapted from the original model [Li et al., 2016], does not always correspond to their respective physical states.

low and green, respectively; while the greys indicate the remaining non-conducting states. In Figure 4.4D, the top panel shows an AP of the ORd-Li model at steady state paced at 1 Hz. When this AP shape is used as an AP-clamp protocol on the Li hERG-channel component, the generated hERG current and its state occupancy are as shown in the middle and bottom panels of Figure 4.4D. Similarly, Figures 4.4E and F are hERG currents and state occupancies of the Lei component under the modified Milnes' protocol and the AP-clamp protocol of the ORd-Lei model, respectively. The ORd-Lei model is the ORd-Li model with its hERG-channel component replaced by the Lei component (Figure 4.4B).

Comparing the Li component and the Lei component under the modified Milnes' protocol shows that the maximum I_{Kr} magnitude of the Li component is lower than the Lei component. The I_{Kr} of the Li component also achieves its maximum faster than the Lei component, almost instantaneously. Despite the Li component showing a lower I_{Kr} magnitude, the percentage of time spent within the open state is higher than the Lei component. At the AP level, the ORd-Li model starts to repolarise slightly earlier than the ORd-Lei model. This is explainable by its higher I_{Kr} magnitude than the ORd-Lei model during the first half of the AP. Similar to I_{Kr} under the modified Milnes' protocol, the I_{Kr} of the Li component peaks at a lower magnitude than the Lei component under the AP-clamp protocol. However, conversely to the modified Milnes' protocol, the maximum open state probability of the Li component is lower than the Lei component. The total time spent (area under the curve) in the open state and the inactivated open state for the Li component is lower than that of the Lei component (data not shown). The two hERG-channel components exhibit distinct hERG current behaviour. This will have an impact on how the SD drug-binding formulation affects the AP.

Based on the SD drug-binding formulation (blue and red parts of Figure 4.4A & B), drugs can bind to the hERG channel only when it is in either the open state or the inactivated state. The total time the channel spends in these states limits the amount of drug that can bind to the channel. Therefore, to achieve the same amount of current inhibition, the drug-binding rate will have to be different for the Li component and the Lei component. The drug-related parameters taken for the Li component from Li et al. [2017b] are recalibrated for the Lei component. The calibration process is described in Section 4.3.2.

4.2.3 Drug-binding formulations to the hERG channel

This section gives a brief overview of the drug-binding formulations used in this study. A detailed explanation is provided in Chapter 2.5.5.

The binding rate (K_b), the untrapping rate (K_n) and the trapping rate (K_t) of the SD formulation are given by

$$K_b = K_u \times K_{\max} \times \frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{4.1}$$

$$K_n = K_t \times \frac{1}{1 + e^{-(V - V_{\mathrm{half-trap}})/6.789}}, \tag{4.2}$$

$$K_t = 3.5 \times 10^{-5}, \tag{4.3}$$

where K_u is the unbinding rate, $K_{\max}$ is the maximum drug response on the binding rate, D is the drug concentration, n is the Hill coefficient and EC_{50} is the concentration at which half maximal drug response is achieved. For Equation 4.2, V is the input voltage and $V_{\mathrm{half-trap}}$ represents the trapping tendency of a drug. K_r (as shown in Figure 4.4A) is defined to satisfy the microscopic reversibility. A drug compound in the SD drug-binding formulation is defined by parameters K_u, $K_{\max}$, n, EC_{50}, and $V_{\mathrm{half-trap}}$. These parameters are referred to as drug-related parameters.

The conductance scaling (CS) drug-binding formulation, on the other hand, is implemented by scaling the conductance to inhibit the hERG current. This is achieved by

$$I_{\mathrm{Kr}} = \alpha g_{\mathrm{Kr}} O(V - E_K), \tag{4.4}$$

where $\alpha \in [0, 1]$ is a scaling factor that lowers the hERG current. α is scaled following the Hill curve of inhibition

$$\alpha = \frac{{\mathrm{IC}_{50}}^m}{D^m + \mathrm{IC}_{50}^m}, \tag{4.5}$$

where IC_{50} is the drug concentration that causes 50% current inhibition and m is the Hill coefficient. Note that IC_{50} and m are not the same as EC_{50} and n in Equation 4.1 of the SD drug-binding formulation. The Hill curve used for the CS formulation describes the relationship between drug concentration and the inhibition of the hERG current, whereas the EC_{50} and n in the SD formulation describe the change in the binding rate in response to the drug concentration.

For any given drug within the SD drug-binding formulation, its equivalent drug effect on the CS drug-binding formulation is captured by the Hill curve of the drug. It is common practice in the industry to use the Hill curve based on the peak current

to characterise a drug. In brief, we obtain the Hill curve of a drug by quantifying the peak I_{Kr} within the SD formulation for a range of drug concentrations. The magnitude of inhibition, as described by the Hill curve, then tunes the conductance of the CS formulation by the scaling factor α in Equation 4.4. The detailed calibration procedure for the CS formulation is described in Section 3.2.3.

4.3 Methods

To summarise from above, we will investigate the impact of the AP model and/or the hERG-channel component on the effect of the drug-binding formulation on the AP. To do so, we embed the drug-binding formulations (SD and CS formulations) into various AP models and/or hERG-channel components, then compare the simulated AP when different drug-binding formulations are used. A schematic of the procedure is shown in Figure 4.2. The AP models chosen for this study are the Grandi model [Grandi et al., 2010], the ten Tusscher model [ten Tusscher and Panfilov, 2006], and the Tomek model [Tomek et al., 2019, 2020]. For each AP model, the hERG-channel component is replaced with the Li hERG-channel component to allow the embedding of the SD drug-binding formulation. For the testing of the role of the hERG-channel component, we take the ORd-Li model, that is, the ORd-CiPAv1.0 model [Li et al., 2017b], and replace its hERG-channel component with the Lei component [Lei et al., 2019b]. Thus, the modified AP models are the Grandi-Li model, the ten Tusscher-Li model, and the Tomek-Li model to investigate the impact of the AP model; and the ORd-Lei model to study the contribution of the hERG-channel component. We can then embed either the SD or the CS drug-binding formulation into these modified AP models and compare the simulated AP under some given drugs.

We first compare the drug-binding formulations within all modified AP models for two example drugs, then repeat the comparison for a wide range of parameter combinations (*virtual drugs*) within the parameter space of the SD formulation. We then identify "drug properties" (parameterisations of the SD formulation) that are affected by the choice of drug-binding formulation.

4.3.1 Calibration of models

Every AP model is made up of its constituent components that are developed based on different experimental approaches and calibration methods, and each is therefore very dependent on the choices made by the authors. Due to many factors, including

those detailed in Chapter 2, any modifications to the AP model require the model to be recalibrated to ensure that the resultant model still makes physiological sense.

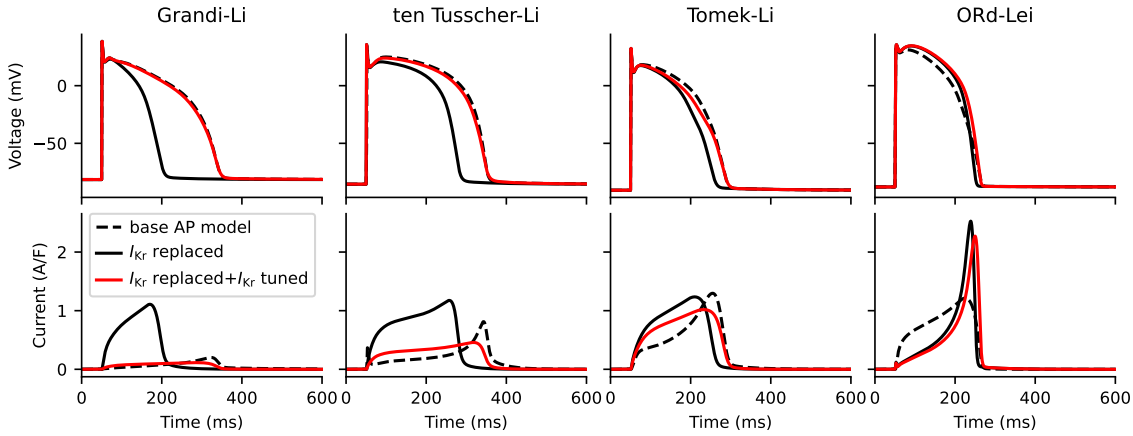


Figure 4.5: **Tuning of I_{Kr} in the modified AP models.** The dashed lines show (top) the AP and (bottom) I_{Kr} of the original AP model. The hERG-channel component of each AP model is replaced with the Li component or the Lei component (solid black lines). The hERG conductance is calibrated so that the APD_{90} is the same as the AP model before modification (solid red lines).

As shown in Section 4.2.1 and Figure 4.3, the hERG-channel component of all AP models exhibits different behaviour. Different weightings are given to the ionic currents that form the AP. Thus, replacement of the hERG-channel component may disrupt the AP. Figure 4.5 shows the AP and hERG current of the original AP models in dashed lines and the modified AP models in solid black lines. For example, the Grandi-Li model (solid black lines) exhibits a very short APD because the large magnitude of the hERG current (its conductance value is the same as in Li et al. [2017b]), accelerates the repolarisation process and shortens the action potential. The resultant APD (~ 150 ms) is shorter than a reasonable APD of a healthy cardiomyocyte, which is at about 280 ms. The modified AP models are calibrated by tuning the hERG conductance to ensure that the simulated APD_{90} (paced at 1 Hz with an amplitude of -80 A/F) is the same as that in the original AP model. The hERG conductance is thus scaled by 0.1012, 0.3877, 0.8231, and 0.8744 for the Grandi-Li model, the ten Tusscher-Li model, the Tomek-Li model, and the ORd-Lei model, respectively. The AP and hERG current after calibration is shown in Figure 4.5 in solid red lines. Due to the change of the hERG current within each AP model, there are slight differences in the repolarisation phase between the original AP models and the calibrated AP models.

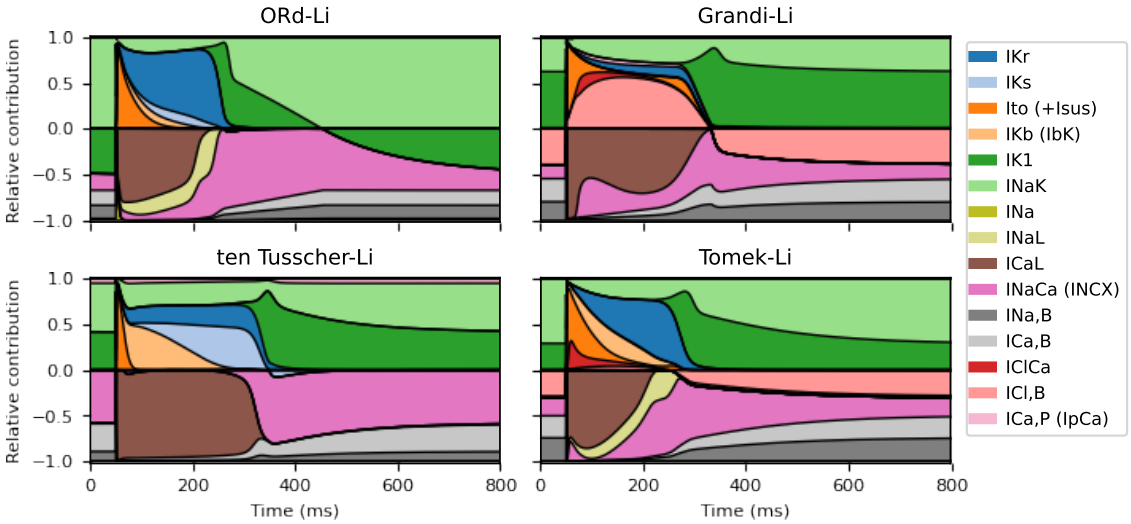


Figure 4.6: **Relative current contributions making up each modified AP model remain different.** The ionic currents of each calibrated AP model are grouped into positive and negative currents and normalised. The relative current contribution of each current to the overall AP is shown for the ORd-Li model [Dutta et al., 2017], the Grandi-Li model, the ten Tusscher-Li model, and the Tomek-Li model. The APs are paced at 1 Hz with an amplitude of −80 A/F till steady state.

Figure 4.6 compares the inward and outward currents of the calibrated AP models, as opposed to the original AP models in Figure 4.3. The constituent ionic currents are similar to the original AP models in Figure 4.3, even though the formulation of the hERG-channel component is different.

4.3.2 Synthetic drugs

The drugs used in this study are dofetilide, verapamil, bepridil, terfenadine, cisapride, ranolazine, quinidine, sotalol, chlorpromazine, ondansetron, diltiazem, and mexiletine, which are the training drugs used for CiPA [Colatsky et al., 2016]. The parameterisation of these drugs into the drug-binding formulations shall be referred to as *synthetic drugs*. Parameters of the drugs for the SD formulation are based on experimental data by Li et al. [2017b]. Dofetilide and verapamil are chosen to represent trapped and nontrapped drugs, respectively, as they have the lowest and highest V_{half} (the trapping tendency parameter), except for cisapride. Cisapride is not used to represent nontrapped drugs because its low binding rate K_{max} limits the inhibition of the hERG current. The hERG current inhibition predicted by the Li-SD model with the parameter abstraction of dofetilide and verapamil is not the actual drug itself.

Therefore, we will refer to the parameter abstraction of dofetilide and verapamil as example drug T and example drug N, for trapped and nontrapped drugs.

In addition to the calibration of the AP models under drug-free conditions, the hERG current inhibition by the synthetic drugs in the Li-SD model is recalibrated for the Lei-SD model due to their very different hERG current behaviour. The calibration is performed by inferring the drug-related parameters from the experimental data published by Li et al. [2017b] for the Lei-SD model. The experimental data are the recorded hERG currents stimulated with a modified Milnes' protocol [Li et al., 2017b], which is designed to induce the trapping behaviour of drugs. The protocol lasts 25 s with a resting potential at −80 mV and a voltage step of 0 mV for 10 s, repeated for 10 pulses. The recorded hERG currents are then normalised against those recorded at drug-free conditions to obtain the *fractional block*. Only fractional block data at the 0 mV step were provided in Li et al. [2017b]. Following the same data processing method from Li et al. [2017b], hERG current data at four drug concentrations are combined in an optimisation run. The parameters are optimised to minimise the root mean squared difference (RMSD) between the simulated and actual normalised current. Parameters K_{max}, K_u and EC_{50}^n are logarithmic-transformed, while parameters n and $V_{\text{half-trap}}$ are optimised as is. The optimisation is performed using the covariance matrix adaptation-evolution strategy (CMA-ES) algorithm [Hansen et al., 2003] with the PINTS software package [Clerx et al., 2019]. The optimisation is repeated 5 times with different initial points, and the best is chosen. The initial guesses are sampled from log-uniform distribution for $K_{\text{max}} \in [10^5, 10^{10}]$, $K_u \in [10^{-7}, 10^{-5}]$, $\text{EC}_{50}^n \in [10^7, 10^9]$ and uniform distribution for $n \in [0, 1]$ and $V_{\text{half-trap}} \in [-100, 0]$. Boundaries were set at $K_{\text{max}} \in [10^{-1}, 10^{10}]$, $K_u \in [10^{-8}, 10]$, $\text{EC}_{50}^n \in [1, 10^{10}]$, $n \in [0, 2]$ and $V_{\text{half-trap}} \in [-200, 0]$ to ensure physiologically reasonable parameter values.

Table 4.1 lists the drug-related parameter values for the Lei-SD model after recalibration. Figure 4.7 compares the drug-related parameter values for the SD formulation in the Li hERG-channel component and the Lei component. The right panels show the distribution of the parameter values of all synthetic drugs. $V_{\text{half-trap}}$ and n remain largely the same for both hERG-channel components, while K_{max}, K_u, and EC_{50}^n are mostly larger for the Lei component than the Li component, with wider distributions too. The RMSD shows the difference between the fitted simulation output and the experimental data across all synthetic drugs. Both the Li and the Lei hERG-channel components are able to fit the experimental data equally well. As noted in Chang et al. [2017], K_{max} and EC_{50}^n are not identifiable individually, but identifiable as an expression.

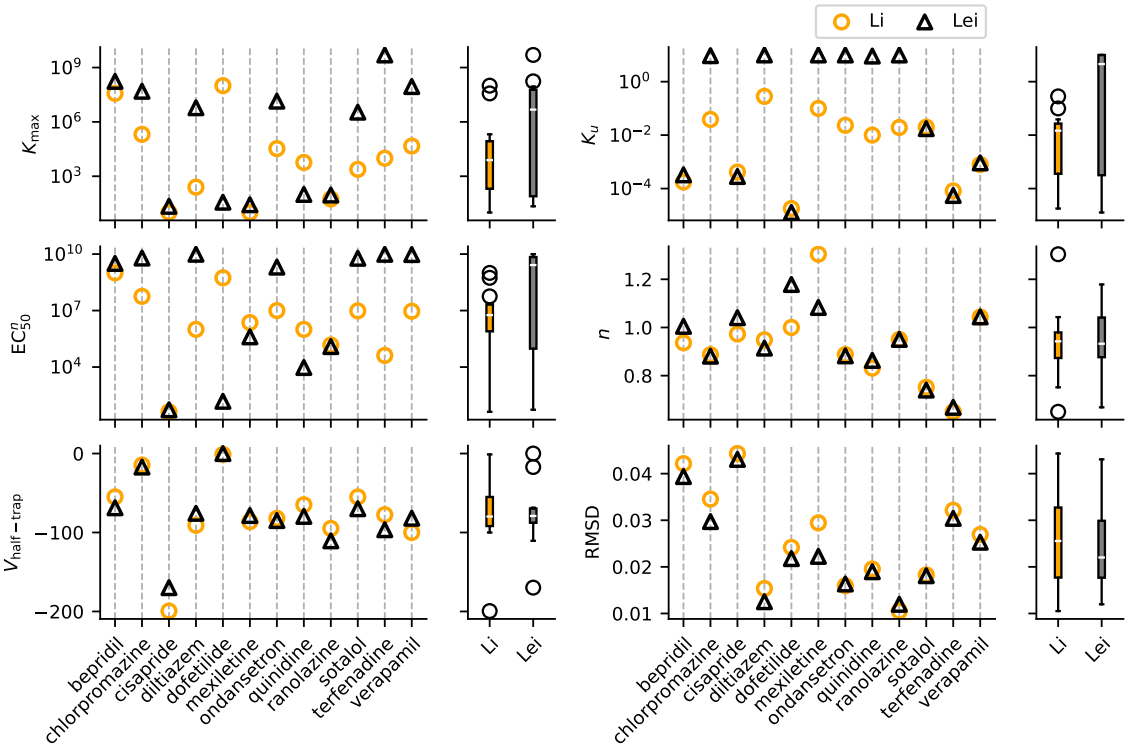


Figure 4.7: **Comparison of the drug-related parameters for the Li-SD model and the Lei-SD model.** Panels on the left plot the parameter values and the root mean squared difference (RMSD) between the simulated data and the experimental data for the 12 synthetic drugs within the Li hERG-channel component and the Lei component. Panels on the right show the distribution of each inferred parameter and the RMSDs.

Table 4.1: Parameter values of the Lei-SD model for all synthetic drugs.

Drug	K_{max}	K_u (ms^{-1})	EC_{50}^n (nM)	n	$V_{\text{half-trap}}$ (mV)
Dofetilide	3.699×10^1	1.272×10^{-5}	1.520×10^2	1.178	-1.048×10^{-4}
Verapamil	8.939×10^7	9.152×10^{-4}	9.461×10^9	1.044	−81.91
Bepridil	1.732×10^8	3.263×10^{-4}	3.191×10^9	1.005	−68.66
Terfenadine	4.962×10^9	5.515×10^{-5}	9.771×10^9	0.6685	−96.46
Cisapride	2.199×10^1	2.830×10^{-4}	5.403×10^1	1.041	−169.8
Ranolazine	9.196×10^1	9.982	1.222×10^5	0.9505	−110.5
Quinidine	9.944×10^1	9.099	9.467×10^3	0.8635	−79.57
Sotalol	3.458×10^6	1.749×10^{-2}	5.987×10^9	0.7402	−69.87
Chlorpromazine	4.925×10^7	9.404	6.157×10^9	0.8808	−17.11
Ondansetron	1.398×10^7	9.981	2.038×10^9	0.8823	−84.84
Diltiazem	6.057×10^6	1.000×10^1	9.813×10^9	0.9143	−75.68
Mexiletine	2.680×10^1	9.991	4.031×10^5	1.083	−78.49

4.3.3 Parameter space exploration and definition of similarity range

The effect of a drug on the hERG current is parameterised by five drug-related parameters within the SD drug-binding formulation — K_u, K_{max}, n, EC_{50} and $V_{\text{half-trap}}$ (Equations 4.1–4.3). The SD formulation can be simplified to three drug-related parameters — K_u, K_{max} and $V_{\text{half-trap}}$, as is done in Chapter 3.2.7. Based on Equation 4.1, the drug concentration D can be nondimensionalised by treating the expression $\frac{D^n}{D^n+\text{EC}_{50}^n}$ as a dimensionless drug concentration input. Given a drug concentration, this expression scales the intensity of the drug effect from a spectrum of no drug effect to maximum effect. This scaling follows a sigmoidal function described by the drug-related parameters EC_{50} and n. Since the full spectrum of drug effects will be covered in the comparisons, EC_{50} and n do not have an influence on the kinetics of the drug-binding mechanism, but affect only the exact drug concentration where the same intensity of drug effect is achieved. The SD formulation can thus be simplified to the three remaining parameters — K_u, K_{max} and $V_{\text{half-trap}}$. To investigate the behaviour of the SD drug-binding formulation, we explore the parameter space of K_u, K_{max}, and $V_{\text{half-trap}}$ (reflecting the unbinding rate, binding rate, and untrapping rate, respectively). The parameter space encompasses the minimum and maximum parameter values of all synthetic drugs (see Section 4.3.2), which is $K_{\text{max}} \in [1, 10^9]$,

$K_u \in [10^{-5}, 1]$ and $V_{\text{half-trap}} \in [-200, -1]$ for the Li-SD model. As mentioned in Section 4.3.2, the drug-related parameters were recalibrated for the Lei-SD model, thus the corresponding parameter space is $K_{\text{max}} \in [10, 10^{10}]$, $K_u \in [10^{-5}, 10]$ and $V_{\text{half-trap}} \in [-170, 0]$ (Table 4.1). Parameter combinations within the parameter space do not necessarily represent any real drug, and are thus referred to as *virtual drugs*. For each virtual drug, the current inhibition caused by the SD formulation is simulated based on the parameter values; while the equivalent current inhibition by the CS formulation is inferred from the Hill curve (detailed in Section 3.2.3). The current inhibition of both the SD and the CS formulation are simulated for the full spectrum of drug effect, from at most 5% peak reduction to at least 95% peak reduction. The APD_{90}s predicted by these formulations are then compared with the metrics described in Section 4.3.4.

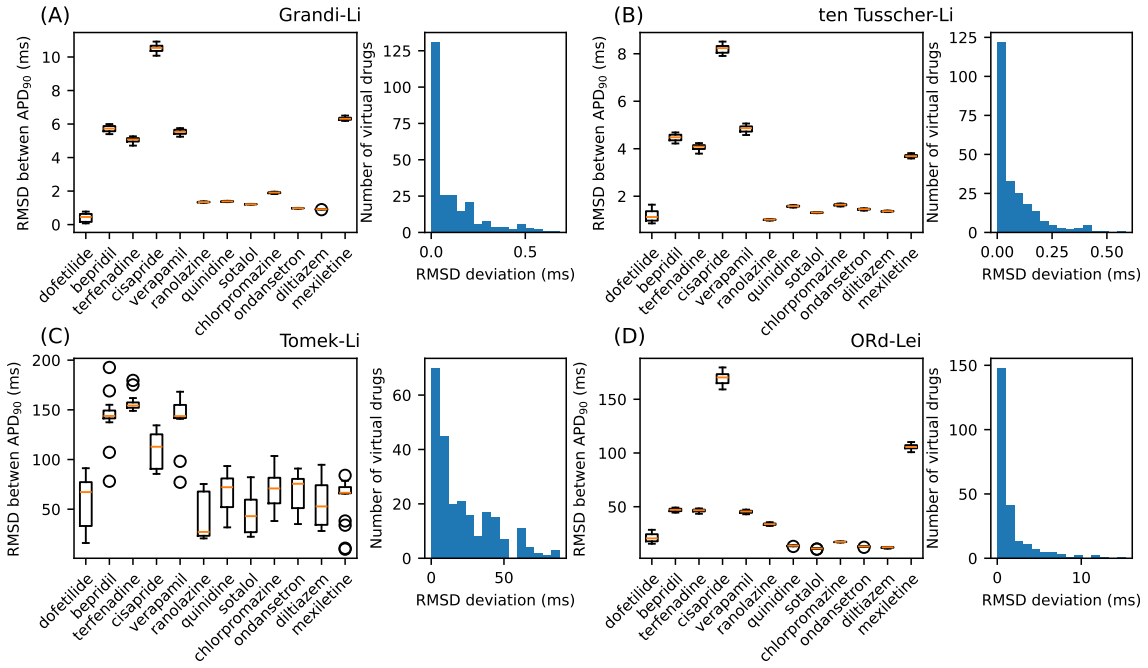


Figure 4.8: **The RMSD of all synthetic drugs with varying parameter n.** For each synthetic drug, its parameter n is varied to give a different parameter combination. (Left panels) The RMSD of the synthetic drugs and their virtual drugs plotted as boxplots for (A) the Grandi-Li model, (B) the ten Tusscher-Li model, (C) the Tomek-Li model and (D) the ORd-Lei model. (Right panels) A histogram of the difference in RMSD between the synthetic drugs and their corresponding virtual drugs.

In addition to measuring the exact difference between the SD and the CS drug-binding formulation, we want to identify situations where it does not matter which

drug-binding formulation is chosen. We therefore define a similarity range where if the difference between the drug-binding formulations lies within this range, they are said to be similar, that is, the use of either drug-binding formulation does not affect the APD_{90} prediction. This is based on the assumption that the parameter values and Hill curves (used to characterise the synthetic drugs for the SD and the CS formulation) derived from experimental data contain variability and uncertainty. This variability and uncertainty will propagate to the output APD_{90}, therefore a small margin of difference in APD_{90} is acceptable.

Since the drug-related parameters EC_{50} and n affect only the actual drug concentration where the AP prolongation, including the early afterdepolarisation (EAD)-like behaviour, is observed, we utilise this property to define the similarity range. Following that, a synthetic drug with its parameter n perturbed should not have a significant difference with its corresponding synthetic drug, and thus should be classified as similar to its corresponding synthetic drug. Therefore, we define the similarity range to be the maximum difference between the synthetic drug and those with n perturbed (Figure 4.8), multiplied by a factor of 1.2. That way, all synthetic drugs with their value of n perturbed will still lie within the similarity range. To obtain this similarity range, we first take each synthetic drug and perturb the parameter n to range from 0.65 to 1.304 (the minimum and maximum of n within all synthetic drugs) to form a group of parameter combinations for that synthetic drug. These parameter combinations are referred to as *virtual drugs* as they might not represent any real drug. These parameter combinations are each incorporated into the AP model by the SD and the CS drug-binding formulation, then their APD_{90}s are compared to measure the difference. This procedure, including forming virtual drugs and comparing APD_{90} between the SD and the CS formulation, is performed for all synthetic drugs and all AP models. The boxplots in Figure 4.8 show the distributions of the difference between the SD and the CS formulations for the groups of virtual drugs. The narrow distributions indicate that the RMSD is not affected by the change in parameter n. The Tomek-Li model has wider distributions because it is highly sensitive to perturbations in the hERG current as its AP is heavily reliant on the contribution from the hERG current (Figure 4.6). However, such magnitudes of RMSD are lower than the difference observed in the parameter space (Figure 4.11). Note that the RMSDs within the boxplot distributions shown in Figure 4.8 are the difference in APD_{90} between the SD and the CS formulation for each virtual drug. Each synthetic drug has its own baseline value of RMSD (the difference between the SD and the CS formulation). The change in parameter n in each synthetic drug creates a virtual drug,

and this virtual drug has an RMSD value that deviates from the baseline RMSD of the synthetic drug. Since the change in parameter n should not affect the difference between the drug-binding formulations, the maximum deviation should be within the similarity range. The similarity range is thus defined to be the maximum deviation observed in all synthetic drugs, multiplied by a factor of 1.2. The histograms in Figure 4.8 compile the deviations observed in all synthetic drugs. The similarity range are thus 0.820 ms, 0.704 ms, 106.1 ms and 18.37 ms for the Grandi-Li model, the ten Tusscher-Li model, the Tomek-Li model and the ORd-Lei model, respectively.

This definition of the similarity range is based on the changes in APD_{90} of each AP model, and is dependent on the AP model's sensitivity to the hERG current. Due to the variability in experimental data, 10 ms is commonly used as an acceptable range for uncertainties, for example, in drug-safety testing [Hammond et al., 2001]. An absolute similarity threshold of 10 ms is also used as comparison to a similarity range that is dependent on the AP model (Section 4.4.3).

4.3.4 Measuring the difference between the drug-binding formulations

Similar to the comparison between drug-binding formulations performed in Chapter 3, the difference between the drug-binding formulations are measured by the signed root mean square different (RMSD) between the drug-binding formulations based on the APD_{90} and the qNet score (Section 3.2.8). AP models without the specific currents included in the qNet score will be treated as having zero-value currents.

Following Section 3.2.8, the signed RMSD is given by

$$\mathrm{RMSD}_{\mathrm{signed}} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(SD_i - CS_i)^2} \times \frac{\mathrm{MD}}{|\mathrm{MD}|}, \tag{4.6}$$

where the mean difference $\mathrm{MD} = (1/N) \cdot \sum_{i=1}^{N}(SD_i - CS_i)$ and, SD_i and CS_i are APD_{90} or qNet scores of the SD and the CS formulations, respectively, at a drug concentration i. The signed RMSD between an AP-IKr-SD model and an AP-IKr-CS model will be represented using the symbol Δ for convenience, giving ΔAPD_{90} and ΔqNet for signed RMSD of APD_{90} and qNet scores, respectively.

In some cases, it may be difficult to compare the actual APD_{90} or qNet score difference, as each AP model has a different sensitivity to perturbations in the hERG current and thus exhibits various magnitudes of AP prolongation. An example shown in Figure 4.9 showcases the different magnitudes of AP prolongations. The top left

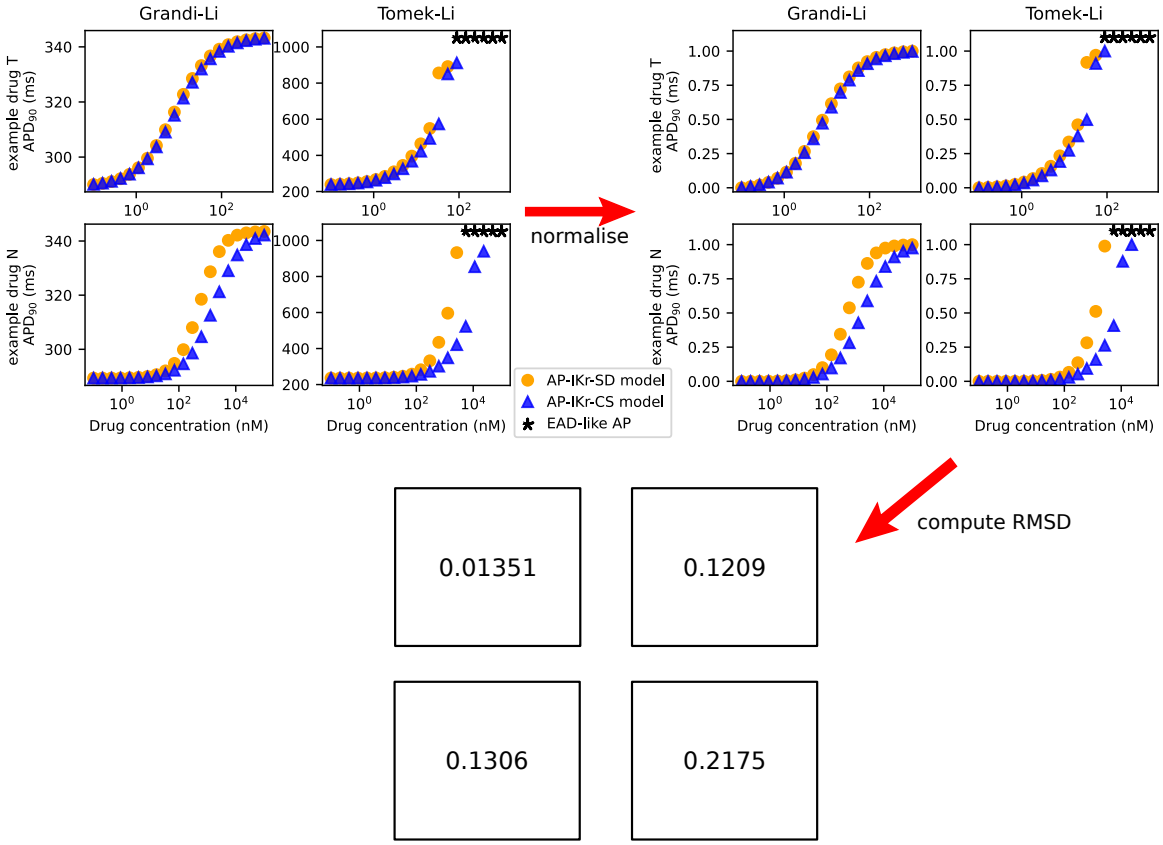


Figure 4.9: **Example of the calculation procedure of $\Delta\widetilde{\text{APD}}_{90}$ for the Grandi-Li model and the Tomek-Li model.** APD_{90} is taken as an example here. The Grandi-Li model and the Tomek-Li model are chosen to emphasise the difference. A similar procedure is performed for the qNet score and other AP models. (Top left) We first simulate the APD_{90}s of example drug T and example drug N with the SD (orange circle) and the CS (blue triangle) drug-binding formulations. The asterisks indicate that the simulated AP exhibits EAD-like behaviour. (Top right) All APD_{90}s within each panel, regardless of the SD and the CS formulations, are normalised to a range of zero to one. The min-max normalisation is performed with the minimum and maximum of all APD_{90}s within each panel. (Bottom) The signed RMSD is then computed from the normalised APD_{90}s by Equation 4.6.

group of panels shows the APD_{90} of example drug T and example drug N for the Grandi-Li model and the Tomek-Li model. These two AP models are chosen for this example to emphasise the difference in APD_{90} magnitude simulated from different AP models. The maximum AP prolongation predicted by the Grandi-Li model for both example drugs is at around 50 ms, whereas the maximum AP prolongation of the Tomek-Li model is close to 800 ms. Calculating the signed RMSD from APD_{90} or qNet score will provide a biased view of the difference between the SD and the CS drug-binding formulations. Therefore, the APD_{90}s and qNet scores are normalised to the same scale so that different AP models are comparable. All data points (including the APD_{90}s of both the SD and the CS drug-binding formulation) within each panel are normalised to a range of zero to one. The normalised APD_{90}s are shown in the top right group of panels. This normalisation method is performed because the APD_{90}s are simulated for the full scale of current inhibition, from less than 5% inhibition to more than 95% inhibition. The APD_{90}s simulated for each drug-binding formulation cover almost the full range of AP prolongation. Mathematically, the normalised values, $\widetilde{SD}_i$ and $\widetilde{CS}_i$, are defined as

$$\widetilde{SD}_i = \frac{SD_i - \min_i\{SD_i, CS_i\}}{\max_i\{SD_i, CS_i\} - \min_i\{SD_i, CS_i\}}, \tag{4.7}$$

$$\widetilde{CS}_i = \frac{CS_i - \min_i\{SD_i, CS_i\}}{\max_i\{SD_i, CS_i\} - \min_i\{SD_i, CS_i\}}, \tag{4.8}$$

where $\min_i\{SD_i, CS_i\}$ and $\max_i\{SD_i, CS_i\}$ are the minimum and maximum values, respectively, of all simulated APD_{90}s or qNet scores for a single virtual drug in the SD and the CS drug-binding formulation. Finally, the signed RMSD is calculated with normalised APD_{90}s or qNet scores according to Equation 4.6, as presented in the bottom group of panels in Figure 4.9. The signed RMSD of normalised APD_{90} or qNet score shall be denoted as $\Delta\widetilde{APD}_{90}$ or $\Delta\widetilde{qNet}$, respectively.

4.3.5 Simulations

All simulations were run with Python 3 using Myokit 1.35 [Clerx et al., 2016] with the CVODE solver [Hindmarsh et al., 2005]. The absolute tolerance and relative tolerance were set to 10^{-7} and 10^{-8}, respectively. A summary of the computational details for all coding implementations of this thesis is provided in Appendix A.

4.4 Results

4.4.1 The relative I_{Kr} contribution in each AP model affects APD – dofetilide and verapamil

We first compared the drug-binding formulations embedded within all AP models (the Grandi-Li model, the ten Tusscher-Li model, the Tomek-Li model, and the ORd-Lei model), for two sets of parameter values that represent a trapped drug and a nontrapped drug. These sets of parameter values were taken from Li et al. [2017b]. Dofetilide is chosen to represent a trapped drug, and will be referred to as example drug T. Verapamil is chosen to represent a nontrapped drug as it has the second lowest trapping tendency, reflected in its low $V_{\mathrm{half-trap}}$ value. Cisapride, which has the lowest trapping tendency, is not chosen because it has a low binding rate, which restricts the hERG current inhibition even at high drug concentrations. Parameter values that represent verapamil will be referred to as example drug N. These two example drugs were incorporated into the AP models with the SD or the CS drug-binding formulation embedded within, then simulated for a range of drug concentrations. APD_{90} and qNet scores were computed from the simulated APs and compared in terms of $\Delta\widetilde{\mathrm{APD}}_{90}$ and $\Delta\widetilde{\mathrm{qNet}}$ (details of the metrics are described in Section 4.3.4).

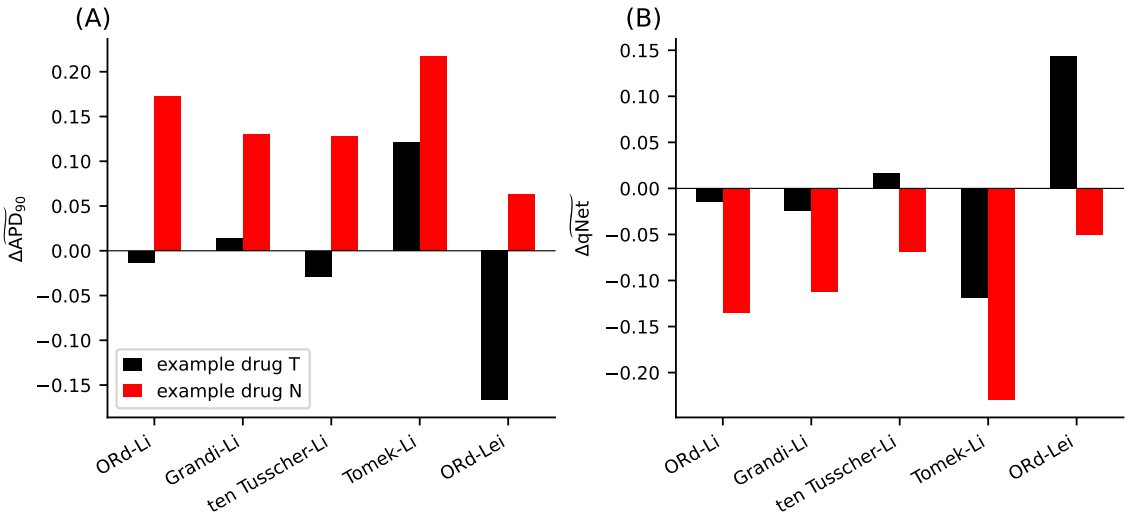


Figure 4.10: **Example drug N shows a larger magnitude of $\Delta\widetilde{\mathrm{APD}}_{90}$ and $\Delta\widetilde{\mathrm{qNet}}$ across all AP-IKr models, except for the ORd-Lei model.** The (A) $\Delta\widetilde{\mathrm{APD}}_{90}$ and (B) $\Delta\widetilde{\mathrm{qNet}}$ between the AP-IKr-SD model and the AP-IKr-CS model for example drug T and example drug N. Positive values indicate that the AP-IKr-SD model has higher APD_{90} or qNet scores than the AP-IKr-CS model.

Figure 4.10 shows $\Delta\widetilde{\text{APD}}_{90}$ and $\Delta\widetilde{\text{qNet}}$ for example drug T and example drug N in all AP models. A positive value indicates that the AP-IKr-SD model has a larger APD_{90} or qNet score than the AP-IKr-CS model. Example drug T (trapped drug) has a smaller magnitude of $\Delta\widetilde{\text{APD}}_{90}$ and $\Delta\widetilde{\text{qNet}}$ than example drug N (nontrapped drug), except for the ORd-Lei model, implying that the effect of example drug T on APD_{90} and qNet score is minimally affected by the choice of drug-binding formulation. ΔAPD_{90} of example drug T lies within the similarity range (defined in Section 4.3.4) for the Grandi-Li model and the Tomek-Li model, whereas the others, including example drug T in the ten Tusscher-Li model and the ORd-Lei model and example drug N, lie outside the similarity range. This suggests that if the Grandi-Li model or the Tomek-Li model is chosen, hERG current inhibition of example drug T by the SD or the CS drug-binding formulation does not affect the outcome. This counters our hypothesis that a trapped drug will have a large $\Delta\widetilde{\text{APD}}_{90}$, while a nontrapped drug has $\Delta\widetilde{\text{APD}}_{90}$ close to zero, or lies within the similarity range. Assuming the SD formulation captures the trapping mechanism accurately and assuming a trapped drug is more pro-arrhythmic, a trapped drug should show a longer APD_{90} than a nontrapped drug. The CS formulation of a trapped drug is equivalent to a trapped drug's potency but without its trapping mechanism. Therefore, using the SD formulation for a trapped drug should have a larger $\Delta\widetilde{\text{APD}}_{90}$ when it is compared to its CS formulation equivalent. A nontrapped drug in the SD formulation, on the other hand, should have similar APD_{90} to its CS formulation equivalent, and thus small $\Delta\widetilde{\text{APD}}_{90}$.

Positive $\Delta\widetilde{\text{APD}}_{90}$s for example drug T are observed only in the Grandi-Li model and the Tomek-Li model. A positive $\Delta\widetilde{\text{APD}}_{90}$ is expected for example drug T following the reasoning described above, where the trapped drug should have a longer APD_{90} in the SD drug-binding formulation than in the CS drug-binding formulation. The drug-trapping mechanism incorporated in the SD formulation attempts to capture the accumulation of drugs in the ion channel, which will cause long-lasting and more intense current inhibition than the CS formulation, thus prolonging the APD_{90} more than the CS formulation. The other models show negative $\Delta\widetilde{\text{APD}}_{90}$ for example drug T, indicating that the AP-IKr-SD model has less AP prolongation on average. On the other hand, $\Delta\widetilde{\text{APD}}_{90}$s for example drug N are all positive. However, example drug N is not expected to be affected by the choice of drug-binding formulation because the inclusion of the trapping mechanism should not have an effect on the output for nontrapped drugs.

The ORd-Lei model has a larger magnitude of $\Delta\widetilde{\text{APD}}_{90}$ for example drug T than for example drug N, different from the rest of the AP models. The SD formulation for example drug T predicts significantly less AP prolongation than the CS formulation. This can be explained by the low binding rate (K_{max}) of example drug T within the SD formulation for the ORd-Lei model (Figure 4.7 and Table 4.1), which slows down the binding of drugs to the channel and lessens the hERG current inhibition and eventually lessens the AP prolongation. The CS formulation, on the other hand, is unaffected by the binding rate and can fully inhibit the hERG current. Consistently, the $\Delta\widetilde{\text{qNet}}$ for example drug T and example drug N in the ORd-Lei model has an opposite direction to $\Delta\widetilde{\text{APD}}_{90}$.

$\Delta\widetilde{\text{qNet}}$s are mainly negatives, except for example drug T in the ten Tusscher-Li model and the ORd-Lei model. $\Delta\widetilde{\text{qNet}}$s are expected to be of opposite sign from $\Delta\widetilde{\text{APD}}_{90}$s because the AP prolongation is caused by a decrease in the hERG current, which lowers the qNet scores (Equation 3.6).

Across all AP models, the $\Delta\widetilde{\text{APD}}_{90}$ and $\Delta\widetilde{\text{qNet}}$ are generally relatively smaller in magnitude for the Grandi-Li model and the ten Tusscher-Li model than the ORd-Li model and the Tomek-Li model. This trend is proportional to the magnitude of the hERG current within each AP model, where the Grandi model and the ten Tusscher model have the lowest hERG current contribution to the AP (Figures 4.6 & 4.5). The inhibition of the hERG current in an already low contribution will not affect the AP as much.

These example drugs, representing trapped drug and nontrapped drug, show that the APD_{90}s and qNet scores are not only dependent on the drug-binding formulation, but also on the AP model and the hERG-channel component. Moreover, the differences are not associated with the trapping tendency of the drugs, as we might have expected. Example drug T does not cause more AP prolongation with the SD drug-binding formulation, while example drug N does not predict similar AP prolongation between the two drug-binding formulations. To investigate whether this outcome is generalisable to the parameter space of the SD drug-binding formulation, we then perform a systematic exploration of the parameter space (detailed methods in Section 4.3.3).

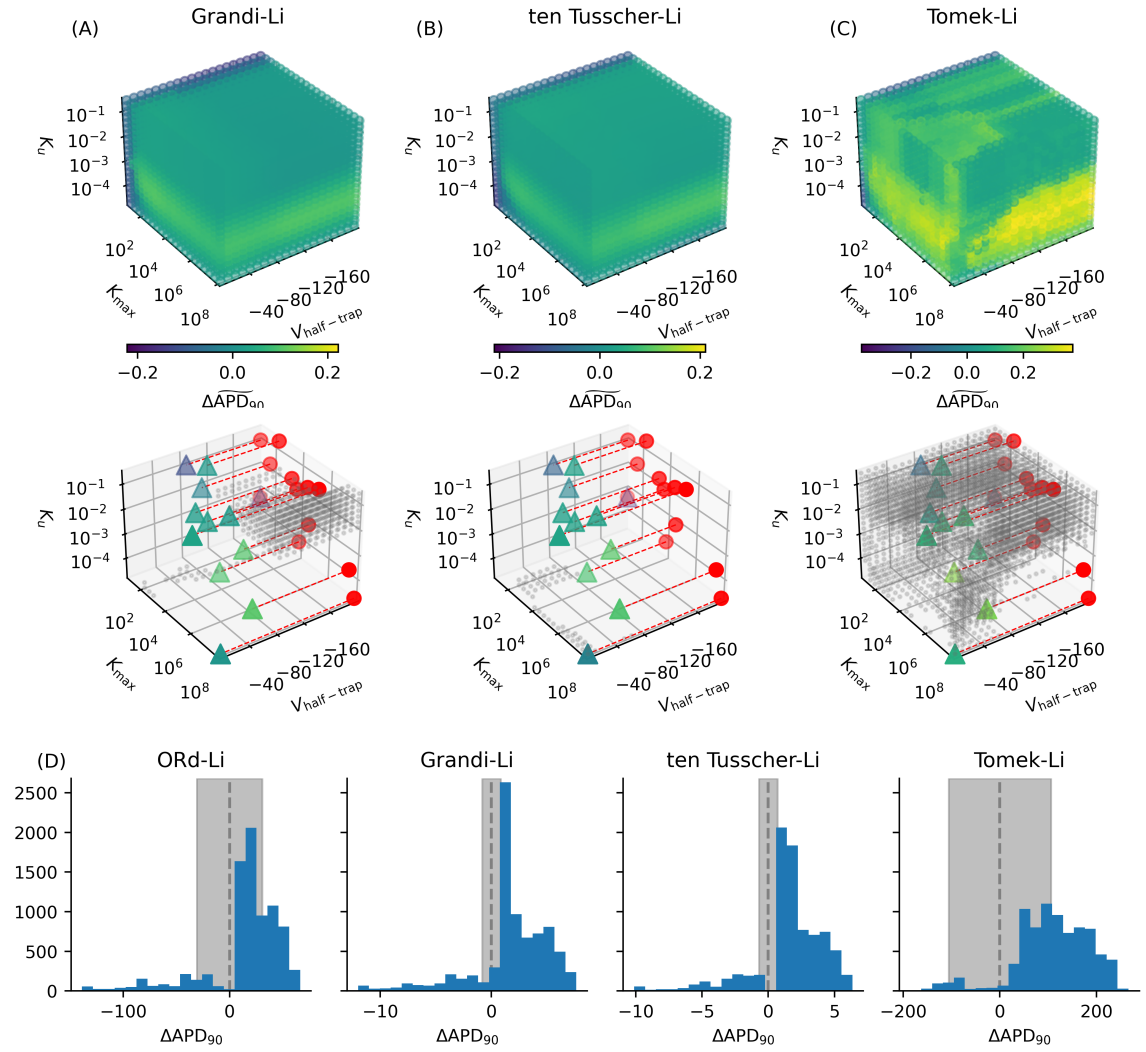


Figure 4.11: **The majority of the drugs in the AP-IKr-SD model have higher APD_{90}s than the AP-IKr-CS model, with the largest difference observed in the Tomek-Li model.** $\widetilde{\Delta APD_{90}}$s within the parameter space of the SD drug-binding formulation for **(A)** the Grandi-Li model, **(B)** the ten Tusscher-Li model, and **(C)** the Tomek-Li model. The colours of the markers indicate the $\widetilde{\Delta APD_{90}}$ of each point. The grey circles in the bottom panels of (A), (B), and (C) are regions where ΔAPD_{90} lies within the similarity range (Section 4.3.3). The triangles are the 12 synthetic drugs listed in Section 4.3.2, colour-coded with their $\widetilde{\Delta APD_{90}}$ values. These triangles are projected to the K_{max} - K_u plane as red circles for better visualisation. **(D)** Histograms of ΔAPD_{90}, without normalisation, for each AP model in (A), (B) and (C). The shaded region indicates the region where the ΔAPD_{90} lies within the similarity range. The dashed vertical line indicates zero ΔAPD_{90}. Positive values indicate that the AP-IKr-SD model has longer APD_{90} than the AP-IKr-CS model.

4.4.2 The binding and unbinding rates remain the primary factors contributing to the differences between the SD and the CS drug-binding formulation

Example drug T and example drug N, taken to represent a trapped drug and a nontrapped drug, respectively, differ not only in their trapping tendency parameter ($V_{\text{half-trap}}$), but also in their binding and unbinding rates (K_{max} and K_u). We therefore further investigate the role of each parameter in the AP prolongation by exploring the parameter space of K_{max}, K_u, and $V_{\text{half-trap}}$ (see Section 4.3.3 for details). For each virtual drug, its drug effect, as captured by the SD drug-binding formulation, is used to calibrate the equivalent drug effect in the CS formulation (a brief description in Section 4.2.3 and a detailed description in Chapter 3.2.3). The virtual drug is then incorporated into the AP models to predict its AP prolongation. The APD_{90}s simulated by the AP-Li-SD model and the AP-Li-CS model are compared using the $\Delta\widetilde{APD}_{90}$ (Section 4.3.4).

In the top panels of Figures 4.11A-C, the $\Delta\widetilde{APD}_{90}$ between the drug-binding formulations for the Grandi-Li model, the ten Tusscher-Li model, and the Tomek-Li model are shown. The dependence of $\Delta\widetilde{APD}_{90}$ on the parameters exhibits similar behaviour to that in the previous chapter (Chapter 3), where the AP-Li-SD model predicts longer APD_{90}s than the AP-Li-CS model in the majority of the cases. Furthermore, the maximum positive $\Delta\widetilde{APD}_{90}$ is achieved at the middle range of K_u values. The AP-Li-SD models predict shorter APD_{90}s than the AP-Li-CS models at low K_{max} values, as indicated by negative $\Delta\widetilde{APD}_{90}$s. $\Delta\widetilde{APD}_{90}$s remain relatively constant as $V_{\text{half-trap}}$ increases, while they increase rapidly then plateau as K_{max} increase. While the Grandi-Li model and the ten Tusscher-Li model show relatively smooth changes in $\Delta\widetilde{APD}_{90}$ across the parameter space, the Tomek-Li model exhibits large fluctuations in the $\Delta\widetilde{APD}_{90}$ due to the high sensitivity of the Tomek model to perturbations in the hERG current and thus exhibits EAD-like behaviour frequently.

The bottom panels of Figures 4.11A-C show the $\Delta\widetilde{APD}_{90}$ of the 12 synthetic drugs in triangles, and the data points with $\Delta\widetilde{APD}_{90}$ within the similarity range (Section 4.3.3) in grey. The similarity range, where the difference in AP between the AP-IKr-SD model and the AP-IKr-CS model are considered to be negligible, is adjusted according to the AP model. $\Delta\widetilde{APD}_{90}$s that lie within the similarity range form the similarity region. The similarity region encompasses a small volume in the Grandi-Li model and in the ten Tusccher-Li model. The seeming dependence of the similarity region on $V_{\text{half-trap}}$ arises from the strict cut-off point of the threshold. Points around the similarity region are within 1 ms of $\Delta\widetilde{APD}_{90}$ in the $V_{\text{half-trap}}$ axis of

the Grandi-Li and the ten Tusscher-Li model. The similarity region for the Tomek-Li model, on the other hand, covers a larger volume, predominantly at high K_u values. This implies that the Tomek-Li model is robust against the choice of the drug-binding formulation with approximately 44% of the virtual drugs within the parameter space cause similar AP prolongation regardless of the drug-binding formulation used. This can be compared to more than 90% of the virtual drugs exhibiting AP prolongation beyond the acceptable margin if the Grandi or the ten Tusscher model is used to implement the SD formulation.

Figure 4.11D shows the histograms of the ΔAPD_{90}s for each AP model. The distributions of the ΔAPD_{90} for all AP models are skewed towards positive values. Although occurrences of negative values (APD_{90}s of the AP-Li-SD model are on average shorter than those of the AP-Li-CS model) are few, their magnitudes reach larger values in all AP models except the Tomek-Li model. The ranges of ΔAPD_{90} observed in the Grandi-Li model and the ten Tusscher-Li model are significantly lower than those observed in the Tomek-Li model (roughly 20 times lower). This discrepancy aligns with the contribution of the hERG current to the AP model (Figures 4.6 and 4.5). The larger hERG current in the Tomek model allows for greater AP prolongation, resulting in a larger difference between the APD_{90}s.

4.4.3 Changes in the open state probability reduce the effect of drug binding

In addition to the replacement of AP models, we replaced the hERG-channel component of the ORd-Li model (the AP model and hERG-channel component combination chosen to develop the SD drug-binding formulation in Li et al. [2017b]) with the Lei component to form the ORd-Lei model. The drug-binding formulations are then embedded within the ORd-Lei model, and their simulated APD_{90}s are compared. As detailed in Section 4.3.2, the parameter values for the synthetic drugs are recalibrated to capture the same drug effect as in the ORd-Li model.

In Section 4.4.1, we briefly compared the APD_{90} and qNet score between the SD and the CS drug-binding formulation for a representative trapped drug and a nontrapped drug (Figure 4.10). Distinct from the other AP models, the ORd-Lei model has a large negative $\Delta\widetilde{APD}_{90}$ for example drug T (trapped drug representative) and a small positive $\Delta\widetilde{APD}_{90}$ for example drug N (nontrapped drug representative). This can be explained by the low binding rate, controlled by K_{max}, of example drug T for the Lei hERG-channel component (Figure 4.7). The low binding rate limits the amount of drug compound that can bind to the channel, which limits the inhibition

of the hERG current and thus results in lower AP prolongation. In contrast to the CS drug-binding formulation, the ion flow can be fully stopped by fixing the conductance value to near zero. Consistently, the $\Delta\widetilde{\text{qNet}}$ for example drug T and example drug N in the ORd-Lei model has an opposite direction to $\Delta\widetilde{\text{APD}}_{90}$.

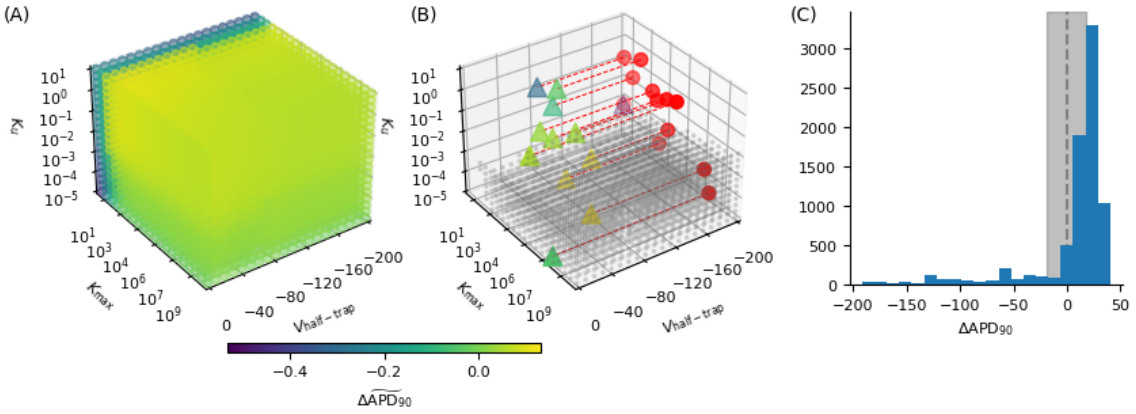


Figure 4.12: **The ORd-Lei-SD model is most likely to return slightly longer APD_{90}s than the ORd-Lei-CS model.** **(A)** $\Delta\widetilde{\text{APD}}_{90}$s within the parameter space of the SD drug-binding formulation for the ORd-Lei model. The colours of the markers indicate the $\Delta\widetilde{\text{APD}}_{90}$ of each virtual drug within the parameter space. **(B)** The grey circles are virtual drugs where $\Delta\widetilde{\text{APD}}_{90}$ is within the similarity range, between -18.37 ms and 18.37 ms. The triangles are the synthetic drugs, colour-coded with their $\Delta\widetilde{\text{APD}}_{90}$ values. These triangles are projected to the K_{max} - K_u plane as red circles for better visualisation. **(C)** Histogram of ΔAPD_{90}, without normalisation, for ORd-Lei model. The shaded region indicates the region where the ΔAPD_{90} lies within the similarity range. The dashed vertical line indicates zero ΔAPD_{90}. Positive values show that the ORd-Lei-SD model has longer APD_{90} than the ORd-Lei-CS model.

Extending the comparison to the parameter space, Figure 4.12A shows the outcome of the parameter space exploration for the ORd-Lei model. The color of each data point in Figure 4.12A indicates the difference between using either the SD drug-binding formulation or the CS formulation for a virtual drug as viewed in terms of $\Delta\widetilde{\text{APD}}_{90}$. Positive values of $\Delta\widetilde{\text{APD}}_{90}$ imply that the SD formulation causes more AP prolongation than the CS formulation, which covers the majority of the parameter space. The difference in APD_{90} is not driven by $V_{\text{half-trap}}$ (the trapping tendency parameter). However, as opposed to the outcome in Figure 4.11, the $\Delta\widetilde{\text{APD}}_{90}$ of the ORd-Lei model is independent of both K_u (unbinding rate) and $V_{\text{half-trap}}$. $\Delta\widetilde{\text{APD}}_{90}$s are negative at small values of K_{max}, covering only a small region of the parameter space while achieving large magnitudes. Similar to the reasoning for example drug T, low values of K_{max} limit the binding of drugs to the ion channel and the inhibition of

the hERG current, which restricts the AP prolongation. The CS formulation, on the other hand, has no restriction in its inhibition of hERG current, and therefore can cause longer AP prolongation than the SD formulation.

Figure 4.12B shows the $\Delta\widetilde{\text{APD}}_{90}$s of the 12 synthetic drugs in triangles and the data points that lie within the similarity range in grey. The data points in grey indicate that these virtual drugs are predicted to cause similar AP prolongation regardless of the drug-binding formulation used. The similarity region is mostly at low K_u values (unbinding rate). The low unbinding rates and high binding rates provide the opportunity for drugs to stay within the ion channel longer and maintain their hERG current inhibition, without limitation from binding rates as before, allowing the SD formulation to exhibit similar current inhibition to the CS formulation. Figure 4.12C summarises $\Delta\widetilde{\text{APD}}_{90}$ from Figure 4.12A into a histogram, showing the difference in APD_{90} without normalisation. The ΔAPD_{90}s are mostly positive, with a long tail in the negative values.

In this study, the similarity range is defined according to each AP model. A similarity range, independent of the AP model, may be more appropriate in certain cases, for example when defining a safety margin for drugs in causing AP prolongation. Figure 4.13 shows the virtual drugs that lie within the absolute similarity range of 10 ms for all AP models. With this absolute similarity range, the ORd-Li model and the Tomek-Li model show that the choice of drug-binding formulation significantly affects the predicted APD90; whereas for the Grandi-Li model or the ten Tusscher-Li model, the predicted APD_{90} is unaffected by the choice of the drug-binding formulation. On the other hand, the similarity region of the ORd-Lei model is similar to that when a relative similarity range is used because the absolute similarity range of 10 ms is close to the ORd-Lei model's relative similarity range of 18.37 ms. The histograms in Figure 4.13B further emphasise this outcome.

4.5 Discussion

This study aims to show the changes observed in AP due to the choice of the drug-binding formulation of the hERG channel. The drug-binding formulations of interest are the SD formulation, where drugs can only bind to the open and inactivated open state of the hERG channel and are able to stay trapped within the hERG channel throughout a voltage-clamp protocol pulse, and the CS formulation, where the inhibition of hERG current caused by the drug is approximated by the conductance reduction following the Hill curve. The AP considered in this study is based on several

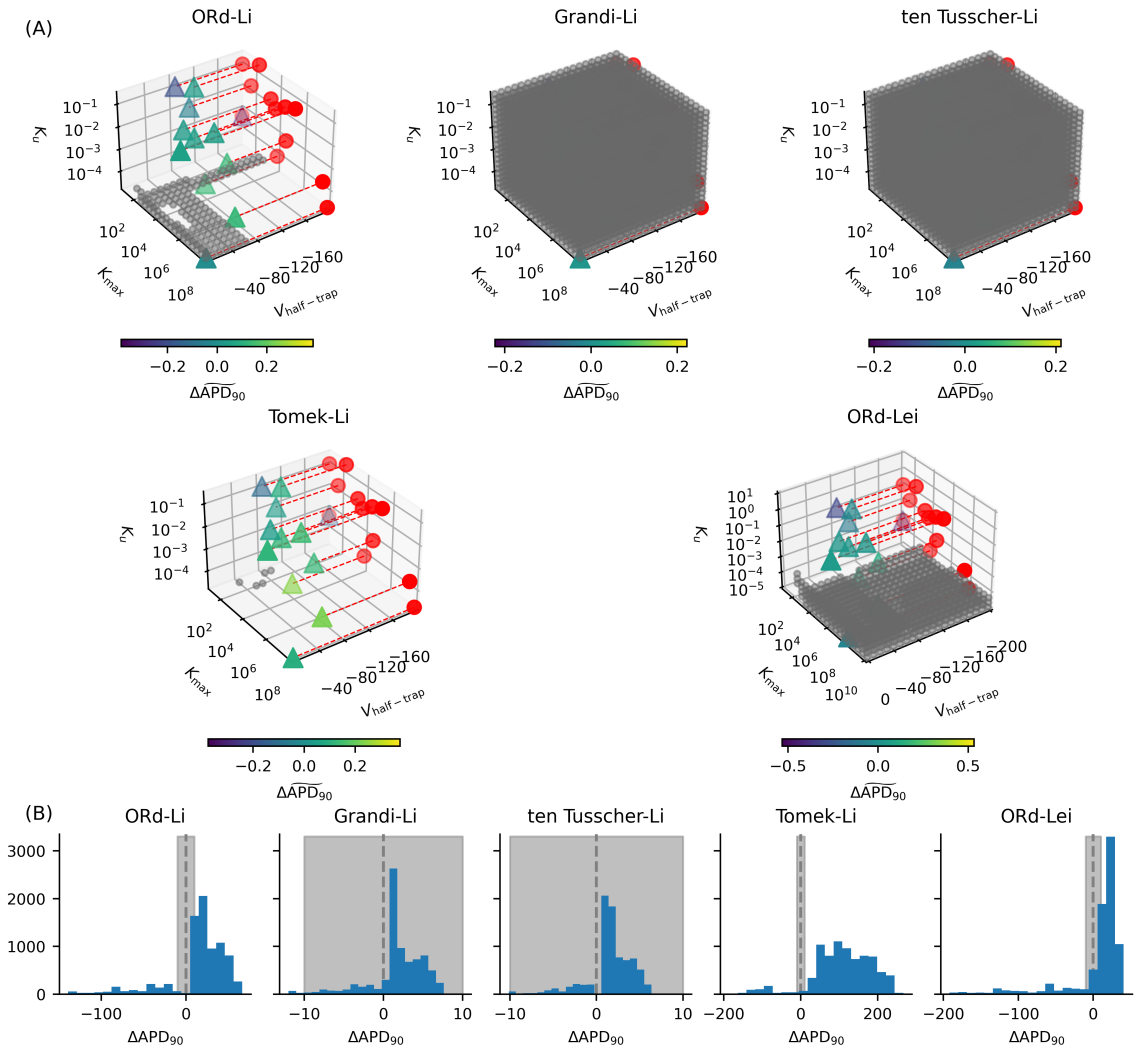


Figure 4.13: **The Grandi-Li model and the ten Tusscher-Li model would effectively predict the same APD_{90} for the whole parameter space regardless of the drug-binding formulation. (A)** $\widetilde{\Delta APD_{90}}$ of the 12 synthetic drugs in coloured triangles for the ORd-Li model, the Grandi-Li model, the ten Tusscher-Li model, the Tomek-Li model, and the ORd-Lei model. These triangles are projected to the K_{max} - K_u plane as red circles for better visualisation. The grey circles are parameter value combinations where $\widetilde{\Delta APD_{90}}$ lies within 10 ms. **(B)** Histograms of ΔAPD_{90}, without normalisation, for all AP models in (A). The shaded region indicates the 10 ms similarity range. The dashed vertical line is at zero ΔAPD_{90}. Positive values indicate that the AP-IKr-SD model has longer APD_{90} than the AP-IKr-CS model.

AP models — the Grandi model, the ten Tusscher model, and the Tomek model. To allow the drug-binding formulations to be embedded within the AP models, the hERG-channel components of the AP models are replaced by the Li hERG-channel component [Li et al., 2017b], to form the AP-Li models. The SD and CS drug-binding formulations were then used to incorporate a drug's effect into the AP-Li models and compared with the metrics $\Delta\widetilde{\text{APD}}_{90}$ and $\Delta\widetilde{\text{qNet}}$. $\Delta\widetilde{\text{APD}}_{90}$ and $\Delta\widetilde{\text{qNet}}$ are the signed RMSD of normalised APD_{90}s and qNet scores between the SD and the CS drug-binding formulation, respectively.

$\Delta\widetilde{\text{APD}}_{90}$s and $\Delta\widetilde{\text{qNet}}$s have smaller magnitude for example drug T (a trapped drug) than for example drug N (a nontrapped drug) (Figure 4.10), except for the ORd-Lei model. The AP prolongation caused by example drug N is affected by whether the SD or the CS drug-binding formulation is used. This is consistent with the conclusion in Chapter 3, where the trapped drug does not invoke more AP prolongation with the SD formulation.

We then explored the parameter space of the SD formulation, which is made up of K_{max}, K_u, and $V_{\text{half}-\text{trap}}$, for all AP-Li models. The parameters correspond to the binding rate, the unbinding rate, and the untrapping rate, respectively. The SD formulation generally predicts longer APD_{90} than the CS formulation (Figure 4.11A-C). The average magnitude of $\Delta\widetilde{\text{APD}}_{90}$ is correlated to the magnitude of the hERG current of each AP model. In all AP-Li models, negative $\Delta\widetilde{\text{APD}}_{90}$s are observed at the extreme low end of the K_{max} range. The low binding rates limit the drug-binding to the hERG channel, lower the effect of hERG current inhibition, and thus have lesser APD_{90} prolongation in the SD formulation than in the CS formulation. The Tomek model has a relatively high hERG current (Figure 4.5) and is more prone to predicting EAD-like behaviour. Since APs showing EAD-like behaviour are omitted from the metric calculation, $\Delta\widetilde{\text{APD}}_{90}$s of the Tomek-Li model in the parameter space are not smooth (Figure 4.11C). Even though the use of the SD formulation changes the APD_{90} of the AP models, these changes are not driven primarily by the trapping tendency parameter (V_{half}), but rather by the binding and unbinding rates (K_{max} and K_u).

Virtual drugs whose $\Delta\widetilde{\text{APD}}_{90}$ is within the similarity range are said to be unaffected by the choice of the drug-binding formulation. This similarity range is defined under the assumption that the parameter n of the SD formulation has a negligible effect on ΔAPD_{90}. The threshold is thus set to be the maximum ΔAPD_{90} caused by the perturbation in the parameter n, and multiplied by a factor of 1.2. The threshold value of the similarity range varies with the AP model and encompasses different

regions within the parameter space (Figure 4.11). It follows from our definition of the similarity range that the threshold value will be model-dependent. It makes sense to have different threshold values for different AP models, as their sensitivities to hERG current inhibition are not identical. After accounting for the difference, the Tomek-Li model has the most virtual drugs that are unaffected by the choice of the drug-binding formulation, while the Grandi-Li model and the ten Tusscher-Li model have a small similarity region. The similarity region indicates drugs where the choice of using either the SD or the CS formulation does not affect the AP prolongations predictions. The implication of the similarity region to a real-world setting is the potential replacement of the SD drug-binding formulation to the CS formulation for drugs within the similarity region to reduce cost in performing complicated experiments and parameterisation of drugs for the SD drug-binding formulation. However, the Tomek-Li model and the ORd-Lei model will have opposing conclusions as their similarity regions encompass different groups of drugs. In most real-world settings, a similarity threshold independent of the AP model may be more appropriate. If an absolute similarity threshold of 10 ms is chosen, the drug-binding formulations do not have an effect on the APD_{90}s of the Grandi-Li model and the ten Tusscher-Li model, while the Tomek-Li model shows the opposite (Figure 4.13).

In addition to altering the AP models, the hERG-channel component of the ORd-Li model is replaced with the Lei component [Lei et al., 2019b]. The ORd-Lei model, although showing a majority of positive $\Delta\widetilde{APD}_{90}$s, exhibits low magnitudes. The negative $\Delta\widetilde{APD}_{90}$s occur at low binding rates, similar to the other AP-Li models. The low binding rates restrict the hERG current inhibition in the SD formulation and therefore cause shorter APD prolongation than the CS formulation. A correlation between the low positive values of $\Delta\widetilde{APD}_{90}$s and the hERG-channel component is complicated by the interaction of several factors. The Lei component has lower open-state and inactivated open-state probabilities under its AP-clamp protocol (Figure 4.4), which would lower the drug-binding opportunities. However, the binding rate and unbinding rate of the 12 synthetic drugs for the Lei component are mostly higher than the Li component. This could be due to compensation against the difference in drug-binding opportunities. The higher binding rate may interact with the lower drug-binding opportunity to achieve a similar drug effect in both drug-binding formulations. The drug-binding opportunity can also be viewed in terms of total time spent in the open-state and inactivated open state. This way, the drug-binding opportunity of the Lei component is higher than that of the Li component. The combination of higher drug-binding opportunity with higher binding rates should inhibit the hERG

current more and prolong the AP more, but the opposite is observed. This observed trend should not be generalised as only a single alternative hERG-channel component has been used.

One of the limitations of this study comes from the calibration of the CS drug-binding formulation. The drug response of the Hill curve, used in the calibration, is measured by the peak of hERG current simulated by the SD formulation using the modified Milnes protocol [Li et al., 2017b]. It is known that the IC_{50} (one of the parameters that defines the Hill curve) is dependent on voltage-clamp protocols [Yao et al., 2005, Kirsch et al., 2004, Stork et al., 2007]. For example, Gomis-Tena et al. [2020] designed protocols to determine the state preference of drug-binding and showed that the IC_{50}s vary with the protocols (as shown in Chapter 3). The dependence of IC_{50} on the voltage-clamp protocol will affect the calibrated drug effect on the CS drug-binding formulation, which then propagates to the simulated APD_{90} and eventually the $\Delta\widetilde{APD}_{90}$. However, it will not affect the qualitative comparison between the SD and the CS formulations, where the trapping tendency of a drug does not play a role in altering $\Delta\widetilde{APD}_{90}$, because the calibration issue changes the reference APD_{90} (of the CS formulation) and not the trapping mechanism of the SD formulation. Moreover, the range of drug concentrations used in the simulations is chosen to encompass the drug effect for the entirety of the Hill curve (from no effect to full inhibition) to ensure that the comparison between the drug-binding formulations is robust against the IC_{50} values. The other limitation of this study is the use of APD_{90} as a surrogate marker to the QT interval when comparing the drug-binding formulations. Finally, this study focuses only on the hERG channel and its drug-binding mechanism. We note that the AP prolongation predicted in this study is not reflective of multi-channel blockers such as verapamil.

4.6 Conclusion

In all the comparisons performed, the binding rate and unbinding rate of the SD drug-binding formulation remain the primary factor affecting the difference in APD_{90}. The trapping component plays virtually no role in the APD_{90} difference between the SD and the CS drug-binding formulations. This study has shown that regardless of the AP model or the hERG-channel component, the role of the trapping component, specifically the parameter $V_{\text{half}-\text{trap}}$, remains minor in affecting the APD_{90} prolongation. The magnitude of the difference in the APD_{90} varies significantly across models, some of which exceeds the accepted range of a drug safety margin. It demonstrates

the importance of choosing a suitable model for specific situations, considering the purpose and intent, as it could lead to inconsistent conclusions. More care is required in designing models replicating the drug-binding mechanism and in extrapolating to out of context-of-use. An experimental exploration is needed to determine the most appropriate choice of AP model and hERG-channel component, and the resulting uncertainty should be taken into account in future drug-binding studies.

The lack of influence of $V_{\text{half}-\text{trap}}$ on the difference in APD_{90} suggests that it may not be identifiable. An identifiability analysis on the SD formulation is required to determine if $V_{\text{half}-\text{trap}}$ can be estimated from the experimental data. It can also inform improvements to the SD formulation and potentially the modelling of drug-trapping. In the next chapter, we look into the practical identifiability of the SD drug-binding formulation and potential ways to improve the drug-binding formulation.

Chapter 5

Identifiability of the trapping tendency parameter and its role within the state-dependent drug-binding formulation

Overview:

In Chapters 3 and 4, we performed sensitivity analyses on the state-dependent (SD) drug-binding formulation. These analyses were done by simulating its action potential (AP) for parameter combinations covering the parameter space and determining the influence of each parameter on the AP duration (APD). The results showed that $V_{\text{half}-\text{trap}}$ of the SD formulation does not influence the APD. This indicates that $V_{\text{half}-\text{trap}}$ may not be identifiable from the experimental data of current recordings. In this chapter, we investigate the practical identifiability of the parameters of the SD formulation and simplify the SD formulation based on the identifiability results. We show that a simpler drug-binding formulation has improved identifiability while still having a performance of risk prediction similar to the SD formulation.

5.1 Introduction

As discussed in Chapter 2, drug-binding to the hERG channel is an area of active ongoing research, with implications contributing to the comprehensive risk assessment of drugs. Mathematical models incorporating drug-binding that better reflect channel behaviour under drug-induced inhibition may potentially provide accurate proarrhythmic risk predictions.

The state-dependent (SD) drug-binding formulation that we have described in detail in Section 2.5.5 was developed to incorporate drug-trapping into a drug-binding

model, and hence to improve the proarrhythmic risk of drugs [Li et al., 2017b, Dutta et al., 2017, Li et al., 2019]. However, drug-binding models cannot fully recapitulate every detail of the drug-binding dynamics due to the lack of complete biological knowledge. Moreover, the inclusion of every detail into the models, even if that were possible, would increase the difficulty of accurately inferring all parameters of the model, leading to parameter identifiability issues.

The SD drug-binding formulation is a relatively complex model that has been developed to capture specifically the drug-trapping to hERG channel and has been used, together with the Torsades Metric Score (TMS), to make predictions of the proarrhythmic risk of drugs [Li et al., 2017b, Dutta et al., 2017, Li et al., 2019]. Trapping is postulated to be observable in the experimental current by the rate of recovery of the current between pulses [Yang et al., 1995]. The modified Milnes' protocol used by Li et al. [2017b], modified from the original Milnes' protocol [Milnes et al., 2003, 2010], is designed with long resting potentials between pulses. The long resting potential is postulated to allow sufficient time for nontrapped drugs to unbind from the channel, while trapped drugs will stay bound within the channel. At the subsequent pulse, these authors assume that trapped drugs that stay within the channel during the resting potential will maintain the current inhibition, whereas nontrapped drugs will unbind from the channel, and the current magnitude will return to that of the previous pulse. This recovery of the current magnitude to the previous pulse is an indicator of a nontrapped drug. In Li et al. [2017b], the trapping behaviour of drugs is quantified by the rate of recovery, which is calculated as the reduction of the peak current between the pulses.

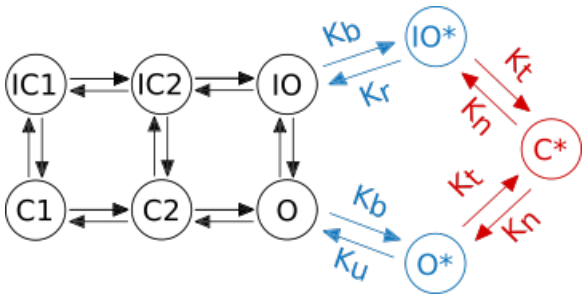


Figure 5.1: **The structure of the SD drug-binding formulation (blue and red) attached to the Li hERG-channel component in black.** The blue and red parts indicate the binding and the trapping components, respectively.

5.1.1 Summary of the SD formulation

For ease of reference, the SD drug-binding formulation is described again briefly. The SD drug-binding formulation characterises the current behaviour with a Markov structure with a binding component and a trapping component (blue and red parts of Figure 5.1, respectively). The SD formulation is given by

$$K_b = K_u \times K_{\max} \times \frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.1}$$

$$K_n = K_t \times \frac{1}{1 + e^{-(V - V_{\mathrm{half-trap}})/6.789}}, \tag{5.2}$$

$$K_t = 3.5 \times 10^{-5}, \tag{5.3}$$

where the binding rate K_b, the unbinding rate K_u, the trapping rate K_t, the untrapping rate K_n and the microscopic reversibility rate K_r describe the transition rates between the drug-bound states (O*, IO* and C*) (see Figure 5.1). The parameters that describe the drug effect can be interpreted as follows: $K_{\max}$ controls the binding rate; K_u defines the unbinding rate; $V_{\mathrm{half-trap}}$ determines the trapping tendency of a drug; n describes the dependency of the current inhibition on drug concentration; and EC_{50} normalises the drug concentration input against the drug concentration that induces 50% of the maximal drug response. Trapped drugs are differentiated from nontrapped drugs by having a slower transition rate out of the trapped state, thus staying in the trapped state longer.

5.1.2 Uncertainty quantification of the SD formulation

Although the SD formulation has eight transitions among its binding and trapping components, the current inhibition behaviour induced by each drug is characterised by only five parameters. Chang et al. [2017] have shown that $K_{\max}$ and EC_{50}^n have large uncertainty, but $K_{\max}/\mathrm{EC}_{50}^n$ is identifiable, therefore the uncertainty is not reflected in the variability in the current behaviour. In the same study, $V_{\mathrm{half-trap}}$ was shown to have a large uncertainty for most drugs, with half of all drugs showing bimodality in the distribution, indicating signs of nonidentifiability. The authors argue that despite the large uncertainty, the trapping tendency parameter still correlates with the expected proarrhythmic risk of drugs, consistent with the trend observed in their previous study [Li et al., 2017b]. It is challenging to elucidate the parameters of the SD formulation because the suggested drug-binding process is a two-step process where the binding of drugs precedes trapping, and it is difficult to isolate the effect of both from the current signal alone.

5.1.3 Research question

The development of the SD formulation was aimed at improving the proarrhythmic risk assessment of drugs. Towards this goal, the authors introduced the qNet and the Torsades Metric Score (TMS) to measure the proarrhythmic risk of drugs, and demonstrated that their predictions outperform other alternatives [Dutta et al., 2017, Li et al., 2019]. In the previous chapters (Chapters 3 & 4), we explored the use of the SD formulation outside its "context-of-use" [Pathmanathan et al., 2017] and tested it against the APD, instead of the proposed TMS. We showed that the prolongation of the AP is not driven by the trapping component of the SD formulation. If $V_{\text{half-trap}}$ does not affect the APD, is it possible to accurately identify it from the experimental data used in calibrating the SD formulation? Further, how will this affect proarrhythmic risk prediction?

5.1.4 Summary

In this chapter, we investigate the practical identifiability of the parameters of the SD formulation on synthetic data, that mimics the calibration data of the SD formulation, for a range of parameter sets that cover the parameter space (Section 5.3). We demonstrate that the trapping parameter $V_{\text{half-trap}}$ has poor identifiability under several noise conditions. We then adjust the trapping rate K_t (Section 5.4.1) and explore several voltage-clamp protocols (Section 5.4.2), finding unexpectedly that this does not improve the identifiability. Trapping seems to be neither detectable nor practically identifiable from I_{Kr} and the AP. Therefore, we propose three alternative drug-binding formulations without the trapping component (Section 5.5). For the first, we combine the binding and trapping components into a single new binding component. For the second, the trapping component is effectively removed by setting the trapping rate K_t to zero. For the third, we keep only the binding kinetics around the open state of the hERG channel. These alternative drug-binding formulations are explored in terms of their parameter identifiability (Section 5.5.2), their fit to *in vitro* experimental data (Section 5.5.3) and their differences against the SD formulation (Section 5.5.4). Finally, all drug-binding formulations are used to categorise the drugs used in Li et al. [2017b] and their performance is compared (Section 5.5.5). All studies performed in this chapter are listed in Table 5.1.

Table 5.1: Summary of the studies performed in this chapter

Studies	Details
Parameter identifiability of the SD formulation (Section 5.3)	• To assess the practical identifiability of the SD formulation • The parameter that controls the trapping tendency of drugs $V_{\text{half}-\text{trap}}$ has poor identifiability
Modifications to the SD formulation (Sections 5.4.1 & 5.4.2)	• To improve the parameter identifiability of the SD formulation • Changed the parameter that controls the trapping tendency and explored several voltage-clamp protocols • Parameter identifiability of the trapping component is not improved
Introduction of alternative drug-binding formulations (Section 5.5)	• To suggest alternative drug-binding formulations without the non-identifiable trapping component of the SD formulation
Assessment of the alternative drug-binding formulations (Sections 5.5.2, 5.5.3 & 5.5.4)	• To assess the performance (parameter identifiability, fit to experimental data and proarrhythmic risk predictions) of the alternative drug-binding formulations against the SD formulation
Performance of all drug-binding formulations (Section 5.5.5)	• To compare the risk prediction of all drug-binding formulations

5.2 Simulation of the SD formulation

We begin by giving a brief recap on the simulation details of the SD formulation, including the choice of drug concentration and the parameter space. As detailed in Section 3.2.7 of Chapter 3, we can reduce the number of parameters of the SD formulation to three parameters. We take $\hat{D} = \frac{D^n}{D^n + \text{EC}_{50}^n}$ as the dimensionless concentration input. The range of $\hat{D}$ used for the simulations within this chapter is based on the drug concentrations used within the paper of the SD formulation [Li et al., 2017b]. The $\hat{D}$ values of the experimental data in Li et al. [2017b] have 5th and 95th quantiles as 1.2×10^{-8} and 0.65 respectively. Li et al. [2017b] performed experiments on four drug concentrations for almost all of the drugs. For each parameter combination in our simulations, we therefore take four values of $\hat{D}$, explored logarithmically within the given quantile range.

The parameter space, consisting of parameters K_u, K_{max}, and $V_{\text{half}-\text{trap}}$, is defined by the minimum and maximum parameter values of all drugs in the CiPA training and validation drug sets [Li et al., 2017b, 2019]. The parameter space is thus a cube of $K_{\text{max}} \in [1, 10^8]$ (unitless), $K_u \in [10^{-5}, 1]$ ms^{-1} and $V_{\text{half}-\text{trap}} \in [-200, -1]$ mV. For each point within the parameter space (a parameter combination), its effect on

I_{Kr} can be simulated with the SD formulation for any given $\hat{D}$ and then compared. These parameter combinations will be referred to as *virtual drugs*, consistent with previous chapters, as they are hypothetical and might not be equivalent to any existing drugs. The parameter space is discretised into five equidistant grids to encompass a wide range of possibilities in the parameter space. K_{max} and K_u are discretised on logarithmic scales.

All simulations were run with Python 3.8 using Myokit 1.33 [Clerx et al., 2016] with the CVODE solver [Hindmarsh et al., 2005]. The absolute tolerance and relative tolerance were set to 10^{-7} and 10^{-8}, respectively. The code and datasets associated with this study are available in a GitHub repository: `https://github.com/FarmHJ/SD-reduction`. A summary of the computational details for all coding implementations of this thesis is provided in Appendix A.

5.3 Identifiability analysis of the SD drug-binding formulation

We now perform a practical identifiability analysis of the SD drug-binding formulation using a modified profile likelihood method on synthetic data. Throughout this chapter, *identifiability*, *practical identifiability* and *parameter identifiability* are used interchangeably. We first describe the generation of the synthetic data, then the modified profile likelihood method.

5.3.1 Generation of synthetic data

We want to generate synthetic data that is as close to the experimental data in Li et al. [2017b] as possible to better replicate actual experimental data. For any given virtual drug (parameter combination), the I_{Kr} of the SD formulation is simulated for the first ten pulses of the modified Milnes' protocol (by Li et al. [2017b]) at a time step of 10 ms. The SD formulation is first paced to steady state under control conditions before the addition of the virtual drug. The drug-inhibited current is then normalised against the control current to obtain the fractional block. The fractional block is defined as

$$B_{t,c} = \frac{I_{inhibited,t,c}}{I_{control,t}}, \tag{5.4}$$

where $B_{t,c}$ is the fractional block at time point t and drug concentration c. The $I_{inhibited,t,c}$ and $I_{control,t}$ currents represent the I_{Kr} current after the addition of drugs

and the I_{Kr} current under control conditions, respectively, at time t and drug concentration c. The fractional block $B_{t,c}$ for a given drug concentration c has a total of 9900 time points in ms, coming from the 9.9 s period of interest of the modified Milnes' protocol (after removing the channel opening phase at the first 0.1 s of the 0 mV step of the modified Milnes' protocol [Li et al., 2017b, Lei et al., 2024]) at a 10 ms time step for 10 pulses.

The normalisation of the drug-inhibited current by the control current ensures the data is relatively consistent throughout the repeats by removing artifacts due to cell size or cell-cell variability. When current magnitudes are close to zero (the channel is closed), the fractional block will be dominated by noise or will achieve large values and the simulations become computationally infeasible. Therefore, the fractional block only at the 0 mV step of the modified Milnes' protocol (the shaded region of Milnes' protocol in Figure 5.7) is used.

There are three noise models used in this study, which we term N1, N2 and N3. The first (N1) is the ideal case where no noise is added, so the fractional block is

$$B_{t,c} = \frac{I_{inhibited,t,c}}{I_{control,t}}. \tag{5.5}$$

The second noise model (N2) adds independent and identically distributed (i.i.d.) Gaussian noise to the simulated I_{Kr} of the control condition and the drug condition individually. The Gaussian noise is assumed to have zero mean and the standard deviation, σ, is chosen to be 6.6×10^{-3} µA, based on experimental data. The referenced experimental data are I_{Kr} under the ramp protocol and the step protocol for three drugs, i.e. dofetilide, cisapride and verapamil, provided by our collaborators at Roche. These protocols are also used in one of the studies in this chapter (Section 5.4.2). The estimated noise level is taken as the standard deviation of recorded currents when the ion channel is in the resting state, i.e. the first 50 ms of the ramp protocol and the first 100 ms of the step protocol. The Gaussian noise is added to the simulated current of the control condition and the drug condition separately, then normalised to obtain the synthetic fractional block. The synthetic fractional block with the N2 noise model is

$$B_{t,c} = \frac{I_{inhibited,t,c} + \epsilon_{t,c}}{I_{control,t} + \xi_{t,c}}, \tag{5.6}$$

where $\epsilon_{t,c} \sim \mathcal{N}(0, \sigma^2)$ and $\xi_{t,c} \sim \mathcal{N}(0, \sigma^2)$.

To mimic the experimentally observed fluctuation in I_{Kr} values, we also consider a third noise model (N3). The N2 noise model did not seem to replicate the intensity of

noise commonly observed in experimental data, where more instabilities and variabilities were observed. Due to the size of cells and cell-cell variability, raw currents can fluctuate by as much as twice this magnitude. Even though these variabilities in current magnitude were accounted for in the fractional block data, where the inhibited current is normalised against the control current, the noise level cannot be completely cancelled out. In the N3 noise model, we scale the magnitude of I_{Kr}, both the control current and the drug-inhibited current, by half and then add Gaussian noise (with the same standard deviation as the N2 noise model), creating a higher noise intensity than the N2 noise model. The noisy drug-inhibited current is then normalised against the noisy control current to obtain the fractional block. The fractional block with the N3 noise model is

$$B_{t,c} = \frac{\frac{I_{inhibited,t,c}}{2} + \epsilon_{t,c}}{\frac{I_{control,t}}{2} + \xi_{t,c}}. \tag{5.7}$$

This is equivalent to doubling the noise intensity.

For reproducibility, generation of $\epsilon_{t,c}$ and $\xi_{t,c}$ were seeded. The realisations of $\xi_{t,c}$ are the same for all noise models and all parameter combinations, while the realisations of $\epsilon_{t,c}$ are the same for all noise models but differ between the parameter combinations.

5.3.2 Error profile

Here, we perform a modified profile likelihood analysis on synthetic data to investigate practical identifiability of the parameters of the SD formulation. Instead of using log-likelihood as an assessment of the optimisation, we use the root mean squared difference (RMSD) between the synthetic data and the simulated fit. It is easier to interpret RMSD because it has the same unit as I_{Kr}. We refer to this modified profile likelihood as the error profile. Parameter identifiability in the error profile is determined by a minimum RMSD, as opposed to the maximum of the log-likelihood for the profile likelihood.

Three sets of synthetic data with known parameter combinations, corresponding to dofetilide, verapamil and cisapride, are used for this analysis. Dofetilide is chosen to represent a trapped drug as it has the highest $V_{\text{half}-\text{trap}}$ (the parameter associated with drug-trapping) value; whereas cisapride and verapamil are chosen to represent nontrapped drugs. Cisapride has the lowest $V_{\text{half}-\text{trap}}$ value, but due to its low K_{max} (the parameter controlling the binding rate) value, cisapride cannot fully inhibit I_{Kr} at high concentrations. Therefore verapamil, with the second lowest $V_{\text{half}-\text{trap}}$, is chosen

as an alternative example drug. All drugs, i.e. dofetilide, cisapride and verapamil, mentioned in this chapter refer to the parameter combinations of those drugs, not the real drug.

The parameter ranges explored for the error profile are $K_{\max} \in [1, 10^{10}]$ (unitless), $K_u \in [1.8 \times 10^{-6}, 1]$ ms^{-1} and $V_{\text{half-trap}} \in [-240, -0.25]$ mV (see Equations 5.1 to 5.3 and Section 5.2). The parameter ranges defined for the error profile were chosen to encompass the parameter ranges observed from the CiPA training and validation drugs [Li et al., 2017b, 2019], which are listed in Section 5.2. These parameter ranges defined for the error profile extend from the experimentally observed ranges because our chosen example drugs have parameter values close to the border of the experimentally observed ranges. Dofetilide has $K_{\max}$ and $V_{\text{half-trap}}$ values close to the maximum among all drugs in Li et al. [2017b, 2019], while cisapride has the lowest $V_{\text{half-trap}}$ value. The extension of the parameter ranges will provide a better picture of the error profile around the neighbourhood of the parameter values of the chosen example drugs. $V_{\text{half-trap}}$ is restricted to negative values to retain the property of the untrapping rate K_n. The highest $V_{\text{half-trap}}$ of all drugs (indicating a highly trapped drug) tested in Li et al. [2017b] and Li et al. [2019] is -1 mV, giving an untrapping rate K_n of roughly half the trapping rate K_t (see Equation 5.2). Whilst mathematically $V_{\text{half-trap}}$ could take positive values, no actual drugs have a positive value (loratadine has the largest $V_{\text{half-trap}}$ at -1 mV among all 28 drugs in Li et al. [2017b, 2019]), and therefore we restrict the range we consider to negative values.

5.3.3 Minimisation of RMSD

The parameter values are optimised by minimising the root mean squared difference (RMSD) between the synthetic data and the model simulation. The objective function is the average RMSD across all drug concentrations tested:

$$\frac{1}{4}\sum_{c}\sqrt{\frac{\sum_t \left(Y_{t,c} - B_{t,c}\right)^2}{T}} \tag{5.8}$$

where $Y_{t,c}$ is the simulated fractional block and $B_{t,c}$ is the synthetic fractional block (refer to Equations 5.5 to 5.7). The RMSD between the fractional block is computed for all time points $T = 9900$ (Section 5.3.1). The RMSD values are then averaged over the four drug concentrations c tested.

Given that the noise is compounded in the N2 and the N3 noise models (Equations 5.6 & 5.7) (from the noise added to the control current and the inhibited current) the resultant noise on the fractional block $\frac{I_{inhibited,t,c}}{I_{control,t}}$ cannot be assumed to follow a

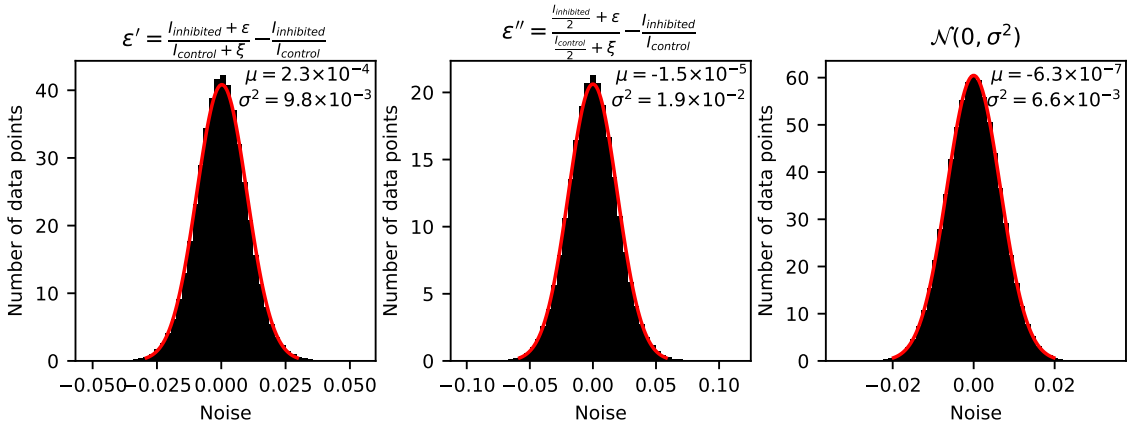


Figure 5.2: **The resultant noise of synthetic data are similar to normal distributions.** The left and the middle panels show the probability density (in black) of $\epsilon'_{t,c}$ and $\epsilon''_{t,c}$, respectively. They are the difference between the noisy synthetic data (N2 and N3 noise models) and the noise-free data of $\frac{I_{inhibited}}{I_{control}}$ (based on Equations 5.9 & 5.10). The right panel shows the distribution of random numbers sampled from a normal distribution for reference. The red solid lines show the probability density function of normal distributions with mean and standard deviation (text in each panel) of the histograms.

normal distribution [Hayya et al., 1975]. We examine the distribution of the resultant noise following the equations:

$$\text{N2 noise model, } \epsilon'_{t,c} = \frac{I_{inhibited,t,c} + \epsilon_{t,c}}{I_{control,t} + \xi_{t,c}} - \frac{I_{inhibited,t,c}}{I_{control,t}}, \tag{5.9}$$

$$\text{N3 noise model, } \epsilon''_{t,c} = \frac{\frac{I_{inhibited,t,c}}{2} + \epsilon_{t,c}}{\frac{I_{control,t}}{2} + \xi_{t,c}} - \frac{I_{inhibited,t,c}}{I_{control,t}}, \tag{5.10}$$

where $\epsilon'_{t,c}$ and $\epsilon''_{t,c}$ are the approximated resultant noise for the N2 and the N3 noise model, respectively. Figure 5.2 shows the distribution of the approximated noise $\epsilon'_{t,c}$ and $\epsilon''_{t,c}$ in the left and the middle panels, respectively. The distributions are plotted by taking all instances of the noisy synthetic data (left-hand sides of Equations 5.9 and 5.10), following the procedure described in Section 5.3.1, and subtracting the noise-free fractional block $\frac{I_{inhibited,t,c}}{I_{control,t}}$ from it, to get instances of $\epsilon'_{t,c}$ and $\epsilon''_{t,c}$. The histogram in the right panel shows samples generated from the normal distribution $\mathcal{N}(0, 6.6 \times 10^{-3}\mu\text{A})$ (the distribution of noise added to the simulated currents (Section 5.3.1)) as a reference. The distributions of $\epsilon'_{t,c}$ and $\epsilon''_{t,c}$ have means of 2.3×10^{-4} µA and -1.5×10^{-5} µA, and standard deviations of 9.8×10^{-3} µA and 1.9×10^{-2} µA, respectively. This further verifies that the intensity level of noise imposed by the N3 noise model on the synthetic data is higher than that of the N2

noise model, with almost twice the standard deviation. The red solid lines overlaying the histograms are the probability density function of the normal distributions with means and standard deviations of the histograms. Even though the distribution of the approximated noise $\epsilon'_{t,c}$ and $\epsilon''_{t,c}$ are not strictly normal, they are very close to a normal distribution, as observed in Figure 5.2. This gives us confidence that our use of least squares optimisation will give accurate recovery of the optimal parameter values while mimicking the approach taken by Li et al. [2017b] in parameterising their original model.

When performing inference, all parameters undergo logarithmic transformation, except for $V_{\text{half-trap}}$, as they span several orders of magnitude. The starting point of each optimisation repeat is randomly sampled from a log-uniform distribution spanning the parameter space (defined in Section 5.2) for $K_{\max}$ and K_u, and a uniform distribution for $V_{\text{half-trap}}$. Bounds are imposed on each parameter as follows, $K_{\max} \in [10^{-1}, 10^{10}]$ (unitless), $K_u \in [10^{-7}, 10^{2}]$ ms^{-1} and $V_{\text{half-trap}} \in [-300, 0)$ mV. The optimisation is performed using the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) from the PINTS 0.5.0 [Clerx et al., 2019] on Python 3.8.

For each parameter combination, five repeats of optimisations are performed with different initial guesses to ensure the parameter space is well explored. The best-performing repeat is presented. To ensure that the inferred parameter values were within a sensible range, constraints were imposed on the range of the parameters. These constraints were taken to be two orders of magnitude from the boundary of the parameter space, except for $V_{\text{half-trap}}$ because it was explored linearly, not logarithmically, where the limit is $[-300, 0)$ mV. The optimal objective function is achieved at least three out of five times, regardless of the noise model (see Section 5.3.1).

5.3.4 Practical identifiability of the SD formulation for three example drugs

Figure 5.3 shows the error profile of the SD formulation for parameters $K_{\max}$, K_u and $V_{\text{half-trap}}$ (see Figure 5.1 and Section 5.2) for the three example drugs (dofetilide, cisapride and verapamil) under the three noise models (Section 5.3.1). The error profile shows the minimum RMSD achievable when one of the parameters is set to a given value and then the remaining parameters are optimised given the value of the fixed parameter (Section 5.3.2) [Cole et al., 2013]. The error profile provides insights into the identifiability of the parameter [Raue et al., 2009].

For the binding-rate parameter $K_{\max}$, an optimum is observed at the true parameter value (first column), shown as vertical dashed lines in Figure 5.3, for all example

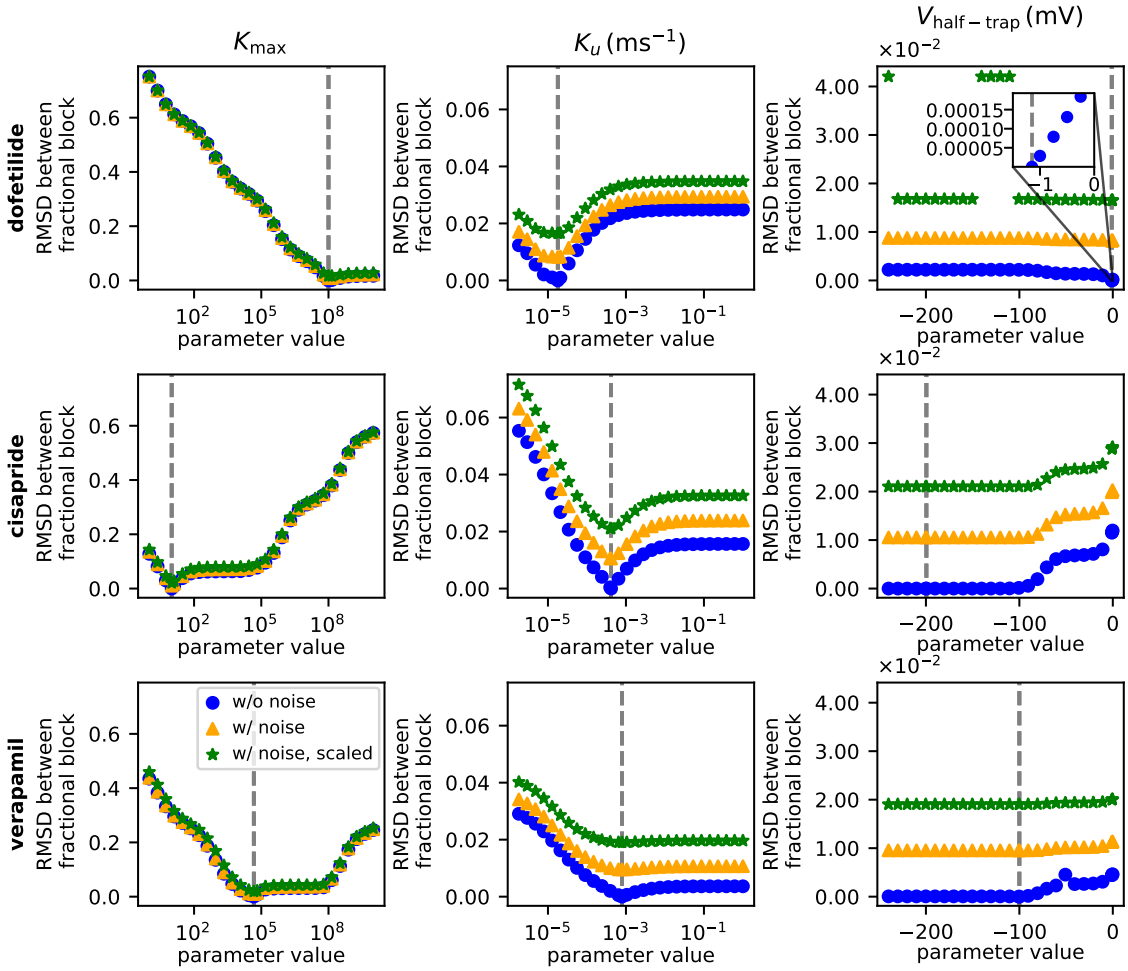


Figure 5.3: **$\mathbf{K_{max}}$ is identifiable in all noise models, the confidence interval for $\mathbf{K_u}$ becomes larger when the noise intensity increases and $\mathbf{V_{half\text{-}trap}}$ is unidentifiable when its value is lower than -100 mV.** The error profile of K_{max}, K_u, and $V_{\text{half}-\text{trap}}$ for example drugs dofetilide, cisapride, and verapamil under three noise models. The noise models are indicated by blue circles for the condition without noise (N1), orange triangles when noise is added (N2), and green stars when the currents are scaled by half and the noise is added (N3). The vertical dashed lines are the true parameter values. The inset in the top right panel for $V_{\text{half}-\text{trap}}$ of dofetilide is a zoomed-in view for $V_{\text{half}-\text{trap}}$ values greater than -1 mV.

drugs. As the noise intensity increases (from the N1 to the N3 noise model, from blue circles to orange triangles to green stars), the error profile of K_{max} remains largely similar, with a slightly flattened out error profile. The estimation of the unbinding rate K_u is more susceptible to the effects of noise. K_u is identifiable under noise-free conditions for all example drugs (second column of Figure 5.3). The local optima become less pronounced as the noise intensity increases. The unbinding rate K_u of verapamil (last row, middle column) is affected significantly by noise. The error profile flattens out at large values when noise is added. The trapping tendency parameter $V_{\text{half-trap}}$ has the worst identifiability of all the parameters (rightmost column). The error profile is flat at $V_{\text{half-trap}}$ smaller than $-100\,\text{mV}$ for all example drugs, indicating that $V_{\text{half-trap}}$ is not identifiable in that range. The only local optimum observed for $V_{\text{half-trap}}$ is for dofetilide when no noise is added, as emphasised in the inset in the top right panel. For cisapride (middle row, right column), the RMSD does not change with $V_{\text{half-trap}}$ when it is at any value below $-100\,\text{mV}$. This implies that $V_{\text{half-trap}}$ values smaller than $-100\,\text{mV}$ cannot be inferred accurately. The error profile of $V_{\text{half-trap}}$ for verapamil (bottom row, right column) becomes flatter even for values greater than $-100\,\text{mV}$ when noise is added. The unidentifiability of $V_{\text{half-trap}}$ is therefore not limited to values smaller than $-100\,\text{mV}$.

The binding rate K_{max} is reasonably identifiable for all example drugs under all noise models; whereas the confidence interval for the unbinding rate K_u gets larger as the noise intensity increases. $V_{\text{half-trap}}$ has poor identifiability in the range smaller than $-100\,\text{mV}$ under noise-free conditions. This range of unidentifiability is extended when noise is added, in particular for verapamil. The magnitude of the RMSD is relatively small for $V_{\text{half-trap}}$ as compared to the other parameters, further verifying its poor identifiability. This suggests that it would be difficult to infer $V_{\text{half-trap}}$ accurately, especially with noisy data.

The conclusions from this error profile analysis are limited as it is restricted to the chosen parameter combination and to a single dimension. For a more holistic view of the parameter identifiability of the SD formulation, we extend the limited number of parameter combinations explored in the error profile to a large number of parameter combinations that covers the parameter space (Section 5.2). Synthetic data with the three noise models are generated for the sampled parameter combinations, then the parameters are inferred from the noisy synthetic data, as done for the error profile (Section 5.3.3). The accuracy of the inferred parameters compared to the actual parameter values is then assessed to get insights into the practical identifiability of the parameters. If all parameter combinations can be recovered from the noisy

synthetic data, then the data contains features that are unique to the parameters and cannot be replicated by compensation between parameters, thus implying that the parameters are identifiable.

5.3.5 Practical identifiability of the SD formulation for the parameter space

The top panels of Figure 5.4 show example realisations of the fit of the SD formulation to synthetic data of the three noise models. The chosen example realisations are the parameter combinations that give the largest RMSD (worst performing) in each noise model. The bottom panels show the RMSDs between the noisy synthetic data and their fitted simulations for all parameter combinations performed. Each pixel in the 3D plots is colour-coded with the RMSD between the fitted simulation and the synthetic data generated by the parameter combination indicated by the axes ($K_{\max}$, K_u and $V_{\text{half}-\text{trap}}$). For each parameter combination, noisy synthetic data is generated according to the procedure described in Section 5.3.1, then the SD formulation is optimised repeatedly five times with different initial values in a search for the best fit to the synthetic data. A detailed description of the method is provided in Section 5.3.3.

The top left panel of Figure 5.4 shows that the fit of the SD formulation to synthetic data is almost perfect when no noise is added, with the RMSD being lower than 2.5×10^{-7}. As expected, we get larger RMSDs when noise is added. The RMSDs range around 10^{-2} for the N2 noise model and double this for the N3 noise model, which are of the same magnitude as those obtained when fitting to actual experimental data (see Section 5.5.3). Qualitatively, the SD formulation can still replicate the fractional block when noise is added (top panels). The 3D plots show a correlation between the optimisation performance and the binding-rate parameter $K_{\max}$. The optimisation performance worsens as $K_{\max}$ increases, but the difference is small.

Figure 5.4 shows that the SD formulation can capture the fractional block even with noisy data. We next explore if the parameters can be inferred accurately in Figure 5.5. Figure 5.5 compares the true parameter value (x-axis) used to generate the synthetic data and the inferred parameter values (y-axis). The diagonal lines serve as references for parameter values that are inferred accurately. The red and green shaded regions in the $V_{\text{half}-\text{trap}}$ column refer to the association between the $V_{\text{half}-\text{trap}}$ values and the risk categorisation of drugs observed by Li et al. [2017b]. It was observed that low-risk drugs have $V_{\text{half}-\text{trap}} < -85$ mV (green region), while high-risk drugs have $V_{\text{half}-\text{trap}} > -65$ mV (red region) [Li et al., 2017b]. Note that

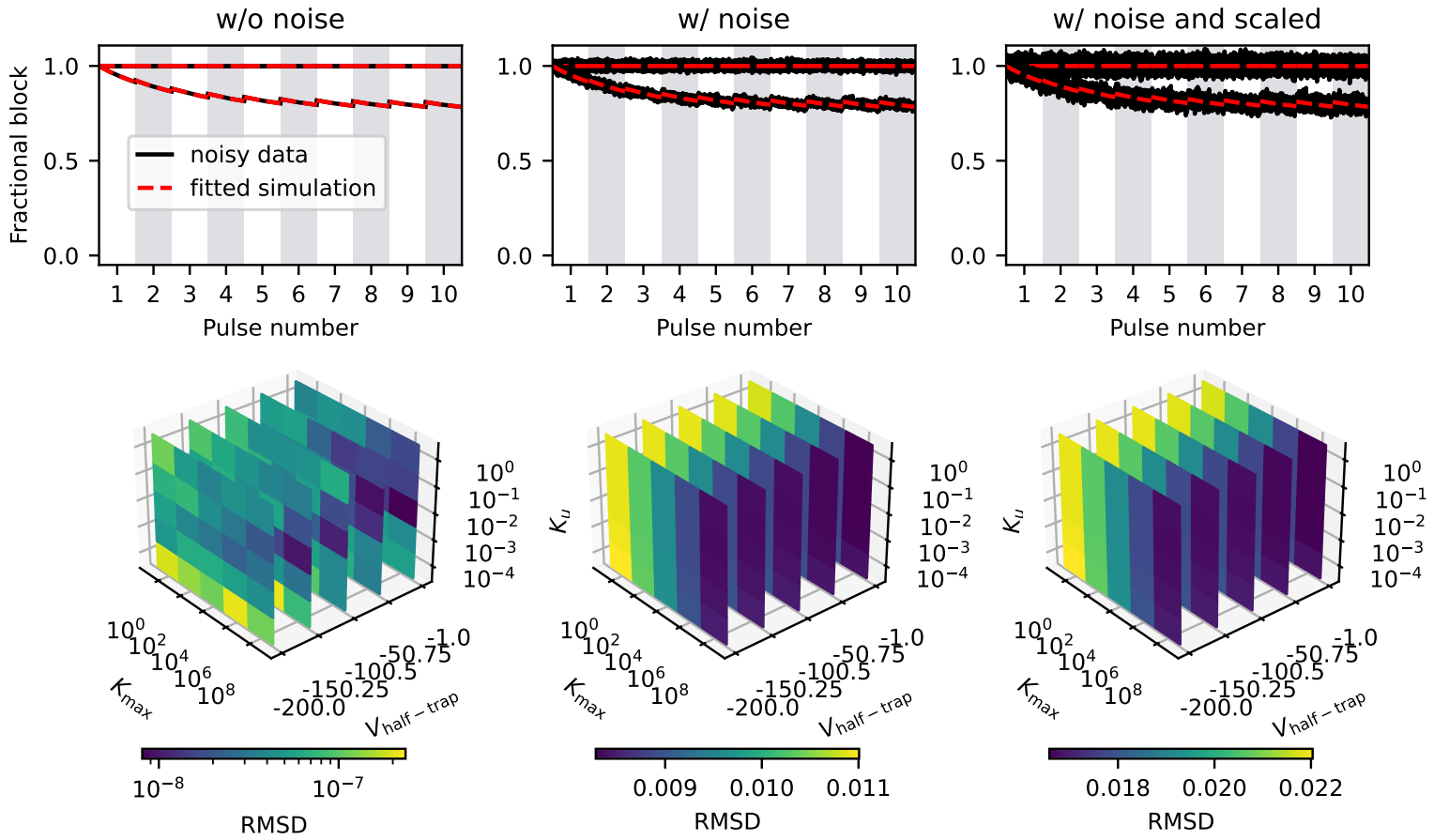


Figure 5.4: **The SD drug-binding formulation fits reasonably well to its own synthetic data.** Top panels show the synthetic data (black solid) and the fitted simulations (red dashed) for a parameter combination out of all tested that gives the largest RMSD. The 3D plot shows the RMSD between synthetic data and the fitted simulations for all tested parameter combinations within the parameter space of $K_{\max}$, K_u and $V_{\text{half}-\text{trap}}$.

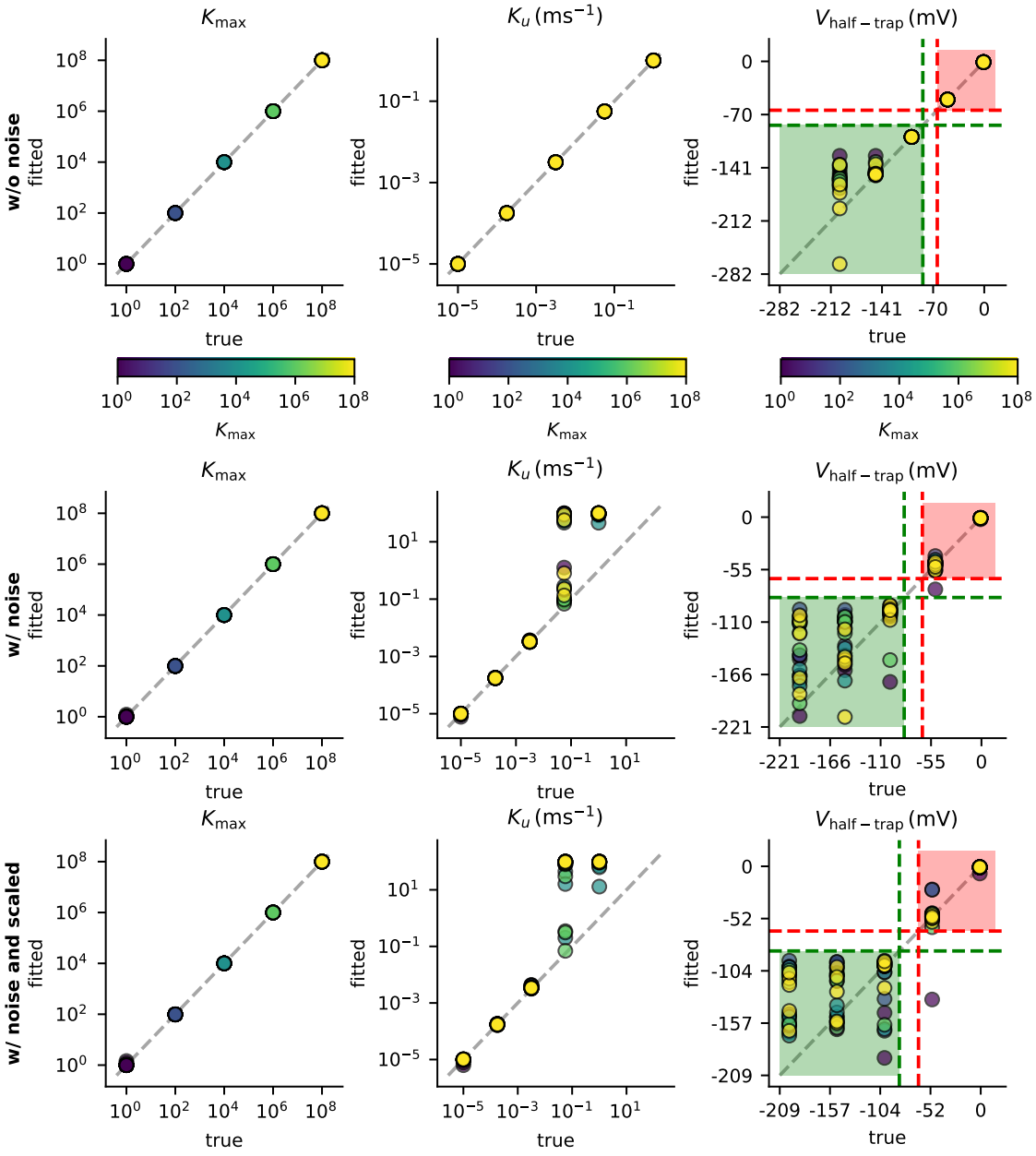


Figure 5.5: **When noise is added, the unbinding rate K_u cannot be inferred accurately at high values, while $V_{half\text{-}trap}$ largely stays within the associated risk category observed by Li et al. [2017b].** Each panel compares the inferred parameter value (y-axis) with the true parameter value (x-axis). The grey dashed line is the $y = x$ line, where the true and inferred parameters are equal. The colour of the data points indicate the value of K_{max}. The red and green shaded regions for $V_{half-trap}$ indicate the high- and low-risk categories, respectively, that are associated with the $V_{half-trap}$ values observed in Li et al. [2017b].

these threshold values of $V_{\text{half-trap}}$ are not metrics used to categorise drugs, simply observations.

Without noise (top row), $K_{\max}$ and K_u (related to binding and unbinding rates, respectively) are inferred accurately as all points lie on the diagonal lines, while $V_{\text{half-trap}}$ cannot be inferred accurately at low values. This is consistent with the error profile in Figure 5.3, where the RMSD is flat when $V_{\text{half-trap}} < -100$ mV. Although $V_{\text{half-trap}}$ is not fully identifiable, the inferred $V_{\text{half-trap}}$ values still lie within the same risk region as that identified by Li et al. [2017b]. We refer to this, where the parameter is not exactly identifiable but is identifiable within a defined threshold range, as *weakly identifiable*.

When noise is added, the unbinding rate K_u loses its identifiability at large values. In the N3 noise model (bottom row), true K_u values larger than $10^{-2}\,\text{ms}^{-1}$ are fitted mostly close to the boundary of K_u at $10^{2}\,\text{ms}^{-1}$. This is also consistent with the error profile in Figure 5.3, where K_u becomes unidentifiable at large values, as indicated by constant RMSDs. The identifiability of $V_{\text{half-trap}}$ gets worse with noise, though the inferred values still remain largely within the same risk region as the true value. The data points in Figure 5.5 were coloured based on their $K_{\max}$ values to investigate the correlation between the RMSD and the $K_{\max}$ values observed in Figure 5.4, where the RMSD increases with $K_{\max}$. However, Figure 5.5 does not show any observable correlation between parameter identifiability and $K_{\max}$.

This identifiability analysis shows that $V_{\text{half-trap}}$ is weakly identifiable (the inferred values remain within the same risk region as their true values, as mentioned previously), even in noise-free conditions. In general, the higher the level of noise, the less likely the parameters will be inferred accurately. The binding-rate parameter $K_{\max}$ can be inferred accurately regardless of the level of noise; the fits of K_u are affected by the level of noise when its values are large; and $V_{\text{half-trap}}$ is strongly affected by the level of noise. Given that $V_{\text{half-trap}}$ cannot be inferred accurately (Figure 5.5), but no difference in the RMSD is shown with regards to $V_{\text{half-trap}}$ (Figure 5.4), implies that $V_{\text{half-trap}}$ has a limited effect on the fractional block. This is consistent with our previous results where $V_{\text{half-trap}}$ does not drive the difference in APD_{90} between the SD formulation and the conductance scaling (CS) formulation (Chapters 3 & 4). All results point to the fact that the synthetic data generated by the SD drug-binding formulation contains insufficient information to allow for the inference of the value of $V_{\text{half-trap}}$. The effect of $V_{\text{half-trap}}$ is not observable in the AP, nor in the I_{Kr} current.

5.4 Adjustments to the SD formulation

5.4.1 Investigation of the range of parameter values in the SD drug-binding formulation – trapping rate K_t

This limited effect of $V_{\text{half-trap}}$ on the AP and I_{Kr} could result from the definition of the SD drug-binding formulation. $V_{\text{half-trap}}$ is defined implicitly through the equation of the untrapping rate K_n (Equation 5.2),

$$K_n = K_t \times \frac{1}{1 + e^{-(V - V_{\text{half-trap}})/6.789}}. \tag{5.11}$$

Moreover, noting that $K_t = 3.5 \times 10^{-5}\,\text{ms}^{-1}$ is fixed as a constant value as defined in Li et al. [2017b] and $V = 0$ from the use of only the 0 mV step of the modified Milnes' protocol as the calibration data (Section 5.3.1), we have

$$K_n = 3.5 \times 10^{-5} \times \frac{1}{1 + e^{V_{\text{half-trap}}/6.789}}. \tag{5.12}$$

From here, we can make use of the parameter values inferred from experimental data for all CiPA drugs published in Li et al. [2017b, 2019], where $V_{\text{half-trap}} \in (-200, -1]$ mV. Using this range for $V_{\text{half-trap}}$ in Equation 5.12 allows us to calculate the range of the untrapping rate K_n as being $(1.8 \times 10^{-5}, 3.5 \times 10^{-5})$ ms^{-1}. Again, referring to the parameter values inferred from experimental data by Li et al. [2017b, 2019], we have the binding rate coefficient $K_b = K_{\text{max}} \times K_u$ and unbinding rate K_u being within ranges $(8.0 \times 10^{-5}, 2.8 \times 10^{4})$ ms^{-1} and $(1.7 \times 10^{-5}, 1)$ ms^{-1}, respectively. Thus, we have that the ranges of binding and unbinding rates (K_b and K_u) are much higher than the ranges of the trapping and untrapping rates (K_t and K_n), by up to eight orders of magnitude. Referring to the structure of the SD formulation (Figure 5.1, Section 5.2), the trapping of drugs within the closed bound (C*) state is preceded by the binding of the drug compounds into the channel either to the open bound (O*) state or the inactivated bound (IO*) state. Therefore, we can hypothesise that drug compounds will bind and unbind from the channel quickly due to the fast binding and unbinding rates but seldom get trapped within the C* state because of the relatively slow trapping rate. The expected effect of trapping on the current might thus not be detectable from the fractional block or differentiable between binding and trapping.

From Equation 5.12, we can see that the expression $\frac{1}{1+e^{V_{\text{half-trap}}/6.789}}$ has an upper bound of 1, which equates to the untrapping rate K_n having an upper limit of K_t at $3.5 \times 10^{-5}\,\text{ms}^{-1}$ (defined in Li et al. [2017b]) and thus limits its effect on the current. The trapping rate K_t was chosen by Li et al. [2017b] to best separate

trapped and nontrapped drugs and best fit the SD drug-binding formulation to the experimental data. To increase the frequency of drug-trapping within the SD formulation, we increased the trapping rate K_t to values comparable to the binding rate, from i.e. $10^{-5}\,\mathrm{ms}^{-1}$ (the same order of magnitude of the actual trapping rate of $3.5 \times 10^{-5}\,\mathrm{ms}^{-1}$) to $10^{3}\,\mathrm{ms}^{-1}$, which is close to the maximum binding rate K_b at around $10^{4}\,\mathrm{ms}^{-1}$. We then observed the effect of K_t on the fractional block and the identifiability of $V_{\mathrm{half-trap}}$. This is performed on the three example drugs of dofetilide, cisapride and verapamil.

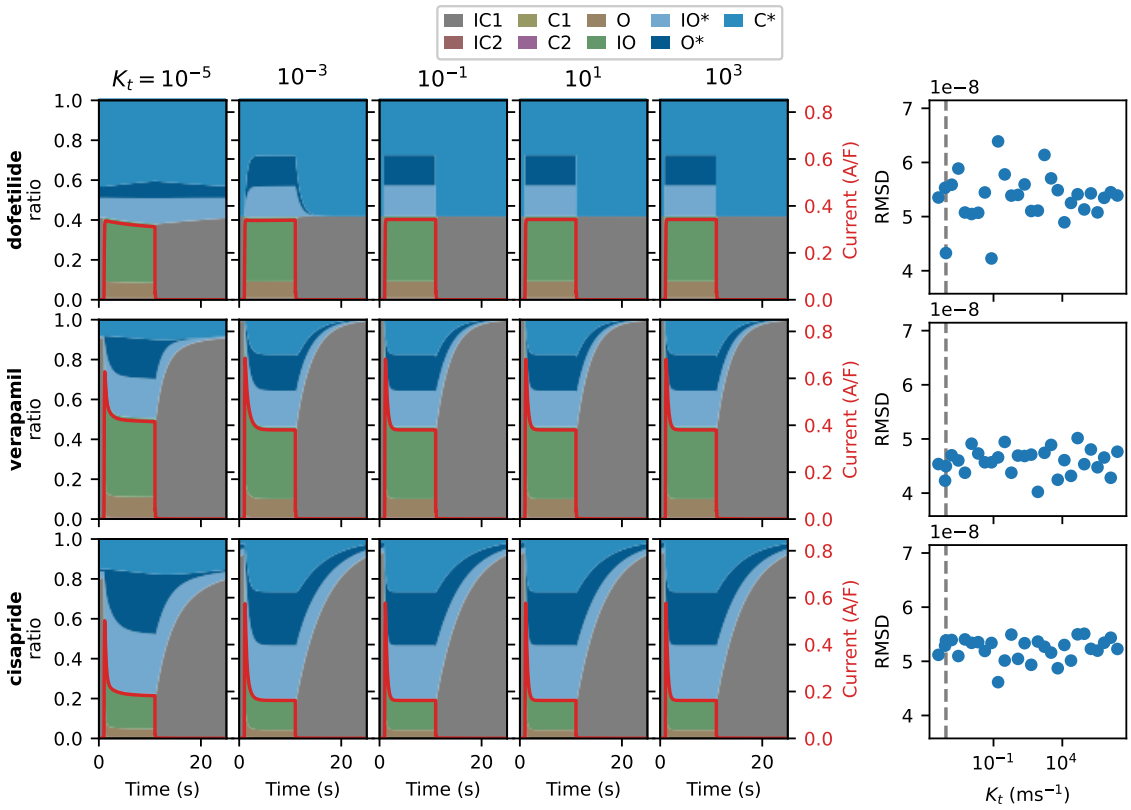


Figure 5.6: **Increasing the trapping rate $\mathbf{K_t}$ affects the steady state $\mathbf{I_{Kr}}$ and state occupancy minimally, with almost no effect when $\mathbf{K_t \geq 10^{-1}}$ ms^{-1}. $\mathbf{K_t}$ therefore is not identifiable.** The stack plots show the relative contribution of each state of the SD formulation (as illustrated in Figure 5.1) and the I_{Kr} current (red solid line) at steady state with increasing trapping rate K_t. The rightmost panels show the error profile (RMSD) of each example drug with respect to the trapping rate K_t. The vertical dashed lines indicate the actual K_t at $3.5 \times 10^{-5}\,\mathrm{ms}^{-1}$. Dofetilide, verapamil and cisapride are chosen as reference points.

The stack plots in Figure 5.6 show the proportion of time of the channel being in each state throughout the modified Milnes' protocol at steady state (1000th pulse

of the protocol). The colours represent different states within the SD formulation (as illustrated in Figure 5.1 and described in Section 5.1.1), and the blues represent the drug-bound states. The states with asterisks (O*, IO* and C*) are the drug-bound states of the SD formulation, whereas the states without asterisks are part of the base hERG-channel model (as illustrated in black in Figure 5.1). The red lines overlaying the stack plot show the steady-state current, which would be measurable in an experimental setting. The K_t values of each example drug increase across the columns from left to right. The stack plots can provide insights into how the change in K_t affects the occupancy of each state during the modified Milnes' protocol.

In the stack plots representing dofetilide (first row), more than half of the channels are in the drug-bound states (blues) throughout the protocol. At the start of the protocol, excluding the drug-bound states, the remaining channels enter the open (O) and the inactivated (IO) state, allowing the flow of current from the open state (indicated by the red lines). At the resting potential (from 11 s onwards), the channels either remain in the drug-bound states or enter the non-current conducting states (IC1, IC2, C1 and C2), stopping all ionic flow across the channels and thus having zero current. The trapping behaviour of dofetilide (a trapped drug) can be observed from the stack plots. The occupancies of the drug-bound states (blues), especially the C* state, are large from the start of the protocol, indicating that the drug compounds remain within the channels from the previous pulse. The drug compounds also remain within the channels until the end of the pulse when the channels are closed. In contrast, the nontrapped drugs verapamil and cisapride enter drug-bound states after the channels are open, inhibiting the current after the channels are open, as can be observed in the sharp decrease in the current after reaching its peak. The drug compounds then unbind from the channels and the channels exit the drug-bound states at the end of the protocol, showing the properties of nontrapped drugs where they cannot stay bound within the channels when the channels are closed.

As the trapping rate K_t increases (from left columns to right), more dofetilide stays within the trapped state C* at the end of the protocol, while verapamil and cisapride, both nontrapped drugs, exit the trapped state C*. As the K_t of the example drugs increases from 10^{-5} to $10^{-3}\,\mathrm{ms}^{-1}$, the I_{Kr} of dofetilide has a slightly decreased peak and a flattened current, while the I_{Kr} of the nontrapped drugs have increased peaks. However, as the trapping rate K_t increases further, no further changes are observable.

To further verify our observations, we performed an error profile analysis of the trapping rate K_t on the example drugs (rightmost panels of Figure 5.6) under ideal

conditions without noise. The error profile shows that the trapping rate K_t is not identifiable. This further demonstrates that the effect of the trapping rate K_t is not reflected in I_{Kr}. The fractional block does not contain sufficient information to elucidate K_t. Li et al. [2017b] may have chosen to fix the trapping rate to address this problem.

The untrapping rate K_n is modified by the value of $V_{\text{half-trap}}$, while also being dependent on the trapping rate K_t and the membrane potential V (Equation 5.3). We have shown in Figure 5.6 that I_{Kr} is unaffected by the trapping rate K_t. The next step is to look into the membrane potential V. The parameters of the SD formulation are inferred from the fractional block recorded at the 0 mV step of the modified Milnes' protocol. The voltage-dependency of the trapping component (Equation 5.3) is not well-explored. We take the equation of the untrapping rate K_n (Equation 5.3) as it is defined in the SD formulation and perturb the membrane potential V by using voltage-clamp protocols with varying voltage steps. This is performed with the intention of investigating whether it is possible to improve the identifiability of $V_{\text{half-trap}}$ by designing an improved protocol.

5.4.2 Investigation of the equations of the SD drug-binding formulation – membrane potential V

We repeat a similar parameter identifiability analysis to the one performed in Section 5.3, using three other protocols — the ramp protocol, the step protocol and the staircase protocol (Figure 5.7).

The first protocol is the modified Milnes' protocol, taken from Li et al. [2017b] and modified from Milnes et al. [2010]. The modified Milnes' protocol is the protocol used to develop the SD drug-binding formulation [Li et al., 2017b] and used in the previous two chapters (Chapters 3 & 4). Please refer to Chapter 3.2.6 for more details.

The other protocols used are the ramp, the step, and the staircase protocol. The ramp protocol and the staircase protocol have varying voltage steps; this allows a better exploration of the membrane potential V. The ramp and the step protocols are used by our collaborators at Roche and recommended by the CiPA initiative to study the binding of drugs to the ion channel, and thus could provide us with experimental data if needed. Due to the use of a fractional block when inferring the parameter values [Li et al., 2017b] (Section 5.3.1), control currents close to zero are not used, as this could cause the data to be dominated by noise or achieve extremely large values, and become computationally infeasible. The time period of interest for the modified Milnes' protocol is $[1100, 11000]$ ms, for the ramp protocol it is $[60, 656]$

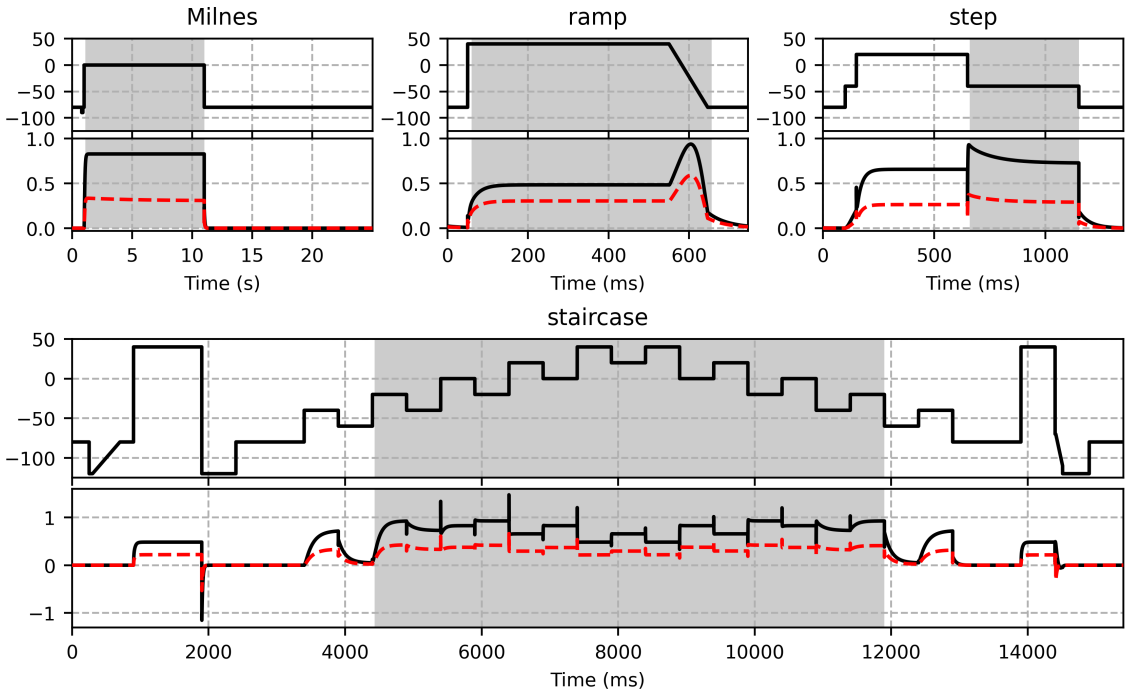


Figure 5.7: **Protocols and the simulated $\mathbf{I_{Kr}}$.** The modified Milnes', the ramp, the step and the staircase protocols and their elicited I_{Kr} (black solid lines in the bottom panels). The red dashed lines show an example of the inhibited current. The example chosen is dofetilide, with parameter values taken from Li et al. [2017b], at a drug concentration of 10 nM. The shaded region indicates the time period of interest for the inference performed on synthetic data to ensure that the fractional block stays within the computationally feasible values.

ms, for the step protocol it is [660, 1150] ms (for the tail current) and for the staircase protocol it is [3410, 12900] ms. The periods of interest are indicated as the grey region in Figure 5.7. The chosen time period for the step protocol is the tail current (the repolarising step at $-40\,\mathrm{mV}$) as it is commonly used to study current behaviour (deactivation of the channel after activation) in the literature [Kodirov, 2020], and to avoid spikes or dips in the current. The noisy synthetic data for this analysis are generated in a similar way as with the modified Milnes' protocol, as detailed in Section 5.3.1.

Similar to the approach of using 10 pulses of the modified Milnes' protocol of Li et al. [2017b], we take the first 10 pulses of the ramp and the step protocols when generating the synthetic data. The trapping tendency of drugs is commonly identified from the recovery of the current to previous pulses [Stork et al., 2007, Windisch et al., 2011], thus several pulses of the protocols (the opening of the channels) are used to differentiate trapped and nontrapped drugs. Experimental data in Figure 5.12 shows that the peak current for a trapped drug dofetilide decreases with each successive pulse, while the peak current of a nontrapped drug cisapride recovers at each pulse. However, the staircase protocol is designed to explore the full dynamics of the hERG channel by exploring many voltage steps within a single pulse [Lei et al., 2019b]. The resultant I_{Kr} of a single pulse therefore contains more information on the channel kinetics than the other protocols. It is also a longer protocol than the ramp and the step protocols. However, to retain the idea of using several pulses to identify trapped and nontrapped drugs, we decided to use more than a single pulse, but fewer than 10 pulses. After running a simple test with three and five pulses, there is no clear difference in the inference accuracy between the two. Therefore, we settled on using three pulses of the staircase protocol, as it is also computationally cheaper.

We perform the practical identifiability analysis on the SD formulation with synthetic data generated from the protocols described above. Figure 5.8 compares the fit of the parameters between the protocols. Each panel shows distributions of the difference between the inferred parameter values and their respective true parameter values. Each distribution contains 125 points, i.e., all the parameter combinations chosen to cover the parameter space (Section 5.3.1). A zero value of the differences (ΔK_{max}, ΔK_u, $\Delta V_{\mathrm{half-trap}}$ and Δtraces) is the target as it indicates that the inferred value is the same as the true value.

In the ideal case without noise (first row), both K_{max} and K_u (parameters related to the binding and unbinding rates respectively) are equally identifiable for all protocols as the distributions are centered close to zero. Even though ΔK_{max} can

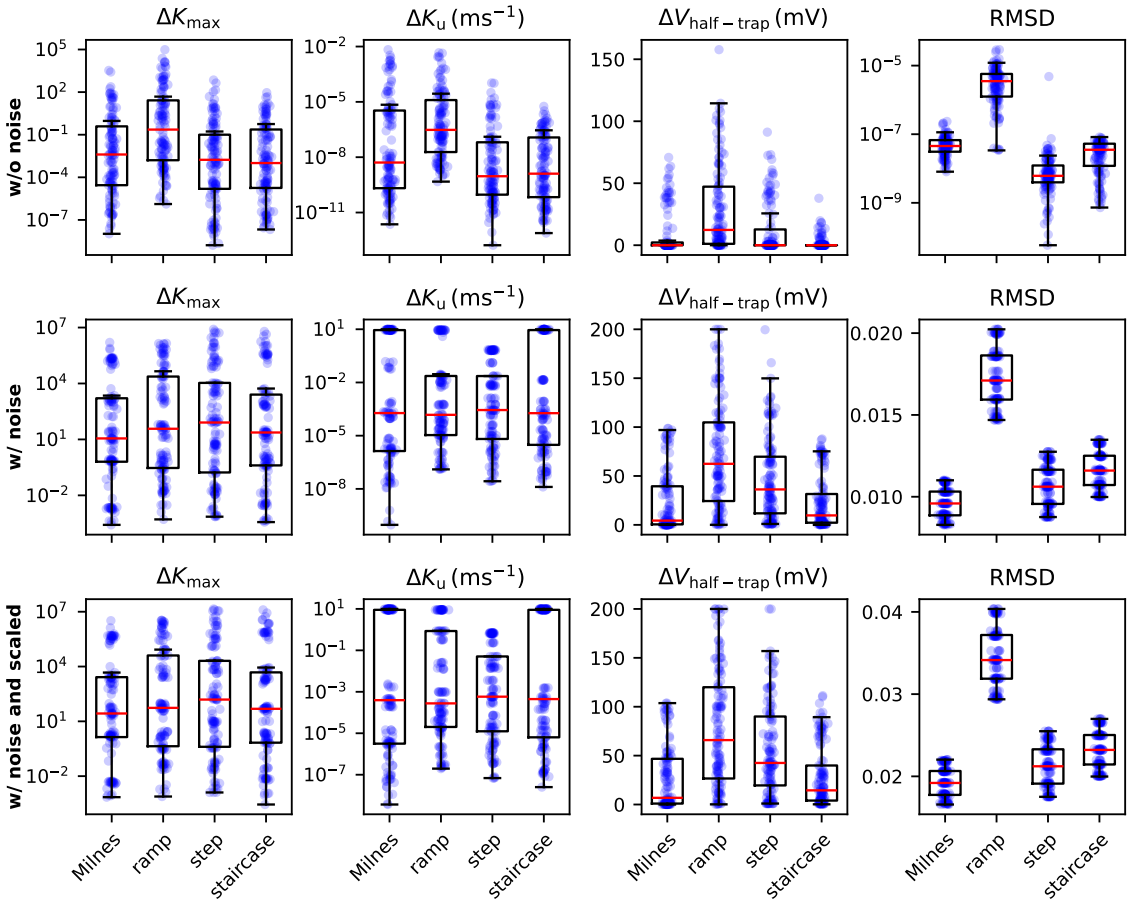


Figure 5.8: **The modified Milnes' protocol provides the highest accuracy for all parameters in noisy data.** Each panel shows boxplots of the difference between the inferred and the true parameter value, and the RMSD between the fitted and the true fractional block (right column) for all four protocols – the modified Milnes' protocol, the ramp protocol, the step protocol, and the staircase protocol. The red lines indicate the median of the distribution.

achieve values close to 10^5 under the ramp protocol, it is relatively small given that the maximum possible difference based on the parameter space is 10^8. Note that the parameter space is defined with ranges $K_{\text{max}} \in [1, 10^8]$, $K_u \in [10^{-5}, 1]$ ms^{-1} and $V_{\text{half-trap}} \in [-200, -1]$ mV. The staircase protocol improves the identifiability of $V_{\text{half-trap}}$ under noise-free conditions, as observed by the median that is closer to zero and a smaller spread of the distribution. No improvement to the identifiability of $V_{\text{half-trap}}$ is observed for the ramp and the step protocol, even though the RMSD is lower with the step protocol than with the modified Milnes' protocol. As the noise intensity increases, ΔK_{max} and ΔK_u of all protocols have similar orders of magnitude. On the other hand, the inference of $V_{\text{half-trap}}$ gets worse for the ramp and the step protocol, with the recovered value of $\Delta V_{\text{half-trap}}$ even reaching 200 mV; this is equivalent to the whole of the defined range of $V_{\text{half-trap}}$. The spread of the distribution for $\Delta V_{\text{half-trap}}$ is similar between the modified Milnes' protocol and the staircase protocol, but the median $\Delta V_{\text{half-trap}}$ and RMSD are higher in the staircase protocol.

For synthetic data without noise (top row), the step protocol and the staircase protocol achieve better accuracy in ΔK_{max}, ΔK_u and RMSD than the other two protocols, and comparable accuracy in $\Delta V_{\text{half-trap}}$ to the modified Milnes' protocol. $V_{\text{half-trap}}$ is best inferred using the staircase protocol if we have noise-free experimental data. As noise is added to the synthetic data, K_{max} and K_u are inferred with similar accuracy regardless of the protocol used. The unbinding rate K_u tends to be overestimated close to the bound when the true values are large (data for the modified Milnes' protocol are shown in Figure 5.5), but that is not the case for the ramp and the step protocol. The fractional block generated from the modified Milnes' protocol, out of all protocols tested, can best predict $V_{\text{half-trap}}$ and best fit the traces when encountered with noisy data. Despite the poor identifiability of $V_{\text{half-trap}}$ from noisy data (Figure 5.5), the modified Milnes' protocol is still best suited to identify $V_{\text{half-trap}}$, out of all of the protocols tested.

In the SD drug-binding formulation, the drug-trapping mechanism is voltage-dependent (Equation 5.3). However, the experimental data used to calibrate $V_{\text{half-trap}}$ has a fixed voltage of $V = 0$ mV. This leads to our attempt to explore the membrane potential V by using protocols with varying voltage steps to better identify $V_{\text{half-trap}}$. Indeed, the staircase protocol (Figure 5.7) that consists of many voltage steps can best identify $V_{\text{half-trap}}$ when no noise is added. However, in noisy conditions, the median of $\Delta V_{\text{half-trap}}$ using the staircase protocol is higher than the modified Milnes' protocol, but the spread of the distribution is narrower (Figure 5.8). In conclusion, the use of other protocols to explore the voltage-dependency defined in the drug-trapping

mechanism of the SD formulation (Equation 5.3) does not improve the identifiability of $V_{\text{half-trap}}$.

We investigated the variables involved in the equation of the untrapping rate K_n (Equation 5.2) by increasing the trapping rate K_t and perturbing the membrane potential V via various voltage-clamp protocols. These adjustments do not improve the identifiability of $V_{\text{half-trap}}$ nor induce any changes in I_{Kr}. This leads to the question of whether the trapping component is required to explain the trapping behaviour observed in the experimental data.

5.5 Alternative drug-binding formulations without the trapping component

The SD drug-binding formulation consists of a binding component and a trapping component (Figure 5.1 and Section 5.2). We have shown that by modifying the trapping component (Equations 5.2 & 5.3), the practical identifiability of $V_{\text{half-trap}}$ remains unchanged. This implies that $V_{\text{half-trap}}$ is not practically identifiable from experimental data, and that the fractional block does not contain sufficient information to elucidate $V_{\text{half-trap}}$.

5.5.1 Introduction to the alternative drug-binding formulations

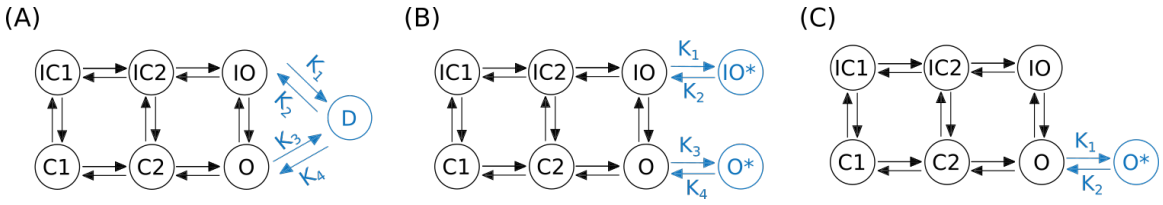


Figure 5.9: **Alternative drug-binding formulations to the SD drug-binding formulation without the trapping component.** The **(A)** C, **(B)** NC and **(C)** O drug-binding formulations are shown in blue, attached to the hERG channel component (black) from Li et al. [2016].

The first option is to combine all three drug-bound states into a single drug-bound state. Since the trapping/untrapping rates are not identifiable, we combine the complicated drug-binding kinetics into an aggregated rate which we could potentially identify while retaining the behaviour of the SD formulation. The structure of this drug-binding formulation is shown in Figure 5.9(A) in blue. It will be referred to as

the C drug-binding formulation after the cycle of the drug-binding formulation. The second alternative is to remove the trapping component (red component in Figure 5.1) altogether and to keep the binding component (blue component in Figure 5.1). This is essentially having the trapping rate $K_t = 0$ ms^{-1}, which also enables us to remove the microscopic reversibility requirement placed on K_r in the SD formulation. Figure 5.9(B) shows the structure of this drug-binding formulation, which will be referred to as the No Cycle (NC) drug-binding formulation. The third option is to have a single drug-bound state and to allow channels to enter this state only from the open state. The reason for this structure is that the measurable current flows only when the channel is open, and any transition between non-open states will have a minimal effect on the measured current. This drug-binding formulation will be referred to as the O formulation, and its structure is shown in Figure 5.9(C). The Markov structures of the NC formulation and the O formulation were used in a comparative study of various drug-binding formulations [Lei et al., 2024], but with different definitions for the binding kinetics. They showed that the prediction of proarrhythmic risk varies with the drug-binding formulations.

We are going to compare the models mathematically to show how the parameters in the SD formulation and the alternative drug-binding formulations map to one another. This will be used in Section 5.5.2 to compare the formulations. To do this, we first define the transition rates of the C formulation by comparing it with the SD formulation. We take the binding rates (K_1 and K_3 in Figure 5.9A) to be the same as defined in the SD formulation, leading to

$$K_1 = K_3 = K_{b,C}\frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.13}$$

where the $K_{b,C}$ is the binding rate of the C formulation. It is equivalent to $K_{\max} \times K_u$ in the SD formulation (Equation 5.1). The C formulation has a cycle that consists of the open (O), inactivated-open (IO), and drug-bound (D) states; thus, the transition rates have to satisfy the microscopic reversibility, where we have

$$K_2 K_{IO\to O} K_3 = K_1 K_4 K_{O\to IO}, \tag{5.14}$$

$$\Rightarrow K_2 = K_4 \frac{K_{O\to IO}}{K_{IO\to O}}, \tag{5.15}$$

where $K_{O\to IO}$ is the transition rate from the O state and the IO state, and similarly for $K_{IO\to O}$. For a comparison between the drug-binding formulations, we take the C formulation as a summarised version of the SD formulation, and thus let

$$[\mathrm{D}] = [\mathrm{O}^*] + [\mathrm{IO}^*] + [\mathrm{C}^*]. \tag{5.16}$$

where $[X]$ is the probability of being in state X. The rate of change of the drug-bound state of the C formulation is

$$\frac{d[\mathrm{D}]}{dt} = K_1[\mathrm{IO}] + K_3[\mathrm{O}] - (K_2 + K_4)[\mathrm{D}]. \tag{5.17}$$

From Equation 5.16, we have

$$\frac{d[\mathrm{D}]}{dt} = \frac{d[\mathrm{O}^*]}{dt} + \frac{d[\mathrm{IO}^*]}{dt} + \frac{d[\mathrm{C}^*]}{dt}. \tag{5.18}$$

The derivatives of the transition rates are defined in the SD formulation:

$$\frac{d[\mathrm{O}^*]}{dt} = K_b[\mathrm{O}] + K_n[\mathrm{C}^*] - (K_u + K_t)[\mathrm{O}^*], \tag{5.19}$$

$$\frac{d[\mathrm{IO}^*]}{dt} = K_b[\mathrm{IO}] + K_n[\mathrm{C}^*] - (K_r + K_t)[\mathrm{IO}^*], \tag{5.20}$$

$$\frac{d[\mathrm{C}^*]}{dt} = K_t[\mathrm{IO}^*] + K_t[\mathrm{O}^*] - 2K_n[\mathrm{C}^*]. \tag{5.21}$$

Combining Equations 5.19–5.21 with Equation 5.18, we get

$$\frac{d[\mathrm{D}]}{dt} = K_b[\mathrm{IO}] + K_b[\mathrm{O}] - K_u[\mathrm{O}^*] - K_n[\mathrm{C}^*] - K_r[\mathrm{IO}^*]. \tag{5.22}$$

We then compare Equation 5.22 and Equation 5.17. We have that $K_1 = K_3 = K_b$ as defined in Equation 5.13. Based on Equation 5.16, we take the maximum value of each term on the right hand side and

$$\text{suppose } [\mathrm{O}^*] = 1 \Rightarrow [\mathrm{IO}^*] = [\mathrm{C}^*] = 0, [\mathrm{D}] = 1 \Rightarrow K_2 + K_4 = K_u, \tag{5.23}$$

$$\text{suppose } [\mathrm{IO}^*] = 1 \Rightarrow [\mathrm{O}^*] = [\mathrm{C}^*] = 0, [\mathrm{D}] = 1 \Rightarrow K_2 + K_4 = K_r, \tag{5.24}$$

$$\text{suppose } [\mathrm{C}^*] = 1 \Rightarrow [\mathrm{O}^*] = [\mathrm{IO}^*] = 0, [\mathrm{D}] = 1 \Rightarrow K_2 + K_4 = K_n. \tag{5.25}$$

The microscopic reversibility rate, termed K_r in the SD formulation, is defined as in Equation 5.15 in the C formulation. Additionally, there is no untrapping rate K_n in the C formulation. Therefore, we are left with the first relation in Equation 5.23

$$K_2 + K_4 = K_u. \tag{5.26}$$

Referring to the microscopic reversibility relation (Equation 5.15) for the C formulation, we have

$$K_2 = \frac{K_u K_{O\to IO}}{K_{IO\to O} + K_{O\to IO}}, \tag{5.27}$$

$$K_4 = \frac{K_u K_{IO\to O}}{K_{IO\to O} + K_{O\to IO}}. \tag{5.28}$$

Note that the relations in Equations 5.27 and 5.28 are for the comparison of the drug-binding kinetics between the SD formulation and the C formulation. It is also valid to have the binding rates of the C formulation to be defined as in Equation 5.13 and the unbinding rates as

$$K_2 = \frac{K_{u,C} K_{O \to IO}}{K_{IO \to O} + K_{O \to IO}}, \tag{5.29}$$

$$K_4 = \frac{K_{u,C} K_{IO \to O}}{K_{IO \to O} + K_{O \to IO}}, \tag{5.30}$$

where $K_{u,C}$ is the unbinding rate of the C formulation.

Similarly, for the NC drug-binding formulation, we take the binding rates to be the same as defined in the SD formulation, which gives

$$K_1 = K_3 = K_{b,NC} \frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.31}$$

where $K_{b,NC}$ is the binding rate of the NC formulation. Since there is no cycle in the NC formulation, there is no requirement on microscopic reversibility, so we take K_2 and K_4 as is. For better readability and consistency in the naming method, we name K_4 as $K_{u,NC}$ for the unbinding rate from the open bound state (O*) and K_2 as $K_{r,NC}$ for the unbinding rate from the inactivated bound state (IO*), though it has nothing to do with the microscopic reversibility.

Finally, for the O drug-binding formulation, we define the transition rates K_1 and K_2 by taking reference from the SD formulation, which gives

$$K_1 = K_{b,O} \frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.32}$$

$$K_2 = K_{u,O}, \tag{5.33}$$

where $K_{b,O}$ and $K_{u,O}$ are the binding rate and unbinding rate of the O formulation, respectively.

Now, we have proposed three alternatives to the SD drug-binding formulation without the trapping component. Figure 5.10 shows the structure of the C formulation, the NC formulation, and the O formulation, with the naming of the transition rates comparable to that of the SD formulation. For easier readability, the defining equations of the alternative drug-binding formulations are reiterated. Based on the

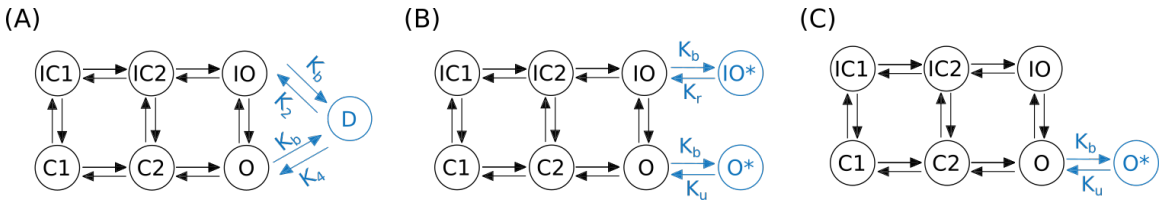


Figure 5.10: **Alternative drug-binding formulations to the SD drug-binding formulation without the trapping component, with matching naming convention.** The **(A)** *C*, **(B)** *NC* and **(C)** *O* drug-binding formulations are shown in blue, attached to the hERG channel component from Li et al. [2016]. The naming of the transition rates matches that of the SD formulation for comparison between drug-binding formulations.

naming convention in Figure 5.10, the transition rates of the C formulation are

$$K_b = K_{b,C}\frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.34}$$

$$K_2 = \frac{K_{u,C}K_{O\to IO}}{K_{IO\to O} + K_{O\to IO}}, \tag{5.35}$$

$$K_4 = \frac{K_{u,C}K_{IO\to O}}{K_{IO\to O} + K_{O\to IO}}; \tag{5.36}$$

the equations of the NC formulation are

$$K_b = K_{b,NC}\frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.37}$$

$$K_u = K_{u,NC}, \tag{5.38}$$

$$K_r = K_{r,NC}; \tag{5.39}$$

and the definition of the O formulation is

$$K_b = K_{b,O}\frac{D^n}{D^n + \mathrm{EC}_{50}^n}, \tag{5.40}$$

$$K_u = K_{u,O}. \tag{5.41}$$

We then compare these formulations to the SD formulation in terms of their parameter identifiability, their fit to the experimental data (how well the formulations capture the experimental data), the properties unique to each formulation and their contribution to the risk metric score.

5.5.2 Practical identifiability of the alternative drug-binding formulations

We have shown in Section 5.3.5 that some of the parameters of the SD drug-binding formulation are unidentifiable, specifically the parameter that controls the trapping

tendency $V_{\text{half-trap}}$. The C formulation and the O formulation each have two parameters, one less than the SD formulation, while the NC formulation has fewer constraints on its transition rates (without the microscopic reversibility and the fixed trapping rate K_t). These alternative drug-binding formulations are therefore expected to have better parameter identifiability than the SD formulation.

To investigate the parameter identifiability of these formulations, we repeat the procedure described in Section 5.3.3 with the parameter space individually defined for each drug-binding formulation. Based on our definition of the binding rate of the C formulation $K_{b,C}$ (Equation 5.13), it is equivalent to $K_{\text{max}} \times K_u$ of the SD formulation. We can define its range by referring to $K_{\text{max}} \times K_u$ of the SD formulation, which would give the range $[10^{-5}, 3 \times 10^4]$ ms^{-1} to encompass the parameter values of all CiPA training and validation drugs [Li et al., 2017b, 2019]. $K_{u,C}$ of the C formulation is comparable to K_u of the SD formulation, as defined in Equation 5.26, thus can be defined to have the same range as $K_u \in [10^{-5}, 1]$ ms^{-1}. However, if the parameter space is explored on equidistant grids, we would have parameter combinations with larger unbinding rates K_2 and K_4 than binding rates K_1 and K_3 (Figure 5.9A and Equations 5.34 to 5.36). That would imply drug compounds almost never stay bound to the channel. Once they bind, they unbind very quickly. The unbinding of drugs is preceded by binding. The reaction would be too fast for the current to be inhibited by the drugs and, therefore, would not contain sufficient information to identify the unbinding rates. The parameters cannot be inferred accurately whenever the binding rates are lower than the unbinding rates (data not shown). The relation $K_b > K_u$ is given in the SD formulation because K_{max} is always greater than 1, giving the binding rate coefficient $(K_{\text{max}} \times K_u) > K_u$. Moreover, the inferred parameters of the C formulation from the experimental data have binding rates greater than unbinding rates (Table 5.4, Figure 5.14). Therefore, we impose a constraint $K_b > K_u$ on the parameter space to ensure the sampled parameter combinations are physiologically feasible. Since the defined range of K_b and K_u of the parameter space overlapped, we would have parameter combinations that violate the constraint if we were to sample the parameter combinations as an equidistant grid. As a simple cut-off, we remove parameter combinations that do not satisfy the constraint $K_{b,C} > K_{u,C}$ and increase the number of points in the grids explored, from five to six, in the parameter space for higher resolution. With this revised definition of the parameter space, we have $(6 \times 6) - 11 = 25$ parameter combinations for the C formulation, with 11 parameter combinations violating the $K_{b,C} > K_{u,C}$ constraint.

For the NC formulation, the ranges of its binding and unbinding rates are $K_{b,NC} \in [10^{-5}, 3 \times 10^{4}]$ ms^{-1} and $K_{u,NC} \in [10^{-5}, 1]$ ms^{-1}, respectively, similar to the C formulation. On the other hand, the range of $K_{r,NC}$ is referenced from K_r of the SD formulation. K_r of the SD formulation was defined to satisfy the microscopic reversibility, to ensure the transition rate along the cycle is the same regardless of the route taken (clockwise or counterclockwise) [Li et al., 2017b]. K_r is $\frac{K_{O\to IO}}{K_{IO\to O}} K_u$, where $\frac{K_{O\to IO}}{K_{IO\to O}}$ is constant at 0 mV of the modified Milnes' protocol, as is the case for Li et al. [2017b]. K_r roughly equals to $3.8672 K_u$. Therefore, we had the range of $K_{r,NC} \in [5 \times 10^{-5}, 4]$ ms^{-1}. Following the same reasoning as in the C formulation in constraining the binding and unbinding rates, we remove parameter combinations with $K_{b,NC} > K_{u,NC}$ and do not impose any restriction on $K_{r,NC}$. After increasing the number of grid points and constraining the binding and unbinding rates, there are $(6 \times 6 \times 6) - 88 = 128$ parameter combinations for the constrained parameter space of the NC formulation, where there are 88 parameter combinations violating the $K_{b,NC} > K_{u,NC}$ constraint.

The binding and unbinding rates of the O formulation are similar to those of the NC formulation. We have the range of the binding and unbinding rates as $K_{b,O} \in [10^{-5}, 3 \times 10^{4}]$ ms^{-1} and $K_{u,O} \in [10^{-5}, 1]$ ms^{-1}, respectively. We imposed the same constraint $K_{b,O} > K_{u,O}$ on the O formulation. Similar to the C formulation, the parameter space is explored on an equidistant grid of six grid points. There are also 25 parameter combinations for the O formulation. The parameter space of all drug-binding formulations is listed in Table 5.2.

The parameter identifiability of the alternative drug-binding formulations is investigated using the same procedure as for the SD formulation, described in Section 5.3.3. Briefly, synthetic data with varying noise intensity are generated, as detailed in Section 5.3.1. The parameter space, as defined in Table 5.2, is explored on an equidistant grid on a logarithmic scale. The alternative formulations are then fitted to the noisy synthetic data. The inferred parameters are compared to their true parameter values.

Figure 5.11 shows the parameter identifiability of all drug-binding formulations under all noise models. The x-coordinates of the points in the left and the middle panels are the true parameter values used to generate the synthetic data. The inferred parameter values correspond to the y-coordinates. The parameter is identifiable if all points lie on the diagonals, where the inferred values equal the true values. The binding rate K_b and the unbinding rate K_u of all drug-binding formulations are comparable based on the definition of the drug-binding formulations, as described in Section 5.5.1. The binding rates $K_{b,C}$, $K_{b,NC}$ and $K_{b,O}$ are comparable to the

Table 5.2: Defined parameter space for all drug-binding formulations.

Drug-binding formulation	Parameter space
SD formulation	$K_{\max} \in [1, 10^8]$ $K_u \in [10^{-5}, 1]\,\mathrm{ms}^{-1}$ $V_{\mathrm{half-trap}} \in [-200, -1]\,\mathrm{mV}$ $\Longrightarrow (5 \times 5 \times 5) = 125\,\mathrm{points}$
C formulation	$K_{b,C} \in [10^{-5}, 3 \times 10^4]\,\mathrm{ms}^{-1}$ $K_{u,C} \in [10^{-5}, 1]\,\mathrm{ms}^{-1}$ with $K_{b,C} > K_{u,C}$ $\Longrightarrow (6 \times 6) - 11 = 25\,\mathrm{points}$
NC formulation	$K_{b,NC} \in [10^{-5}, 3 \times 10^4]\,\mathrm{ms}^{-1}$ $K_{u,NC} \in [10^{-5}, 1]\,\mathrm{ms}^{-1}$ $K_{r,NC} \in [5 \times 10^{-5}, 4]\,\mathrm{ms}^{-1}$ with $K_{b,NC} > K_{u,NC}$ $\Longrightarrow (6 \times 6 \times 6) - 88 = 128\,\mathrm{points}$
O formulation	$K_{b,O} \in [10^{-5}, 3 \times 10^4]\,\mathrm{ms}^{-1}$ $K_{u,O} \in [10^{-5}, 1]\,\mathrm{ms}^{-1}$ with $K_{b,O} > K_{u,O}$ $\Longrightarrow (6 \times 6) - 11 = 25\,\mathrm{points}$

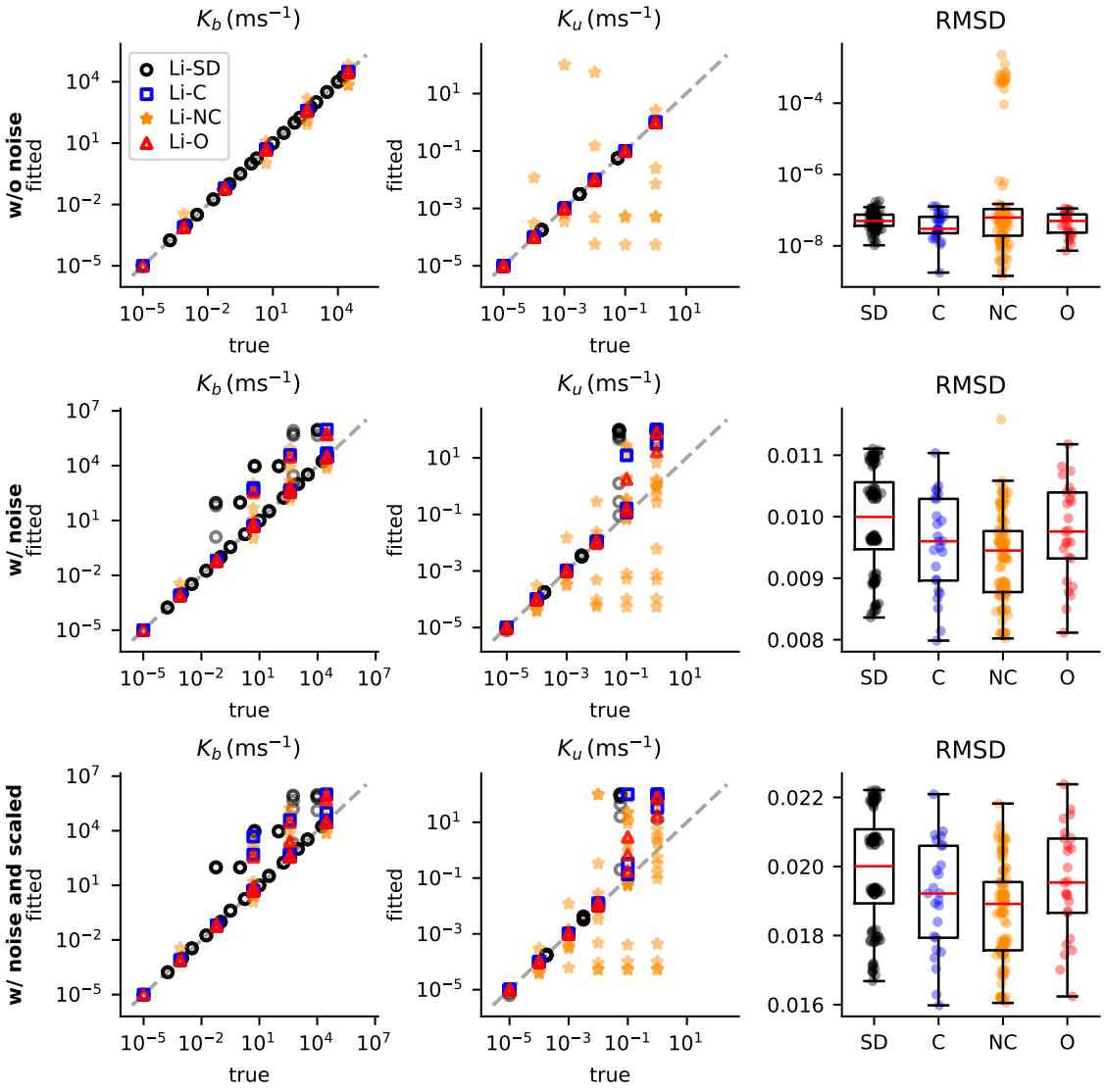


Figure 5.11: **The O formulation with the least number of parameters has the best parameter identifiability, comparable with the C formulation.** The left and middle columns show the inferred values of the binding rate K_b and the unbinding rate K_u, respectively, against the true values. The results of the SD formulation (black circles), the D formulation (blue squares), the NC formulation (orange stars), and the O formulation (red triangles) are shown. The diagonal line serves as a reference where the fitted parameters equal the true parameter values. The right-most column gives the RMSDs between the synthetic data and the fitted simulations.

binding-rate coefficient $K_{\max} \times K_u$ of the SD formulation, while the unbinding rates $K_{u,C}$, $K_{u,NC}$ and $K_{u,O}$ can be compared to K_u of the SD formulation. Please see Figures 5.1 and 5.10 for references to the transition rates. The right column shows the distribution of the RMSDs of the fit to the synthetic data for all the drug-binding formulations.

In the noise-free condition (top row), the binding rates K_b and unbinding rates K_u of all drug-binding formulations, except the NC formulation, are inferred perfectly to the true values. The NC formulation has both K_b and K_u inferred inaccurately, with K_u being inferred further from the true values. This could be due to the lack of information within the current signal regarding the transition rates between the inactivated states, and the additional degree of freedom allows for compensation between parameters to capture the current behaviour. This will be further discussed in Section 5.6. The RMSDs of all drug-binding formulations have similar medians. When noise is added (second row), K_b and K_u of the SD formulation are overestimated when the unbinding rate K_u is large. The overestimation of K_b of the SD formulation results from the overestimation of K_u as K_b here refers to $K_{\max} \times K_u$ of the SD formulation (please also refer to Figure 5.4). The C formulation loses some degree of identifiability at larger K_bs and K_us, with both parameters being overestimated. The parameter identifiability of the NC formulation gets worse with noise. While K_b is mainly overestimated, K_u can be both overestimated and underestimated in the NC formulation. The mismatch between the inferred K_u and the true K_u becomes greater as the true K_u gets larger. Even though a wide range of K_u, covering almost the whole parameter space, is inferred from larger values of actual K_u in the NC formulation, there is a subset of parameter combinations that have K_u inferred accurately. K_b and K_u of the O formulation are misidentified when the true values are large. This is similar to the C formulation, but the O formulation has K_u values inferred closer to the true value. On the other hand, K_u values of the SD formulation are inferred at the upper limit of the parameter space for almost all parameter combinations with large K_u. The RMSD between the synthetic data and the fitted simulations of the NC formulation has the lowest median, with its distribution placed relatively lower. At the highest noise intensity (the N3 noise model), the parameter identifiability of all drug-binding formulations is the worst. In the SD formulation, more K_us are inferred close to the upper limit of the parameter space when the true K_u is large. That is similar to the C formulation. In the NC formulation, more binding rates K_b are overestimated and K_u is inferred with even lesser accuracy. Similarly, in the O formulation, more of the binding rates K_b and more of the unbinding rates K_u are overestimated. The

NC formulation has the lowest median for its RMSDs, indicating a good fit to the noisy synthetic data, but has poor parameter identifiability, as observed in the left and middle columns.

For all drug-binding formulations, the parameter identifiability gets worse as the noise intensity increases. On the whole, the binding rate K_b has better identifiability than the unbinding rate K_u, regardless of the noise intensity or the drug-binding formulation. The binding/unbinding rates of all drug-binding formulations tend to be overestimated, except for the K_u of the NC formulation, where such a tendency is not observed. The NC formulation has the overall best fit to noisy synthetic data, as shown by the lower RMSDs, but it has the worst parameter identifiability. Moreover, the inferred unbinding rates K_u span the range of the parameter space. Even though the SD formulation has good parameter identifiability at low values of K_u, the identifiability deteriorates quickly when K_u is large, with K_u values mostly inferred close to the upper limit of the parameter space. Overall, the O formulation has the best parameter identifiability, and the C formulation comes next.

The O formulation and the C formulation have the fewest parameters, only two (K_b and K_u), instead of three parameters for the SD and the NC formulations. Although the NC formulation has a simpler Markov structure than the SD formulation and does not have a trapping component, the NC formulation has the same number of parameters as the SD formulation due to the freedom of independent unbinding rates K_u and K_r (transition rate from the open-bound O* state and from the inactivated-bound IO* state). Moreover, the SD formulation has a fixed trapping rate K_t, which leaves only one free parameter to control the trapping/untrapping mechanism. Therefore, rather than having two additional degrees of freedom from having the trapping component (trapping and untrapping rate) in the SD formulation, one degree of freedom is removed by fixing the trapping rate K_t, while the other degree of freedom is taken by the fixed microscopic reversibility rate K_r in the SD formulation, and the additional degree of freedom in the unbinding rates in the NC formulation. The flexibility in the unbinding rates of the NC formulation potentially allows for compensation between the parameters. The O formulation is similar to the NC formulation, but having zero binding and unbinding rates from and to the IO state. Without the IO* state, the identifiability (of the O formulation) is improved significantly. This implies that the fractional block generated from the NC formulation does not contain information to elucidate the binding and unbinding rates. On the other hand, the SD formulation has two parameters controlling the binding/unbinding rates (K_b and K_u), instead of three for the NC formulation, and one parameter controlling the

trapping/untrapping rates whose effect is limited by the fixed trapping rate K_t, thus the SD formulation has better identifiability with the binding and unbinding rates K_b and K_u.

5.5.3 Inference performance of all drug-binding formulations on experimental data

The experimental data used in this study were published by Li et al. [2017b], and are the data these authors used in developing the SD formulation that we have considered throughout this thesis. The modified Milnes' protocol described in Li et al. [2017b] is used. Current stability was ensured prior to the addition of drugs by ensuring that the difference between the current amplitudes over 10 consecutive traces is less than 5% [Li et al., 2017b]. The first ten pulses after the drug wash-in are recorded at 10 ms timestep. The fractional block, instead of the raw recorded current, at the 0 mV step of the modified Milnes' protocol is made available by Li et al. [2017b] on the GitHub repository: `https://github.com/FDA/CiPA`. Each drug concentration is repeated several times, and the mean fractional block across all repeats is treated as the data for the optimisation of the drug-binding formulations. When fitting to experimental data, we performed ten repeats of each optimisation.

We assess the performance of the C, the NC, and the O drug-binding formulations in real experimental settings by fitting these formulations to the experimental data published by Li et al. [2017b]. This is different to the synthetic data studies as there are no ground truths with experimental data, and it is impossible for any noise models to fully replicate the noise in the experimental data due to the various sources of noise, such as cell-cell variability and experimental artefacts.

The optimisation procedure is detailed in Section 5.3.3. Briefly, the experimental data are fractional blocks at the 0 mV step of the first 10 pulses of the modified Milnes' protocol (Figure 5.7, Section 5.4.2), averaged over several repeats. The optimisations to the experimental data are performed 10 times with different initial guesses. The optimisation run with the lowest RMSD is presented. The optimal objective function is achieved more than once for all drug-binding formulations.

For better visibility, the fits of the drug-binding formulations to the experimental data for 12 CiPA training drugs [Li et al., 2017b] are split into two figures of the SD and the O formulation (Figure 5.12) and the C and the NC formulation (Figure 5.13). The solid grey lines in each panel are the mean fractional block, with the shaded region indicating the standard deviation of all repeats. The drug concentrations used for each drug are listed as text in each panel. The higher the concentrations, the greater

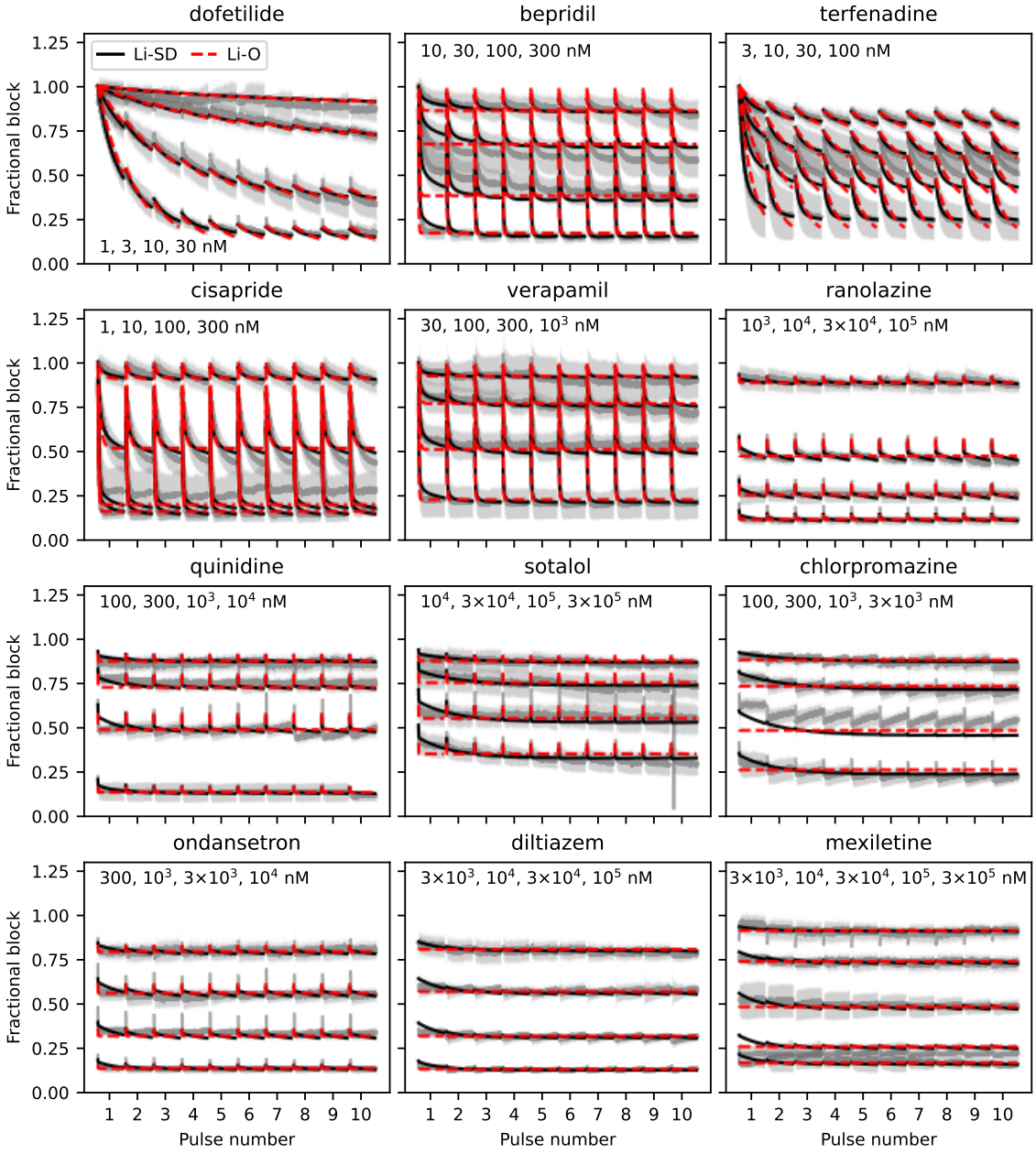


Figure 5.12: **The SD drug-binding formulation fits the experimental data (published in Li et al. [2017b]) better than the O formulation.** The grey lines show the mean fractional block, with the shaded region showing the standard deviation. The drug concentrations used for each drug are listed as text in each panel. The fitted fractional block of the SD (black solid lines) and the O (red dashed lines) formulations are overlaid on the experimental data.

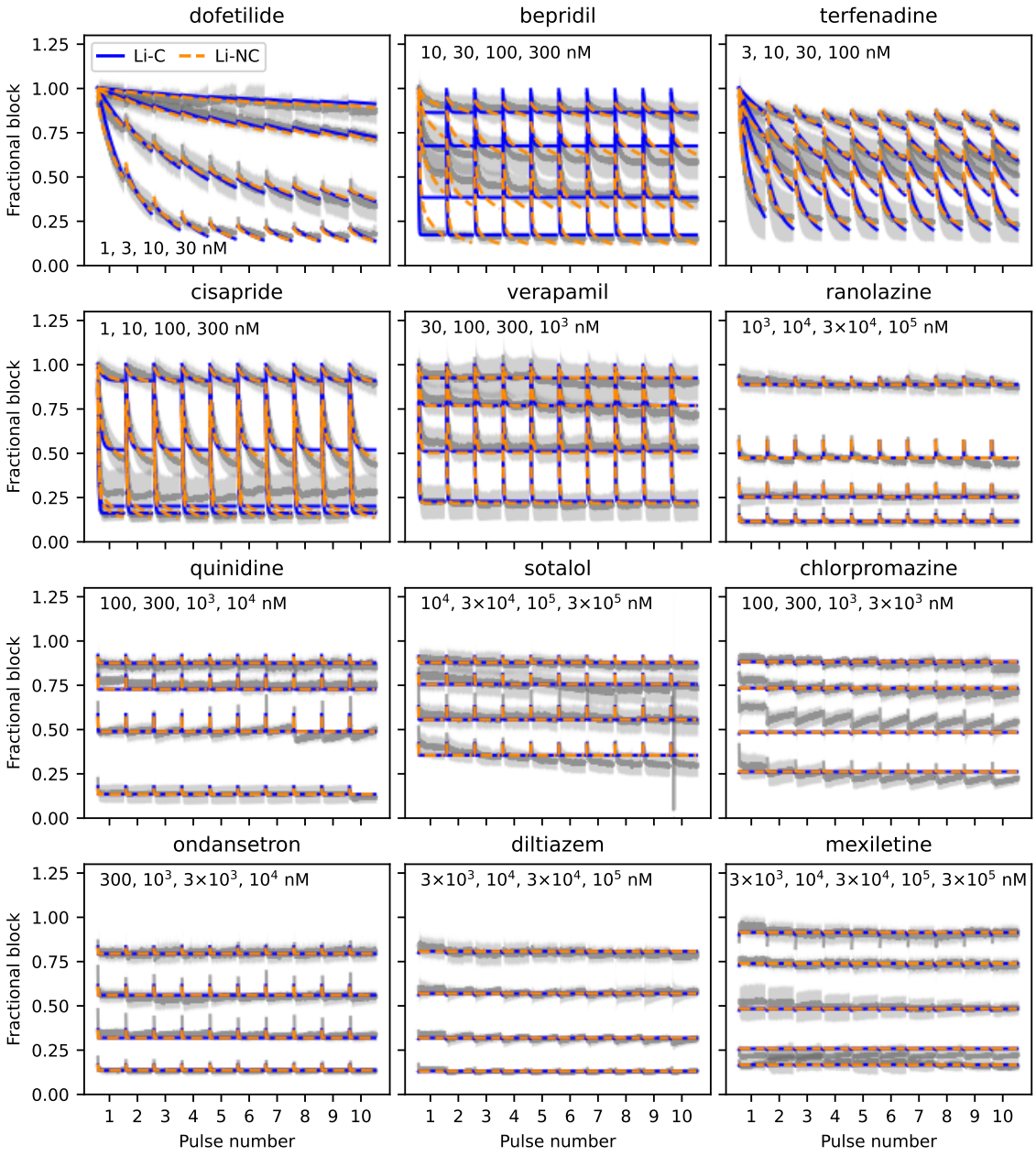


Figure 5.13: **The C drug-binding formulation tends to achieve in-pulse steady state, while the NC formulation is flexible in exhibiting slopes of the fractional blocks.** The grey lines show the mean fractional block, with the shaded region showing the standard deviation. The drug concentrations used for each drug are listed as text in each panel. The fitted fractional block of the C (blue solid lines) and the NC (orange dashed lines) formulations are overlaid on the experimental data.

the inhibition on the current, the lower the magnitude of the fractional block. Some drugs have larger standard deviations than others, showing more variability between the repeats, such as terfenadine and verapamil. Each trace shows 10 pulses of the fractional block only at the 0 mV step of the modified Milnes' protocol next to each other. During the 0 mV step, the hERG channel activates, and drug compounds bind to the open channel and inhibit the current. This gives the decreasing traces within each pulse. In between pulses at the resting potential of −80 mV (see Section 3.2.6 and Figure 5.7 for details of the modified Milnes' protocol), the membrane voltage returns to its resting potential and closes the channel. To reiterate the descriptions from Section 5.1 and Chapter 3, in this state, some drugs would stay bound in the closed channel, which is termed the *trapped* drug; others would unbind from the closed channel, which are *nontrapped* drugs. Therefore, at the subsequent opening of the channel, trapped drugs would still be bound to the channel and maintain the current inhibition from the previous pulse, with potentially more binding of drug compounds to the channel, whereas nontrapped drug are cleared from the channel and the current would recover to its previous magnitude. The rate of this recovery of current characterises the trapping phenotype of drugs. The rate of recovery is said to be slow when the current at the start of each pulse is similar to the current at the end of the previous pulse. Examples of drugs showing a slow recovery rate are dofetilide and terfenadine, while drugs showing a fast recovery rate are cisapride and verapamil. For trapped drugs, i.e., a slow recovery rate, the peak of the trace at the tenth pulse is lower than the peak at the first pulse, whereas nontrapped drugs, i.e., a fast recovery rate, have similar peaks for all 10 pulses.

In Figures 5.12 and 5.13, the fits of the SD, O, C, and NC formulations to the experimental fractional block are shown in black solid lines, red dashed lines, blue solid lines, and orange dashed lines, respectively. The SD formulation is fitted to the experimental data with the optimisation method described in Section 5.3.3, and not taken directly from Li et al. [2017b], for a fair comparison between all drug-binding formulations. Figure 5.14 compares the fitted K_b, the fitted K_u, and the RMSD between the fitted simulations and the experimental data of all drug-binding formulations. Figures 5.12, 5.13 and 5.14 will be viewed collectively for comparison. We will focus on three behaviours of the fractional block for easier explanation and comprehension. The first feature is the peak of the fractional block of each pulse, a property that is used to characterise the trapping behaviour. The next feature is the decreasing curve within each pulse. The fractional block can decrease linearly such as dofetilide, or decrease in a way that resembles an exponential decay such as bepridil.

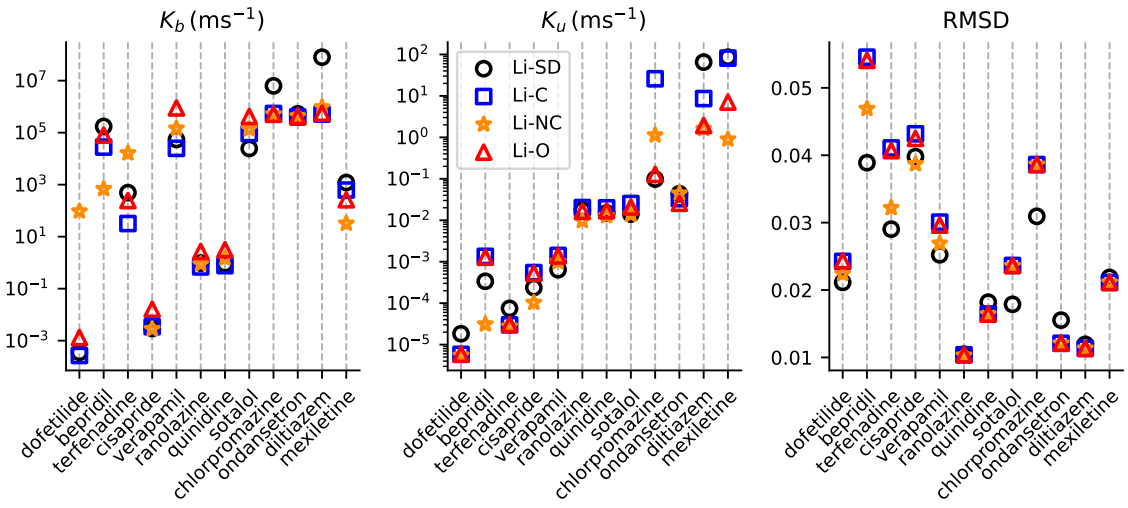


Figure 5.14: **Comparison of fitted K_b, K_u and the RMSD between all drug-binding formulations. The SD formulation has the overall lowest RMSD, while the alternative drug-binding formulations outperform the SD formulation for some of the drugs.**

Table 5.3: Inferred parameter values of the SD drug-binding formulation for all CiPA training drugs.

Drug	K_b	K_u (ms^{-1})	EC50^n (nM)	N	$V_{\text{half-trap}}$ (mV)	RMSD
Dofetilide	3.553×10^{-4}	1.828×10^{-5}	1.39×10^{2}	1.185	0	3.889×10^{-2}
Verapamil	5.453×10^{4}	6.408×10^{-4}	1.854×10^{10}	1.047	−73.01	2.523×10^{-2}
Bepridil	1.730×10^{5}	3.371×10^{-4}	2.02×10^{10}	1.009	−14.02	3.889×10^{-2}
Terfenadine	4.926×10^{2}	7.474×10^{-5}	2.713×10^{7}	0.667	−75.77	2.903×10^{-2}
Cisapride	2.938×10^{-3}	2.376×10^{-4}	5.7×10^{1}	1.034	−204.1	3.979×10^{-2}
Ranolazine	9.616×10^{-1}	1.768×10^{-2}	1.509×10^{5}	0.9544	−88.80	1.047×10^{-2}
Quinidine	9.087×10^{-1}	1.495×10^{-2}	1.238×10^{4}	0.8719	−74.32	1.819×10^{-2}
Sotalol	2.468×10^{4}	1.405×10^{-2}	7.548×10^{9}	0.7539	−10.47	1.788×10^{-2}
Chlorpromazine	6.309×10^{6}	9.855×10^{-2}	1.803×10^{10}	0.8991	−7.168	3.094×10^{-2}
Ondansetron	5.288×10^{5}	4.379×10^{-2}	4.046×10^{9}	0.9030	−79.48	1.553×10^{-2}
Diltiazem	7.999×10^{7}	6.529×10^{1}	4.847×10^{9}	0.9359	−69.39	1.194×10^{-2}
Mexiletine	1.209×10^{3}	8.594×10^{1}	4.569×10^{5}	1.089	−72.04	2.187×10^{-2}

Table 5.4: Inferred parameter values of the C drug-binding formulation for all CiPA training drugs.

Drug	K_b	K_u (ms^{-1})	$EC50^n$ (nM)	N	RMSD
Dofetilide	2.691×10^{-4}	5.831×10^{-6}	1.468×10^{2}	1.175	2.425×10^{-2}
Verapamil	2.496×10^{4}	1.418×10^{-3}	7.084×10^{9}	1.048	3.007×10^{-2}
Bepridil	2.784×10^{4}	1.348×10^{-3}	1.222×10^{9}	0.9942	5.455×10^{-2}
Terfenadine	3.197×10^{1}	2.997×10^{-5}	3.589×10^{6}	0.6385	4.108×10^{-2}
Cisapride	3.429×10^{-3}	5.464×10^{-4}	5.808×10^{1}	1.016	4.318×10^{-2}
Ranolazine	6.86×10^{-1}	2.024×10^{-2}	1.683×10^{5}	0.9419	1.04×10^{-2}
Quinidine	7.68×10^{-1}	1.995×10^{-2}	1.345×10^{4}	0.8629	1.645×10^{-2}
Sotalol	9.266×10^{4}	2.536×10^{-2}	2.368×10^{10}	0.7464	2.365×10^{-2}
Chlorpromazine	5.462×10^{5}	2.59×10^{1}	8.93×10^{6}	0.8889	3.862×10^{-2}
Ondansetron	4.028×10^{5}	3.317×10^{-2}	8.162×10^{9}	0.9117	1.21×10^{-2}
Diltiazem	5.166×10^{5}	8.638	4.687×10^{8}	0.9466	1.134×10^{-2}
Mexiletine	6.202×10^{2}	8.164×10^{1}	4.953×10^{5}	1.1	2.112×10^{-2}

The third feature is that in some instances, the current attains a steady value within a single pulse, which we term in the foregoing the *steady value*. For example, quinidine and mexiletine achieve a steady value instantaneously at the start of each pulse, while dofetilide and terfenadine do not achieve a steady value within each pulse, and their traces decrease throughout the pulse.

For dofetilide, all four drug-binding formulations have similar simulated outputs with a slight difference in the peak and the decreasing curve within each pulse, especially at higher concentrations, where the alternative drug-binding formulations (red dashed lines for the O formulation, blue solid lines for the C formulation, and orange dashed lines for the NC formulation) predict a more linear decrease. The RMSDs of the drug-binding formulations (Figure 5.14) are close to each other, but the SD formulation has the lowest RMSD, indicating that the SD formulation fits closer to the experimental data. Bepridil and terfenadine show the largest difference among the drug-binding formulations in terms of their RMSDs. For bepridil, the C formulation (blue solid) struggles to capture the decreasing curve of the traces, especially at 30 nM and 100 nM. The prediction of the C formulation achieves a steady value instantaneously for all concentrations from the first pulse onwards, which does not necessarily replicate the experimental data. At 30 nM, the C formulation matches the steady value of the first pulse, while at 100 nM, the C formulation matches the steady value of the last few pulses. However, this feature of the C formulation matches the

Table 5.5: Inferred parameter values of the NC drug-binding formulation for all CiPA training drugs.

Drug	K_b	K_u (ms^{-1})	K_r (ms^{-1})	$EC50^n$ (nM)	N	RMSD
Dofetilide	9.472×10^1	5.467×10^{-6}	1.938×10^{-3}	9.256×10^6	1.025	2.242×10^{-2}
Verapamil	1.452×10^5	9.95×10^{-4}	1.769×10^{-2}	1.582×10^{10}	1.067	2.691×10^{-2}
Bepridil	6.896×10^2	3.102×10^{-5}	8.329×10^{-4}	1.467×10^8	0.9881	4.69×10^{-2}
Terfenadine	1.625×10^4	2.753×10^{-5}	1.124×10^{-3}	5.111×10^8	0.6552	3.219×10^{-2}
Cisapride	2.814×10^{-3}	1.029×10^{-4}	9.582×10^{-4}	5.423×10^1	1.053	3.87×10^{-2}
Ranolazine	8.855×10^{-1}	9.687×10^{-3}	4.74×10^{-2}	1.672×10^5	0.9424	1.029×10^{-2}
Quinidine	1.496	1.27×10^{-2}	8.386×10^{-2}	1.34×10^4	0.863	1.642×10^{-2}
Sotalol	1.389×10^5	1.365×10^{-2}	6.707×10^{-2}	2.419×10^{10}	0.7464	2.361×10^{-2}
Chlorpromazine	5.137×10^5	1.123	1.535	1.524×10^8	0.889	3.862×10^{-2}
Ondansetron	4.323×10^5	4.45×10^{-2}	3.536×10^{-2}	7.882×10^9	0.912	1.209×10^{-2}
Diltiazem	9.566×10^5	1.606	9.996×10^1	1.018×10^9	0.9465	1.134×10^{-2}
Mexiletine	3.202×10^1	8.971×10^{-1}	9.492×10^1	4.945×10^5	1.1	2.113×10^{-2}

Table 5.6: Inferred parameter values of the O drug-binding formulation for all CiPA training drugs.

Drug	K_b	K_u (ms^{-1})	$EC50^n$ (nM)	N	RMSD
Dofetilide	1.317×10^{-3}	5.902×10^{-6}	1.487×10^{2}	1.173	2.435×10^{-2}
Verapamil	8.959×10^{5}	1.375×10^{-3}	5.429×10^{10}	1.05	2.963×10^{-2}
Bepridil	8.003×10^{4}	1.279×10^{-3}	7.593×10^{8}	0.9944	5.408×10^{-2}
Terfenadine	2.448×10^{2}	3.072×10^{-5}	5.594×10^{6}	0.6378	4.068×10^{-2}
Cisapride	1.628×10^{-2}	5.266×10^{-4}	5.847×10^{1}	1.016	4.25×10^{-2}
Ranolazine	2.696	1.644×10^{-2}	1.6783×10^{5}	0.942	1.041×10^{-2}
Quinidine	3.161	1.695×10^{-2}	1.339×10^{4}	0.863	1.643×10^{-2}
Sotalol	4.195×10^{5}	2.091×10^{-2}	2.671×10^{10}	0.7464	2.364×10^{-2}
Chlorpromazine	5.073×10^{5}	1.28×10^{-1}	3.446×10^{8}	0.8889	3.863×10^{-2}
Ondansetron	4.226×10^{5}	2.563×10^{-2}	2.272×10^{9}	0.9115	1.212×10^{-2}
Diltiazem	5.609×10^{5}	1.93	4.677×10^{8}	0.9465	1.134×10^{-2}
Mexiletine	2.621×10^{2}	7.087	4.945×10^{5}	1.1	2.113×10^{-2}

experimental data at 300 nM where the fractional block is replicated nicely for all pulses. Although this feature does not reflect the properties of the data at lower concentrations, the C formulation replicates the data nicely at the highest drug concentration of 300 nM. The O formulation (red dashed), with the same degree of freedom as the C formulation, performs similarly to the C formulation (blue solid) in replicating the experimental data of bepridil. However, the SD (black solid) and the NC formulations (orange dashed) perform better at capturing the decreasing curve, especially at the lowest drug concentration of 10 nM. The NC formulation (orange dashed), in particular, does not achieve a steady value (flatten out) at the end of each pulse, but keeps its decreasing trend for all drug concentrations. Again, this property does not replicate the experimental data, especially not at the highest concentration of 300 nM. Of all drug-binding formulations, the SD formulation (black solid) appears to fit the trace best at 100 nM, while none of the formulations can replicate the trace at 30 nM well. All drug-binding formulations failed to capture the in-pulse steady value and the decreasing curve simultaneously. For terfenadine, another trapped drug, both the SD (black solid) and the NC (orange dashed) formulations capture the decreasing curve better than the O (red dashed) and the C (blue solid) formulations. This is apparent at 100 nM. All drug-binding formulations, except the SD formulation (black solid), fail to achieve the steady value observed within each pulse. At 30 nM, all drug-binding formulations underestimate the experimen-

tal data. The mismatch of all drug-binding formulations at one drug concentration could indicate that the experimental data at that concentration is an outlier to the expression $\frac{D^n}{D^n+\mathrm{EC}_{50}^n}$ (Equation 5.1) that defines the binding rate. This expression represents the efficacy of the drug and describes how the change in drug concentration affects the intensity of the current inhibition. Among the trapped drug examples, all drug-binding formulations show a decrease in peak current throughout the pulses, except for the C formulation (blue solid) for bepridil, indicating that all drug-binding formulations are capable of exhibiting the trapping behaviour (associated with the reduction of the peak current [Li et al., 2017b]).

For the nontrapped drug cisapride, the simulated output of the O (red dashed) and the C (blue solid) formulation reaches a steady value within each pulse almost instantaneously, but at a slower rate than for bepridil. The NC formulation (orange dashed) replicates the decreasing curve of each pulse at 10 nM better than the SD formulation. The fractional block at 100 nM of cisapride is underestimated by all drug-binding formulations, predicting magnitudes that are close to 300 nM. We speculate that the drug effect at 100 nM does not follow the expression $\frac{D^n}{D^n+\mathrm{EC}_{50}^n}$, and thus it is unable to capture the variability of the drug effect accurately. This is similar to terfenadine at 30 nM, as mentioned earlier. This mismatch from all drug-binding formulations could be due to model misspecification or inconsistent data. However, given the reasonable fit for all concentrations of other drugs, it is unlikely that the model is misspecified. Moreover, the inconsistencies in this set of data dominate the mean fractional block, conforming to our speculation that this dataset is dominated by noise. The experimental dataset giving the mean fractional block of cisapride at 100 nM is calculated from three complete repeats and one repeat with only the first five pulses. Additionally, these three complete repeats have very different magnitudes (see Figure B.1 for the individual traces). This is also apparent from the large standard deviation in the mean fractional block (Figures 5.12 and 5.13). We do not address this outlier data because of the many variational inconsistencies observed in the other drugs, adding to the complexity of the pre-processing of the data. A detailed description of the data inconsistencies and the complexity it brings is discussed in Appendix B. For the other nontrapped drug verapamil at 300 nM, the SD formulation (black solid) captures the decreasing curve at the first pulse and subsequently achieves steady values that are slightly lower than the experimental data. On the other hand, the O (red dashed), C (blue solid), and the NC (orange dashed) formulations achieve instantaneous steady values from the first pulse onwards. This

property matches the experimental current behaviour of verapamil at 10^3 nM and the later pulses of the other drug concentrations.

In the remaining drugs, the O (red dashed), the C (blue solid), and the NC (orange dashed) formulations tend to predict an instantaneous steady value within each pulse and remain relatively constant throughout the pulses. This can be reflective of the experimental data in some cases, for example, ondansetron and diltiazem. In these examples, the SD (black solid) formulation always predicts a decreasing curve at the first few pulses, never achieving a steady value instantaneously, regardless of the experimental data. In conclusion, the O formulation, similar to the C formulation, cannot capture the decreasing curve of the fractional block; the NC formulation cannot capture the in-pulse steady value unless it is achieved instantaneously; and the SD formulation cannot achieve the in-pulse steady value instantaneously at the first few pulses. None of the drug-binding formulations can exhibit a decreasing curve and achieve instantaneous steady value simultaneously for a single drug.

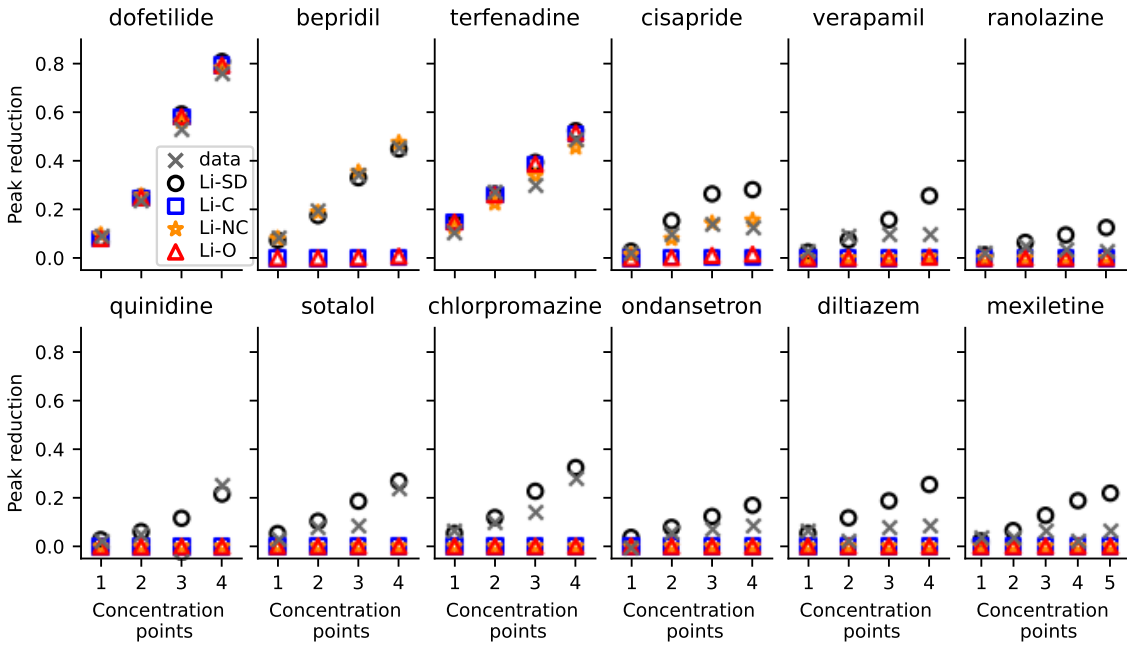


Figure 5.15: **The alternative drug-binding formulations have the flexibility to exhibit the concentration-dependent increase in peak reduction, and the constant low peak reduction as the drug concentration increases, whereas the peak reduction increases only for the SD formulation.** Peak reduction at the tenth pulse from the first pulse of the experimental data and simulations of all drug-binding formulations. The x-axis concentration points are concentrations used for each drug (listed in Figures 5.12 and 5.13) in increasing order.

The main motivation for the development of the SD formulation is to describe the trapping behaviour of drugs, which is characterised by the rate of recovery of current from previous pulses. Li et al. [2017b] uses the reduction of the peak current at the tenth pulse from the first pulse as an indicator of trapping behaviour. In Figures 5.12 and 5.13, the alternative drug-binding formulations (the C, the NC, and the O formulations) demonstrate flexibility in presenting a reduction of peak, such as dofetilide and terfenadine, and a constant peak throughout the pulses, such as cisapride and diltiazem. This implies that the alternative drug-binding formulations are capable of distinguishing the trapping behaviour, as characterised by the peak reduction. Figure 5.15 shows the peak reduction between the first pulse and the tenth pulse for the experimental data and all drug-binding formulations. The x-axis refers to the sequence of concentrations used for each drug (as listed in the panels of Figures 5.12 and 5.13) in increasing order. The peak reduction of the known trapped drugs (dofetilide, bepridil, and terfenadine) shows an increasing trend with the drug concentration, an expected feature of the trapping behaviour. The SD and the NC formulations are comparable in matching this trend, whereas the C and the O formulations fail to predict the peak reduction of bepridil. For the nontrapped drugs (cisapride, verapamil, and ranolazine), the peak reduction of the experimental data remains relatively constant as the concentration increases. The C, the NC, and the O formulations are better fits to the trapping behaviour of nontrapped drugs, i.e., maintaining similar peak reduction as drug concentration increases, than the SD formulation, even though they do not match the exact value of peak reduction. The SD formulation predicts an increasing trend of the peak reduction even when the experimental data shows constant peak reduction. This is observable in the panels of the bottom row in Figure 5.15. On the other hand, the alternative drug-binding formulations show the same peak reduction for all drug concentrations, regardless of the trapping behaviour exhibited by the experimental data. This agrees with our observation in the fit to experimental data (Figures 5.12 & 5.13), where the SD formulation (black solid) shows decreasing peaks at the first few pulses for all drugs, while the alternative drug-binding formulations show similar peak values throughout all pulses for most of the drugs (for example, quinidine, sotalol, ondansetron and diltiazem). Figure 5.15 supports our observation that the alternative drug-binding formulations are capable of exhibiting the spectrum of trapping behaviour, from the lack of recovery seen in dofetilide to the full recovery of the peak in ranolazine.

Overall, the RMSDs (Figure 5.14) show that the SD formulation fits the experimental data with the lowest RMSD for half of the drugs, whereas the C and the O

formulations perform the worst in most cases. This is to be expected since they have fewer free parameters. However, in some cases, the alternative drug-binding formulations perform better than the SD formulation (quinidine and ondansetron), and at times as well as the SD formulation (ranolazine, diltiazem and mexiletine). When the RMSDs are similar between the drug-binding formulations, the binding and unbinding rates (K_b and K_u, respectively) can take on different values. For example, diltiazem and mexiletine have similar RMSD, but their K_bs and K_us can be up to two orders of magnitude smaller than that of the SD formulation. The greatest difference in K_b is observed with dofetilide, which has at least four orders of magnitude difference, but this translates to a reasonably close RMSD.

Even though the drug-binding formulations have their limitations in replicating the experimental data, it could be partially due to the inconsistencies within the experimental data. In addition to the inconsistencies in cisapride and terfenadine, as described previously, the fractional block of chlorpromazine also shows some unexpected behaviour. The fractional block increases within each pulse instead of the expected decrease that is caused by the inhibition from drug compounds. This behaviour is also observed in quinidine at 10^3 nM. The inconsistencies in the experimental data make it challenging for us to perform our own pre-processing on the data and therefore we decided to use the experimental data as it is published and we discuss the inconsistencies further in Appendix B.

5.5.4 Comparisons to the SD formulation

The drug-binding formulations each have their limitations in fully recapitulating the current behaviour exhibited in the experimental data, although all are capable of capturing the current behaviour specific to drug-trapping (Section 5.5.3). In the fit to the experimental data in the previous section (Section 5.5.3), the SD formulation shows a largest peak reduction of 80.59%, while the C formulation has a highest peak reduction of 79.03%, the NC formulation shows 78.06% maximum peak reduction and the O formulation shows 79.07% maximum peak reduction. In addition to large peak reduction, the C, the NC and the O formulations can exhibit low peak reduction, as shown in Figures 5.12 and 5.13, with a minimum peak reduction of close to 0, illustrating their ability to encompass the spectrum of trapping behaviour. The SD formulation, however, shows a minimum peak reduction of 12.54%.

We further investigate the features of the trapping component within the SD formulation that are unique to it and cannot be replicated by the alternative drug-binding formulations. In particular, we want to investigate the difference in the

current behaviour and the binding kinetics between the alternative drug-binding formulations and the SD formulation. We have previously shown that the SD formulation is better at capturing both the decreasing curve and the in-pulse steady value, but struggles to achieve steady values instantaneously from the first pulse.

To do so, we generate noisy synthetic data from the SD formulation, as described in Section 5.3.1. Then, we fit the alternative drug-binding formulations to the noisy data generated by the SD formulation. We have described in Section 5.5.1 how the parameters of the C, the NC, and the O formulations can be comparable to the parameters of the SD formulation. We take the binding rate coefficient $K_{b,SD}$ of the SD formulation to be $K_{\max} \times K_u$. $K_{b,SD}$ can be compared to the binding rates of the alternative drug-binding formulations ($K_{b,C}$ for the C formulation, $K_{b,NC}$ for the NC formulation, and $K_{b,O}$ for the O formulation). The unbinding rate $K_{u,SD}$ of the SD formulation can be compared to the unbinding rates $K_{u,C}$ of the C formulation, $K_{u,NC}$ of the NC formulation, and $K_{u,O}$ of the O formulation.

Figure 5.16 compares the binding and unbinding rates of all drug-binding formulations inferred from synthetic data generated by the SD formulation. In the left and middle columns, the x-axis indicates the true parameter value of the SD formulation used to generate the synthetic data, while the y-coordinates refer to the inferred parameter values for all drug-binding formulations. Data points on the diagonal line imply that the binding/unbinding rates of the fitted drug-binding formulation match those of the SD formulation that was used to generate the synthetic data. Since model misspecification is inevitably introduced in this analysis, we do not expect all parameters to lie on the diagonal line.

However, since the binding rates $K_{b,SD}$, $K_{b,C}$, $K_{b,NC}$ and $K_{b,O}$ have the same definitions (Sections 5.2 & 5.5), they are expected to fall close to the diagonal line. On the other hand, the unbinding rate K_u is hypothesised to be smaller in the alternative drug-binding formulations as they do not have any trapping component, and the channel would have to stay within the D state or the O* state (refer to their structure in Figure 5.9) longer to mimic the stay within the C* state in the SD formulation and the delay of the unbinding process.

The left column of Figure 5.16 shows that K_b of the C formulation (blue square), the NC formulation (orange star), and the O formulation (red triangle) are recovered close to that of the SD formulation. Under ideal conditions without noise, more parameter combinations in the C formulation (blue squares) are predicted to have binding rates K_b lower than the true K_b of the SD formulation, whereas the NC and the O formulations mostly overestimate the K_b values of the SD formulation.

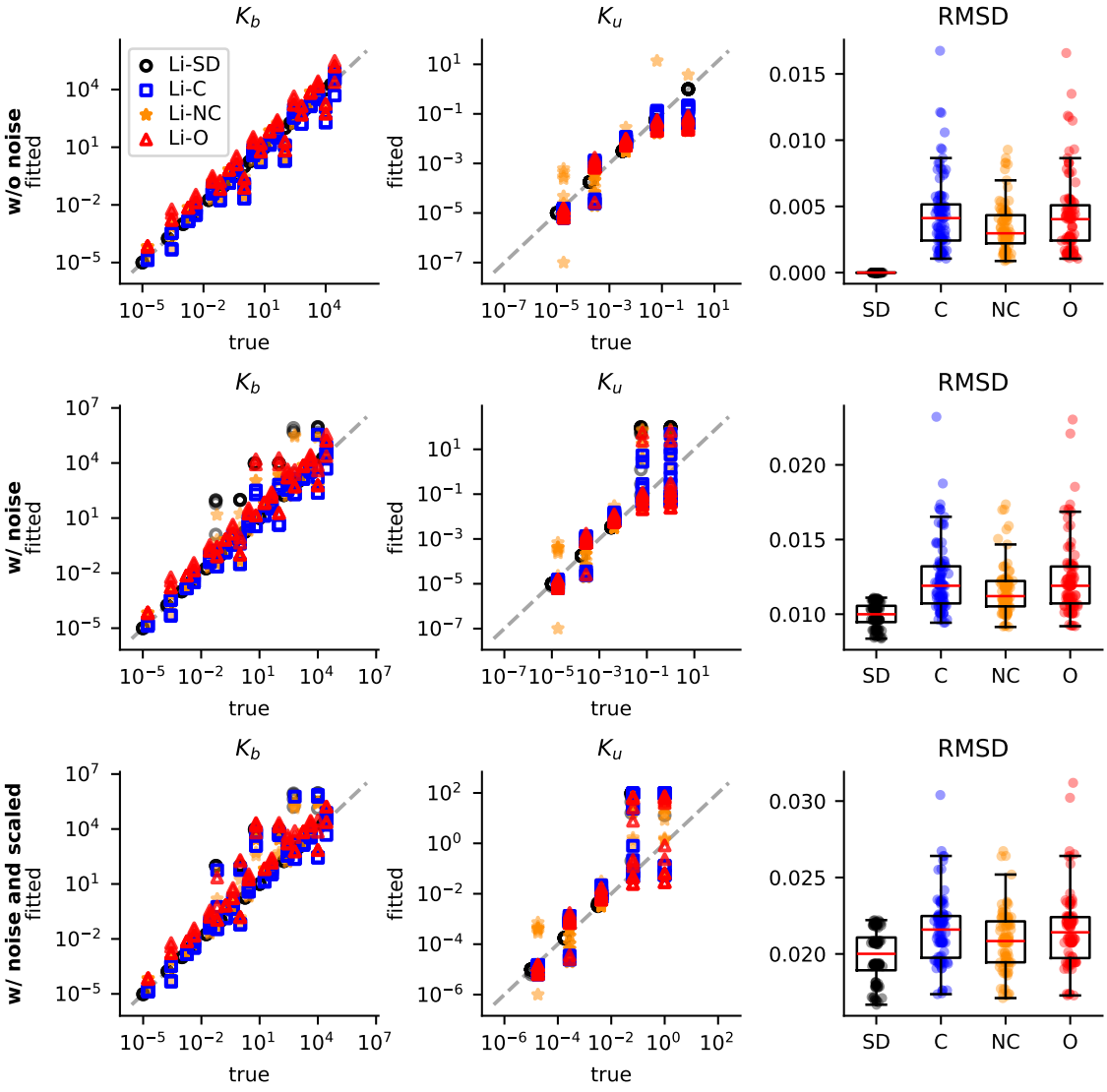


Figure 5.16: **The same binding rate K_b values of the alternative drug-binding formulations can replicate that of the SD formulation, while the unbinding rates K_u are predicted to have different values from the SD formulation.** The left and middle columns show the inferred values of the binding rate K_b and the unbinding rate K_u, respectively, against the true values. The results of the C formulation (blue squares), the NC formulation (orange stars), and the O formulation (red triangles) are shown. The results of the SD formulation (black circles) are the same as in Figure 5.4, and shown here for reference. The diagonal line serves as a reference where the fitted parameters equal the true parameter values. The right-most column gives the RMSDs between the synthetic data and the fitted simulations.

In the noisy conditions (the N2 and the N3 noise models), the alternative drug-binding formulations mostly infer K_b values to be higher than the true K_b, but at values comparable to the inferred K_b of the SD formulation. Even within the SD formulation, where there is no model misspecification, the predicted K_b is not perfect (also discussed in Section 5.3 and shown in Figure 5.11).

The inferred K_u values of the alternative drug-binding formulations do not match well with the K_u of the SD formulation. When no noise is added, the alternative drug-binding formulations (all markers except black circles) mostly infer K_u lower than the K_u of the SD formulation, as expected, except for the smallest K_u, where some of them are overestimated in the NC formulation (orange star). As the noise intensity increases (middle and bottom rows), K_u values seem to be inferred higher than the true values, especially when the true K_u values are large. However, the unbinding rates K_u of the alternative drug-binding formulations are mainly predicted to be lower than the inferred K_u of the SD formulation. Note that the K_u values of the SD formulation are also inferred to be higher than their true values when the true values are large.

The RMSD distribution in the right column shows that, when no noise is added (top row), the simulated output of the SD formulation fits perfectly to the synthetic data, whereas there are some mismatches between the alternative drug-binding formulations and the SD formulation. Among the alternative drug-binding formulations, the NC formulation (orange) can best fit to the synthetic data generated by the SD formulation, while the C formulation (blue) and the O formulation (red) shows similar distribution of the RMSD values. The magnitude of the RMSD increases with the noise intensity. Given the model misspecification in this analysis, the RMSDs of the alternative drug-binding formulations are only slightly higher than the SD formulation. Even though the RMSD between all drug-binding formulations is comparable, the alternative drug-binding formulations present different kinetics from the SD formulation, which are also evident in the experimental setting (Section 5.5.3).

5.5.5 Proarrhythmic risk categorisation of drugs using different drug-binding formulations

Although we have shown in Sections 5.5.2 and 5.5.3 that all drug-binding formulations can recapitulate the current behaviour with reasonable accuracy, they can only do so using very different parameter values. The SD formulation was developed to improve the prediction of the proarrhythmic risk of a drug by including a trapping component (red component of Figure 5.1). This was combined with the introduction of new

metric scores, the qNet and the Torsades Metric Score (TMS) [Dutta et al., 2017, Li et al., 2019]. The qNet computes the resultant ionic charge of several major currents across the channel during an action potential. These currents include the L-type calcium current I_{CaL}, the late sodium current I_{NaL}, I_{Kr}, the slow rectifier potassium current I_{Ks}, the inwardly rectifying potassium current I_{K1} and the transient outward potassium current I_{to}. This simultaneous use of several major currents takes into account the compensatory effect of the currents on the action potential and on the risk profile of the drugs [Dutta et al., 2017, Weiss et al., 2010, Volders et al., 2000]. The TMS is calculated by taking the average of qNet values for drug concentrations of 1–4 $\times C_{\mathrm{max}}$. A drug with a TMS that is less than 0.0579 C/F is classified as a high-risk drug, whereas a drug with a TMS that is greater than 0.0689 C/F is a low-risk drug, with any drug in between categorised as an intermediate-risk drug [Li et al., 2019].

Given that the drug-binding formulations each exhibit different features in the current behaviour (Section 5.5.3), and that the drug-binding formulations propose different drug-binding kinetics for the same current behaviour, we want to investigate whether the drug-binding formulations affect the risk categorisation of drugs. We therefore calculate the TMS of the 12 CiPA training drugs for all drug-binding formulations. The current inhibition on currents other than I_{Kr} is based on the Hill curve coefficients published in Li et al. [2017b]. The current inhibition on I_{Kr} is based on the drug-binding formulations we have, with the parameter values inferred from the experimental data. The parameter values for the drugs are listed in Tables 5.3 to 5.6. The parameters of the SD formulation we used are our inferred parameters, optimised using the procedure described in Section 5.3.3, to ensure that we use the same optimisation procedure throughout all drug-binding formulations for a fair comparison. The conductance scaling (CS) formulation, which we have used extensively in Chapters 3 and 4, is included in this comparison. Briefly, the CS formulation inhibits the current by scaling the maximal conductance following a Hill curve, which effectively reduces the magnitude of the current. The coefficients of the Hill curve are taken from Li et al. [2017b].

Figure 5.17 shows the TMS of the 12 CiPA drugs for all drug-binding formulations. The colour of the data points follows from Dutta et al. [2017], where red, blue, and green represent drugs in the high, intermediate, and low risk categories, respectively. The horizontal lines refer to the threshold values classifying the drugs into different risk categories, taken from Li et al. [2019] for reference. The TMS of the drugs calculated using the SD formulation remain within their risk categories,

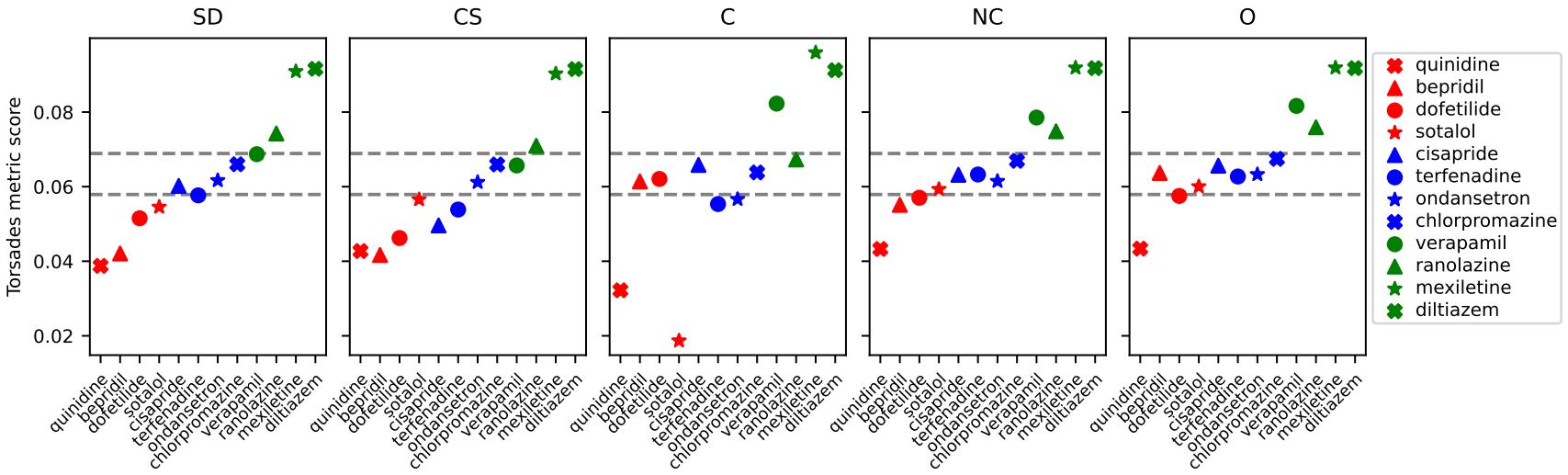


Figure 5.17: **The Torsades Metric Scores of drugs using the NC drug-binding formulation are well-separated into high-, intermediate-, and low-risk categories.** TMS for all drug-binding formulations: SD, conductance scaling (CS), C, NC, and O. The red, blue, and green markers refer to high-, intermediate-, and low-risk drugs, respectively. The horizontal lines indicate the threshold values for the risk categorisation of drugs from Li et al. [2019].

except for terfenadine (blue circle) and verapamil (green circle), whose TMS are close to the thresholds of a higher risk category. This is consistent with the uncertainty range published in Li et al. [2019], where terfenadine (blue circle) has an uncertainty range close to the higher risk category. The TMS of drugs using the SD formulation (Figure 5.17) is not the same as presented by Li et al. [2019] because the parameter values we used are re-inferred with our own optimisation procedure (Section 5.3.3). This is to ensure a fair comparison between all drug-binding formulations by using the same optimisation procedure throughout. The correlation between the risk of drugs and the TMS remains; that is, drugs can be categorised into their risk categories by their TMS with recalibrated threshold values. This implies that the threshold values categorising the risk of drugs are sensitive to the parameter values in the SD formulation.

The separation of drugs by their TMS using the CS formulation is not as clear as that of the SD formulation. The TMS for low-risk verapamil (green circle) is comparable to that of intermediate-risk chlorpromazine (blue cross). Moreover, the TMS of high-risk sotalol (red star) lies between the intermediate-risk drugs (blues), a higher score than the intermediate-risk cisapride (blue triangle) and terfenadine (blue circle). Therefore, using the TMS, the CS formulation does not give the correct risk categorisation for the drugs.

We have shown in Figures 5.13 and 5.14 that the C formulation performs the worst in capturing the decreasing curve within each pulse of the experimental data. This is because the transition through the cycle of the C formulation, i.e. from the inactivated (IO) state to the drug-bound state (D) and then to the open state (O) (refer to Figure 5.9A for the model structure), and the reverse route, happens much more frequently and easily in the C formulation than in the SD formulation. Recall that the trapping rate of the SD formulation is much slower than the binding/unbinding rates, with a difference of up to eight orders of magnitude (Section 5.4.1), thus the transition through the cycle (IO to IO* to C* to O* and to O) (refer to Figure 5.2 for the structure of the SD formulation) happens very slowly. In the C formulation, the channel transits through the drug-bound state more frequently. Drug compounds bind and unbind quickly and do not stay bound within the channel. Therefore, the C formulation lacks the ability to delay the unbinding process to mimic the drug-trapping behaviour. The C formulation may fail to capture the behaviour of I_{Kr} under drug inhibition. This will then be propagated to the TMS. In Figure 5.17, for the C formulation, the TMS of the low-risk drugs (greens) are separated from

the intermediate-risk drugs (blues), but that is not the case between intermediate-risk drugs (blues) and high-risk drugs (red). The TMS of low-risk ranolazine (green triangle) is close to the intermediate-risk cisapride (blue triangle) and lies within the intermediate-risk threshold. The TMS of high-risk drugs bepridil (red triangle) and dofetilide (red circle) blend in with the intermediate-risk drugs (blues), making it impossible to differentiate high-risk drugs from intermediate-risk drugs using the TMS.

The NC formulation performs consistently better than the C formulation in its fit to the experimental data, while performing better than the SD formulation for some drugs (Figures 5.13 & 5.14), including cisapride, ranolazine, quinidine, ondansetron, diltiazem and mexiletine. From that, we hypothesised that the NC formulation can differentiate the drugs into their risk categories, showing an association between the risk categories and the TMS, similar to the SD formulation. Figure 5.17 shows that the TMS using the NC formulation are clearly separated between the low-risk and intermediate-risk drugs, following the threshold values defined by Li et al. [2019] using the SD formulation. The high-risk drugs (reds) are also separated from the intermediate-risk drugs (blues), although with a smaller margin. Although high-risk sotalol (red star) falls into the intermediate-risk category defined by Li et al. [2019], a recalibration of the threshold values for the NC formulation would accurately classify all drugs into their respective risk categories.

Although the NC and the SD formulations have different TMS, they both can classify the drugs into their respective risk categories. Among all drugs with similar TMS between the NC and the SD formulation, such as diltiazem (green cross) and ranolazine (green triangle), the RMSDs between the simulated current and the experimental data of the SD and the NC formulations are similar (Figure 5.14), but the simulated current exhibits different properties, as evident in the fractional block shown in Figures 5.12 and 5.13 and their respective binding kinetics shown in Figure 5.14. This implies that the risk categorisation by the TMS can be replicated with different binding kinetics and without the trapping component.

While the O formulation struggles to fit to the decreasing curve of drug inhibition, similar to the C formulation (Figures 5.12 & 5.13), the TMS of all drugs computed using the O formulation are ordered based on their risk categories, except for the high-risk bepridil (red triangle). The TMS for high-risk bepridil (red triangle) is slightly higher than the intermediate-risk drugs terfenadine (blue circle) and ondansetron (blue star). The O formulation can be viewed as the NC formulation without the binding of drugs to the inactivated-open (IO) state (Figure 5.10). The idea for the

O formulation stems from the fact that the current recordings from patch-clamp experiments are measurable when the ion channel is open. It is a challenge to design protocols that extract information or allow elucidation of the kinetics surrounding the inactivated-open state. This is apparent in the identifiability analysis where the unbinding rates of the NC formulation are not identifiable (Figure 5.11).

We have shown in Figure 5.15 that the alternative drug-binding formulations are capable of exhibiting the full spectrum of the trapping behaviour. However, for a trapped drug bepridil, the C and the O formulations fail to match the increasing trend of peak reduction with drug concentration. Bepridil (red triangle in Figure 5.17) is also the only drug that lies outside the association between the TMS and its risk category. Since the O formulation is capable of capturing the peak reduction but fails to do so for bepridil, we next explore whether, following the approach of Li et al. [2017b], incorporating the contribution of the trapping phenotype, quantified by the peak reduction, into the optimisation procedure, will allow us to test if the binding kinetics is capable of capturing the drug-trapping properties.

5.5.6 Inclusion of trapping behaviour into the objective function

As introduced in Section 5.1, the trapping behaviour of drugs can be quantified by the reduction of the peak of current between pulses of the modified Milnes' protocol [Li et al., 2017b]. In Li et al. [2017b], the objective function used to infer the parameter for the SD formulation includes two terms quantifying the trapping phenotype. These terms are the relative peak reduction between the first and the second pulse and the relative peak reduction between the first and the tenth pulse. These terms are added to the objective function only when 0.5 of the fractional block is achieved for a given dose.

We added these two terms to the objective function in our optimisation procedure, in addition to the traces, following the approach taken by Li et al. [2017b], as follows

$$\sum_{c}\left[\sum_{t}\left(Y_{t,c}-B_{t,c}\right)^{2}+0.2w_{c}T\frac{P_{1,c}-P_{10,c}}{P_{1,c}}+0.2w_{c}T\frac{P_{1,c}-P_{2,c}}{P_{1,c}}\right], \tag{5.42}$$

where the weight $w_c = 1$ if $B_{t,c} < 0.5$ for some t for a given dose c, else $w_c = 0$. $Y_{t,c}$ is the simulated fractional block, $B_{t,c}$ is the experimental fractional block, $P_{n,c}$ is the peak at the nth pulse for drug concentration c, and T is the total time points. Here, sum of squared error is used to compare the traces, following Li et al. [2017b], instead of the root mean squared error used throughout this chapter. Figure 5.18

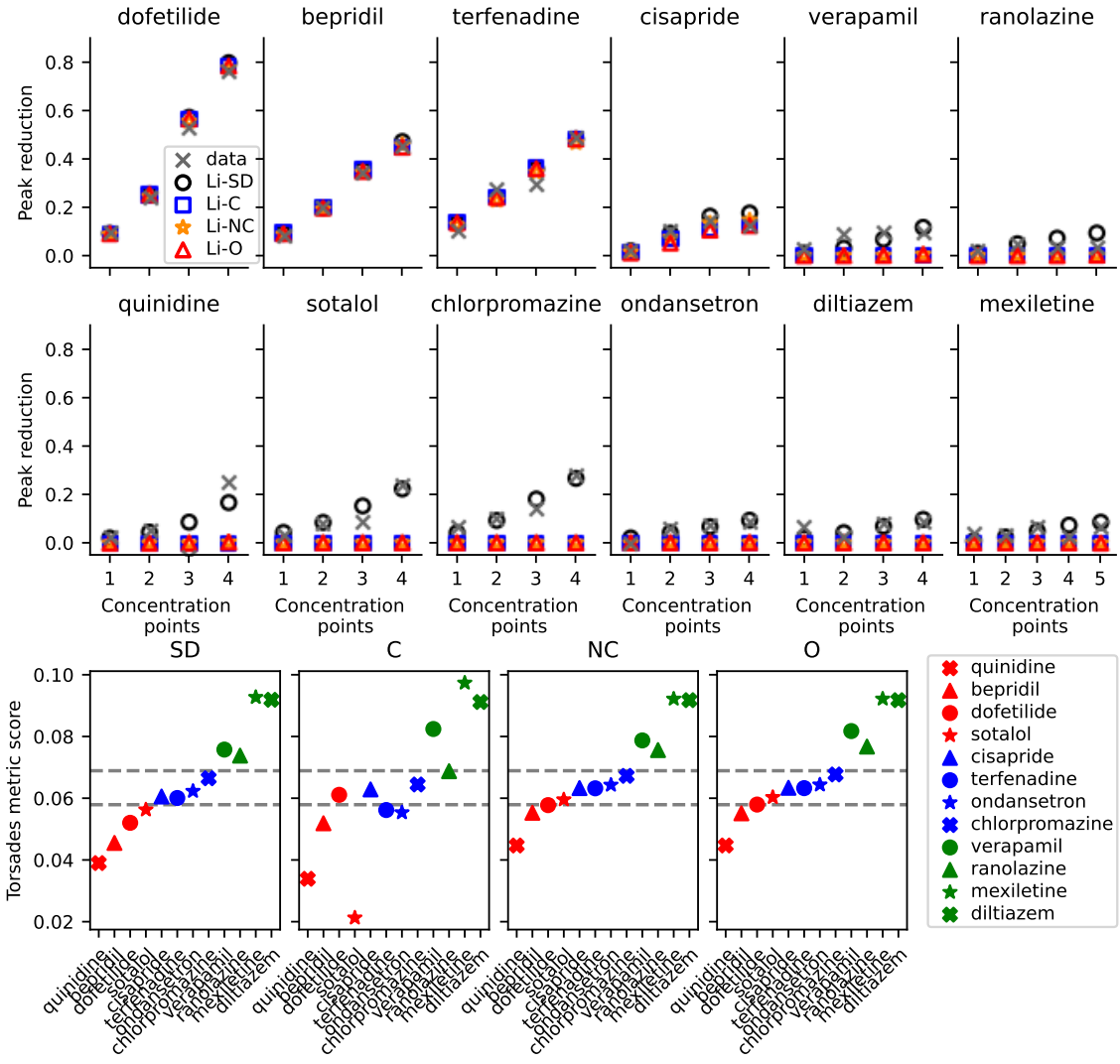


Figure 5.18: **The drug-binding formulations can fit to the peak reduction of trapped drugs (dofetilide, bepridil and terfenadine) when the trapping phenotype is included in the objective function. The SD, the NC and the O formulations can separate the drugs into their risk categories.** Top panels (top and middle rows) show the peak reduction of the experimental data and all drug-binding formulations. The y-axis is the relative peak reduction between the first and the tenth pulse for each drug concentration. Two additional terms which quantify the peak reduction are included in the objective function when inferring the parameters for the drugs. The last row shows the TMS for all drug-binding formulations: SD, C, NC and O formulations. The TMS is computed with parameters that are optimised to match the peak reduction, in addition to traces. The red, blue and green markers refer to high-, intermediate- and low-risk drugs, respectively. The horizontal lines indicate the threshold values for the risk categorisation of drugs from Li et al. [2019].

(first two rows) shows the relative peak reduction between the first and the tenth pulse of the experimental data and the drug-binding formulations when the peak reduction terms are added to the objective function. Compared to the peak reduction when only traces were included (Figure 5.15), the inclusion of the peak reduction in the objective function improves the fit of the peak reduction (quantification of the trapping phenotype) of bepridil for the C and the O formulation, and cisapride for the SD, the C and the O formulation. The SD formulation has improved fit to the peak reduction for most of the drugs, such as verapamil, ondansetron and diltiazem (refer to Figure 5.15 for comparison). The peak reduction becomes more consistent as the drug concentration increases, matching the experimental data. Other than the drugs mentioned, the peak reduction of the alternative drug-binding formulations are unaffected by the adjusted objective function. With this adjustment to the objective function, there is a clear separation in the trend of the peak reduction between known trapped drugs (dofetilide, bepridil and terfenadine) and known nontrapped drugs (cisapride, verapamil and ranolazine). The peak reduction of trapped drugs increases with drug concentration, while the peak reduction stays relatively constant for nontrapped drugs as the drug concentration increases.

When the terms quantifying the peak reduction were included in the objective function, the parameters of the drug-binding formulations were optimised to different parameter values. The TMS will therefore also change. The bottom row of Figure 5.18 shows the TMS calculated using the drug-binding formulations with parameters optimised based on the adjusted objective function. With the objective function following closely to Li et al. [2017b], the TMS of the SD formulation once again falls into the risk categorisation threshold from Li et al. [2019], replicating their results. The C formulation has some minor changes in the TMS, specifically bepridil (red triangle), from being among the intermediate-risk drugs to being below them (see Figure 5.17 for the TMS without the peak reduction). The TMS calculated using the C formulation still does not correlate with the risk categories of drugs and thus cannot be used as a marker to predict a drug's risk category. For the NC formulation, the TMS correlates with the drug's risk categories, with a similar sequence as the SD formulation. A recalibration of the threshold values by a regression analysis, as is done in Chang et al. [2017], would have all drugs classified into the same risk categories as those assigned to them in Colatsky et al. [2016], but that is out of scope for this study.

The TMS of the O formulation has a major change for bepridil (red triangle). Its TMS is much lower than the previous one (Figure 5.17) and falls among the high-risk drugs after the peak reduction is incorporated into the objective function. Another

difference for the O formulation is cisapride, where its TMS lowers, but still remain within the intermediate-risk drugs threshold. When the peak reduction is included into the objective function, using the O formulation, the drugs can be separated into the same risk categories using the TMS as those computed using the SD formulation.

In conlusion, the O drug-binding formulation, with interpretable parameters (i.e., the binding and unbinding rates), can classify the CiPA training drugs accurately when the trapping phenotype is taken into consideration in the characterisation of the drugs. The SD formulation, with its more complex and non-identifiable trapping component, is therefore not necessary to classify drugs using the TMS.

5.6 Discussion

This has been a complex chapter in which we have introduced multiple studies to allow us to fully explore the different aspects of the SD formulation. In this discussion section, we bring together all of our findings and conclusions, repeating in places to make it easy to follow.

In previous chapters (Chapters 3 & 4), we showed that the trapping component of the SD formulation has minimal effect on the changes in the APD_{90}s. In this chapter, we showed that this is due to the non-identifiability of the trapping tendency parameter $V_{\text{half}-\text{trap}}$ (see Figure 5.1 and Section 5.2 for description of the $V_{\text{half}-\text{trap}}$) (Section 5.3). Adjustments in the defining equation of the trapping component (Equation 5.2) did not improve the identifiability of $V_{\text{half}-\text{trap}}$ (Sections 5.4.1 & 5.4.2). Therefore, we introduced three alternative drug-binding formulations which do not include the unidentifiable trapping component (Section 5.5). The C formulation combines all drug-bound states of the SD formulation into a single drug-bound state; the NC formulation removes the trapping component and is equivalent to having zero trapping/untrapping rates of the SD formulation; and the O formulation only retains a single drug-bound state that connects to the open state (Figure 5.9). We have shown that the C formulation and the O formulation give improved parameter identifiability compared to the SD formulation, and the NC formulation has poor identifiability for both the unbinding rates, K_u and K_r (Section 5.5.2). For the same current behaviour, we have shown that the alternative drug-binding formulations have lower unbinding rates than the SD formulation, which we have speculated compensates for the missing trapping component (Sections 5.5.3 & 5.5.4). All drug-binding formulations had their limitations in capturing the current behaviour: the SD formulation could not achieve an instantaneous steady value at the first two pulses, the C

and the O formulations failed to replicate the decreasing curve within each pulse, and the NC formulation could not capture the decreasing curve and replicate the in-pulse steady value simultaneously (Section 5.5.3). Although none of the drug-binding formulations could fully capture the full spectrum of current behaviour observed in the experimental data, all exhibited a range of recovery rate of the current, indicating that they can express different trapping behaviours. Moreover, the Torsades Metric Score (TMS) calculated using the NC formulation classified the drugs into their respective risk categories, showing that the trapping component is not necessary for obtaining the correct classification (Section 5.5.5). The addition of the trapping phenotype into the objective function improved the classification of drugs for the O formulation (Section 5.5.6), which has only two parameters that are both interpretable and practically identifiable (Sections 5.5 & 5.5.2). In conclusion, the trapping component of the SD formulation was not necessary to categorise the drugs accurately. A drug-binding formulation with only binding/unbinding rates, and inclusion of the trapping phenotype in the objective function, can replicate the risk classification of drugs, as with the SD formulation.

5.6.1 Reasoning for the alternative drug-binding formulations

The trapping component of the SD formulation is neither measurable nor inferable from the fractional block measured from the current-conducting state of the channel. We removed the trapping component in two ways and suggested three alternative drug-binding formulations. The idea for the construction of the C formulation (refer to Figure 5.9 for the structure of the C formulation) was to summarise all drug-bound states and transition rates into one, retaining only the binding kinetics instead of separating them into two components of binding and trapping. This formulation reduced the number of parameters from three in the SD formulation (ignoring the drug concentration parameters of n and EC_{50}) to two, namely the binding and unbinding rates. However, these two parameters were unable to capture the full range of current behaviour observed in the experimental data. The binding rate controls the speed of the decrease within each pulse, while the unbinding rate affects the minimum fractional block within each pulse, and simultaneously the rate of recovery in the subsequent pulse. Therefore, at least three parameters are required to individually capture the speed of decrease, the within-pulse steady-state, and the rate of recovery. Moreover, instead of being retained within the trapped state as expected, drug compounds remained within the O* state and the IO* state to mimic the stay within the

trapped state and seldom transited through the cycle of the SD formulation (refer to Figure 5.1 for the structure of the SD formulation). When we prevented drugs from entering the trapped state by fixing the trapping rate to zero, the fractional block still showed the decreasing curve similar to those observed in the experimental data. Therefore, we constructed the NC formulation to have zero trapping rate and without the trapping component (refer to Figure 5.9 for the structure of the NC formulation).

Since there is no cycle in the structure of the NC formulation, we were free to define the transition rates without the need to satisfy the microscopic reversibility. No constraints were imposed on the unbinding rates ($K_{r,NC}$ and $K_{u,NC}$) of the NC formulation. The NC formulation has three parameters, which is the same as the SD formulation, which are $K_{b,NC}$, $K_{u,NC}$, $K_{r,NC}$ for the NC formulation and $K_{b,SD}$, $K_{u,SD}$, $V_{\text{half}-\text{trap}}$ for the SD formulation. It was therefore not surprising that the NC formulation with more parameters has poorer parameter identifiability than the C formulation with two parameters. The SD formulation has two additional constraints: microscopic reversibility and the fixed trapping rate K_t. The definition of the trapping component of the SD formulation also added another layer of complexity to the drug-binding formulation, namely the choice to define the trapping component as voltage-dependent, the decision to define it in a similar fashion to the activation gate of the hERG channel, and the choice to fix the trapping rate (Equation 5.2). The complexity of the definition made it difficult to interpret the parameters in physiological terms and verify them in any way or constrain them into feasible and reasonable ranges. One thing to note is that by imposing these constraints, Li et al. [2017b] assumed that all drugs conform to the constraints and that these constraints are intrinsic properties of the ion channel. Li et al. [2017b] assumed a fixed trapping rate K_t for all drugs. This trapping rate K_t was chosen for an optimised fit for this particular dataset. A different value would likely be needed for a different dataset. Despite the additional complexity, the SD formulation has better parameter identifiability than the NC formulation. It is speculated that the constraints imposed on the SD formulation allow little room for compensation between parameters and thus have better parameter identifiability. Moreover, the poor parameter identifiability of the NC formulation could be due to the lack of information in the current recordings to elucidate the transition rates.

The current recordings measured in a patch-clamp experiment are indicative of open ion channels. Given that the experimental recording do not provide direct information on the inactivated state, we remove the binding to the inactivated state from the NC formulation to create the O formulation (see Figure 5.9 for the structure

of both drug-binding formulations), even though several studies have suggested higher binding affinity associated with the inactivation of the channel [Ficker et al., 1998, Perrin et al., 2008, Mitcheson et al., 2000a, Mitcheson, 2008, Vandenberg et al., 2012]. The O formulation allows drugs to bind and unbind only from the open channel. Developing a drug-binding formulation that incorporates the real-world condition but without informative data to elucidate would only allow compensation between parameters and obtain the wrong characterisation of the drug-binding mechanism. We can further suggest other drug-binding formulations based on our observations, such as extending the drug-bound state of the D formulation with an additional state that transits into and out of the drug-bound state D.

5.6.2 Implications and future work

Our work has shown that the trapping phenotype has to be included in the objective function when inferring the parameters of the O formulation to accurately categorise the drugs using the TMS. The parameters of the O formulation have better interpretability and identifiability. This implies that the binding kinetics, without the trapping kinetics as defined in the SD formulation, are sufficient to capture the drug's property that plays a role in the risk prediction. A follow-up is to compute the TMS of the CiPA validation drugs [Li et al., 2019] using the O formulation to further validate our conclusion. The improved practical identifiability of the O formulation may make it easier to measure or infer the parameter values from experimental data and improve the reliability of its risk prediction.

While there is no way to determine the ground truth of these parameter values, the parameter values that incorporate the trapping phenotype classify the proarrhythmic risk of drugs better. This suggests the need to have experimental data recorded over a few pulses (to quantify the peak reduction) if we are to design alternative shorter protocols. Ten pulses of the modified Milnes' protocol used for the development of the SD formulation are extremely long (lasts for 250 s in total). With a simpler drug-binding formulation, with fewer parameters, it is possible to design protocols that are easier to perform while also serving our purpose in predicting the risk of drugs.

5.7 Conclusion

We conclude that the trapping component of the SD formulation is not necessary for classifying drugs into their respective risk category. In previous chapters, we showed that the trapping component of the SD formulation plays a minor role in affecting

the action potential duration. We showed in this chapter that the restriction is due to the non-identifiability of the trapping component defined within the SD formulation. This issue cannot be improved by loosening the constraints or extending the range of values considered for the model variables, particularly the membrane voltage V in Equation 5.2. We then developed three alternative drug-binding formulations by modifying the model structure and its defining equations to remove the trapping component. We have a drug-binding formulation that summarises all drug-bound states in the SD formulation into a single drug-bound state (C formulation), the other one that eliminates the trapping component (NC formulation) and finally a drug-binding formulation that keeps only the binding kinetics at the open state (O formulation). The C and the O formulations have better identifiability than the SD formulation, but do not have sufficient degrees of freedom to capture the drug-binding from the experimental data. The NC formulation has better interpretability and less constraints than the SD formulation, although the parameters are not as identifiable as those of the SD formulation. The NC formulation has comparable fit to the experimental data as the SD formulation and can classify drugs with similar accuracy as the SD formulation. The analysis of the O formulation showed that the inclusion of only the binding kinetics of drugs (without the trapping behaviour of drugs) within the drug-binding formulation and the incorporation of the peak reduction (quantification of the trapping behaviour of drugs) into the objective function are sufficient to classify drugs into their risk categories.

Chapter 6

Conclusion and discussion

In this conclusion chapter, we summarise the main results presented in this thesis and the contributions of these results to the aim of the thesis in assessing the credibility of the drug-binding formulation.

Computational models of cardiac cells have been utilised to establish a comprehensive proarrhythmic risk assessment for translational drug safety evaluation [Brennan et al., 2009]. The most recent drug safety guidelines are beginning to include risk predictions of computational models as part of a more comprehensive drug risk assessment [ICH E14/S7B Q&A, 2022]. Therefore, it is of importance for computational models, also as part of the CiPA initiative, to provide robust and reliable risk predictions of drugs for the benefit of patients while also preventing unpredicted and nonjustifiable proarrhythmic risk. A credible computational model should be able to make predictions with high confidence while also acknowledging the limits and uncertainties surrounding the model predictions.

In this thesis, we investigated and assessed the predictive reliability of the state-dependent (SD) drug-binding formulation [Li et al., 2017b] in predicting the proarrhythmic risk of drugs. The SD drug-binding formulation was developed by research teams in the FDA as part of the CiPA initiative to provide a proarrhythmic risk prediction model. In Chapter 2, we introduced the concepts required to understand the thesis, which includes the physiology of the heart, the mathematical models describing the action potentials, the translation of experimental data to mathematical models, and the relevance of these models in cardiac safety testing, with a specific focus on models of drug-binding to the hERG channel and the assessment of the credibility of these models.

In Chapter 3, we assessed the robustness and reliability of the SD formulation by performing a sensitivity analysis. We investigated the contribution of each component

(the binding component and the trapping component) of the SD drug-binding formulation to the prolongation of the action potential (AP). The AP prolongation of the SD formulation was compared to that of the conductance-scaling (CS) drug-binding formulation. We concluded that the SD formulation differs from the CS formulation in predicting the AP duration; however, the main driver of the difference is not the trapping component of the SD formulation. This runs counter to the claim that the trapping behaviour contributes to the risk of drugs, or at least shows that the trapping component defined by Li et al. [2017b] does not contribute to AP prolongation in this particular context.

In Chapter 4, we further consider the choice of model components on the reliability of the predictions of the SD formulation. To investigate the influence of different AP models in the prediction of AP prolongation, we embedded the SD formulation into several combinations of AP models and hERG-current components and compared their predicted AP prolongation. The predicted AP prolongation differs significantly among the AP model. However, the qualitative result remains the same. The trapping component of the SD formulation does not drive the difference between the SD formulation and the CS formulation. This is in agreement with our previous result in Chapter 3. The trapping component of the SD formulation contributes little to the AP and possibly the hERG current as well, regardless of the AP model.

In Chapter 5, we investigated the practical identifiability of the SD formulation. We showed that the $V_{\text{half}-\text{trap}}$ parameter of the trapping component is unidentifiable, but it remains correlated to the proarrhythmic risk prediction of the drug. Adjustments to the SD formulation, which include changing the fixed trapping rate and using multiple different voltage-clamp protocols, did not improve the identifiability of $V_{\text{half}-\text{trap}}$.

The practical non-identifiability of $V_{\text{half}-\text{trap}}$ led us to simplify the SD formulation by removing the trapping component. We suggested three alternative drug-binding formulations of the SD formulation and compared them to the SD formulation that has the trapping component. Not surprisingly, the drug-binding formulations with fewer parameters have better parameter identifiability but lack the degrees of freedom to capture the current behaviour fully. We showed that the trapping component is not necessary when categorising the proarrhythmic risk of drugs using the Torsades Metric Score (TMS). Moreover, the binding kinetics, without the trapping kinetics, is sufficient to classify the proarrhythmic risk of drugs accurately. This may provide some insight into how the trapping component should be defined or how the binding component can be updated to encapsulate the trapping behaviour of drugs.

On the whole, this thesis demonstrates the importance of conducting a detailed assessment of the credibility of a computational model. The SD formulation, by incorporating drug-trapping, works well in categorising drugs according to the TMS [Chang et al., 2017]; however, it does not imply a direct association between drug-trapping and proarrhythmic risk. The defined trapping component does not result in the expected prolongation of the action potential, regardless of the hERG-channel component and the AP model, implying that a drug-binding formulation without the trapping component might work just as well. The introduction of redundant components into the drug-binding formulation may lead to overfitting and wrongly link the trapping component to the risk prediction of the SD formulation. We further verified that the trapping component is not necessary for classifying the risk of drugs, as a reduced model without the trapping component can separate the drugs into their risk categories as well as the SD formulation.

6.1 Future work

There are other aspects to consider, which are not covered in the thesis, when developing computational models for robust and reliable predictions. A straightforward next step will be to complete the model assessment of the O formulation, similar to those performed on the SD formulation [Li et al., 2019, Chang et al., 2017]. This includes a regression analysis to recalibrate the TMS threshold for the O formulation, verification of the risk prediction of the O formulation to the CiPA validation drugs, and an uncertainty quantification to characterise the confidence interval of the risk categorisation of the O formulation. For the SD formulation, the TMS of the CiPA validation drugs did not fall within the pre-defined threshold calibrated from the training drugs [Li et al., 2019]. The authors argue that a ranking score, instead of a strict cut-off threshold, may lead to better performance due to the relatively small training set [Li et al., 2019]. The verification of the O formulation with CiPA validation drugs will show if the TMS threshold calibrated using the CiPA training drugs will also hold for the validation drugs. Moreover, the uncertainty quantification may further inform whether the binding kinetics, without drug-trapping, are sufficient for the risk prediction of drugs.

Extending from the O formulation, the simulation studies performed in this thesis can be validated using appropriate experimental data. Validation of the simulation outputs is an important step to establish the credibility of computational models.

Throughout this thesis, several comparisons between computational models were performed at various scales. In Chapters 3 and 5, several drug-binding formulations were compared, while in Chapter 4, the AP models were compared in the context of the drug-binding formulations. A validation of these results may inform the suitability of these computational models in proarrhythmic risk prediction or other contexts-of-use. This can be done by comparing the simulation results with experimental data, such as AP prolongation under drug inhibition, the reverse use dependency of drugs, or the risk prediction of drugs not within the training set. To perform validation against experimental data, the computational models have to first be calibrated to reflect the properties of an AP. One example is the calibration of an AP model performed by Tomek et al. [2019]. This thesis focuses on the role of the trapping component on the AP, not the exact AP duration. Therefore, such detailed calibration was not performed. We hypothesise that it will not affect our conclusion that the AP prolongation does not depend on the trapping component.

Another avenue for future work is to obtain additional experimental data to provide insights into drug-trapping or the drug-binding process, or to validate the computational models. The additional experimental datasets could be experimental repeats performed using voltage-clamp protocols other than the modified Milnes' protocol [Li et al., 2017b]. The voltage-clamp protocols could be designed to provide insights into developing a new mathematical formulation of drug-trapping or to validate a given mathematical formulation of drug-trapping. Although it was shown in Chapter 5 that the use of other voltage-clamp protocols did not improve the parameter identifiability of the SD formulation, the protocols were not designed to specifically improve the parameter identifiability of the SD formulation. Protocols designed based on the SD formulation may improve its identifiability. Protocols could also be optimised for the O formulation. The simpler O formulation with only two parameters may allow shorter and easier-to-implement experiments to elucidate its binding kinetics. The experimental procedure used for the SD formulation is difficult to perform [Kramer et al., 2020]. The modified Milnes' protocol, which lasts for 25 s, is repeated 10 times to capture the trapping behaviour. It was challenging to maintain signal stability throughout this long period [Kramer et al., 2020]. High-quality signal recordings are required, given the sensitivity of the parameters to noise. Although the experimental data used for the SD formulation [Li et al., 2017b] was performed with manual patch-clamp, which generally produces high-quality recordings compared to high-throughput patch clamp experiments, the inconsistencies within the data, as discussed in Appendix B, may affect the risk prediction outcome. When designing an optimal experiment, the

practicality and feasibility of the data acquisition process and the ability to identify parameters from the data should be taken into account. Our conclusion in Chapter 5 showed that the trapping behaviour has to be well-characterised when inferring the parameters, which may require protocols to be repeated over several pulses. This prior knowledge may be utilised to constrain the experimental design space.

All of the alternative drug-binding formulations, introduced in Chapter 5, are proposed based on our observations from the evaluation of the SD formulation. The alternative drug-binding formulations each have their own strengths and weaknesses. While the O formulation has good parameter identifiability and can categorise the risk of drugs accurately, it has a poor fit to the experimental data. It struggles to capture the delay in current inhibition (the decreasing curve feature observed in the calibration data for the SD formulation [Li et al., 2017b], mentioned in Section 5.5.3). To avoid this drawback, we could consider and assess other possibilities for the drug-binding formulation. For example, we can add a trapped state to the drug-bound (D) state of the C formulation (see Figure 5.10 for the structure of the C formulation), where drugs can bind to the open state or the inactivated-open state and transition to this trapped state. The proposed structure is shown in blue on the left of Figure 6.1, with the additional state shown in red. This is to mimic the delay of staying in the drug-bound states in the SD formulation, and slow down the cyclic transition between the open state, the inactivated-open state, and the drug-bound state, which is observed in the C formulation but not the SD formulation (Section 5.6.1). Another example will be to add a trapped state to the drug-bound (O*) state of the O formulation to allow for the delayed current inhibition observed in the experimental data (Figure 6.1 on the right). However, the addition of drug-bound states to the drug-binding formulation may lead to identifiability issues. Additional experimental data may be required to differentiate between the drug-binding formulations to select the most suitable model for a given purpose [Patten-Elliott et al., 2025].

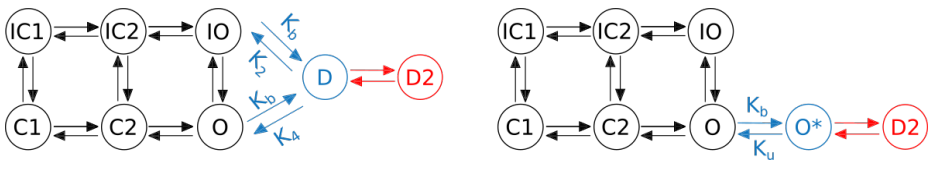


Figure 6.1: **Markov structure of the proposed drug-binding formulations in blue and the additional proposed component in red.**

6.2 Concluding Remarks

The results of this thesis lead to larger questions in computational modelling of biological systems. It is impossible to develop a comprehensive computational model that fully captures a biological system. Therefore, considerations should be given to the level of detail to be included in the computational model, which is the main focus of this thesis. It is based on the context of use (or a list of contexts of use) for the model, and an understanding of the data available. On the one hand, the additional detail may provide a more accurate prediction of the reference data set, but on the other, it could be an effect of overfitting. The effect of the additional component may not be detectable from the experimental data, or it may have a negligible effect on the model outcome. If the computational models are used in safety-critical situations, it becomes more important to establish the credibility of the model. The improvement in the robustness and accuracy of the prediction provided by the additional detail has to be justifiable for the resources required to elucidate or measure the additional complexity.

This thesis has shown that the trapping component does not contribute to the robustness and accuracy of the prediction, and thus, the inclusion of the trapping component in the drug-binding formulation cannot be justified. However, it is not straightforward to assess whether the inclusion of a particular additional component in a complex mathematical model is justifiable. The main challenge is to capture a particular experimental observation within a mathematical formulation that reflects the underlying biological mechanism. A further challenge is then to devise experimental protocols that allow this mathematical formulation to be parameterised. The focus of this thesis has been on one such mathematical formulation of Li et al. [2017b], and we have demonstrated just how complex this modelling process can be. Further factors that must be taken into consideration in this process are: the choice of base model, the parameter sensitivity, and parameter identifiability. The forensic analysis of a particular computational model performed in this thesis demonstrates the importance of model evaluation within this model development process. Such a model evaluation acts as an essential foundation in assessing the credibility of model predictions. This approach should find wider application in mathematical and computational modelling of biological systems, and is particularly important when assessing models that are used in safety-critical decision making.

Moreover, our work has highlighted the importance of fully considering the intended "context-of-use" in this model assessment process. In our case, when com-

paring the effects of varying the base AP model, the ion channel model, and the drug-binding model, on proarrhythmic risk prediction of drugs, we found that all combinations had an influence on the particular measured output (in this case, APD), but did not change our overall conclusions. This may not be the case in other situations, and in mature areas of systems biology, which greatly complicates the task of the modeller. We hope that the approach taken in this thesis will provide a blueprint for other researchers undertaking such complex modelling projects.

Appendix A

Software availability and details

For reproducibility and replicability of the research work in this thesis, the codes are available on GitHub repositories:

1. Chapter 3: https://github.com/FarmHJ/importance-of-binding-mechanism
2. Chapter 4: https://github.com/FarmHJ/binding-mechanism-extend
3. Chapter 5: https://github.com/FarmHJ/SD-reduction

All code is written in Python language. Simulations of the cardiac models were performed using Myokit 1.35.4 or later versions [Clerx et al., 2016]. The numerical solver used within Myokit to solve for the systems of ODE is SUNDIALS [Hindmarsh et al., 2005]. The cardiac models were taken from CellML repository [Lloyd et al., 2008], and modified accordingly. The optimisations were performed using the PINTS software package [Clerx et al., 2019]. High performance computing provided by University of Oxford Advanced Research Computing (ARC) facility was used to perform the computations in Chapter 4 and Chapter 5. https://doi.org/10.5281/zenodo.22558

Appendix B

Data inconsistencies

In Section 5.5.3, we compared the performance of the alternative drug-binding formulations in terms of their fit to the experimental data. In the Figures 5.12 and 5.13, it was observed that each drug-binding formulation has their own properties and limitations. As mentioned in Section 5.5.3, the SD formulation cannot achieve instantaneous steady state from the first pulse; the C formulation cannot capture the decreasing curve unless it is relatively linear; the NC formulation struggles to replicate the decreasing curve and the in-pulse steady state simultaneously; the O formulation also fails to capture the decreasing curve, similar to the C formulation. However, there were instances where none of the drug-binding formulations were able to capture the experimental data, for example cisapride at 100 nM and chlorpromazine at 10^3 nM (Figures 5.12 & 5.13). A similar situation occurred for terfenadine at 30 nM. All drug-binding formulations underestimated the fractional block from the second pulse onwards. All drug-binding formulations also overestimated the inhibition by mexiletine at 10^5 nM.

These could be due to inconsistencies within the experimental data. For example, the fractional block of chlorpromazine at 10^3 nM reached a minimum and then slowly increased within each pulse, but the fractional block was expected to decrease throughout a single pulse. Moreover, the poor fit of terfenadine and cisapride for one single drug concentration may imply that the experimental data is inconsistent with the expression $\frac{D^n}{D^n+\mathrm{EC}_{50}^n}$ (Equation 5.1). We speculated that these inconsistencies of the experimental data arise from noise and variability between repeats and is amplified by the pre-processing.

The experimental data published by Li et al. [2017b] was fractional block, where the drug-inhibited current was normalised against the current under control conditions, not the raw recordings of current. The traces were then averaged over all

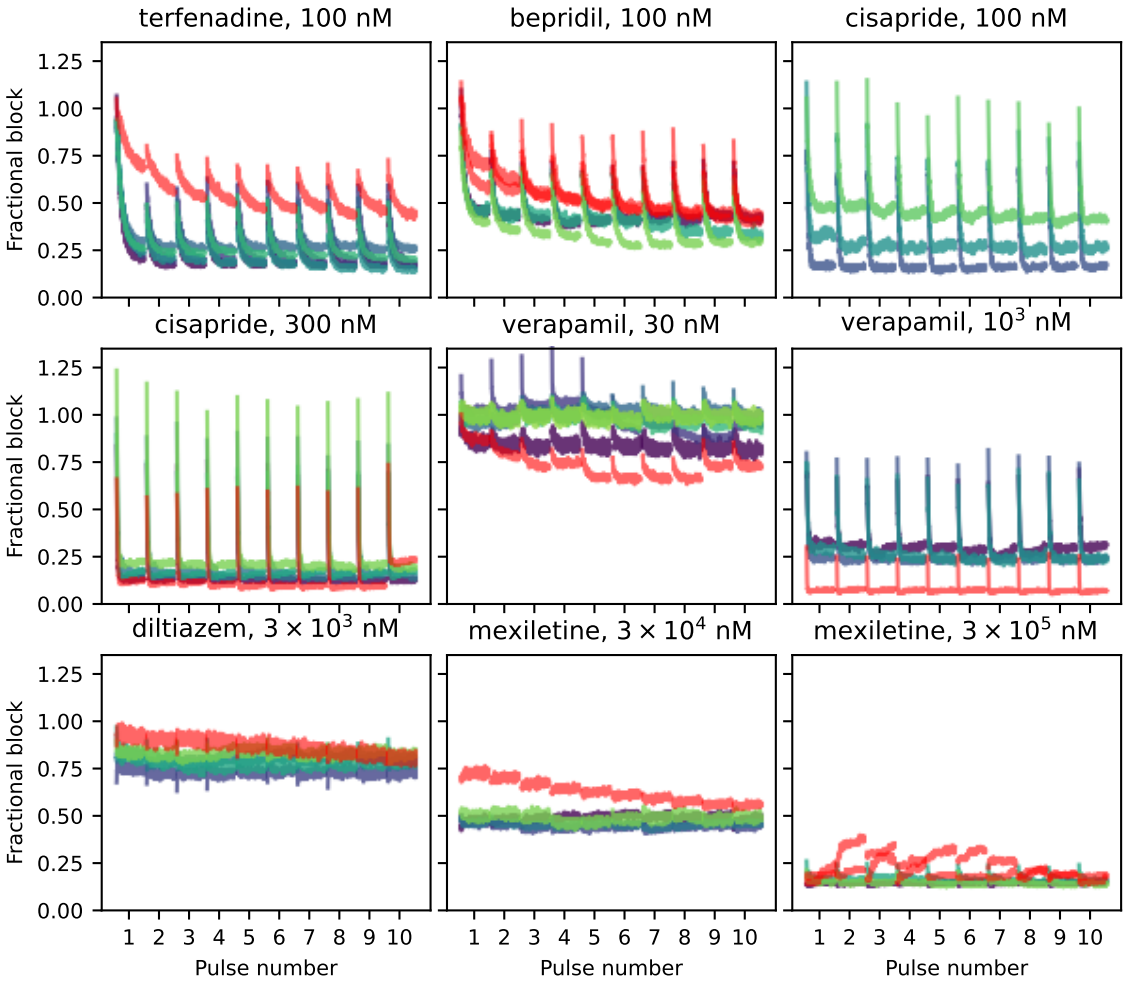


Figure B.1: **Individual traces of chosen drug at a chosen concentration to show examples of the outliers.** Individual traces of terfenadine at 100 nM, bepridil at 100 nM, cisapride at 100 nM, verapamil at 30 nM, 100 nM and 10^3 nM, diltiazem at 3×10^3 nM and mexiletine at 3×10^4 nM and 3×10^5 nM. The red traces highlight the outlier.

repeats. The inconsistencies were evident from the wide standard deviations, for example terfenadine at 100 nM, cisapride at 100 nM and verapamil at 30 nM. Figure B.1 shows the individual traces of a few drugs chosen to illustrate the data inconsistencies. There were some clear outliers in the data, such as one of the repeat of terfenadine at 100 nM has higher magnitude than the rest of the repeats and the opposite was observed in verapamil at 10^3 nM. The red traces highlight the inconsistencies. Other than the current magnitude, the current behaviour also varied with repeats. Such behaviour included the fractional block decreasing across the pulses for all repeats except for one or two repeats where it stays consistent throughout the pulses, and vice versa. This was observed in bepridil at 100 nM, diltiazem at 3×10^3 nM and mexiletine at 3×10^4 nM. Another unexpected current behaviour was observed in the fractional block where it did not decrease monotonically across all pulses nor stay consistent throughout all pulses, such as verapamil at 30 nM and mexiletine at 3×10^5 nM. In addition, we observed an increase in the fractional block within each pulse, instead of an expected decrease. This behaviour was observed in chlorpromazine at 10^3 nM and 3×10^3 nM (see Figure 5.12). We also observed inconsistencies in cisapride at 100 nM where the repeats span a wide range of magnitudes. The range of its magnitude covered the range from the lower drug concentration at 10 nM to the higher drug concentration at 300 nM, as is evident from the overlapped standard deviation (Figure 5.13). Finally, the 10th pulse of cisapride at 300 nM (Figure B.1)was significantly higher than the rest, leading to a sudden jump at the tenth pulse in the average traces after pre-processing. There is no clear pattern of inconsistencies within the data. It is therefore difficult for us to come up with a principled way of excluding subsets of data to ensure that all data are consistent. We will leave this for future work, where we could also potentially look into different sets of data using protocols other than the modified Milnes' protocol.

Bibliography

A. Agrawal, K. Wang, L. Polonchuk, J. Cooper, M. Hendrix, D. J. Gavaghan, G. R. Mirams, and M. Clerx. Models of the cardiac L-type calcium current: A quantitative review. *WIREs Mechanisms of Disease*, 15:e1581, 2023. doi: 10.1002/wsbm.1581.

B. Amuzescu, R. Airini, F. B. Epureanu, S. A. Mann, T. Knott, and B. M. Radu. Evolution of mathematical models of cardiomyocyte electrophysiology. *Mathematical Biosciences*, 334:108567, 2021. ISSN 0025-5564. doi: 10.1016/j.mbs.2021.108567.

C. M. Armstrong. Interaction of tetraethylammonium ion derivatives with the potassium channels of giant axons. *Journal of General Physiology*, 58(4):413–437, 10 1971. ISSN 0022-1295. doi: 10.1085/jgp.58.4.413.

T. Asai, N. Adachi, T. Moriya, H. Oki, T. Maru, M. Kawasaki, K. Suzuki, S. Chen, R. Ishii, K. Yonemori, S. Igaki, S. Yasuda, S. Ogasawara, T. Senda, and T. Murata. Cryo-EM structure of K^+-bound hERG channel complexed with the blocker astemizole. *Structure*, 29:203–212.e4, 3 2021. ISSN 0969-2126. doi: 10.1016/j.str.2020.12.007.

C. Audet and W. Hare. *Derivative-Free and Blackbox Optimization*. Springer International Publishing, 2017. ISBN 978-3-319-68912-8. doi: 10.1007/978-3-319-68913-5.

J. R. Balser, H. B. Nuss, D. W. Orias, D. C. Johns, E. Marban, G. F. Tomaselli, and J. H. Lawrence. Local anesthetics as effectors of allosteric gating. Lidocaine effects on inactivation-deficient rat skeletal muscle Na channels. *The Journal of Clinical Investigation*, 98(12):2874–2886, 12 1996. doi: 10.1172/JCI119116.

K. A. Beattie, A. P. Hill, R. Bardenet, Y. Cui, J. I. Vandenberg, D. J. Gavaghan, T. P. de Boer, and G. R. Mirams. Sinusoidal voltage protocols for rapid characterisation of ion channel kinetics. *The Journal of Physiology*, 596:1813–1828, 5 2018. ISSN 0022-3751. doi: 10.1113/JP275733.

G. Bellu, M. P. Saccomani, S. Audoly, and L. D'Angiò. Daisy: A new software tool to test global identifiability of biological and physiological systems. *Computer Methods and Programs in Biomedicine*, 88(1):52–61, 2007. ISSN 0169-2607. doi: 10.1016/j.cmpb.2007.07.002.

G. C. L. Bett, Q. Zhou, and R. L. Rasmusson. Models of HERG gating. *Biophysical Journal*, 101:631–642, 8 2011. ISSN 0006-3495. doi: 10.1016/j.bpj.2011.06.050.

S. F. Bifulco, N. Akoum, and P. M. Boyle. Translational applications of computational modelling for patients with cardiac arrhythmias. *Heart*, 107(6):456–461, 2021. ISSN 1355-6037. doi: 10.1136/heartjnl-2020-316854.

C. I. Bliss. The toxicity of poisons applied jointly. *Annals of Applied Biology*, 26(3): 585–615, 1939. doi: 10.1111/j.1744-7348.1939.tb06990.x.

G. E. P. Box, J. S. Hunter, and W. G. Hunter. *Statistics for experimenters : an introduction to design, data analysis, and model building*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley, 1978. ISBN 0471093157.

S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2014. ISBN 0521833787.

S. Braakman, P. Pathmanathan, and H. Moore. Evaluation framework for systems models. *CPT: Pharmacometrics & Systems Pharmacology*, 11(3):264–289, 2022. doi: 10.1002/psp4.12755.

T. Brennan, M. Fink, and B. Rodriguez. Multiscale modelling of drug-induced effects on cardiac electrophysiological activity. *European Journal of Pharmaceutical Sciences*, 36:62–77, 2009. ISSN 0928-0987. doi: 10.1016/j.ejps.2008.09.013.

J. P. Byers and J. G. Sarver. Chapter 10 - pharmacokinetic modeling. In M. Hacker, W. Messer, and K. Bachmann, editors, *Pharmacology*, pages 201–277. Academic Press, San Diego, 2009. ISBN 978-0-12-369521-5. doi: 10.1016/B978-0-12-369521-5.00010-5.

B. Carbonell-Pascual, E. Godoy, A. Ferrer, L. Romero, and J. M. Ferrero. Comparison between Hodgkin–Huxley and Markov formulations of cardiac ion channels. *Journal of theoretical Biology*, 399:92–102, 2016. ISSN 0022-5193. doi: 10.1016/j.jtbi.2016.03.039.

E. Carmeliet. Use-dependent block and use-dependent unblock of the delayed rectifier K^+ current by almokalant in rabbit ventricular myocytes. *Circulation Research*, 73 (5):857–868, 1993. doi: 10.1161/01.RES.73.5.857.

K. C. Chang, S. Dutta, G. R. Mirams, K. A. Beattie, J. Sheng, P. N. Tran, M. Wu, W. W. Wu, T. Colatsky, D. G. Strauss, and Z. Li. Uncertainty quantification reveals the importance of data variability and experimental design considerations for *in silico* proarrhythmia risk assessment. *Frontiers in Physiology*, 8, 2017. ISSN 1664-042X. doi: 10.3389/fphys.2017.00917.

J. Chen, G. Seebohm, and M. C. Sanguinetti. Position of aromatic residues in the S6 domain, not inactivation, dictates cisapride sensitivity of HERG and eag potassium channels. *Proceedings of the National Academy of Sciences*, 99(19):12461–12466, 2002. doi: 10.1073/pnas.192367299.

O. T. Chis, J. R. Banga, and E. Balsa-Canto. Structural identifiability of systems biology models: A critical comparison of methods. *PLOS ONE*, 6(11):1–16, 11 2011. doi: 10.1371/journal.pone.0027755.

A. P. Clark, S. Wei, D. Kalola, T. Krogh-Madsen, and D. J. Christini. An in silico–in vitro pipeline for drug cardiotoxicity screening identifies ionic pro-arrhythmia mechanisms. *British Journal of Pharmacology*, 179(20):4829–4843, 2022. doi: 10.1111/bph.15915.

C. W. Clarkson and L. M. Hondeghem. Evidence for a specific receptor site for lidocaine, quinidine, and bupivacaine associated with cardiac sodium channels in guinea pig ventricular myocardium. *Circulation Research*, 56(4):496–506, 1985. doi: 10.1161/01.RES.56.4.496.

R. H. Clayton, Y. Aboelkassem, C. D. Cantwell, C. Corrado, T. Delhaas, W. Huberts, C. L. Lei, H. Ni, A. V. Panfilov, C. Roney, and R. W. dos Santos. An audit of uncertainty in multi-scale cardiac electrophysiology models. *Philosophical Transactions of the Royal Society A*, 378(2173):20190335, 2020. doi: 10.1098/rsta.2019.0335.

M. Clerx, P. Collins, E. de Lange, and P. G. A. Volders. Myokit: A simple interface to cardiac cellular electrophysiology. *Progress in Biophysics and Molecular Biology*, 120(1–3):100–114, 2016. ISSN 0079-6107. doi: 10.1016/j.pbiomolbio.2015.12.008.

M. Clerx, M. Robinson, B. Lambert, C. L. Lei, S. Ghosh, G. R. Mirams, and D. J. Gavaghan. Probabilistic Inference on Noisy Time Series (PINTS). *Journal of Open Research Software*, 7(1):23, 2019. doi: 10.5334/jors.252.

T. Colatsky, B. Fermini, G. Gintant, J. B. Pierson, P. Sager, Y. Sekino, D. G. Strauss, and N. Stockbridge. The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — update on progress. *Journal of Pharmacological and Toxicological Methods*, 81:15–20, 9 2016. ISSN 1056-8719. doi: 10.1016/J.VASCN.2016.06.002.

S. R. Cole, H. Chu, and S. Greenland. Maximum likelihood, profile likelihood, and penalized likelihood: A primer. *American Journal of Epidemiology*, 179(2):252–260, 10 2013. ISSN 0002-9262. doi: 10.1093/aje/kwt245.

A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to derivative-free optimization*. MPS-SIAM series on optimization ; 8. Society for Industrial and Applied Mathematics SIAM, 2009. ISBN 0-89871-876-7.

J. Cooper, R. J. Spiteri, and G. R. Mirams. Cellular cardiac electrophysiology modeling with Chaste and CellML. *Frontiers in Physiology*, Volume 5 - 2014, 2015. ISSN 1664-042X. doi: 10.3389/fphys.2014.00511.

C. Corrado, A. Avezzù, A. W. C. Lee, C. Mendoca Costa, C. H. Roney, M. Strocchi, M. Bishop, and S. A. Niederer. Using cardiac ionic cell models to interpret clinical data. *WIREs Mechanisms of Disease*, 13(3):e1508, 2021. doi: 10.1002/wsbm.1508.

M. E. Curran, I. Splawski, K. W. Timothy, G. M. Vincen, E. D. Green, and M. T. Keating. A molecular basis for cardiac arrhythmia: hERG mutations cause long QT syndrome. *Cell*, 80:795–803, 3 1995. ISSN 0092-8674. doi: 10.1016/0092-8674(95)90358-5.

M. R. Davies, H. B. Mistry, L. Hussein, C. E. Pollard, J.-P. Valentin, J. Swinton, and N. Abi-Gerges. 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:H1466–H1480, 12 2011. ISSN 0363-6135. doi: 10.1152/ajpheart.00808.2011.

M. R. Davies, K. Wang, G. R. Mirams, A. Caruso, D. Noble, A. Walz, T. Lavé, F. Schuler, T. Singer, and L. Polonchuk. Recent developments in using mechanistic cardiac modelling for drug safety evaluation. *Drug Discovery Today*, 21(6):924–938, 2016. ISSN 1359-6446. doi: 10.1016/j.drudis.2016.02.003.

G. Y. Di Veroli, M. R. Davies, H. Zhang, N. Abi-Gerges, and M. R. Boyett. High-throughput screening of drug-binding dynamics to hERG improves early drug safety assessment. *American Journal of Physiology-Heart and Circulartory Physiology*, 304:H104–H117, 10 2012. ISSN 0363-6135. doi: 10.1152/ajpheart.00511.2012.

G. Y. Di Veroli, M. R. Davies, H. Zhang, N. Abi-Gerges, and M. R. Boyett. 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*, 25:197–207, 2 2014. ISSN 1045-3873. doi: 10.1111/jce.12289.

C. Y. Du, A. El Harchi, M. J. McPate, C. H. Orchard, and J. C. Hancox. Enhanced inhibitory effect of acidosis on hERG potassium channels that incorporate the hERG1b isoform. *Biochemical and Biophysical Research Communications*, 405 (2):222–227, 2011. ISSN 0006-291X. doi: 10.1016/j.bbrc.2011.01.014.

A. Dunnink, T. R. G. Stams, A. Bossu, V. M. F. Meijborg, J. D. M. Beekman, S. C. Wijers, Jacques M.T. De Bakker, and Marc A. Vos. Torsade de pointes arrhythmias arise at the site of maximal heterogeneity of repolarization in the chronic complete atrioventricular block dog. *EP Europace*, 19(5):858–865, 10 2016. ISSN 1099-5129. doi: 10.1093/europace/euw087.

S. Dutta, K. C. Chang, K. A. Beattie, J. Sheng, P. N. Tran, W. W. Wu, M. Wu, D. G. Strauss, T. Colatsky, and Z. Li. Optimization of an in silico cardiac cell model for proarrhythmia risk assessment. *Frontiers in Physiology*, 8, 2017. ISSN 1664-042X. doi: 10.3389/fphys.2017.00616.

N. El-Sherif, E. B. Caref, H. Yin, and M. Restivo. The electrophysiological mechanism of ventricular arrhythmias in the long QT syndrome. *Circulation Research*, 79(3): 474–492, 1996. doi: 10.1161/01.RES.79.3.474.

N. El-Sherif, M. Chinushi, E. B. Caref, and M. Restivo. Electrophysiological mechanism of the characteristic electrocardiographic morphology of torsade de pointes tachyarrhythmias in the long-QT syndrome. *Circulation*, 96(12):4392–4399, 1997. doi: 10.1161/01.CIR.96.12.4392.

R. C. Elkins, M. R. Davies, S. J. Brough, D. J. Gavaghan, Y. Cui, N. Abi-Gerges, and G. R. Mirams. Variability in high-throughput ion-channel screening data and consequences for cardiac safety assessment. *Journal of Pharmacological and Toxicological Methods*, 68:112–122, 2013. ISSN 1056-8719. doi: 10.1016/j.vascn.2013.04.007.

F. Escobar, J. Gomis-Tena, J. Saiz, and L. Romero. Automatic modeling of dynamic drug-herg channel interactions using three voltage protocols and machine learning techniques: A simulation study. *Computer Methods and Programs in Biomedicine*, 226:107148, 2022. ISSN 0169-2607. doi: 10.1016/j.cmpb.2022.107148.

H. J. Farm, M. Clerx, F. Cooper, L. Polonchuk, K. Wang, D. J. Gavaghan, and C. L. Lei. Importance of modelling herg binding in predicting drug-induced action potential prolongations for drug safety assessment. *Frontiers in Pharmacology*, 14, 2023. ISSN 1663-9812. doi: 10.3389/fphar.2023.1110555.

B. Fermini, J. C. Hancox, N. Abi-Gerges, M. Bridgland-Taylor, K. W. Chaudhary, T. Colatsky, K. Correll, W. Crumb, B. Damiano, G. Erdemli, G. Gintant, J. Imredy, J. Koerner, J. Kramer, P. Levesque, Z. Li, A. Lindqvist, C. A. Obejero-Paz, D. Rampe, K. Sawada, D. G. Strauss, and J. I. Vandenberg. A new perspective in the field of cardiac safety testing through the Comprehensive in vitro Proarrhythmia Assay paradigm. *Journal of Biomolecular Screening*, 21(1):1–11, 2015. doi: 10.1177/1087057115594589.

E. Ficker, W. Jarolimek, J. Kiehn, A. Baumann, and A. M. Brown. Molecular determinants of dofetilide block of HERG K^+ channels. *Circulation Research*, 82(3): 386–395, 1998. doi: 10.1161/01.RES.82.3.386.

M. Fink and D. Noble. Markov models for ion channels: versatility versus identifiability and speed. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367:2161–2179, 6 2009. doi: 10.1098/rsta.2008.0301.

M. Fink, S. A. Niederer, E. M. Cherry, F. H. Fenton, J. T. Koivumäki, G. Seemann, R. Thul, H. Zhang, F. B. Sachse, D. Beard, E. J. Crampin, and N. P. Smith. Cardiac cell modelling: Observations from the heart of the cardiac physiome project. *Progress in Biophysics and Molecular Biology*, 104:2–21, 2011. ISSN 0079-6107. doi: 10.1016/j.pbiomolbio.2010.03.002.

D. Franz, H. L. Olsen, O. Klink, and J. Gimsa. Automated and manual patch clamp data of human induced pluripotent stem cell-derived dopaminergic neurons. *Scientific data*, 4:170056, 4 2017. ISSN 2052-4463. doi: 10.1038/sdata.2017.56.

D. C. Gadsby. Ion channels versus ion pumps: the principal difference, in principle. *Nature Reviews Molecular Cell Biology*, 10:344–352, 2009. ISSN 1471-0080. doi: 10.1038/nrm2668.

J. Gallaher, K. Larripa, M. Renardy, B. Shtylla, N. Tania, D. White, K. Wood, L. Zhu, C. Passey, M. Robbins, N. Bezman, S. Shelat, J. C. Hearn, and H. Moore. Methods for determining key components in a mathematical model for tumor–immune dynamics in multiple myeloma. *Journal of Theoretical Biology*, 458:31–46, 2018. ISSN 0022-5193. doi: 10.1016/j.jtbi.2018.08.037.

N. Gamper, J. D. Stockand, and M. S. Shapiro. The use of Chinese hamster ovary (CHO) cells in the study of ion channels. *Journal of Pharmacological and Toxicological Methods*, 51:177–185, 2005. ISSN 1056-8719. doi: 10.1016/j.vascn.2004.08.008.

A. Garny, D. P. Nickerson, J. Cooper, R. W. dos Santos, A. K. Miller, S. McKeever, P. M. F. Nielsen, and P. J. Hunter. CellML and associated tools and techniques. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1878):3017–3043, 2008. doi: 10.1098/rsta.2008.0094.

A. H. Glassman and J. T. Bigger. Antipsychotic drugs: Prolonged QTc interval, Torsade de Pointes, and sudden death. *American Journal of Psychiatry*, 158:1774–1782, 11 2001. doi: 10.1176/appi.ajp.158.11.1774.

D. E. Goldman. Potential, impedance, and rectification in membranes. *Journal of General Physiology*, 27:37–60, 9 1943. ISSN 0022-1295. doi: 10.1085/jgp.27.1.37.

J. Gomis-Tena, B. M. Brown, J. Cano, B. Trenor, P.-C. Yang, J. Saiz, C. E. Clancy, and L. Romero. When does the IC_{50} accurately assess the blocking potency of a drug? *Journal of Chemical Information and Modeling*, 60:1779–1790, 3 2020. ISSN 1549-9596. doi: 10.1021/acs.jcim.9b01085.

T. González, C. Arias, R. Caballero, I. Moreno, E. Delpón, J. Tamargo, and C. Valenzuela. Effects of levobupivacaine, ropivacaine and bupivacaine on HERG channels: stereoselective bupivacaine block. *British Journal of Pharmacology*, 137(8):1269–1279, 2002. doi: 10.1038/sj.bjp.0704978.

E. Grandi, F. S. Pasqualini, and D. M. Bers. A novel computational model of the human ventricular action potential and Ca transient. *Journal of Molecular and Cellular Cardiology*, 48:112–121, 1 2010. ISSN 0022-2828. doi: 10.1016/j.yjmcc.2009.09.019.

A. O. Grant. Cardiac ion channels. *Circulation: Arrhythmia and Electrophysiology*, 2:185–194, 4 2009. doi: 10.1161/CIRCEP.108.789081.

A. O. Grant, L. J. Strauss, A. G. Wallace, and H. C. Strauss. The influence of pH on th electrophysiological effects of lidocaine in guinea pig ventricular myocardium. *Circulation Research*, 47(4):542–550, 1980. doi: 10.1161/01.RES.47.4.542.

D. M. Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, 32:135–154, 1994. ISSN 1573-2959. doi: 10.1007/BF00547132.

T. G. Hammond, L. Carlsson, A. S. Davis, W. G. Lynch, I. MacKenzie, W. S. Redfern, A. T. Sullivan, and A. J. Camm. Methods of collecting and evaluating non-clinical cardiac electrophysiology data in the pharmaceutical industry: results of an international survey. *Cardiovascular Research*, 49:741–750, 3 2001. ISSN 0008-6363. doi: 10.1016/S0008-6363(00)00310-2. 10

N. Hansen, S. D. Müller, and P. Koumoutsakos. Reducing the time complexity of the derandomized evolution strategy with covariance matrix adaptation (CMA-ES). *Evolutionary Computation*, 11:1–18, 3 2003. ISSN 1063-6560. doi: 10.1162/106365603321828970.

J. Hayya, D. Armstrong, and N. Gressis. A note on the ratio of two normally distributed variables. *Management Science*, 21(11):1338–1341, 1975. doi: 10.1287/mnsc.21.11.1338.

M. Heinrich, R. Arutjunjan, and J. Timmer. On the different flavours of practical identifiability. *Current Opinion in Systems Biology*, page 100556, 2025a. ISSN 2452-3100. doi: 10.1016/j.coisb.2025.100556.

M. Heinrich, M. Rosenblatt, F.-G. Wieland, H. Stigter, and J. Timmer. On structural and practical identifiability: Current status and update of results. *Current Opinion in Systems Biology*, 41:100546, 2025b. ISSN 2452-3100. doi: 10.1016/j.coisb.2025.100546.

E. K. Heist and J. N. Ruskin. Drug-induced arrhythmia. *Circulation*, 122(14):1426–1435, 2010. doi: 10.1161/CIRCULATIONAHA.109.894725.

S. Heitmann, J. I. Vandenberg, and A. P. Hill. Assessing drug safety by identifying the axis of arrhythmia in cardiomyocyte electrophysiology. *eLife*, 2023. doi: 10.7554/elife.90027.2.

S. Heitmann, J. I. Vandenberg, and A. P. Hill. Independent derivations of the axis of arrhythmia for predicting drug-induced torsades de pointes. *British Journal of Pharmacology*, 182(18):4299–4313, 2025. doi: 10.1111/bph.70078.

A. P. Hill, M. D. Perry, N. Abi-Gerges, J.-P. Couderc, B. Fermini, J. C. Hancox, B. C. Knollmann, G. R. Mirams, J. Skinner, W. Zareba, and J. I. Vandenberg. Computational cardiology and risk stratification for sudden cardiac death: one of the grand challenges for cardiology in the 21st century. *The Journal of Physiology*, 594:6893–6908, 12 2016. ISSN 0022-3751. doi: 10.1113/JP272015.

A. V. Hill. The possible effects of the aggregation of the molecules of haemoglobin on its dissociation curves. *The Journal of Physiology*, 40(suppl):iV–vii, 1910. doi: 10.1113/jphysiol.1910.sp001386.

C. L. Hill and G. J. Stephens. *An Introduction to Patch Clamp Recording*, pages 1–19. Springer US, 2021. ISBN 978-1-0716-0818-0. doi: 10.1007/978-1-0716-0818-0_1.

B. Hille. Ionic channels in excitable membranes. current problems and biophysical approaches. *Biophysical journal*, 22:283–94, 5 1978. ISSN 0006-3495. doi: 10.1016/S0006-3495(78)85489-7.

B. Hille. *Ion Channels of Excitable Membranes*. 2001. ISBN 9780197583074.

Alan C. Hindmarsh, Peter N. Brown, Keith E. Grant, Steven L. Lee, Radu Serban, Dan E. Shumaker, and Carol S. Woodward. SUNDIALS: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software*, 31(3):363–396, sep 2005. ISSN 0098-3500. doi: 10.1145/1089014.1089020.

K. E. Hines, T. R. Middendorf, and R. W. Aldrich. Determination of parameter identifiability in nonlinear biophysical models: A Bayesian approach. *Journal of General Physiology*, 143(3):401–416, 02 2014. ISSN 0022-1295. doi: 10.1085/jgp.201311116.

A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*, 117:500–544, 8 1952. ISSN 0022-3751. doi: 10.1113/jphysiol.1952.sp004764.

A. L. Hodgkin and B. Katz. The effect of sodium ions on the electrical activity of the giant axon of the squid. *The Journal of Physiology*, 108:37–77, 3 1949. ISSN 0022-3751. doi: 10.1113/jphysiol.1949.sp004310.

L. M. Hondeghem. Antiarrhythmic agents: modulated receptor applications. *Circulation*, 75:514–520, 3 1987. ISSN 0009-7322. doi: 10.1161/01.CIR.75.3.514.

L. M. Hondeghem. Thorough QT/QTc not so thorough: Removes torsadogenic predictors from the T-wave, incriminates safe drugs, and misses profibrillatory drugs. *Journal of Cardiovascular Electrophysiology*, 17(3):337–340, 2006. doi: 10.1111/j.1540-8167.2006.00347.x.

M. Huang, Z. Liao, X. Li, Zhen Z. Yang, X. Fan, Y. Li, Z. Zhao, S. Lang, L. Cyganek, X. Zhou, I. Akin, M. Borggrefe, and I. El-Battrawy. Effects of antiarrhythmic drugs on hERG gating in human-induced pluripotent stem cell-derived cardiomyocytes from a patient with short QT syndrome type 1. *Frontiers in Pharmacology*, Volume 12, 2021. ISSN 1663-9812. doi: 10.3389/fphar.2021.675003.

W. Huberts, C. de Jonge, W. P. M. van der Linden, M. A. Inda, K. Passera, J. H. M. Tordoir, F. N. van de Vosse, and E. M. H. Bosboom. A sensitivity analysis of a personalized pulse wave propagation model for arteriovenous fistula surgery. Part B: Identification of possible generic model parameters. *Medical Engineering & Physics*, 35(6):827–837, 2013a. ISSN 1350-4533. doi: 10.1016/j.medengphy.2012.08.012.

W. Huberts, C. de Jonge, W. P. M. van der Linden, M. A. Inda, J. H. M. Tordoir, F. N. van de Vosse, and E. M. H. Bosboom. A sensitivity analysis of a personalized pulse wave propagation model for arteriovenous fistula surgery. Part A: Identification of most influential model parameters. *Medical Engineering & Physics*, 35(6):810–826, 2013b. ISSN 1350-4533. doi: 10.1016/j.medengphy.2012.08.013.

J. R. Hume and A. O. Grant. *Agents Used in Cardiac Arrhythmias*. McGraw-Hill, New York, NY, 2015.

ICH E14. E14 the clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs. https://database.ich.org/sites/default/files/E14_Guideline.pdf, 2005. [Accessed 11-October-2025].

ICH E14/S7B Q&A. Questions and answers: Clinical and nonclinical evaluation of QT/QTc interval prolongation and proarrhythmic potential. https://database.ich.org/sites/default/files/E14-S7B_QAs_Step4_2022_0221.pdf, 2022. [Accessed 11-October-2025].

ICH S7B. S7B the non-clinical evaluation of the potential for delayed ventricular repolarization (QT interval prolongation) by human pharmaceuticals. https://database.ich.org/sites/default/files/S7B_Guideline.pdf, 2005. [Accessed 11-October-2025].

M. W. Jennings, P. Nithiarasu, and S. Pant. Quantifying the efficacy of voltage protocols in characterising ion channel kinetics: A novel information-theoretic approach. *International Journal for Numerical Methods in Biomedical Engineering*, 40(5):e3815, 2024. doi: 10.1002/cnm.3815.

E. M. C. Jones, E. C. Roti Roti, J. Wang, S. A. Delfosse, and G. A. Robertson. Cardiac I_{Kr} channels minimally comprise hERG 1a and 1b subunits. *Journal of Biological Chemistry*, 279:44690–44694, 10 2004. ISSN 00219258. doi: 10.1074/jbc.M408344200.

N. K. Jurkiewicz and M. C. Sanguinetti. Rate-dependent prolongation of cardiac action potentials by a methanesulfonanilide class III antiarrhythmic agent. Specific block of rapidly activating delayed rectifier K^+ current by dofetilide. *Circulation Research*, 72(1):75–83, 1993. doi: 10.1161/01.RES.72.1.75.

S. Kalyaanamoorthy and K. H. Barakat. Development of safe drugs: The hERG challenge. *Medicinal Research Reviews*, 38(2):525–555, 2018. doi: 10.1002/med.21445.

K. Kamiya, R. Niwa, J. S. Mitcheson, and M. C. Sanguinetti. Molecular determinants of hERG channel block. *Molecular Pharmacology*, 69(5):1709–1716, 2006. ISSN 0026-895X. doi: 10.1124/mol.105.020990.

J. Kiehn, A. E. Lacerda, B. Wible, and A. M. Brown. Molecular physiology and pharmacology of HERG. *Circulation*, 94(10):2572–2579, 1996. doi: 10.1161/01.CIR.94.10.2572.

J. Kiehn, A. E. Lacerda, and A. M. Brown. Pathways of HERG inactivation. *American Journal of Physiology-Heart and Circulatory Physiology*, 277(1):H199–H210, 1999. doi: 10.1152/ajpheart.1999.277.1.H199.

G. E. Kirsch, E. S. Trepakova, J. C. Brimecombe, S. S. Sidach, H. D. Erickson, M. C. Kochan, L. M. Shyjka, A. E. Lacerda, and A. M. Brown. Variability in the measurement of hERG potassium channel inhibition: Effects of temperature and stimulus pattern. *Journal of Pharmacological and Toxicological Methods*, 50: 93–101, 9 2004. ISSN 1056-8719. doi: 10.1016/J.VASCN.2004.06.003.

S. A. Kodirov. Tale of tail current. *Progress in Biophysics and Molecular Biology*, 150:78–97, 2020. ISSN 0079-6107. doi: 10.1016/j.pbiomolbio.2019.06.002.

J. Kramer, H. M. Himmel, A. Lindqvist, S. Stoelzle-Feix, K. W. Chaudhary, D. Li, G. A. Bohme, M. Bridgland-Taylor, S. Hebeisen, J. Fan, M. Renganathan, J. Imredy, E. S. A. Humphries, N. Brinkwirth, T. Strassmaier, A. Ohtsuki, T. Danker, C. Vanoye, L. Polonchuk, B. Fermini, J. B. Pierson, and G. Gintant. Cross-site and cross-platform variability of automated patch clamp assessments of drug effects on human cardiac currents in recombinant cells. *Scientific Reports*, 10:5627, 3 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-62344-w.

C. Kreutz, A. Raue, D. Kaschek, and J. Timmer. Profile likelihood in systems biology. *The FEBS Journal*, 280(11):2564–2571, 2013. doi: 10.1111/febs.12276.

O. I. Krivorotko, D. V. Andornaya, and S. I. Kabanikhin. Sensitivity analysis and practical identifiability of some mathematical models in biology. *Journal of Applied and Industrial Mathematics*, 14:115–130, 2020. ISSN 1990-4797. doi: 10.1134/S1990478920010123.

T. Krogh-Madsen, A. F. Jacobson, F. A. Ortega, and D. J. Christini. Global optimization of ventricular myocyte model to multi-variable objective improves predictions of drug-induced Torsades de Pointes. *Frontiers in Physiology*, 8, 2017. ISSN 1664-042X. doi: 10.3389/fphys.2017.01059.

W. Lee, S. A. Mann, M. J. Windley, M. S. Imtiaz, J. I. Vandenberg, and A. P. Hill. *In silico* assessment of kinetics and state dependent binding properties of drugs causing acquired LQTS. *Progress in Biophysics and Molecular Biology*, 120(1): 89–99, 2016. ISSN 0079-6107. doi: 10.1016/j.pbiomolbio.2015.12.005.

W. Lee, M. J. Windley, M. D. Perry, J. I. Vandenberg, and A. Hill. Protocol-dependent differences in IC50 values measured in hERG assays occur in a predictable way and can be used to quantify state preference of drug binding. *Molecular Pharmacology*, page mol.118.115220, 1 2019. doi: 10.1124/mol.118.115220.

W. Lee, B. Ng, M. M. Mangala, M. D. Perry, R. N. Subbiah, J. I. Vandenberg, and A. P. Hill. Action potential morphology accurately predicts proarrhythmic risk for drugs with potential to prolong cardiac repolarization. *Circulation: Arrhythmia and Electrophysiology*, 16(7):399–410, 2023. doi: 10.1161/CIRCEP.122.011574.

J. P. Lees-Miller, Y. Duan, G. Q. Teng, and H. J. Duff. Molecular determinant of high-affinity dofetilide binding to HERG1 expressed in xenopus oocytes: Involvement of S6 sites. *Molecular Pharmacology*, 57(2):367–374, 2000. ISSN 0026-895X. doi: 10.1016/S0026-895X(24)23209-0.

C. L. Lei, M. Clerx, K. A. Beattie, D. Melgari, J. C. Hancox, D. J. Gavaghan, L. Polonchuk, K. Wang, and G. R. Mirams. Rapid characterization of hERG channel kinetics II: Temperature dependence. *Biophysical Journal*, 117:2455–2470, 2019a. ISSN 0006-3495. doi: 10.1016/j.bpj.2019.07.030.

C. L. Lei, M. Clerx, D. J. Gavaghan, L. Polonchuk, G. R. Mirams, and K. Wang. Rapid characterization of hERG channel kinetics I: using an automated high-throughput system. *Biophysical Journal*, 117:2438–2454, 12 2019b. ISSN 0006-3495. doi: 10.1016/J.BPJ.2019.07.029.

C. L. Lei, S. Ghosh, D. G. Whittaker, Y. Aboelkassem, K. A. Beattie, C. D. Cantwell, T. Delhaas, C. Houston, G. M. Novaes, A. V. Panfilov, M. Riabiz P. Pathmanathan, R. W. dos Santos, J. Walmsley, K. Worden, G. R. Mirams, and R. D. Wilkinson. Considering discrepancy when calibrating a mechanistic electrophysiology model. *Philosophical Transactions of the Royal Society A*, 378(2173):20190349, 2020. doi: 10.1098/rsta.2019.0349.

C. L. Lei, D. G. Whittaker, and G. R. Mirams. The impact of uncertainty in hERG binding mechanism on in silico predictions of drug-induced proarrhythmic risk. *British Journal of Pharmacology*, 181:987–1004, 4 2024. ISSN 0007-1188. doi: 10.1111/bph.16250.

T. Li, G. Lu, E. Y. Chiang, T. Chernov-Rogan, J. L. Grogan, and J. Chen. High-throughput electrophysiological assays for voltage gated ion channels using SyncroPatch 768PE. *PLOS ONE*, 12:e0180154–, 7 2017a. doi: 10.1371/journal.pone.0180154.

Z. Li, S. Dutta, J. Sheng, P. N. Tran, W. Wu, and T. Colatsky. A temperature-dependent in silico model of the human ether-à-go-go-related (hERG) gene channel. *Journal of Pharmacological and Toxicological Methods*, 81:233–239, 2016. ISSN 1056-8719. doi: 10.1016/j.vascn.2016.05.005.

Z. Li, S. Dutta, J. Sheng, P. N. Tran, W. Wu, K. Chang, T. Mdluli, D. G. Strauss, and T. Colatsky. Improving the in silico assessment of proarrhythmia risk by combining hERG (human Ether-à-go-go-Related Gene) channel-drug binding kinetics

and multichannel pharmacology. *Circulation: Arrhythmia and Electrophysiology*, 10, 2 2017b. ISSN 19413084. doi: 10.1161/CIRCEP.116.004628.

Z. Li, B. J. Ridder, X. Han, W. W. Wu, J. Sheng, P. N. Tran, M. Wu, A. Randolph, R. H. Johnstone, G. R. Mirams, Y. Kuryshev, J. Kramer, C. Wu, W. J. Crumb Jr., and D. G. Strauss. Assessment of an *in silico* mechanistic model for proarrhythmia risk prediction under the CiPA initiative. *Clinical Pharmacology & Therapeutics*, 105:466–475, 2 2019. ISSN 0009-9236. doi: 10.1002/cpt.1184.

Z. Li, G. R. Mirams, T. Yoshinaga, B. J. Ridder, X. Han, J. E. Chen, N. L. Stockbridge, T. A. Wisialowski, B. Damiano, S. Severi, P. Morissette, P. R. Kowey, M. Holbrook, G. Smith, R. L. Rasmusson, M. Liu, Z. Song, Z. Qu, D. J. Leishman, J. Steidl-Nichols, B. Rodriguez, A. Bueno-Orovio, X. Zhou, E. Passini, A. G. Edwards, S. Morotti, H. Ni, E. Grandi, C. E. Clancy, J. Vandenberg, A. Hill, M. Nakamura, T. Singer, L. Polonchuk, A. Greiter-Wilke, K. Wang, S. Nave, A. Fullerton, E. A. Sobie, M. Paci, F. M. Tshinanu, and D. G. Strauss. General principles for the validation of proarrhythmia risk prediction models: An extension of the CiPA *in silico* strategy. *Clinical Pharmacology & Therapeutics*, 107:102–111, 1 2020. ISSN 0009-9236. doi: 10.1002/cpt.1647.

T. S. Ligon, F. Fröhlich, O. T. Chiş, J. R. Banga, E. Balsa-Canto, and Jan J. Hasenauer. GenSSI 2.0: multi-experiment structural identifiability analysis of SBML models. *Bioinformatics*, 34(8):1421–1423, 11 2017. ISSN 1367-4803. doi: 10.1093/bioinformatics/btx735.

L. Ljung and T. Glad. On global identifiability for arbitrary model parametrizations. *Automatica*, 30(2):265–276, 1994. ISSN 0005-1098. doi: 10.1016/0005-1098(94)90029-9.

C. M. Lloyd, J. R. Lawson, P. J. Hunter, and P. F. Nielsen. The CellML model repository. *Bioinformatics*, 24(18):2122–2123, 07 2008. ISSN 1367-4803. doi: 10.1093/bioinformatics/btn390.

Y.-C. Lo and C.-C. Kuo. Temperature dependence of the biophysical mechanisms underlying the inhibition and enhancement effect of amiodarone on hERG channels. *Molecular Pharmacology*, 96(3):330–344, 2019. doi: 10.1124/mol.119.116400.

T. Maiwald, H. Hass, B. Steiert, J. Vanlier, R. Engesser, A. Raue, F. Kipkeew, H. H. Bock, D. Kaschek, C. Kreutz, and J. Timmer. Driving the model to its limit:

Profile likelihood based model reduction. *PLOS ONE*, 11(9):1–18, 09 2016. doi: 10.1371/journal.pone.0162366.

K. E. Mangold, W. Wang, E. K. Johnson, D. Bhagavan, J. D. Moreno, J. M. Nerbonne, and J. R. Silva. Identification of structures for ion channel kinetic models. *PLOS Computational Biology*, 17(8):1–26, 08 2021. doi: 10.1371/journal.pcbi.1008932.

G. Margaria, E. Riccomagno, M. J. Chappell, and H. P. Wynn. Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. *Mathematical Biosciences*, 174(1):1–26, 2001. ISSN 0025-5564. doi: 10.1016/S0025-5564(01)00079-7.

S. Marino, I. B. Hogue, C. J. Ray, and D. E. Kirschner. A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology*, 254(1):178–196, 2008. ISSN 0022-5193. doi: 10.1016/j.jtbi.2008.04.011.

R. L. Martin, J. S. McDermott, H. J. Salmen, J. Palmatier, B. F. Cox, and G. A. Gintant. The utility of hERG and repolarization assays in evaluating delayed cardiac repolarization: Influence of multi-channel block. *Journal of Cardiovascular Pharmacology*, 43:369–379, 2004. ISSN 0160-2446.

J. R. R. A. Martins and A. Ning. *Engineering Design Optimization*. Cambridge University Press, Cambridge, UK, January 2022. ISBN 9781108833417. doi: 10.1017/9781108980647.

G. Massonis and A. F. Villaverde. Finding and breaking lie symmetries: Implications for structural identifiability and observability in biological modelling. *Symmetry*, 12(3), 2020. ISSN 2073-8994. doi: 10.3390/sym12030469.

N. Meshkat, C. E.z. Kuo, and J III DiStefano. On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and combos: A novel web implementation. *PLOS ONE*, 9(10), 10 2014. doi: 10.1371/journal.pone.0110261.

H. Miao, X. Xia, A. S. Perelson, and H. Wu. On identifiability of nonlinear ode models and applications in viral dynamics. *SIAM Review*, 53:3–39, 1 2011. ISSN 0036-1445. doi: 10.1137/090757009.

J. T. Milnes, O. Crociani, A. Arcangeli, J. C. Hancox, and H. J. Witchel. Blockade of HERG potassium currents by fluvoxamine: incomplete attenuation by S6 mutations at F656 or Y652. *British Journal of Pharmacology*, 139:887–898, 7 2003. ISSN 0007-1188. doi: 10.1038/sj.bjp.0705335.

J. T. Milnes, H. J. Witchel, J. L. Leaney, D. J. Leishman, and J. C. Hancox. Investigating dynamic protocol-dependence of hERG potassium channel inhibition at 37 °C: Cisapride versus dofetilide. *Journal of Pharmacological and Toxicological Methods*, 61:178–191, 3 2010. ISSN 1056-8719. doi: 10.1016/J.VASCN.2010.02.007.

G. R. Mirams. *Computational Cardiac Safety Testing*, pages 1–33. Springer International Publishing, 2022. ISBN 978-3-030-73317-9. doi: 10.1007/978-3-030-73317-9_137-1.

G. R. Mirams, Y. Cui, A. Sher, M. Fink, J. Cooper, B. M. Heath, N. C. McMahon, D. J. Gavaghan, and D. Noble. Simulation of multiple ion channel block provides improved early prediction of compounds' clinical torsadogenic risk. *Cardiovascular Research*, 91:53–61, 7 2011. ISSN 0008-6363. doi: 10.1093/cvr/cvr044.

G. R. Mirams, M. R. Davies, Y. Cui, P. Kohl, and D. Noble. Application of cardiac electrophysiology simulations to pro-arrhythmic safety testing. *British Journal of Pharmacology*, 167:932–945, 11 2012. ISSN 0007-1188. doi: 10.1111/j.1476-5381.2012.02020.x.

G. R. Mirams, P. Pathmanathan, R. A. Gray, P. Challenor, and R. H. Clayton. Uncertainty and variability in computational and mathematical models of cardiac physiology. *The Journal of Physiology*, 594:6833–6847, 2016. doi: 10.1113/JP271671.

J. S. Mitcheson. hERG potassium channels and the structural basis of drug-induced arrhythmias. *Chemical Research in Toxicology*, 21:1005–1010, 5 2008. ISSN 0893-228X. doi: 10.1021/tx800035b.

J. S. Mitcheson, J. Chen, M. Lin, C. Culberson, and M. C. Sanguinetti. A structural basis for drug-induced long QT syndrome. *Proceedings of the National Academy of Sciences*, 97(22):12329–12333, 2000a. doi: 10.1073/pnas.210244497.

J. S. Mitcheson, J. Chen, and M. C. Sanguinetti. Trapping of a methanesulfonanilide by closure of the hERG potassium channel activation gate. *The Journal of general physiology*, 115:229–240, 3 2000b. ISSN 0022-1295. doi: 10.1085/jgp.115.3.229.

Y. Miyashita, T. Moriya, T. Kato, M. Kawasaki, S. Yasuda, N. Adachi, K. Suzuki, S. Ogasawara, T. Saito, T. Senda, and T. Murata. Improved higher resolution cryo-EM structures reveal the binding modes of hERG channel inhibitors. *Structure*, 32:1926–1935.e3, 11 2024. ISSN 0969-2126. doi: 10.1016/j.str.2024.08.021.

J. D. Moreno and J. R. Silva. Emerging methods to model cardiac ion channel and myocyte electrophysiology. *Biophysics Reviews*, 4(1):011315, 03 2023. ISSN 2688-4089. doi: 10.1063/5.0127713.

J. D. Moreno, T. J. Lewis, and C. E. Clancy. Parameterization for in-silico modeling of ion channel interactions with drugs. *PLOS ONE*, 11(3):1–22, 03 2016. doi: 10.1371/journal.pone.0150761.

S. Nattel and L. Carlsson. Innovative approaches to anti-arrhythmic drug therapy. *Nature Reviews Drug Discovery*, 5:1034–1049, 12 2006. ISSN 1474-1776. doi: 10.1038/nrd2112.

S. A. Niederer, M. Fink, D. Noble, and N. P. Smith. A meta-analysis of cardiac electrophysiology computational models. *Experimental Physiology*, 94:486–495, 5 2009. ISSN 0958-0670. doi: 10.1113/expphysiol.2008.044610.

S. A. Niederer, J. Lumens, and N. A. Trayanova. Computational models in cardiology. *Nature Reviews Cardiology*, 16:100–111, 2019. ISSN 1759-5010. doi: 10.1038/s41569-018-0104-y.

D. Noble and R. W. Tsien. Outward membrane currents activated in the plateau range of potentials in cardiac purkinje fibres. *The Journal of Physiology*, 200(1): 205–231, 1969. doi: 10.1113/jphysiol.1969.sp008689.

D. Noble, A. Garny, and P. J. Noble. How the Hodgkin–Huxley equations inspired the Cardiac Physiome Project. *The Journal of Physiology*, 590:2613–2628, 6 2012. ISSN 0022-3751. doi: 10.1113/jphysiol.2011.224238.

T. O'Hara, L. Virág, A. Varró, and Y. Rudy. Simulation of the undiseased human cardiac ventricular action potential: Model formulation and experimental validation. *PLOS Computational Biology*, 7:e1002061, 5 2011. doi: 10.1371/journal.pcbi.1002061.

E. Passini, O. J. Britton, H. R. Lu, J. Rohrbacher, A. N. Hermans, D. J. Gallacher, R. J. H. Greig, A. Bueno-Orovio, and B. Rodriguez. Human in silico drug trials demonstrate higher accuracy than animal models in predicting clinical pro-arrhythmic cardiotoxicity. *Frontiers in Physiology*, Volume 8, 2017. ISSN 1664-042X. doi: 10.3389/fphys.2017.00668.

E. Passini, C. Trovato, P. Morissette, F. Sannajust, A. Bueno-Orovio, and B. Rodriguez. Drug-induced shortening of the electromechanical window is an effective biomarker for in silico prediction of clinical risk of arrhythmias. *British Journal of Pharmacology*, 176(19):3819–3833, 2019. doi: 10.1111/bph.14786.

P. Pathmanathan and R. A. Gray. Ensuring reliability of safety-critical clinical applications of computational cardiac models. *Frontiers in Physiology*, 4, 2013. ISSN 1664-042X. doi: 10.3389/fphys.2013.00358.

P. Pathmanathan and R. A. Gray. Verification of computational models of cardiac electro-physiology. *International Journal for Numerical Methods in Biomedical Engineering*, 30(5):525–544, 2014. doi: 10.1002/cnm.2615.

P. Pathmanathan and R. A. Gray. Validation and trustworthiness of multiscale models of cardiac electrophysiology. *Frontiers in Physiology*, 9, 2018. ISSN 1664-042X. doi: 10.3389/fphys.2018.00106.

P. Pathmanathan, M. O. Bernabeu, S. A. Niederer, D. J. Gavaghan, and D. Kay. Computational modelling of cardiac electrophysiology: explanation of the variability of results from different numerical solvers. *International Journal for Numerical Methods in Biomedical Engineering*, 28(8):890–903, 2012. doi: 10.1002/cnm.2467.

P. Pathmanathan, M. S. Shotwell, D. J. Gavaghan, J. M. Cordeiro, and R. A. Gray. Uncertainty quantification of fast sodium current steady-state inactivation for multi-scale models of cardiac electrophysiology. *Progress in Biophysics and Molecular Biology*, 117(1):4–18, 2015. ISSN 0079-6107. doi: 10.1016/j.pbiomolbio.2015.01.008. Multi-scale Systems Biology.

P. Pathmanathan, R. A. Gray, V. J. Romero, and T. M. Morrison. Applicability analysis of validation evidence for biomedical computational models. *Journal of Verification, Validation and Uncertainty Quantification*, 2, 9 2017. ISSN 2377-2158. doi: 10.1115/1.4037671.

F. Patten-Elliott, C. L. Lei, S. P. Preston, R. D. Wilkinson, and G. R. Mirams. Optimizing experimental designs for model selection of ion channel drug-binding mechanisms. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 383:20240227, 3 2025. doi: 10.1098/rsta.2024.0227.

R. A. Pearlstein, K. A. MacCannell, G. Erdemli, S. Yeola, G. Helmlinger, Q.-Y. Hu, R. Farid, W. Egan, S. Whitebread, C. Springer, J. Beck, H.-R. Wang, M. Maciejewski, L. Urban, and J. S. Duca. Implications of dynamic occupancy, binding kinetics, and channel gating kinetics for hERG blocker safety assessment and mitigation. *Current Topics in Medicinal Chemistry*, 16(16):1792–1818, 2016. ISSN 1568-0266/1873-4294. doi: 10.2174/1568026616666160315142156.

M. J. Perrin, P. W. Kuchel, T. J. Campbell, and J. I. Vandenberg. Drug binding to the inactivated state is necessary but not sufficient for high-affinity binding to human ether-à-go-go-Related gene channels. *Molecular Pharmacology*, 74(5):1443–1452, 2008. ISSN 0026-895X. doi: 10.1124/mol.108.049056.

M. Pia Saccomani, S. Audoly, and L. D'Angiò. Parameter identifiability of nonlinear systems: the role of initial conditions. *Automatica*, 39(4):619–632, 2003. ISSN 0005-1098. doi: 10.1016/S0005-1098(02)00302-3.

A. Pironet, P. D. Docherty, P. C. Dauby, J. G. Chase, and T. Desaive. Practical identifiability analysis of a minimal cardiovascular system model. *Computer Methods and Programs in Biomedicine*, 171:53–65, 2019. ISSN 0169-2607. doi: 10.1016/j.cmpb.2017.01.005.

H. Pohjanpalo. System identifiability based on the power series expansion of the solution. *Mathematical Biosciences*, 41(1):21–33, 1978. ISSN 0025-5564. doi: 10.1016/0025-5564(78)90063-9.

S. P. Preston, R. D. Wilkinson, R. H. Clayton, M. J. Chappell, and G. R. Mirams. Think before you fit: parameter identifiability, sensitivity and uncertainty in systems biology models. 2025. doi: 10.48550/arXiv.2508.18853.

A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 06 2009. ISSN 1367-4803. doi: 10.1093/bioinformatics/btp358.

A. Raue, C. Kreutz, F. J. Theis, and J. Timmer. Joining forces of bayesian and frequentist methodology: a study for inference in the presence of non-identifiability. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1984):20110544, 2013. doi: 10.1098/rsta.2011.0544.

A. Raue, J. Karlsson, M. P. Saccomani, M. Jirstrand, and Jens Timmer. Comparison of approaches for parameter identifiability analysis of biological systems. *Bioinformatics*, 30(10):1440–1448, 01 2014. ISSN 1367-4803. doi: 10.1093/bioinformatics/btu006.

U. Ravens, E. Wettwer, and O. Hála. Pharmacological modulation of ion channels and transporters. *Cell Calcium*, 35:575–582, 2004. ISSN 0143-4160. doi: 10.1016/j.ceca.2004.01.011.

W. S. Redfern, L. Carlsson, A. S. Davis, W. G. Lynch, I. MacKenzie, S. Palethorpe, P. K. S. Siegl, I. Strang, A. T. Sullivan, R. Wallis, A. J. Camm, and T. G. Hammond. 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: 32–45, 4 2003. ISSN 0008-6363. doi: 10.1016/S0008-6363(02)00846-5.

M. R. Rivaud, J. D. Bayer, M. Cluitmans, J. van der Waal, L. R. Bear, B. J. Boukens, C. Belterman, L. Gottlieb, F. Vaillant, E. Abell, R. Dubois, V. M. F. Meijborg, and R. Coronel. Critical repolarization gradients determine the induction of reentry-based torsades de pointes arrhythmia in models of long QT syndrome. *Heart Rhythm*, 18:278–287, 2 2021. ISSN 1547-5271. doi: 10.1016/j.hrthm.2020.09.020.

Y. Rudy and J. R. Silva. Computational biology in the study of cardiac ion channels and cell electrophysiology. *Quarterly Reviews of Biophysics*, 39:57–116, 2006. doi: 10.1017/S0033583506004227.

P. T. Sager, G. Gintant, J. R. Turner, S. Pettit, and N. Stockbridge. Rechanneling the cardiac proarrhythmia safety paradigm: A meeting report from the cardiac safety research consortium. *American Heart Journal*, 167(3):292–300, 2014. ISSN 0002-8703. doi: 10.1016/j.ahj.2013.11.004.

F. Sahli-Costabal, K. Seo, E. Ashley, and E. Kuhl. Classifying drugs by their arrhythmogenic risk using machine learning. *Biophysical Journal*, 118:1165–1176, 3 2020. ISSN 0006-3495. doi: 10.1016/j.bpj.2020.01.012.

A. Saltelli, S. Tarantola, and K. P.-S. Chan. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, 41(1):39–56, 1999. doi: 10.1080/00401706.1999.10485594.

M. C. Sanguinetti, C. Jiang, M. E. Curran, and M. T. Keating. A mechanistic link between an inherited and an acquired cardiac arrhythmia: hERG encodes the I_{Kr} potassium channel. *Cell*, 81:299–307, 1995. ISSN 0092-8674. doi: 10.1016/0092-8674(95)90340-2.

A. X. Sarkar, D. J. Christini, and E. A. Sobie. Exploiting mathematical models to illuminate electrophysiological variability between individuals. *The Journal of Physiology*, 590(11):2555–2567, 2012. doi: 10.1113/jphysiol.2011.223313.

A. Sertkaya, T. Beleche, A. Jessup, and B. D. Sommers. Costs of drug development and research and development intensity in the US, 2000-2018. *JAMA Network Open*, 7:e2415445–e2415445, 6 2024. ISSN 2574-3805. doi: 10.1001/jamanetworkopen.2024.15445.

A. A. Sher, K. Wang, A. Wathen, P. J. Maybank, G. R. Mirams, D. Abramson, D. Noble, and D. J. Gavaghan. A local sensitivity analysis method for developing biological models with identifiable parameters: Application to cardiac ionic channel modelling. *Future Generation Computer Systems*, 29(2):591–598, 2013. ISSN 0167-739X. doi: 10.1016/j.future.2011.09.006.

J. G. Shuttleworth, C. L. Lei, D. G. Whittaker, M. J. Windley, A. P. Hill, S. P. Preston, and G. R. Mirams. Empirical quantification of predictive uncertainty due to model discrepancy by training with an ensemble of experimental designs: An application to ion channel kinetics. *Bulletin of Mathematical Biology*, 86:2, 2023. ISSN 1522-9602. doi: 10.1007/s11538-023-01224-6.

D. Simon. *Evolutionary optimization algorithms*. Wiley-Blackwell, 2013. ISBN 0470937416.

R. C. Smith. *Uncertainty Quantification: Theory, Implementation, and Applications*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013. doi: 10.1137/1.9781611973228.

D. J. Snyders and A. Chaudhary. High affinity open channel block by dofetilide of herg expressed in a human cell line. *Molecular Pharmacology*, 49:949–955, 6 1996. ISSN 0026895X. doi: 10.1016/S0026-895X(25)09019-4.

C. F. Starmer, A. O. Grant, and H. C. Strauss. Mechanisms of use-dependent block of sodium channels in excitable membranes by local anesthetics. *Biophysical Journal*, 46:15–27, 7 1984. ISSN 0006-3495. doi: 10.1016/S0006-3495(84)83994-6. doi: 10.1016/S0006-3495(84)83994-6.

O. Stein. *Unconstrained Optimization*, pages 21–106. Springer Berlin Heidelberg, 2024. ISBN 978-3-662-69741-2. doi: 10.1007/978-3-662-69741-2.

N. Stockbridge, J. Zhang, C. Garnett, and M. Malik. Practice and challenges of thorough QT studies. *Journal of Electrocardiology*, 45(6):582–587, 2012. ISSN 0022-0736. doi: 10.1016/j.jelectrocard.2012.07.020.

D. Stork, E. N. Timin, S. Berjukow, C. Huber, A. Hohaus, M. Auer, and S. Hering. State dependent dissociation of hERG channel inhibitors. *British Journal of Pharmacology*, 151:1368–1376, 8 2007. ISSN 0007-1188. doi: 10.1038/sj.bjp.0707356.

D. G. Strauss, G. Gintant, Z. Li, W. Wu, K. Blinova, J. Vicente, J. R. Turner, and P. T. Sager. Comprehensive in Vitro Proarrhythmia Assay (CiPA) update from a Cardiac Safety Research Consortium / Health and Environmental Sciences Institute / FDA meeting. *Therapeutic Innovation & Regulatory Science*, 53:519–525, 2019. ISSN 2168-4804. doi: 10.1177/2168479018795117.

D. G. Strauss, W. W. Wu, Z. Li, J. Koerner, and C. Garnett. Translational models and tools to reduce clinical trials and improve regulatory decision making for QTc and proarrhythmia risk (ICH E14/S7B updates). *Clinical Pharmacology & Therapeutics*, 109:319–333, 2 2021. ISSN 0009-9236. doi: 10.1002/cpt.2137.

K. H. W. J. ten Tusscher and A. V. Panfilov. Alternans and spiral breakup in a human ventricular tissue model. *American Journal of Physiology-Heart and Circulatory Physiology*, 291:H1088–H1100, 9 2006. ISSN 0363-6135. doi: 10.1152/ajpheart.00109.2006.

K. H. W. J. ten Tusscher, D. Noble, P. J. Noble, and A. V. Panfilov. A model for human ventricular tissue. *American Journal of Physiology-Heart and Circulatory Physiology*, 286:H1573–H1589, 4 2004. ISSN 0363-6135. doi: 10.1152/ajpheart.00794.2003.

K. H. W. J. ten Tusscher, O. Bernus, R. Hren, and A. V. Panfilov. Comparison of electrophysiological models for human ventricular cells and tissues. *Progress*

in Biophysics and Molecular Biology, 90:326–345, 1 2006. ISSN 0079-6107. doi: 10.1016/J.PBIOMOLBIO.2005.05.015.

D. Thomas, C. Karle, and J. Kiehn. Modulation of hERG potassium channel function by drug action. *Annals of Medicine*, 36(sup1):41–46, 2004. doi: 10.1080/17431380410032580.

A. Tivay, X. Jin, A. K.-Y. Lo, C. G. Scully, and J.-O. Hahn. Practical use of regularization in individualizing a mathematical model of cardiovascular hemodynamics using scarce data. *Frontiers in Physiology*, 11, 2020. ISSN 1664-042X. doi: 10.3389/fphys.2020.00452.

J. Tomek, A. Bueno-Orovio, E. Passini, X. Zhou, A. Minchole, O. Britton, C. Bartolucci, S. Severi, A. Shrier, L. Virag, A. Varro, and B. Rodriguez. Development, calibration, and validation of a novel human ventricular myocyte model in health, disease, and drug block. *eLife*, 8:e48890, 2019. ISSN 2050-084X. doi: 10.7554/eLife.48890.

J. Tomek, A. Bueno-Orovio, and B. Rodriguez. ToR-ORd-dynCl: an update of the ToR-ORd model of human ventricular cardiomyocyte with dynamic intracellular chloride. *bioRxiv*, 2020. doi: 10.1101/2020.06.01.127043.

N. A. Trayanova, A. Lyon, J. Shade, and J. Heijman. Computational modeling of cardiac electrophysiology and arrhythmogenesis: toward clinical translation. *Physiological Reviews*, 104(3):1265–1333, 2024. doi: 10.1152/physrev.00017.2023.

K. Tsujimae, S. Suzuki, S. Murakami, and Y. Kurachi. Frequency-dependent effects of various I_{Kr} blockers on cardiac action potential duration in a human atrial model. *American Journal of Physiology-Heart and Circulatory Physiology*, 293:H660–H669, 7 2007. ISSN 0363-6135. doi: 10.1152/ajpheart.01083.2006.

J. I. Vandenberg, A. Varghese, Y. Lu, J. A. Bursill, M. P. Mahaut-Smith, and C. L.-H. Huang. Temperature dependence of human *ether-à-go-go*-related gene K^+ currents. *American Journal of Physiology-Cell Physiology*, 291(1):C165–C175, 2006. doi: 10.1152/ajpcell.00596.2005.

J. I. Vandenberg, M. D. Perry, M. J. Perrin, S. A. Mann, Y. Ke, and A. P. Hill. hERG K^+ channels: Structure, function, and clinical significance. *Physiological Reviews*, 92:1393–1478, 7 2012. ISSN 0031-9333. doi: 10.1152/physrev.00036.2011.

J. I. Vandenberg, E. Perozo, and T. W. Allen. Towards a structural view of drug binding to hERG K^+ channels. *Trends in Pharmacological Sciences*, 38:899–907, 10 2017. ISSN 1873-3735. doi: 10.1016/j.tips.2017.06.004.

N. Vandersickel, T. P. de Boer, M. A. Vos, and A. V. Panfilov. Perpetuation of torsade de pointes in heterogeneous hearts: competing foci or re-entry? *The Journal of Physiology*, 594(23):6865–6878, 2016. doi: 10.1113/JP271728.

P. G. A. Volders, M. A. Vos, B. Szabo, K. R. Sipido, S. H. M. de Groot, A. P. M. Gorgels, H. J. J. Wellens, and R. Lazzara. Progress in the understanding of cardiac early afterdepolarizations and torsades de pointes: time to revise current concepts. *Cardiovascular Research*, 46(3):376–392, 06 2000. ISSN 0008-6363. doi: 10.1016/S0008-6363(00)00022-5.

S. Wang, S. Liu, M. J. Morales, H. C. Strauss, and R. L. Rasmusson. A quantitative analysis of the activation and inactivation kinetics of HERG expressed in xenopus oocytes. *The Journal of Physiology*, 502(1):45–60, 1997a. doi: 10.1111/j.1469-7793.1997.045bl.x.

S. Wang, M. J. Morales, S. Liu, H. C. Strauss, and R. L. Rasmusson. Modulation of HERG affinity for E-4031 by $[K^+]_o$ and C-type inactivation. *FEBS Letters*, 417(1): 43–47, 1997b. ISSN 0014-5793. doi: 10.1016/S0014-5793(97)01245-3.

W. Wang and R. MacKinnon. Cryo-EM structure of the open human *ether-à-go-go-related* K^+ channel hERG. *Cell*, 169:422–430.e10, 4 2017. ISSN 0092-8674. doi: 10.1016/j.cell.2017.03.048.

Y. Wang, J. Guo, L. L. Perissinotti, J. Lees-Miller, G. Teng, S. Durdagi, H. J. Duff, and S. Yu. Noskov. Role of the pH in state-dependent blockade of hERG currents. *Scientific Reports*, 6:32536, 2016. ISSN 2045-2322. doi: 10.1038/srep32536.

J. N. Weiss, A. Garfinkel, H. S. Karagueuzian, P.-S. Chen, and Z. Qu. Early afterdepolarizations and cardiac arrhythmias. *Heart Rhythm*, 7(12):1891–1899, 2010. ISSN 1547-5271. doi: 10.1016/j.hrthm.2010.09.017.

D. G. Whittaker, M. Clerx, C. L. Lei, D. J. Christini, and G. R. Mirams. Calibration of ionic and cellular cardiac electrophysiology models. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 12(4):e1482, 2020. doi: 10.1002/wsbm.1482.

F.-G. Wieland, A. L. Hauber, M. Rosenblatt, C. Tönsing, and J. Timmer. On structural and practical identifiability. *Current Opinion in Systems Biology*, 25:60–69, 2021. ISSN 2452-3100. doi: 10.1016/j.coisb.2021.03.005.

A. Windisch, E. N. Timin, T. Schwarz, D. Stork-Riedler, T. Erker, G. F. Ecker, and S. Hering. Trapping and dissociation of propafenone derivatives in hERG channels. *British Journal of Pharmacology*, 162(7):1542–1552, 2011. doi: 10.1111/j.1476-5381.2010.01159.x.

M. J. Windley, S. A. Mann, J. I. Vandenberg, and A. P. Hill. Temperature effects on kinetics of $K_V11.1$ drug block have important consequences for in silico proarrhythmic risk prediction. *Molecular Pharmacology*, 90:1–11, 7 2016. ISSN 0026895X. doi: 10.1124/mol.115.103127.

B. Wiśniowska and S. Polak. The role of interaction model in simulation of drug interactions and QT prolongation. *Current Pharmacology Reports*, 2:339–344, 2016. ISSN 2198-641X. doi: 10.1007/s40495-016-0075-9.

O. J. Wouters, M. McKee, and J. Luyten. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. *JAMA*, 323: 844–853, 3 2020. ISSN 0098-7484. doi: 10.1001/jama.2020.1166.

T. Yang, D. J. Snyders, and D. M. Roden. Ibutilide, a methanesulfonanilide antiarrhythmic, is a potent blocker of the rapidly activating delayed rectifier K^+ current (I_{Kr}) in AT-1 cells. *Circulation*, 91(6):1799–1806, 1995. doi: 10.1161/01.CIR.91.6.1799.

J. A. Yao, X. Du, D. Lu, R. L. Baker, E. Daharsh, and P. Atterson. Estimation of potency of hERG channel blockers: Impact of voltage protocol and temperature. *Journal of Pharmacological and Toxicological Methods*, 52:146–153, 7 2005. ISSN 1056-8719. doi: 10.1016/J.VASCN.2005.04.008.

J. W. T. Yates, N. D. Evans, and M. J. Chappell. Structural identifiability analysis via symmetries of differential equations. *Automatica*, 45(11):2585–2591, 2009. ISSN 0005-1098. doi: 10.1016/j.automatica.2009.07.009.

G. Yellen. The moving parts of voltage-gated ion channels. *Quarterly Reviews of Biophysics*, 31, 8 1998. ISSN 0033-5835. doi: 10.1017/S0033583598003448.

T. Yu, C. M. Lloyd, D. P. Nickerson, M. T. Cooling, A. K. Miller, A. Garny, J. R. Terkildsen, J. Lawson, R. D. Britten, P. J. Hunter, and P. M. F. Nielsen. The Physiome Model Repository 2. *Bioinformatics*, 27(5):743–744, 01 2011. ISSN 1367-4803. doi: 10.1093/bioinformatics/btq723.

M. Zaniboni, A. E. Pollard, L. Yang, and K. W. Spitzer. Beat-to-beat repolarization variability in ventricular myocytes and its suppression by electrical coupling. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(3):H677–H687, 2000. doi: 10.1152/ajpheart.2000.278.3.H677.

S. Zhang, Z. Zhou, Q. Gong, J. C. Makielski, and C. T. January. Mechanism of block and identification of the verapamil binding domain to hERG potassium channels. *Circulation Research*, 84:989–998, 5 1999. doi: 10.1161/01.RES.84.9.989.

Pz. Zhou, J. Babcock, Lq. Liu, M. Li, and Zb. Gao. Activation of human ether-a-go-go related gene (hERG) potassium channels by small molecules. *Acta Pharmacologica Sinica*, 32:781–788, 2011. ISSN 1745-7254. doi: 10.1038/aps.2011.70.

Z. Zhou, Q. Gong, B. Ye, Z. Fan, J. C. Makielski, G. A. Robertson, and C. T. January. Properties of HERG channels stably expressed in HEK 293 cells studied at physiological temperature. *Biophysical Journal*, 74:230–241, 1 1998. ISSN 0006-3495. doi: 10.1016/S0006-3495(98)77782-3.

D. P. Zipes and H. J. J. Wellens. Sudden cardiac death. *Circulation*, 98(21):2334–2351, 1998. doi: 10.1161/01.CIR.98.21.2334.

B. J. Zünkler. Human ether-a-go-go-related (HERG) gene and ATP-sensitive potassium channels as targets for adverse drug effects. *Pharmacology & Therapeutics*, 112(1):12–37, 2006. ISSN 0163-7258. doi: 10.1016/j.pharmthera.2006.03.002.