IISE Transactions on Healthcare Systems Engineering

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/uhse21

Model-guided concurrent data assimilation for calibrating cardiac ion-channel kinetics

Haedong Kim, Hui Yang, Andrew R. Ednie & Eric S. Bennett

To cite this article: Haedong Kim, Hui Yang, Andrew R. Ednie & Eric S. Bennett (2024) Model-guided concurrent data assimilation for calibrating cardiac ion-channel kinetics, IISE Transactions on Healthcare Systems Engineering, 14:2, 153-166, DOI: 10.1080/24725579.2023.2239271

To link to this article: https://doi.org/10.1080/24725579.2023.2239271

Published online: 15 Aug 2023.

Submit your article to this journal

Article views: 54

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=uhse21

IISE TRANSACTIONS ON HEALTHCARE SYSTEMS ENGINEERING
2024, VOL. 14, NO. 2, 153–166
https://doi.org/10.1080/24725579.2023.2239271

Taylor & Francis
Taylor & Francis Group

Check for updates

Model-guided concurrent data assimilation for calibrating cardiac ion-channel kinetics

Haedong Kim[a], Hui Yang[a], Andrew R. Ednie[b], and Eric S. Bennett[b]

[a]Complex Systems Monitoring, Modeling, and Control Laboratory, The Pennsylvania State University, University Park, PA, USA; [b]Department of Neuroscience, Cell Biology, and Physiology, Boonshoft School of Medicine and College of Science & Mathematics, Wright State University, Dayton, OH, USA

ABSTRACT

Potassium channels (K_v) are responsible for repolarizing the action potential in cardiomyocytes. There is a variety of K_v isoforms and corresponding currents (e.g. I_{Kto}, I_{Kslow1}, I_{Kslow2}) that contribute to different phases of repolarization. Because only the sum of their activities can be measured in the form of currents (I_{Ksum}), there is a need to delineate individual K^+ currents. Most existing studies make inference of K_v activities *via* curve-fitting procedures but encounter certain limitations as follows: (1) curve-fitting decomposition only relies on the shape of K^+ current traces, which does not discern the underlying kinetics; (2) I_{Ksum} traces can only be fitted for one clamp voltage at each time, and then analyzed in a population-averaged way later. This paper presents a novel concurrent data assimilation method to calibrate biophysics-based models and delineate kinetics of K_v isoforms with multiple voltage-clamp responses simultaneously. The proposed method is evaluated and validated with whole-cell I_{Ksum} recordings from wild-type and chronically glycosylation-deficient cardiomyocytes. Experimental results show that the proposed method effectively handles multiple-response data and describes glycosylation-conferred perturbations to K_v isoforms. Further, we develop a graphical-user-interface (GUI) application that provides an enabling tool to biomedical scientists for data-driven modeling and analysis of K_v kinetics in various heart diseases.

KEYWORDS

Data assimilation; simulation modeling; model calibration; cardiac potassium channels; congenital disorders of glycosylation

1. Introduction

Potassium channels (K_v) play critical roles in the electrical conduction system of the heart, particularly in the repolarization phase of the action potential (AP). The AP is a change of membrane potential over time, representing the net electrical activity in a cardiomyocyte, and the AP shape and duration are primarily determined by K_v isoforms. The ability of the heart to pump blood through the body in an appropriate rhythm is controlled by electrical signaling. Hence, even modest changes in K_v activities can significantly affect the AP duration and the QT interval, which lead to fatal heart diseases (Ravens & Cerbai, 2008). As shown in Fig. 1, the different phases of AP repolarization in mouse cardiomyocytes are the result of the coordinated activity of a variety of K_v isoforms (e.g. K_v4.2, K_v1.5, and K_v2.1) and their corresponding currents (e.g. I_{Kto}, I_{Kslow1}, and I_{Kslow2}). There are other major voltage-gated ion channels (VGICs), such as for Na^+ (Na_v) and Ca^{2+} (Ca_v), that also contribute to the AP. The Na_v are primarily responsible for the AP upstroke in the depolarization phase, while the Ca_v control the cellular contraction. Aberrant activities of ion channels can significantly impact the AP and lead to fatal arrhythmias. Using whole-cell voltage-clamp recording methods, only the collective activities of K_v isoforms can be measured reliably as the sum of all K^+ currents (I_{Ksum}), and these K_v isoforms have overlapping biophysical properties (i.e. voltage dependence of gating and kinetics), which complicate the ability to separate one type of I_K from another (Brouillette et al., 2004). K_v activities can be altered in various cardiomyopathies (Giudicessi & Ackerman, 2012; Tristani-Firouzi et al., 2001), thus there is an urgent need to decompose I_{Ksum} more rigorously into individual K^+ currents and estimate gating kinetics of K_v isoforms to understand their pathological roles by comparing their activities in healthy versus diseased hearts and cardiomyocytes.

Transgenic mouse models are the most commonly used animal models that have provided insights into cardiac research despite the differences between rodents and humans; the isoforms expressed in both small and large mammals share many similarities (Milani-Nejad & Janssen, 2014). Most existing studies using mouse models separate individual K^+ currents mathematically *via* a curve-fitting procedure (Costantini et al., 2005; Ednie et al., 2019; Teng et al., 2022), assuming the I_{Ksum} current trace represents the summation of exponential decay functions (two or three, usually) and a constant term; each represents the shape of individual K^+ currents. However, the ability to scrutinize the kinetics and dynamics of individual K_v isoforms is limited by this traditional method because: (1) curve-fitting decomposition only relies on the shape of K^+ current traces

CONTACT Hui Yang huy25@psu.edu Complex Systems Monitoring, Modeling, and Control Laboratory, The Pennsylvania State University, University Park, PA, USA.

© 2023 "IISE"

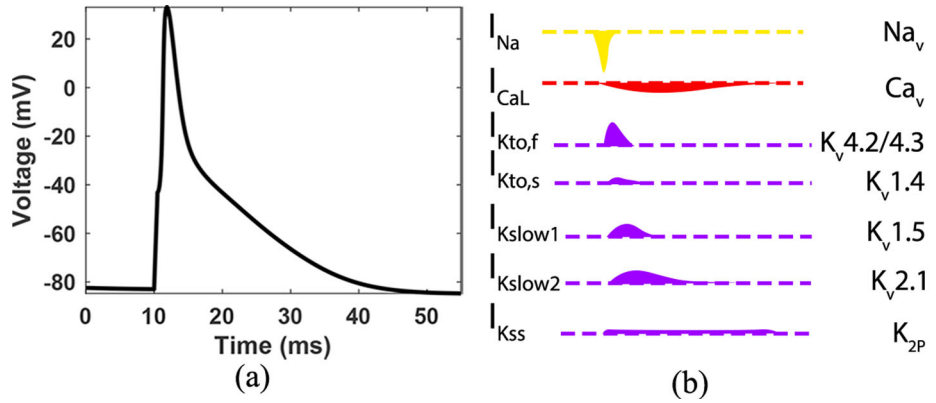

Figure 1. (a) Ventricular action potential in mouse cardiomyocytes and (b) diverse K_v isoforms and their currents contributing to the AP shape.

as discreet exponential functions which does not discern the underlying kinetics and interactions of K_v isoforms; (2) I_{Ksum} traces can only be decomposed for one clamp voltage at a time, and estimated information from each fitting is analyzed in a population-averaged way later. New methods are urgently needed to simultaneously delineate the underlying kinetics of multiple I_{Ksum} traces from the same cell concurrently for a better understanding of K_v channels and their roles as well as cellular variability in diseased cardiomyocytes.

This paper presents a new approach of subject-specific concurrent data assimilation for model-guided learning of K_v isoforms. First, computer models of K_v and their currents are designed with parameters that control the kinetic rates and, in turn, determine the currents. It is worth mentioning that K^+ currents are generated based on the mathematically simulated gating mechanism. Second, we perform a sensitivity analysis of the kinetic parameters using fractional factorial designs to identify the parameters that have significant effects on the current generation. Several markers are defined that best represent the characteristics of the K^+ currents in different perspectives. Third, a calibration routine is proposed to adjust parameters selected by the sensitivity analysis to couple *in-silico* models with *in-vitro* data of I_{Ksum} recordings from the same cell concurrently. Last, low-dimensional embedding is utilized to visualize calibrated parameters in a collective way across the cells. The proposed methodology is evaluated and validated by comparing I_{Ksum} recordings on healthy control versus chronically glycosylation-deficient cardiomyocyte. Glycosylation is a co/post-translational modification that is critical for protein functions including activities of K_v and other VGICs (Ednie & Bennett, 2012; Ohtsubo & Marth, 2006). Our in-vitro experiments showed that chronic reduction in cardiomyocyte N-glycosylation modulates VGIC activities and contribute to both electrical and contractile dysfunction (Ednie et al., 2013; Ednie & Bennett, 2015; Ednie et al., 2019). In fact, these studies also showed that reduced cardiomyocyte complex N-glycosylation is sufficient to cause dilated cardiomyopathy (DCM), which is the third most common cause of heart failure and the major reason for heart transplantation (Weintraub et al., 2017).

Experimental results show that the proposed method provides reliable estimations of K_v currents with underlying kinetics and successfully captures cellular variability. For example, it verifies the experimental data that showed the significant reduction in the current magnitude and elongated decaying for the glycosylation-deficient cardiomyocyte group. Additionally, computational modeling estimates channel kinetics, which was not possible in the electrophysiological experiments alone. Cellular variability is visualized by low-dimensional embedding of calibration parameters into 3D space, which shows one of the powers of systematic computer models that allow one to predict how changes in VGICs impact cellular functions. Further, we develop a graphical-user-interface (GUI) application to make the proposed method accessible to biomedical scientists for investigating K_v-related channelopathy. The proposed method shows strong potential for modeling the kinetics and gating mechanism of K_v to study heart diseases. Our contributions are summarized as follows:

- We model the underlying kinetics of K_v isoforms from I_{Ksum} recordings, which enables to simulate individual K^+ currents based on biophysical principles.
- We develop a subject-specific concurrent data assimilation routine that learns I_{Ksum} recordings measured from the same cell at multiple membrane potentials simultaneously rather than a single I_{Ksum} trace independently to model cellular-level dynamics.
- The prediction uncertainty is quantified and the distribution of kinetic parameters is visualized using low-dimensional embedding.
- A GUI application is provided to make the proposed method an enabling tool for biomedical scientists.

The software package of the GUI application, tutorial, and reproducible MATLAB codes for modeling and analysis results are available at https://github.com/haedong31/Kv_data_assim.

2. Research background

2.1. Data assimilation and calibration of cardiac models

The functionality of the heart to pump blood through the body is controlled by electrical signals generated by the organ itself. Cardiac electrophysiology has contributed significantly

to our understanding of the heart and disease-related modifications, such as congenital disorders of glycosylation (CDG). Advancements in laboratory techniques allow researchers to study molecular-level activities in cardiomyocytes *via* measuring ionic currents conducted through VGICs. For example, whole-cell current recordings are used to show the functions of a certain gene that encodes α subunits of Na_v in human cardiomyocytes, leading to a small but inherent and chronic risk of acquired arrhythmia (Splawski et al., 2002); pathophysiological roles of reduced sialylation impacting Na_v and K_v activities in mouse cardiomyocytes (Ednie et al., 2015; Ednie & Bennett, 2015); modulation of O-glycosylation causing aberrant K_v activities (Schwetz et al., 2011). However, in-vitro experiments alone are limited to investigate detailed channel activities. There is a need to assimilate experimental data to estimate information that are not able to be observed directly. Traditionally, a curve-fitting method is used to make inferences, which assumes a functional form of information to estimate (e.g. individual currents of K_v isoforms) (Liu et al., 2011).

In contrast, computer models consist of mathematical equations describing biophysical properties and physiological functions compatible with experimental observations. Mathematical and computational modeling allows for studying the heart in a quantitative and predictive way (Whittaker et al., 2020). Computer models, coupled with experimental observations, provide integrative insights into the data by calibrating model parameters (Winslow et al., 2011). For example, models can be used to interpolate processes not directly observed in experiments and extrapolate to novel conditions such as disease-related perturbations (Rodriguez et al., 2010). From an engineering perspective, computer models of ion channels are dynamic system simulations that represent continuously changing gating kinetics. The dynamic nature of the models complicates the calibration process. Complex nonlinear structures of cardiac computer models also make calibration difficult. We developed statistical metamodeling and sequential design (Du et al., 2016), nonlinear optimization algorithms (Du et al., 2014, 2018), and a heuristic optimization method (Kim et al., 2022) in our previous in-silico studies to cope with the complexity of cardiac models. However, all these studies are based on current-shape parameters *via* curve fitting, or characteristic curves/statistics from additional curve fitting applied to estimated currents. In addition, only population-average models are derived and calibrated *via* learning the entire dataset for each healthy and diseased group. Little work has been done to leverage computer models to calibrate and delineate kinetic dynamics of K_v isoforms with multiple voltage-clamp responses simultaneously.

2.2. Congenital disorders of glycosylation

Protein glycosylation is one of the most abundant and diverse forms of co/posttranslational modifications that impact essential protein functions, such as modulation of receptor or ion channel activities (Ednie & Bennett, 2012; Ohtsubo & Marth, 2006). A growing number of studies have shown the association between altered glycosylation and heart diseases, such as dilated cardiomyopathy (DCM) and hypertrophic cardiomyopathy (Ednie et al., 2019; Ohtsubo & Marth, 2006). It is reported that up to 20% of patients with congenital disorders of glycosylation (CDG), who commonly show modest reductions in protein glycosylation, present with cardiac deficits, including idiopathic DCM (Marques-da Silva et al., 2017). However, uncovering the underlying pathological mechanisms still remains elusive. We have investigated how regulated glycosylation contributes to heart failure in the context of electrophysiology. Electrical signaling is orchestrated activities of a variety of ion channels and transporters. VGICs are heavily glycosylated, with 30% of the channel mass consisting of N-/O-linked glycans (Ednie & Bennett, 2012). Glycosylation is a multi-step process and usually ends with sialic acid added. We reported that a saturating, electrostatic effect of negatively charged sialic attached to the terminal of N-/O-glycan branches significantly altered electrical signaling in Na_v (Ednie et al., 2013, 2015) as well as K_v (Ednie & Bennett, 2015). Computational modeling has been used to further investigate the functional role of reduced sialylation in Na_v and K_v (Du et al., 2016, 2018).

3. Research methodology

3.1. Computer models of potassium channel isoforms

Because of the relative ease at which their genetics can be altered, mice have been useful models for studying cardiac electric signaling (Milani-Nejad & Janssen, 2014; Nerbonne, 2004). In ventricular and atrial mouse cardiomyocytes, there are three major components of I_{Ksum}: rapidly inactivating transient outward K^+ currents ($I_{Kto,f}$ and/or $I_{Kto,s}$), which are conducted through $K_v4.2$ and $K_v1.4$ respectively, slowly inactivating delayed rectifier K^+ currents (I_{Kslow1} and I_{Kslow2}), which are conducted through $K_v1.5/K_v2.1$, and the non-inactivating steady-state K^+ current (I_{Kss}) (Xu et al., 1999). $I_{Kto,s}$ ($K_v1.4$) is primarily found in septal cardiomyocytes (Bondarenko et al., 2004). In this study, we only include $I_{Kto,f}$ in I_{Kto}, because our in-vitro studies focused on the ventricular apex myocytes. However, I_{Kto} can be easily modified according to the region of ventricular cardiomyocytes. We model these four dominant K^+ currents: I_{Kto}, I_{Kslow1}, I_{Kslow2}, and I_{Kss}. Figure 2(b) illustrates the shape of the primary K^+ currents and their contribution to the whole-cell I_K trace (I_{Ksum}) given the protocol in Fig. 2(a) that applies 0 mV voltage pulse from holding potential −70 mV from time $t^{(h)}$ to $t^{(e)}$. Figure 2(c,d) shows a range of voltage steps from −30 mV to 50 mV in 10 mV increments and consequent changes in I_{Ksum}. Because of their shape, I_{Kss} is assumed as a constant and the other currents an exponential function in the traditional data-driven curve-fitting approach. In this approach, if all four current models are included, I_{Ksum} is defined by (1) (i.e. tri-exponential fitting) with the shape parameters such as amplitude A_i and time constant τ_i for $i \in \{\text{Kto}, \text{Kslow1}, \text{Kslow2}, \text{Kss}\}$. In some cases, it is reduced to a bi-exponential function combining

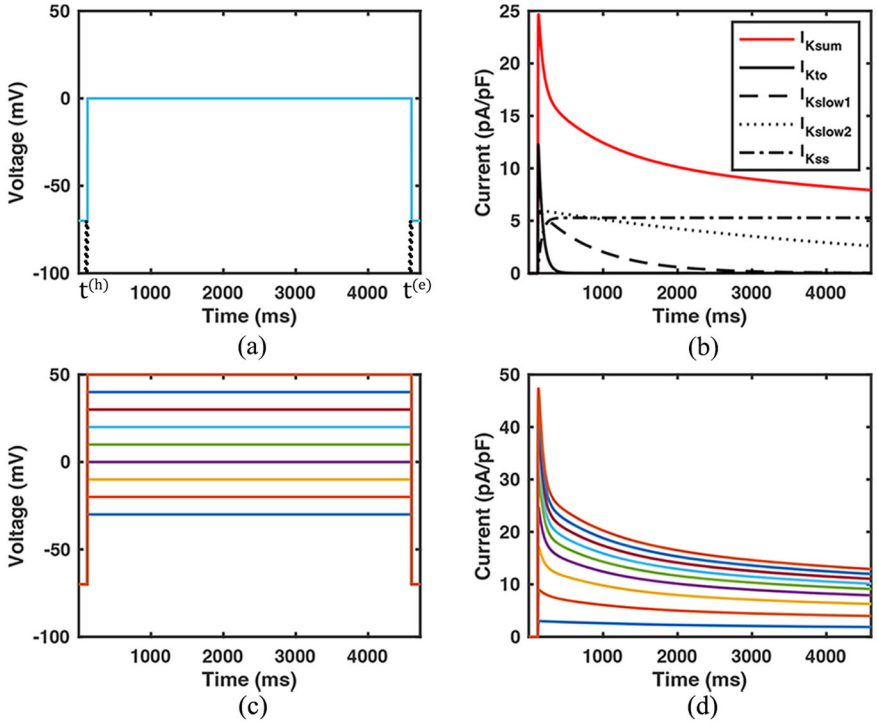

Figure 2. Example of voltage-clamp protocol and K^+ current traces. (a) Clamp-voltage pulse of 0 mV from the holding potential -70 mV. (b) Dominant K^+ currents and their contributions to I_{Ksum}. (c) Series of clamp-voltage pulses (-30–50 mV in 10 mV increments) from the holding potential -70 mV and (d) consequent I_{Ksum} traces.

I_{Kslow1} and I_{Kslow2}.

$$I_{Ksum} = A_{Kto}e^{-t/\tau_{Kto}} + A_{Kslow1}e^{-t/\tau_{Kslow1}} + A_{Kslow2}e^{-t/\tau_{Kslow2}} + A_{Kss}. \quad (1)$$

We developed mouse K_v models based on (Asfaw et al., 2020; Bondarenko, 2014), using the Hodgkin-Huxley modeling scheme that has been used for various species, such as humans (Ten Tusscher et al., 2004) and rabbits (Mahajan et al., 2008). This type of model consists of two gating variables controlling the channel conductance, and its canonical form is defined by (2)

$$I_K = G_K a^n i^m (V - E_K), \quad (2)$$

where G_K is the maximum conductance, a^n and i^m are the gating variables for $n, m \in \mathbb{N}$, V is the transmembrane potential, and E_K is the K^+ Nernst potential. $V - E_K$ implies the driving force of the ion movement. Important components in (2) are the gating variables a and i, representing the fraction of activation and recovery from inactivation of the channel where $a, i \in [0, 1]$. These processes are governed by first-order kinetics and voltage-dependent transition rates α and β. α is the rate at which a gate in a closed state opens, whereas β is the rate at which a gate in an open state closes. Equation (3) shows a schematic relationship of this gating kinetics. n and m represent the numbers of activation and inactivation gates, which are dependent on a particular K_v isoform and determined by specific kinetics and properties of that channel. In general, the number of gates depends on the complexity of the channel behavior and the level of detail required in the model.

$$(1-a) \underset{\beta_a}{\overset{\alpha_a}{\rightleftharpoons}} a \qquad (1-i) \underset{\beta_i}{\overset{\alpha_i}{\rightleftharpoons}} i. \quad (3)$$

These two biophysical processes a and i can be expressed using differential equations in two ways (4) and (5)

$$\frac{da}{dt} = \alpha_a(1-a) - \beta_a a \qquad \frac{di}{dt} = \alpha_i(1-i) - \beta_i i \quad (4)$$

$$\frac{da}{dt} = \frac{a_\infty - a}{\tau_a} \qquad \frac{di}{dt} = \frac{i_\infty - i}{\tau_i}, \quad (5)$$

where a_∞ and i_∞ are the steady-state values to which a and b converge; τ_a and τ_i are time constants determine the convergence speed defined by

$$a_\infty = \frac{\alpha_a}{\alpha_a + \beta_a} \qquad i_\infty = \frac{\alpha_i}{\alpha_i + \beta_i} \quad (6)$$

$$\tau_a = \frac{1}{\alpha_a + \beta_a} \qquad \tau_i = \frac{1}{\alpha_i + \beta_i}. \quad (7)$$

The steady-state values and time constants can be defined directly by functions of voltage V without transition rates in some cases. These voltage-dependent functions, such as transition rates, steady states, or time constants, have parameters p_s that control the behavior of the kinetics of an ion channel.

3.1.1. In-silico modeling of I_{Kto}

The rapidly inactivating transient outward current I_{Kto}, which is conducted through $K_v4.2$, is characterized by a sharp upstroke during activation and subsequent rapid inactivation. It mainly contributes to the peak at the very beginning of activation in I_{Ksum}. I_{Kto} is defined by

$$I_{Kto} = G_{Kto} a_{Kto}^3 i_{Kto} (V - E_K) \quad (8)$$

$$\frac{da_{\mathrm{Kto}}}{dt} = \alpha_a(1 - a_{\mathrm{Kto}}) - \beta_a a_{\mathrm{Kto}} \tag{9}$$

$$\frac{di_{\mathrm{Kto}}}{dt} = \alpha_i(1 - i_{\mathrm{Kto}}) - \beta_i i_{\mathrm{Kto}} \tag{10}$$

$$\alpha_a = p_7 e^{p_5(V+p_1)} \tag{11}$$

$$\beta_a = p_8 e^{-p_6(V+p_1)} \tag{12}$$

$$\alpha_i = \frac{p_9 e^{-(V+p_2)/p_4}}{1 + p_{10} e^{-(V+p_2+p_3)/p_4}} \tag{13}$$

$$\beta_i = \frac{p_{11} e^{(V+p_2+p3)/p_4}}{1 + p_{12} e^{(V+p_2+p_3)/p_4}}. \tag{14}$$

There are two gating variables a_{Kto} and i_{Kto} responsible for activation and inactivation, respectively. Their kinetics are governed by transition-rate functions from (11) to (14). Parameters p_s for $s = \{1, 2, ..., 12\}$ in these equations act like knobs, allowing to control the behavior of the $\mathrm{I_{Kto}}$ model.

3.1.2. In-silico modeling of I_{Kslow1}, I_{Kslow2}, and I_{Kss}

There are two major delayed rectifier currents in mouse ventricular cardiomyocytes that are rapidly activating and slowly inactivating: $\mathrm{I_{Kslow1}}$ and $\mathrm{I_{Kslow2}}$, which are conducted through $\mathrm{K_v}1.5$ and $\mathrm{K_v}2.1$, respectively. As shown in Fig. 2(b), both rectifier currents inactivate slower and have smaller magnitude than $\mathrm{I_{Kto}}$, and $\mathrm{I_{Kslow2}}$ decays more gradually than $\mathrm{I_{Kslow1}}$ (Liu et al., 2011). The non-inactivating steady-state current $\mathrm{I_{Kss}}$, which is likely conducted through $\mathrm{K_v}$ isoforms $\mathrm{K_{2P}}$ family (Feliciangeli et al., 2015), remains constant during the voltage-clamp recording. These three currents contribute to most part of the decaying portion of $\mathrm{I_{Ksum}}$. To keep the models as simple as possible to reduce the structural risk of overfitting, we assume that $\mathrm{I_{Kslow1}}$ and $\mathrm{I_{Kslow2}}$ have the same activation gating variable, and they have a similar inactivation pattern; $\mathrm{I_{Kss}}$ has similar activation behavior with the two delayed rectifier currents but a slightly different rate.

$\mathrm{I_{Kslow1}}$, $\mathrm{I_{Kslow2}}$, and $\mathrm{I_{Kss}}$ are modeled without transition-rate functions as opposed to $\mathrm{I_{Kto}}$, and their steady-state and time-constant functions are directly defined. First, the gating variables of $\mathrm{I_{Kslow1}}$, activation a_{Kslow1}, and inactivation i_{Kslow1}, are defined by (16) and (17). The steady-state functions (a_{ss} and i_{ss}) in (18) and (19) will be shared with the other two current models. A full description of $\mathrm{I_{Kslow1}}$ is given as follows:

$$\mathrm{I_{Kslow1}} = G_{\mathrm{Kslow1}} a_{\mathrm{Kslow1}} i_{\mathrm{Kslow1}} (V - E_{\mathrm{K}}) \tag{15}$$

$$\frac{da_{\mathrm{Kslow1}}}{dt} = \frac{a_{ss} - a_{\mathrm{Kslow1}}}{\tau_a^{(\mathrm{Kslow1})}} \tag{16}$$

$$\frac{di_{\mathrm{Kslow1}}}{dt} = \frac{i_{ss} - i_{\mathrm{Kslow1}}}{\tau_i^{(\mathrm{Kslow1})}} \tag{17}$$

$$a_{ss} = \frac{1}{1 + e^{-(V+p_1)/p_4}} \tag{18}$$

$$i_{ss} = \frac{1}{1 + e^{(V+p_2)/p_5}} \tag{19}$$

$$\tau_a^{(\mathrm{Kslow1})} = \frac{p_7}{e^{p_6(V+p_3)} + e^{-p_6(V+p_3)}} + p_9 \tag{20}$$

$$\tau_i^{(\mathrm{Kslow1})} = p_{10} - p_8 i_{ss}. \tag{21}$$

As a structural regularization, $\mathrm{I_{Kslow2}}$ has the same activation variable with $\mathrm{I_{Kslow1}}$ as in (23), and the time-constant function of the inactivation i_{Kslow2} shares the same steady-state function i_{ss} (19) with $\mathrm{I_{Kslow1}}$. As a result of this modeling strategy, mathematical equations of $\mathrm{I_{Kslow2}}$ are given as follows:

$$\mathrm{I_{Kslow2}} = G_{\mathrm{Kslow2}} a_{\mathrm{Kslow2}} i_{\mathrm{Kslow2}} (V - E_{\mathrm{K}}) \tag{22}$$

$$a_{\mathrm{Kslow2}} = a_{\mathrm{Kslow1}} \tag{23}$$

$$\frac{di_{\mathrm{Kslow2}}}{dt} = \frac{i_{ss} - i_{\mathrm{Kslow2}}}{\tau_i^{(\mathrm{Kslow2})}} \tag{24}$$

$$\tau_i^{(\mathrm{Kslow2})} = p_2 - p_1 i_{ss}. \tag{25}$$

$\mathrm{I_{Kss}}$ does not have an inactivation variable because it is non-inactivating current. It shares the same steady-state function for activation a_{ss} (18) with the other two delayed rectifier currents but have a separate time-constant function (28) to address the different activation rate. $\mathrm{I_{Kss}}$ is modeled as

$$\mathrm{I_{Kss}} = G_{\mathrm{Kss}} a_{\mathrm{Kss}} (V - E_{\mathrm{K}}) \tag{26}$$

$$\frac{da_{\mathrm{Kss}}}{dt} = \frac{a_{ss} - a_{\mathrm{Kss}}}{\tau_a^{(\mathrm{Kss})}} \tag{27}$$

$$\tau_a^{(\mathrm{Kss})} = \frac{p_2}{e^{p_1(V+p_3')} + e^{-p_1(V+p_3')}} + p_3. \tag{28}$$

Note that p_3' is equal to p_3 in $\mathrm{I_{Kslow1}}$.

3.2. Concurrent assimilation of functional data

Data assimilation is a systematic procedure to find the optimal configuration and state of computational/mathematical models by coupling them with experimental data. Experimental data $\mathcal{D}$ are observations of a real process $\mathcal{R}$ that represents scientific phenomena under investigation. The output of physical experiments $y^{\mathcal{D}}(x)$, given input x, inevitably contains errors for various reasons, such as noise in measurement or experimental environment. Suppose $\mathcal{D}$ and $\mathcal{R}$ can be related as follows in (29), where ϵ is the error term.

$$y^{\mathcal{R}}(x) = y^{\mathcal{D}}(x) + \epsilon \tag{29}$$

Let $y^{\mathcal{M}}(x|\theta)$ denote the output from a computer model $\mathcal{M}$, given parameters θ. Assume that there are discrepancies $\delta(x|\theta)$ for the current states of parameters as follows in (30):

$$y^{\mathcal{D}}(x) \quad = y^{\mathcal{M}}(x|\theta) + \delta(x|\theta), \text{ so} \tag{30}$$

$$y^{\mathcal{R}}(x) = y^{\mathcal{M}}(x|\theta) + \delta(x|\theta) + \epsilon. \tag{31}$$

Our goal in data assimilation is to calibrate θ to find the best model states that minimize $\delta(x|\theta)$, while satisfying biophysical constraints. By doing that, *in-silico* models $\mathcal{M}$ are coupled with *in-vitro* experimental data $\mathcal{D}$, which provides two complementary angles to study the real process $\mathcal{R}$.

From this perspective, bi-/tri-exponential function in (1) serves as $y^{\mathcal{M}}(t, v|\theta)$ in the curve-fitting approach, where $\theta = \{\mathrm{A}_i(v), \tau_i(v)\}$ for $i \in \{\mathrm{Kto}, \mathrm{Kslow1}, \mathrm{Kslow2}, \mathrm{Kss}\}$. v represents voltage. Note that $\mathrm{A}_i(v)$ and $\tau_i(v)$ are dependent on input data v, so for each voltage, we need to perform a data assimilation procedure separately. In general, multiple input voltage steps are applied to a cardiomyocyte producing a set of $\mathrm{I}_{\mathrm{Ksum}}$ recordings to study the voltage-dependent characteristics. Suppose the sum of in-silico models of K^+ currents I_i for $i \in \{\mathrm{Kto}, \mathrm{Kslow1}, \mathrm{Kslow2}, \mathrm{Kss}\}$ serve as a computer model for data assimilation as in (32):

$$y^{\mathcal{M}}(t, v|\theta) = \sum_i \mathrm{I}_i, \tag{32}$$

where θ is the union of kinetic parameters p_s for each I_i, which are constants. In contrast, this computer model generates $\mathrm{I}_{\mathrm{Ksum}}$ for different input voltages given one set of parameters, because the in-silico models are designed by biophysical principles encoding voltage dependence. Therefore, δ is defined by the summation of root-mean-square errors (RMSEs) as given in (33). RMSE evaluates the goodness-of-fit between experimental $\mathrm{I}_{\mathrm{Ksum}}$ and model prediction. RMSEs for a set of $\mathrm{I}_{\mathrm{Ksum}}$ traces are summed up to guide the optimization procedure calibrating the models in a concurrent way for clamp voltages $v = 1, 2, \ldots, n$.

$$\delta = \sum_{v=1}^{n} \sqrt{\int_{t_v^{(h)}}^{t_v^{(e)}} \frac{\left(y_v^{\mathcal{D}}(t) - y^{\mathcal{M}}(t, v|\theta)\right)^2}{t_v^{(e)} - t_v^{(h)}} dt} \tag{33}$$

Cardiac models enable this concurrent data assimilation, because they generate multiple current traces with different input voltages by simulating underlying gating kinetics. Note that the proposed method calibrates computer models directly to $\mathrm{I}_{\mathrm{Ksum}}$ recordings, while the previous studies use statistics estimated from the data *via* curve-fitting (Du et al., 2014, 2016, 2018; Kim et al., 2022). We develop the box-constrained nonlinear optimization routine with multi-random initial points to minimize δ. Box constraints mean that θ has a lower and upper bound for each dimension, so solution space is constrained in a hypercube. In this way, the optimization loop can be controlled by users, allowing them to blend their domain knowledge into the modeling. The multi-random-start scheme helps escape local optima and find the solution as close to the optimum as possible. Latin hypercube designs and parallel computations are used to sample initial points and run them on multicores to compensate for the increased computational burden. This work is implemented in MATLAB R2022a.

3.3. Sensitivity analysis and model regularization

The principle of parsimony is critical in model calibration to enhance fitting accuracy, prevent overfitting, and improve interpretability of θ. Excessive flexibility has a risk of overfitting that occurs when the model fits data too closely, even including noise and random effects in data. Besides, as the number of parameters increases, it becomes complicated to interpret the calibration results. It is worth emphasizing that the presented models are designed with structural regularization, in which some parameters and equations are recycled in multiple places to simplify the model structure.

We also perform a sensitivity analysis to identify a subset of the parameters that have significant impacts on the model output and only calibrate these sensitive parameters. Factorial designs are developed in which parameters vary at two levels contrasting their effect on the model output. As illustrated in Fig. 3, six markers are defined that capture characteristics of K^+ current traces in voltage-clamp experiments. Each marker represents: (a) the current magnitude of 10 ms after applying a voltage step, which measures the activation rate; (b) 25% of the total recording time has elapsed, (c) 50%, and (d) 75%, which collectively estimate the inactivation rate over time; (e) the peak magnitude; and (f) the time when current has decayed $(1 - e^{-1})\%$ (almost 63%) from the peak. Marker f will be equal to the total recording time if current does not decline enough as in Fig. 3(c,d).

A fractional factorial design of 1024 runs is adopted for $\mathrm{I}_{\mathrm{Kto}}$, which results in resolution VIII. The resolution ensures that the main effects and 2-/3-factor interactions are strongly clear. Full factorial designs are used for the other three currents. The marker points λ^k defined as in Fig. 3 are evaluated at each design sample and factorial effects are calculated *via* the linear model in (34) for each current. The least squares method is used to estimate β_s^k. Then, half-normal plots are drawn to test significance of the estimated factorial effects.

$$\lambda^k = \beta_o^k + \sum_s \beta_s^k p_s^k + \epsilon, k \in \{a, b, c, d, e, f\}. \tag{34}$$

3.4. Low-dimensional embedding

Most cardiac electrophysiology studies utilize statistical tests for the mean to support hypotheses; in turn, the majority of ion channel models are based on population-averaged data. However, subject-specific analyses that consider the cellular-level variability can provide new insights into data. For example, it is possible there are differences between control and experimental groups but also within the groups. Hence, there is an urgent need to develop a tool to investigate cell-specific characteristics. Because the proposed approach calibrates models using the dataset for each cell, it allows us to quantify cellular variability in the tuned parameters. We adopt low-dimensional embedding that transforms high-dimensional data into a plane or 3D space while preserving

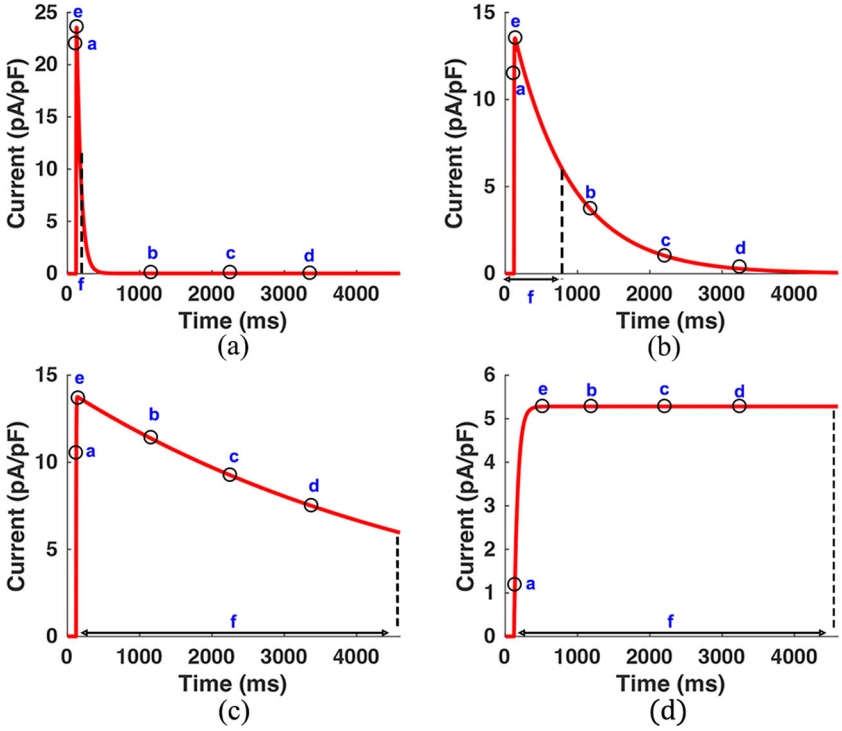

Figure 3. Illustration of the six markers of voltage-clamp K^+ currents that quantify characteristics of the current shape of (a) I_{Kto}, (b) I_{Kslow1}, (c) I_{Kslow2}, and (d) I_{Kss}. All currents are simulated for illustration, and the labels refer to (a) the current magnitude 10 ms after voltage is applied, (b) 25% of the total recording time has elapsed, (c) 50%, (d) 75%, (e) the peak magnitude, and (f) the time when current has decayed $(1 - e^{-1})$% (almost 63%) from the peak.

relative locations of data points to visualize how distributional differences in calibration parameters collectively impact inter/intra-cell variability in healthy and diseased groups.

For this purpose, t-distributed stochastic neighbor embedding (t-SNE) is used (Van der Maaten & Hinton, 2008). It has proven to be an effective method for visualizing high-dimensional data. t-SNE is a statistical method that constructs two sets of probability distributions p_{ij} and q_{ij} over pairs of data points i and j in a high-/low-dimensional space, respectively. These are probabilities of similarities such that neighboring points have a higher probability while dissimilar points have a lower probability. We first define the conditional probability of j given i:

$$p_{j|i} = \begin{cases} \dfrac{\exp\left(-d(x_i, x_j)^2/(2\sigma_i^2)\right)}{\sum_{k\neq i}\exp\left(-d(x_i, x_k)^2/(2\sigma_i^2)\right)}, j \neq i, \\ 0, j = i, \end{cases} \tag{35}$$

where $d(\cdot,\cdot)$ is a distance function such as Euclidean distance, and $\sum_j p_{j|i} = 1$ for all i. Then p_{ij} can be defined by the symmetric property of the joint probabilities:

$$p_{ij} = \frac{p_{j|i} + p_{i|j}}{2N}, \tag{36}$$

where N is the number of total data points, and $\sum_{i,j} p_{ij} = 1$. It can be calculated from data once the standard deviation σ_i is given. σ_i is set in a way that the perplexity of the conditional probability distribution over other data points given x_i equals a prefixed value that is a hyperparameter of t-SNE. Let P_i denote the conditional probability distribution, then the perplexity of the distribution is

$$\text{perplexity}(P_i) = 2^{H(P_i)}, \tag{37}$$

where $H(P_i)$ is the Shannon entropy of P_i defined by

$$H(P_i) = -\sum_j p_{j|i}\log_2(p_{j|i}). \tag{38}$$

Then the probability distribution q_{ij} is defined by the similarity of data points y_i and y_j in a low-dimensional space:

$$q_{ij} = \begin{cases} \dfrac{(1 + ||z_i - z_j||^2)^{-1}}{\sum_k \sum_{l\neq k}(1 + ||z_k - z_l||^2)^{-1}}, j \neq i, \\ 0, j = i, \end{cases} \tag{39}$$

where $\sum_{i,j} q_{ij} = 1$. Note that q_{ij} is modeled by a heavy-tailed Student's t-distribution with one degree of freedom, from which the name "t-SNE" originates. The objective of t-SNE is to learn y that minimizes discrepancies between P and Q, so that the low-dimensional distribution preserves the structure of p_{ij} constructed from the original high-dimensional data. To learn y, t-SNE maps y by minimizing the Kullback-Leibler (KL) divergence, which measures the similarity between two probability distributions:

$$\text{KL}(P||Q) = \sum_j \sum_{i\neq j} pij \log \frac{p_{ij}}{q_{ij}}. \tag{40}$$

A gradient descent method is used for the minimization of $\text{KL}(P||Q)$ with respect to y.

4. Experimental design and results

We apply the proposed framework to our data for investigating the pathophysiology of electrical signaling altered by reduced glycosylation. Recently, we showed that preventing hybrid/complex N-glycosylation in mouse cardiomyocytes was sufficient to cause DCM, achieved through genetic ablation of the *MGAT1* gene (MGAT1KO model), which encodes a critical glycosyltransferase, GlcNAcT1 (Ednie et al., 2019). MGAT1KO mice developed DCM that deteriorated into heart failure, and 100% died early, presumably from ventricular arrhythmias leading to sudden cardiac death. To further investigate the role of altered glycosylation in pathogenesis and disease progression of the heart, we conducted whole-cell patch-clamp experiments that showed reductions in N-glycosylation significantly impact electrical signaling in mouse cardiomyocytes (Ednie et al., 2019). To be specific, whole-cell I_{Ksum} traces were measured in left ventricular apex cardiomyocytes of ~14-week-old control (wild type; WT) and MGAT1KO mice, elicited by 4.5 s 10 mV voltage steps (-30 to +50 mV) from holding potential of −70 mV. There were 31 sets of whole-cell I_K recordings from different WT cells and 30 from MGAT1KO cells. Animals were used and cared for as outlined by the NIH's Guide for the Care and Use of Laboratory Animals. All animal protocols were reviewed and approved by the Wright State University Institutional Animal Use and Care Committee.

It was observed in the experiment that K^+ currents were reduced in MGAT1KO ventricular myocytes (Ednie et al., 2019). Because of the overlapping inactivation rates of the K_v isoforms responsible for I_{Kslow1} and I_{Kslow2} (K_v1.5 and K_v2.1), a curve-fitting method of bi-exponential, combining I_{Kslow1} and I_{Kslow2} of tri-exponential in (1), was applied to decompose I_{Ksum} traces into component currents. All three component currents were reduced, but the rectifier current I_{Kslow} was significantly reduced and slowed notably. Although this in-vitro investigation discovered aberrant reductions in K^+ currents with chronic glycosylation deficiency, it was difficult to determine channel kinetics rigorously. For example, curve fitting was not able to provide reliable results due to small current magnitude of the MGAT1KO model, particularly at the lower voltage steps. Therefore, here we leverage the suggested framework for modeling K_v isoforms kinetics from our experimental data of I_{Ksum} recordings. This new approach allows dissecting whole-cell K^+ current traces into isoform components and modeling their underlying kinetics concurrently.

4.1. Parameter screening

Figure 4 shows the half-normal plots of factorial effects of parameters on the six markers. The red straight lines on the plots serve as a criterion to identify the parameters that are sensitive to the marker points. The farther parameters fall above the straight lines, the more significant impacts they have on the marker. We picked the parameters to be calibrated as a union of the sets of parameters falling above the straight lines on the six half-normal plots. For I_{Kto} $\{p_1, p_2, p_3, p_4, p_5, p_7, p_{11}\}$ are selected, for I_{Kslow1} $\{p_1, p_2, p_4, p_5, p_9, p_{10}\}$, for I_{Kslow2} $\{p_1\}$, and for I_{Kss} $\{p_1, p_2, p_3\}$. In addition, all the maximum conductance variables are included in the calibration parameter set.

Figure 5 shows the selected parameters, highlighted in different colors according to their functional roles in channel kinetics. We categorized the calibration parameters into four classes: The red represents the voltage-threshold parameters and the green voltage slopes, controlling the voltage dependence, the blue scale factors of kinetic functions, and purple time-constant shifters. Note that the voltage-dependence parameters in red and green appear multiple times across different equations. This parsimonious model design is intended to maximize the structural regularization to minimize overfitting. In general, voltage-threshold and time-constant-shifting parameters impact the current traces more than others.

4.2. Model fitness

We tested various nonlinear optimization algorithms in the proposed model calibration routine that support box constraints, and the BFGS algorithm provided the most reliable results. A separate accuracy metric is used to measure the goodness-of-fit and potential prediction power of the calibration results rather than reporting the final objective function value. The objective function (i.e. sum of RMSEs) does not provide a straightforward interpretation that makes it difficult to evaluate the calibration performance. R^2, defined in (41), is the most common prediction accuracy metric ranging from 0 to 1 for linear regression models that represents the proportion of variation in data that is explained by the model. However, this interpretation is not applicable to nonlinear models because the variance decomposition in (42) no longer holds.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{41}$$

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{i=1}^{n}(\widehat{y}_i - \bar{y})^2 + \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2 \tag{42}$$

Therefore, a pseudo-R-squared measure, proposed in (Li & Wang, 2019) is used to evaluate the model fitness in this study. This nonlinear R^2 measure is based on the concept of linear correction of prediction function, allowing the variance decomposition for nonlinear functions, which guarantees straightforward interpretation and normalized scores between 0 and 1 as in the classical R^2. Figure 6(a,c) presents averages of nonlinear R^2 values over voltage steps from −30 to 50 mV by 10 mV step size for WT and MGAT1KO, respectively. In most cases, models explain more than 90 % of variances on average (WT: 0.9624±0.0051; MGAT1KO: 0.8938±0.0077). The model fitness measures in MGAT1KO are lower than WT in general. It is expected that the small current magnitude of MGAT1KO challenges the calibration process. We randomly select a cell from each group and compare the actual model predictions and experimental I_{Ksum} traces in Fig. 6(b,d) to

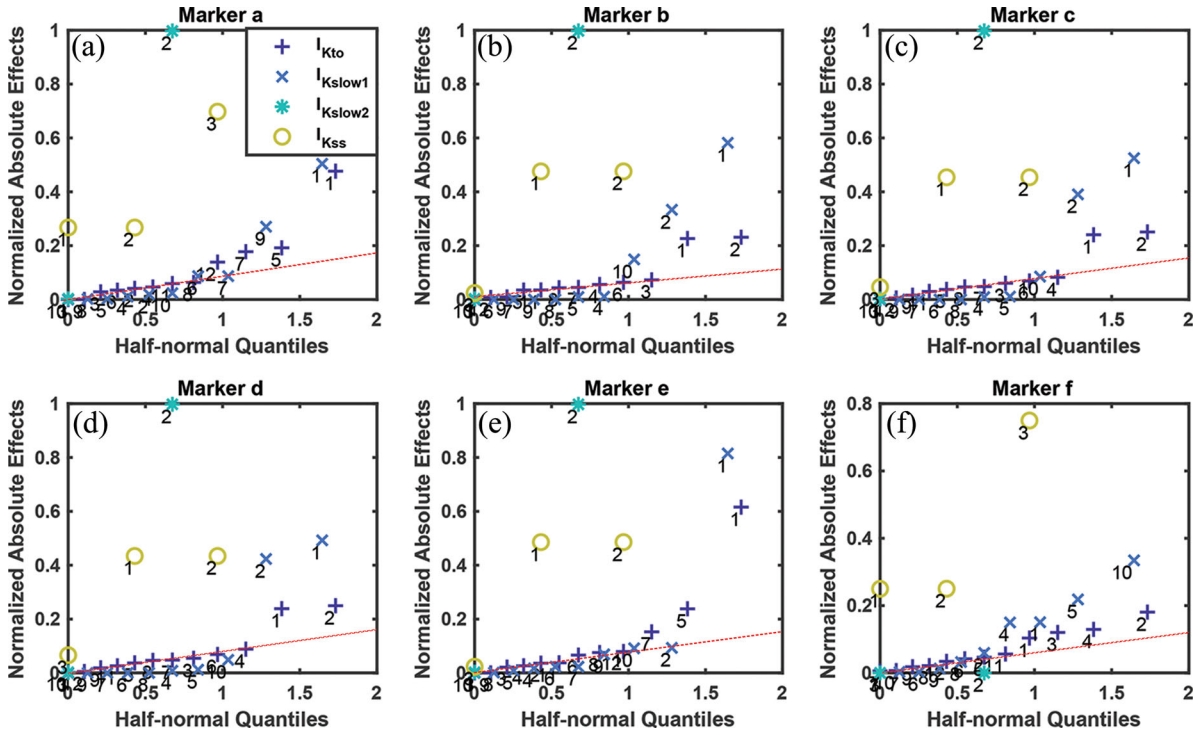

Figure 4. Half-normal plots of factorial effects for identifying significant parameters in the four current models on (a) Marker a, (b) Marker b, (c) Marker c, (d) Marker d, (e) Marker e, and (d) Marker d.

$$\alpha_a = p_7 e^{p_5(V+p_1)} \qquad \text{(a) } I_{Kto}$$

$$\beta_a = p_8 e^{-p_6(V+p_1)}$$

$$\alpha_i = \frac{p_9 e^{-(V+p_2)/p_4}}{1 + p_{10} e^{-(V+p_2+p_3)/p_4}}$$

$$\beta_i = \frac{p_{11} e^{(V+p_2+p_3)/p_4}}{1 + p_{12} e^{(V+p_2+p_3)/p_4}}$$

$$a_{ss} = \frac{1}{1 + e^{-(V+p_1)/p_4}} \qquad \text{(b) } I_{Kslow1}$$

$$i_{ss} = \frac{1}{1 + e^{(V+p_2)/p_5}}$$

$$\tau_a^{(1)} = \frac{p_7}{e^{p_6(V+p_3)} + e^{-p_6(V+p_3)}} + p_9$$

$$\tau_i^{(1)} = p_{10} - p_8 i_{ss}$$

$$\text{(c) } I_{Kslow2}$$

$$\tau_i^{(2)} = p_2 - p_1 i_{ss}$$

$$\text{(d) } I_{Kss}$$

$$\tau_a^{(3)} = \frac{p_2}{e^{p_1(V+p_3')} + e^{-p_1(V+p_3')}} + p_3$$

Figure 5. Selected parameters by the sensitivity analysis of (a) I_{Kto}, (b) I_{Kslow1}, (c) I_{Kslow2}, and I_{Kss}. Parameters are highlighted in different colors according to their functional roles in channel kinetics.

validate the model calibration results further. As shown in Fig. 6(d), MGAT1KO at −30 mV clamp voltage, small and noisy current trace, results in decreased R^2, 0.8335. The fitting was performed on Intel 4-core Xeon E7-4830 processors. Without parallel computing on multiple cores, each cell took an average of 438.1±6.8s to complete the data assimilation procedure for 5 random initial points. It took an average of 333.8±6.0s with multicore processing, which is 31 % efficient.

4.3. Prediction of K_v activities

The in-silico modeling predicts that chronic reduction in cardiomyocyte N-glycosylation results in significant changes in channel steady states and kinetic behaviors. The most notable result in the in-vitro experimental data was the significant reduction in the current density in response to a 50-mV test potential. To verify this observation, Fig. 7 presents the prediction results of the current-density relationship of the four K^+ currents. The model predicts the compatible result showing the I_{Kto}, I_{Kslow1}, and I_{Kslow2} densities are significantly reduced in MGAT1KO cardiomyocytes, while there are no remarkable differences in I_{Kss}. Note that in the in-vitro experiments, bi-exponential fitting was used, in which the two delayed rectifier currents are combined (I_{Kslow}), because it is difficult to reliably fit a tri-exponential function with a relatively short voltage step (4.5 s here) (Liu et al., 2011). On the contrary, although the proposed model-guided data assimilation distinguishes $K_v1.5$ and $K_v2.1$, it shows compatible prediction results with the

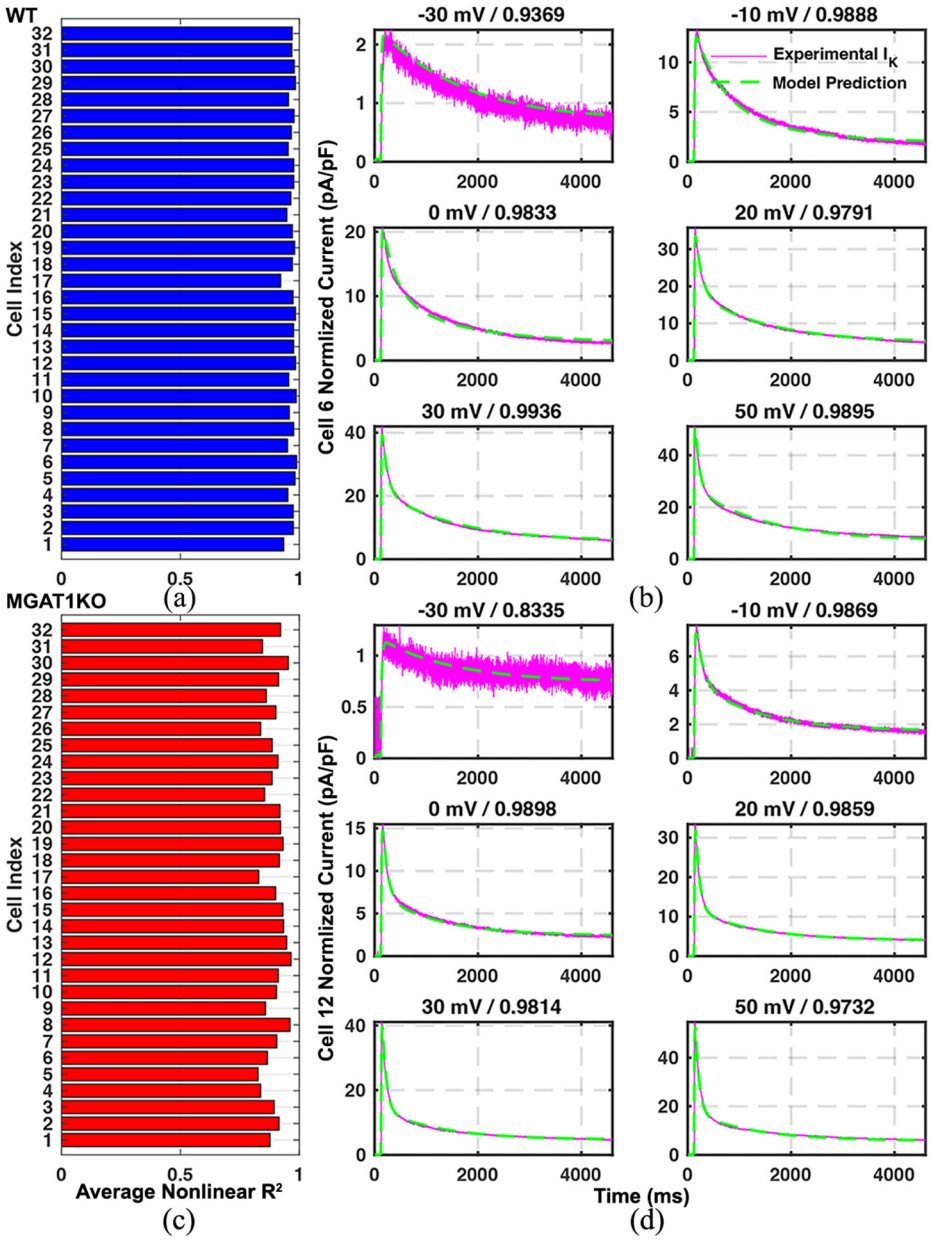

Figure 6. Bar graphs of the averages of nonlinear R^2s for nine voltage steps from −30 mV to 50 mV by the 10 mV step size for (a) WT and (c) MGAT1KO. Exemplary actual fitness between model predictions (in green) and experimental I_{Ksum} recordings (in pink) at six voltage steps of (b) cell 6 in WT and (d) cell 12 in MGAT1KO.

experimental data, as well as high fitness accuracy (∼90 % R^2) (Ednie et al., 2019).

Another aberrant activity of K^+ currents was slower inactivation of delayed rectifier currents in MGAT1KO cardiomyocytes. Figure 8 shows the prediction results of the inactivation time constants of I_{Kslow1} and I_{Kslow2}. The computer models successfully capture this trend, resulting in the time constant of inactivation in for I_{Kslow1} ($\tau_i^{(Kslow1)}$) being estimated to be significantly slowed in MGAT1KO compared to WT, i.e. ∼348.9 ms, as shown in Fig. 8(a). I_{Kslow2} is also slowed (see Fig. 8(b)). Inactivation time constants are important because they determine how slowly I_{Kslow1} and I_{Kslow2} inactivate and, in turn, significantly affect the repolarization of the AP. $\tau_i^{(Kslow1)}$ show higher variability than other currents to address most of the decaying portion of I_{Ksum}. This is likely due to the fact that while $K_v4.2$ and $K_v2.1$ are O-glycosylated, $K_v1.5$ possess a single occupied N-linked glycosylation. Hence, I_{Kslow1} is more affected by reduced N-glycosylation. A plot is not provided here, but there is no significant effect on I_{Kto} for membrane potential greater than 0 mV. It is worth mentioning that the proposed method was able to extrapolate $\tau_i^{(Kslow2)}$ that is greater than 5 s from the 4.5-s-pulse protocol, which shows the generalization power of the in-silico modeling.

4.4. Cellular variability

The calibrated parameters are in a 61×20 matrix (i.e. 31 observations in WT and 30 in MGAT1KO, and there are 20 calibration parameters), and we apply t-SNE to encapsulate

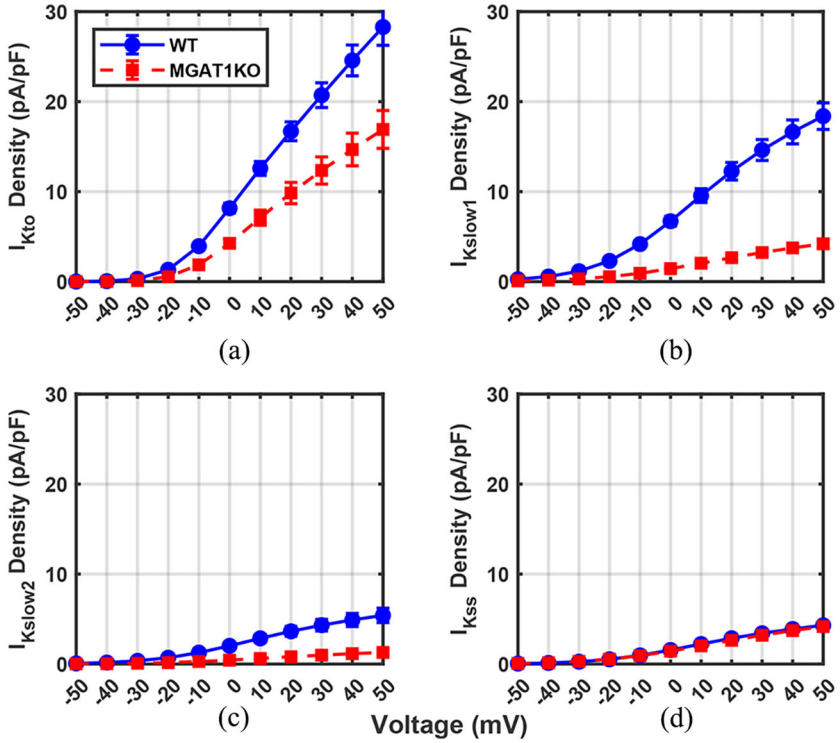

Figure 7. Current-density relationships of the major K_+ currents: (a) I_{Kto} (b) I_{Kslow1}, I_{Kslow2}, and I_{Kss}.

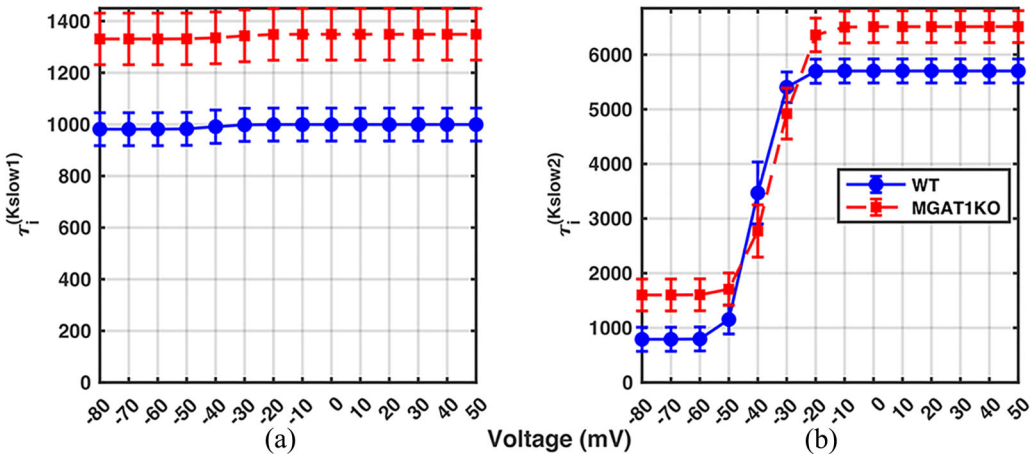

Figure 8. Inactivation time constants of (a) I_{Kslow1}, and (b) I_{Kslow2}.

these high-dimensional data in a three-dimensional space to visualize each data point. Figure 9(a) presents the visualization of the t-SNE embedding of the calibration results. It shows not only clear differences between the two groups but also variances of cells (i.e. data points) within the groups across the entire data, indicating that MGAT1KO cells have higher variance in the aggregated kinetic parameter space. Determinants of the covariance matrices of the 3D embedding data to quantify the variability of WT and MGAT1KO in the kinetic parameter space. It turns out that WT has 38 % greater variability than MGAT1KO.

To further investigate it, we generated the histograms of the maximum conductance of I_{Kslow1} (G_{Kslow1}), which is the most significant parameter determining the magnitude of the current (see Fig. 9(b)), and experimental I_K density from the in-vitro data. There is clear skewness in MGAT1KO I_{Ksum} density (small variance), while WT I_K density is more disperse (high variance). Calibrated parameters, for example G_{Kslow1}, likely show smaller variations in MGAT1KO to address this trend in the experimental data. It is likely due to significant I_K reduction in the diseased group combined with the elimination of complex N-glycosylation, thereby minimizing potential variability of N-glycan structures

5. Discussion and conclusions

Repolarization is a complex process that involves various K_v isoforms. It is critical to understand the unique properties and functional roles of each K_v to investigate pathological

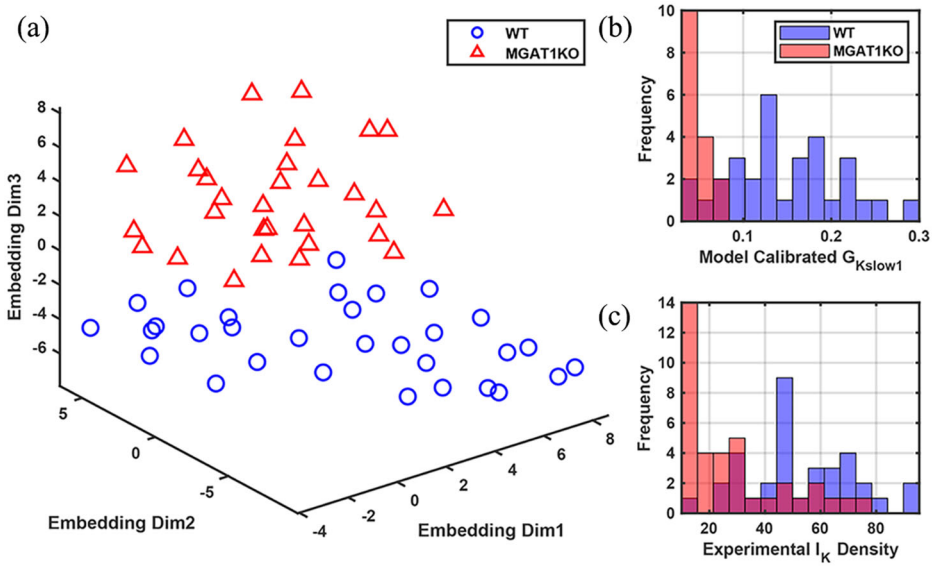

Figure 9. (a) Low-dimensional embedding of calibration parameters into 3D space. Blue circles - WT myocytes ($n = 31$); red triangles - MGAT1KO myocytes ($n = 30$). (b) Histogram of the calibrated kinetic parameter for maximum conductance of I_{Kslow1} (G_{Kslow1}), which is the most significant factor determining the magnitude of I_{Kslow1}. (c) Histogram of experimental I_{Ksum} density.

mechanisms of diseased cardiomyocytes that contribute to fatal heart diseases. However, current laboratory techniques, such as whole-cell patch-clamp recording, are not able to measure individual activities of K_v isoforms reliably that activate, inactivate, and close at overlapping times during recordings, except through the attempt to remove K_v isoform activity through less-than-fully-specific pharmacological intervention. Thus, only the sum of the different K_v isoform activities can be measured as a single current trace, I_{Ksum}. Hence, it is necessary to decompose I_{Ksum} into individual K^+ currents and estimate their channel activity *via* data assimilation. This paper presents a subject-specific concurrent data assimilation method for learning K_v activities using multiple I_{Ksum} recordings simultaneously for each cell. A case study is provided using our in-vitro experimental data of mouse cardiomyocyte I_K in control conditions (WT) and under conditions of reduced complex N-glycosylation (MGAT1KO). We evaluate the calibration results using an adjusted R^2 measure for nonlinear models that preserves the interpretability of the classical R^2 based on the variance decomposition for linear models. Experimental results show the proposed method explains more than 90% of variances by calibrated models in most cases (WT: 0.9624 ± 0.0051; MGAT1KO: 0.8938 ± 0.0077).

In addition to achieving a high degree of goodness-of-fit, it is important to determine the source of uncertainty and variability in model predictions to build trustworthy models. Until recently, the conventional approach to developing cardiac models and fitting model parameters has involved using single values. As a result, most mathematical models currently in use only provide point estimates without quantifying uncertainty (Johnstone et al., 2016). Although variances across cells are provided in the model predictions of current density and inactivation times in Figs. 7 and 8, it remains elusive whether the variability arises from model uncertainty or true cellular variability in the data. Statistical methods for uncertainty quantification (UQ) in cardiac models, such as Bayesian inferences or Gaussian process emulators, have been proposed (Coveney & Clayton, 2020; Johnstone et al., 2016). Exploring the incorporation of UQ techniques into data assimilation is a topic for future research. A GUI application is provided as an enabling tool for biomedical scientists without full expertise in modeling and computational analysis.

The estimation of kinetics provides novel insights into potential mechanisms by which specific K_v isoforms contribute to the overall reduction in I_K observed in in-vitro experiments following chronic reductions in N-glycosylation. Thanks to the cell-specific approach, prediction uncertainty is quantified, and error bars are provided in the prediction results. Further, calibrated parameters are visualized *via* low-dimensional embedding that allows for encapsulating calibrated parameters in 3D space and, in turn, visualizing the variability across cells. WT cells show higher variability than MGAT1KO myocytes, which is likely due to significant I_K reduction in the diseased group, combined with the elimination of complex N-glycosylation, thereby minimizing potential variability of N-glycan structures. The proposed method and pertinent software show strong potential for studying K_v kinetics in various heart diseases. However, there are still some limitations and challenges that need to be addresses in future work. One of the main limitations of this study is that it does not address the interactions of K_v with other major VGICs, namely Na_v and Ca_v, and cellular-level processes. Thus, causal inference of the effects of chronic glycosylation deficiency is limited. We plan to develop an integrative modeling method for learning activities of major VGICs together.

Consent and approval

This research does not involve any human participants or data. Therefore, no consent or approval was required for this study.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This work is supported by the National Science Foundation (MCB-1856132 to HY and HK), (MCB-1856199 to EB and AE), (IOS-1146882 and IOS-1660926 to EB), and the American Heart Association (Postdoctoral Fellowship 15POST25710010 to AE). HY and HK would also like to thank the NSF I/UCRC Center for Health Organization Transformation (CHOT) award NSF IIP-1624727 for the support of their research work. Any opinions, findings, or conclusions found in this research are those of the authors and do not necessarily reflect the views of the sponsors.

ORCID

Hui Yang http://orcid.org/0000-0001-5997-6823

References

Asfaw, T. N., Tyan, L., Glukhov, A. V., & Bondarenko, V. E. (2020). A compartmentalized mathematical model of mouse atrial myocytes. *American Journal of Physiology. Heart and Circulatory Physiology*, *318*(3), H485–H507. https://doi.org/10.1152/ajpheart.00460.2019

Bondarenko, V. E. (2014). A compartmentalized mathematical model of the β1-adrenergic signaling system in mouse ventricular myocytes. *PLoS One*, *9*(2), e89113. https://doi.org/10.1371/journal.pone.0089113

Bondarenko, V. E., Szigeti, G. P., Bett, G. C., Kim, S.-J., & Rasmusson, R. L. (2004). Computer model of action potential of mouse ventricular myocytes. *American Journal of Physiology. Heart and Circulatory Physiology*, *287*(3), H1378–H1403. https://doi.org/10.1152/ajpheart.00185.2003

Brouillette, J., Clark, R. B., Giles, W. R., & Fiset, C. (2004). Functional properties of k^+ currents in adult mouse ventricular myocytes. *Journal of Physiology*, *559*(Pt 3), 777–798. https://doi.org/10.1113/jphysiol.2004.063446

Costantini, D. L., Arruda, E. P., Agarwal, P., Kim, K.-H., Zhu, Y., Zhu, W., Lebel, M., Cheng, C. W., Park, C. Y., Pierce, S. A., Guerchicoff, A., Pollevick, G. D., Chan, T. Y., Kabir, M. G., Cheng, S. H., Husain, M., Antzelevitch, C., Srivastava, D., Gross, G. J., … Bruneau, B. G. (2005). The homeodomain transcription factor irx5 establishes the mouse cardiac ventricular repolarization gradient. *Cell*, *123*(2), 347–358. https://doi.org/10.1016/j.cell.2005.08.004

Coveney, S., & Clayton, R. H. (2020). Sensitivity and uncertainty analysis of two human atrial cardiac cell models using gaussian process emulators. *Frontiers in Physiology*, *11*, 364. https://doi.org/10.3389/fphys.2020.00364

Du, D., Yang, H., Ednie, A. R., & Bennett, E. S. (2016). Statistical metamodeling and sequential design of computer experiments to model glyco-altered gating of sodium channels in cardiac myocytes. *IEEE Journal of Biomedical and Health Informatics*, *20*(5), 1439–1452. https://doi.org/10.1109/JBHI.2015.2458791

Du, D., Yang, H., Ednie, A. R., & Bennett, E. S. (2018). In-silico modeling of the functional role of reduced sialylation in sodium and potassium channel gating of mouse ventricular myocytes. *IEEE Journal of Biomedical and Health Informatics*, *22*(2), 631–639. https://doi.org/10.1109/JBHI.2017.2664579

Du, D., Yang, H., Norring, S. A., & Bennett, E. S. (2014). In-silico modeling of glycosylation modulation dynamics in herg ion channels and cardiac electrical signals. *IEEE Journal of Biomedical and Health Informatics*, *18*(1), 205–214. https://doi.org/10.1109/JBHI.2013.2260864

Ednie, A. R., & Bennett, E. S. (2012). Modulation of voltage-gated ion channels by sialylation. *Comprehensive Physiology*, *2*(2), 1269–1301.

Ednie, A. R., & Bennett, E. S. (2015). Reduced sialylation impacts ventricular repolarization by modulating specific k^+ channel isoforms distinctly. *Journal of Biological Chemistry*, *290*(5), 2769–2783. https://doi.org/10.1074/jbc.M114.605139

Ednie, A. R., Deng, W., Yip, K.-P., & Bennett, E. S. (2019). Reduced myocyte complex n-glycosylation causes dilated cardiomyopathy. *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology*, *33*(1), 1248–1261. https://doi.org/10.1096/fj.201801057R

Ednie, A. R., Harper, J. M., & Bennett, E. S. (2015). Sialic acids attached to n-and o-glycans within the nav1. 4 d1s5–s6 linker contribute to channel gating. *Biochimica et Biophysica Acta*, *1850*(2), 307–317. https://doi.org/10.1016/j.bbagen.2014.10.027

Ednie, A. R., Horton, K.-K., Wu, J., & Bennett, E. S. (2013). Expression of the sialyltransferase, st3gal4, impacts cardiac voltage-gated sodium channel activity, refractory period and ventricular conduction. *Journal of Molecular and Cellular Cardiology*, *59*, 117–127. https://doi.org/10.1016/j.yjmcc.2013.02.013

Ednie, A. R., Parrish, A. R., Sonner, M. J., & Bennett, E. S. (2019). Reduced hybrid/complex n-glycosylation disrupts cardiac electrical signaling and calcium handling in a model of dilated cardiomyopathy. *Journal of Molecular and Cellular Cardiology*, *132*, 13–23. https://doi.org/10.1016/j.yjmcc.2019.05.001

Feliciangeli, S., Chatelain, F. C., Bichet, D., & Lesage, F. (2015). The family of k2p channels: Salient structural and functional properties. *Journal of Physiology*, *593*(12), 2587–2603. https://doi.org/10.1113/jphysiol.2014.287268

Giudicessi, J. R., & Ackerman, M. J. (2012). Potassium-channel mutations and cardiac arrhythmias—diagnosis and therapy. *Nature Reviews. Cardiology*, *9*(6), 319–332. https://doi.org/10.1038/nrcardio.2012.3

Johnstone, R. H., Chang, E. T., Bardenet, R., De Boer, T. P., Gavaghan, D. J., Pathmanathan, P., Clayton, R. H., & Mirams, G. R. (2016). Uncertainty and variability in models of the cardiac action potential: Can we build trustworthy models? *Journal of Molecular and Cellular Cardiology*, *96*, 49–62. https://doi.org/10.1016/j.yjmcc.2015.11.018

Kim, H., Yang, H., Ednie, A. R., & Bennett, E. S. (2022). Simulation modeling of reduced glycosylation effects on potassium channels of mouse cardiomyocytes. *Frontiers in Physiology*, *13*, 816651. https://doi.org/10.3389/fphys.2022.816651

Li, G., & Wang, X. (2019). Prediction accuracy measures for a nonlinear model and for right-censored time-to-event data. *Journal of the American Statistical Association*, *114*(528), 1815–1825. https://doi.org/10.1080/01621459.2018.1515079

Liu, J., Kim, K.-H., London, B., Morales, M. J., & Backx, P. H. (2011). Dissection of the voltage-activated potassium outward currents in adult mouse ventricular myocytes: I(to,f), i(to,s), i(k,slow1), i(k,slow2), and i(ss). *Basic Research in Cardiology*, *106*(2), 189–204. https://doi.org/10.1007/s00395-010-0134-z

Mahajan, A., Shiferaw, Y., Sato, D., Baher, A., Olcese, R., Xie, L.-H., Yang, M.-J., Chen, P.-S., Restrepo, J. G., Karma, A., Garfinkel, A., Qu, Z., & Weiss, J. N. (2008). A rabbit ventricular action potential model replicating cardiac dynamics at rapid heart rates. *Biophysical Journal*, *94*(2), 392–410. https://doi.org/10.1529/biophysj.106.98160

Marques-da Silva, D., Francisco, R., Webster, D., dos Reis Ferreira, V., Jaeken, J., & Pulinilkunnil, T. (2017). Cardiac complications of congenital disorders of glycosylation (cdg): A systematic review of the literature. *Journal of Inherited Metabolic Disease*, *40*(5), 657–672. https://doi.org/10.1007/s10545-017-0066-y

Milani-Nejad, N., & Janssen, P. M. (2014). Small and large animal models in cardiac contraction research: Advantages and disadvantages. *Pharmacology & Therapeutics*, *141*(3), 235–249. https://doi.org/10.1016/j.pharmthera.2013.10.007

Nerbonne, J. M. (2004). Studying cardiac arrhythmias in the mouse – A reasonable model for probing mechanisms? *Trends in Cardiovascular Medicine*, *14*(3), 83–93. https://doi.org/10.1016/j.tcm.2003.12.006

Ohtsubo, K., & Marth, J. D. (2006). Glycosylation in cellular mechanisms of health and disease. *Cell*, *126*(5), 855–867. https://doi.org/10.1016/j.cell.2006.08.019

Ravens, U., & Cerbai, E. (2008). Role of potassium currents in cardiac arrhythmias. *Europace: European Pacing, Arrhythmias, and Cardiac Electrophysiology:Jjournal of the Working Groups on Cardiac Pacing, Arrhythmias, and Cardiac Cellular Electrophysiology of the European Society of Cardiology*, *10*(10), 1133–1137. https://doi.org/10.1093/europace/eun193

Rodriguez, B., Burrage, K., Gavaghan, D., Grau, V., Kohl, P., & Noble, D. (2010). The systems biology approach to drug development: Application to toxicity assessment of cardiac drugs. *Clinical Pharmacology and Therapeutics*, *88*(1), 130–134. https://doi.org/10.1038/clpt.2010.95

Schwetz, T. A., Norring, S. A., Ednie, A. R., & Bennett, E. S. (2011). Sialic acids attached to o-glycans modulate voltage-gated potassium channel gating. *Journal of Biological Chemistry*, *286*(6), 4123–4132. https://doi.org/10.1074/jbc.M110.171322

Splawski, I., Timothy, K. W., Tateyama, M., Clancy, C. E., Malhotra, A., Beggs, A. H., Cappuccio, F. P., Sagnella, G. A., Kass, R. S., & Keating, M. T. (2002). Variant of scn5a sodium channel implicated in risk of cardiac arrhythmia. *Science (New York, N.Y.)*, *297*(5585), 1333–1336. https://doi.org/10.1126/science.1073569

Ten Tusscher, K. H., Noble, D., Noble, P.-J., & Panfilov, A. V. (2004). A model for human ventricular tissue. *American Journal of Physiology. Heart and Circulatory Physiology*, *286*(4), H1573–H1589. https://doi.org/10.1152/ajpheart.00794.2003

Teng, A. C. T., Gu, L., Di Paola, M., Lakin, R., Williams, Z. J., Au, A., Chen, W., Callaghan, N. I., Zadeh, F. H., Zhou, Y.-Q., Fatah, M., Chatterjee, D., Jourdan, L. J., Liu, J., Simmons, C. A., Kislinger, T., Yip, C. M., Backx, P. H., Gourdie, R. G., Hamilton, R. M., & Gramolini, A. O. (2022). Tmem65 is critical for the structure and function of the intercalated discs in mouse hearts. *Nature Communications*, *13*(1), 6166. https://doi.org/10.1038/s41467-022-33303-y

Tristani-Firouzi, M., Chen, J., Mitcheson, J. S., & Sanguinetti, M. C. (2001). Molecular biology of k^+ channels and their role in cardiac arrhythmias. *American Journal of Medicine*, *110*(1), 50–59. https://doi.org/10.1016/s0002-9343(00)00623-9

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, *9*(86), 2579–2605.

Weintraub, R. G., Semsarian, C., & Macdonald, P. (2017). Dilated cardiomyopathy. *Lancet (London, England)*, *390*(10092), 400–414. https://doi.org/10.1016/S0140-6736(16)31713-5

Whittaker, D. G., Clerx, M., Lei, C. L., Christini, D. J., & Mirams, G. R. (2020). Calibration of ionic and cellular cardiac electrophysiology models. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *12*(4), e1482.

Winslow, R. L., Cortassa, S., O'Rourke, B., Hashambhoy, Y. L., Rice, J. J., & Greenstein, J. L. (2011). Integrative modeling of the cardiac ventricular myocyte. *Wiley Interdisciplinary Reviews. Systems Biology and Medicine*, *3*(4), 392–413. https://doi.org/10.1002/wsbm.122

Xu, H., Guo, W., & Nerbonne, J. M. (1999). Four kinetically distinct depolarization-activated k^+ currents in adult mouse ventricular myocytes. *Journal of General Physiology*, *113*(5), 661–678. https://doi.org/10.1085/jgp.113.5.661