# Deciphering the allosteric modulation of the purinergic P2X receptors using Markov models

Laurent MacKay

Masters of Science

Department of Physics

McGill University

Montreal,Quebec

August 15, 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Masters of Science

## ABSTRACT

The family of Purinergic P2X receptors are ligand-gated cation channels which are affected by a number of allosteric modulators. Inducing conformational changes throughout the whole protein, allosteric modulators modify the efficacy and potency of a receptor's orthosteric ligands. We briefly review the literature of mathematical models of P2X receptors and develop modelling paradigms for allosteric modulation of orthosteric systems based on previous models of P2X receptors and the experimental protocols that they depend on. Formalisms for Bayesian inference of model parameters of ion channels using Markov chain Monte Carlo (MCMC) simulations have been tested. These formalisms are based on existing theories of Bayesian inference for single channel data, which we adapt in this thesis to whole cell recording data of cells expressing P2X receptors. We highlight some of the issues that arise when using them and present an alternative MCMC methodology to test its performance in fitting mathematical models to ion channel kinetics. Finally, we analyze in detail allosteric modulation of P2X4 receptors by the pharmacological agent Ivermectin and present two plausible types of models. Using the MCMC methodology developed herein we compare the models and the mechanisms by which they reproduce experimental data, concluding that sensitization and desensitization of receptors are not independent of each other, and that the later can occur subsequent to the former. To capture this behaviour, we develop a two layer Markov model that is more compatible with our understanding of allostery and experimentally observed data.

# ABRÉGÉ

La famille des récepteurs Purinergic P2X sont des canaux cationiques ligand-dépendants qui sont affectées par un certain nombre de modulateurs allostériques. Induisant des changements conformationnels à travers la protéine entière, les modulateurs allostériques modifient l'efficacité et la puissance de ligands d'un récepteur orthostériques. Nous examinons brièvement la littérature de modèles mathématiques de récepteurs P2X et développons des paradigmes de modélisation pour la modulation allostérique des systèmes orthostériques basés sur les modèles précédents de récepteurs P2X et les protocoles expérimentaux dont ils dépendent. Des formalismes pour l'inférence bayésienne des paramètres du modèle de canaux ioniques à l'aide de simulations chaînes de Markov Monte Carlo (MCMC) ont été testés. Ces formalismes sont basées sur les théories existantes de l'inférence bayésienne pour les données monocanal, que nous adaptons dans cette thèse aux données cellules entières de cellules exprimant des récepteurs P2X. Nous soulignons certains des problèmes qui se produisent lors de l'utilisation des formalismes et présentons une méthodologie MCMC alternative pour tester ses performances avec les modèles mathématiques cinétique de canaux d'ions. Enfin, nous analysons en détail la modulation allostérique des récepteurs P2X4 par l'agent pharmacologique Ivermectine et présentons deux types de modèles plausibles. En utilisant la méthodologie MCMC développée ici nous comparons les modèles et les mécanismes par lesquels ils reproduisent des données expérimentales, concluant que la sensibilisation et la désensibilisation des récepteurs ne sont pas indépendants les uns des autres, et que le premier peut se produire à la suite du

deuxième. Pour capturer ce comportement, nous développons un modèle de Markov à deux couches qui est plus compatible avec notre compréhension de l'allostérie et des données observées expérimentalement.

## ACKNOWLEDGEMENTS

Firstly, I must acknowledge the incredible source of support and inspiration my parents have been throughout all of my academic endeavours. Secondly, I would like to acknowledge the patience displayed by my supervisor throughout the trials and tribulations of this project as well as all of the things he has taught me. I would also like to acknowledge the insightful conversations we had and the times spent together with my labmates.

I must also acknowledge our experimental collaborators who not only provided us with experimental data but participated in discussions which were of enormous use to my modeling efforts Finally, I acknowledge the financial support which I received from my supervisor through his NSERC award as well as my personal Quebec AFE award.

Contents

List of Figures

List of Tables

# Chapter 1
# Physiology

## 1.1 Introduction

Purinergic P2X receptors (P2XRs) are a family of ligand-gated non-specific cation channels which are activated primarily by the extracellular presence of the purine adenosine 5′-triphosphate (ATP). There exist seven distinct subunit proteins of this family, labeled as P2X1-7. Early molecular biological evidence suggests that P2XRs aggregate to form functional trimers [2]. This has been confirmed by atomic force and electron microscopy for P2X2R as well as by crystallographic studies of zebrafish P2X4R [3, 4, 5, 6]. Many cell types express more than one P2X subunit and it is known that functional receptors may be formed by either using only one type (homotrimer) of subunit or by a mixture of more than one type of subunit (heterotrimer) [7]. Heterotrimers such as a receptor formed from both P2X1 and P2X2 subunits are referred to as P2X1/2R. Receptors formed as a homotrimer of P2X1 subunits are referred to as P2X1R. P2X6 is known to only form functional heterotrimers, whereas P2X7 can only form functional homotrimers [8, 3]. Many cell types have been identified with both homo and heterotrimers, with heterotrimers inheriting most of the properties of the constitutive subunits [9].

While each subunit in the P2X family differs in a number of ways, the general topology of the domains in subunits of P2XRs is conserved. Both the N and C termini of

P2X receptors possess consensus binding motifs for protein kinases and are located cytosolically. The N terminal is connected to a first transmembrane domain (TM1) involved in channel gating. TM1 then is followed by a large extracellular loop commonly referred to as the ectodomain. When coordinated in a trimer, the ectodomain forms fenestrations which are lined by negatively charged amino acids which attract cations, giving the P2X receptors their selectivity for cations [5]. The interface between adjacent ectodomains are postulated to form binding pockets for ATP or other agonists [10]. It is thought that once ATP binds to the extracellular binding site it induces conformational changes in the ectodomain and transmembrane domains leading to opening of the channel pore once two or more ATP molecules bind [11, 9]. Each ectodomain also contains a hydrophobic region close to the conducting pore of the fenestrations. Metals and other cationic modulators ($Mg^{2+}$, $Ca^{2+}$, $Zn^{+}$, $Cu^{2+}$, and $H^{+}$) may bind to this region and alter receptor behavior [9, 12]. The extracellular loop leads back into the cellular membrane to form a second transmembrane domain (TM2) which lines the pore that allows ions to pass through the membrane. TM2 then connects to the C terminal of the protein.

Aside from its widely known role as an intracellular exchanger of molecular energy, ATP was first noted to have effects on renal vascular resistance when applied extracellularly [13]. In 1972, Burnstock presented evidence for a novel class of nerves whose principal component of activation are purine nucleotides, which he termed purinergic nerves [14]. The receptors responsible for this sensitivity to purines were

eventually termed purinergic receptors. Differentiated by their transduction mechanisms, P2X receptors are the subclass of purinergic receptors that are ligand-gated as opposed to P2Y receptors which belong to the G-protein coupled receptor family.

Purinergic signaling is considered one of the most ancient and widespread signaling systems used in living tissue [15]. Purinergic receptors are implicated in both short-term signaling related to cell functioning and long term signaling related to cell development, such as neurotransmission and programmed cell death respectively. As such, it is not surprising that P2X receptors are a ubiquitous family of receptors which appear in a multitude of cell types in complex organisms. In order to emphasize that purinergic signaling is a primitive system involved in both neuronal and non-neuronal mechanisms, we will briefly summarize some physiological functions associated with P2X receptors.

While the mechanism remains unclear, P2X4 knockout mice have reduced long-term potentiation in the hippocampus [16]. Moreover, hippocampal P2X4R show a decreased sensitivity to ATP when exposed to levels of ethanol associated with extreme intoxication [17]. In taste perception, lack of P2X2 and P2X3 abolishes neural responses to tastants but not mechanical or chemical stimuli [18, 19]. Also, behavioral responses to sweetness, glutamate, and bitter substances were either abolished or reduced in double knockout mice [20]. Furthermore, P2X7 is associated with apoptotic processes involved in the rapid regeneration of taste cells and acinar cell death in auto-immune Sjögren's syndrom [21, 22]. In hearing, P2X2 is involved in the modulation of endocochlear potential [23]. In the bladder, ATP and an analog $\alpha\beta$meATP increase the excitability of bladder afferent nerves [24]. In P2X3 knockout mice, the

bladder must be filled to a greater extent than the wild-type mice in order to induce bladder voiding [25].

P2X receptors have been extensively implicated in pain. Some pain related behavior is blocked by a P2X3 selective antagonist [26]. Subsequent to spinal cord ligation, P2X4 expression is increased in microglia of the dorsal horn of the spinal cord. Moreover, introduction of microglia which had previously been stimulated by ATP into the spinal cord is associated with increased pain sensitivity [27]. This suggests that P2X4 is likely to be part of the activation mechanism of the dorsal horn microglia and neuropathic pain. In pain perception, both through the use of P2X7 specific antagonists and inhibition of genetic expression, it has been shown that P2X7 is involved in neuropathic pain in rats and neuropathic and chronic inflammatory pain in mice [28, 29].

Being associated with the capacity to clear airways of mucus, a novel P2X receptor unique to airway ciliated epithelial cells has been identified and dubbed $P2X_{cilia}$ [30]. It has a pharmacological profile similar to P2X7R but lacking the capacity to dilate and is potentiated rather than inhibited by zinc [31]. P2X receptors have also been implicated in control of respiration, where they are thought to regulate sensitivity to $CO_2$ and thus the adaptive responses of the respiratory system [32]. P2X1, P2X4, and P2X7 are all coexpressed in most immune cells [12]. However, only P2X7 has been confirmed to play a physiological role in these cells. P2X7R is known to induce release of interleukin-1 beta (IL-1$\beta$), while inhibition of P2X7 can abolish the immune inflammatory responses mediated by IL-1$\beta$ [33, 34]. P2X7 is also known to play a role in the release of insulin and the secreation of interleukin-1 receptor antagonist

(IL-1Ra) which reduces the inflammation and β-cell damage associated with type I diabetes [35].

Clearly, P2XRs are implicated in a large number of extremely varied physiological processes. In order to develop novel therapeutics for pathologies associated with any one of these physiological functions, it is important to understand how the specific P2XRs operate in a variety of conditions. This can give us insight into how the dual behaviors of P2X receptors are harnessed by differentiated cells in order to produce their specific physiological functions, which in turn can motivate potential therapies. As it stands, many aspects of P2XR gating are incompletely understood. The task of resolving this solely from experiment is hindered by pharmacological difficulties [36]. Nonetheless, in recent years mathematical modeling has begun to shed light onto many aspects of P2XRs and to guiding experimental designs in order to arrive at a more complete understanding of P2XRs [37, 1, 38, 39, 40, 41].

## 1.2 Experimental Characteristics of P2XRs

P2XR gating can be broken down into three distinguishable phases. We describe these based on their kinetics of current production in whole cell recordings. The activation phase is a rapid phase of increasing inward current subsequent to application of agonist. This is followed by the desensitization phase, a slower decay of current amplitude in the presence of an agonist and whose onset is slower than that of activation. After agonist is removed from the surrounding medium, we observe a relatively rapid decrease in current amplitude which is referred to as the deactivation phase. While all characterized P2XRs display these phases, they differ in both their sensitivity to agonists as well as in the kinetics of the phases described above.

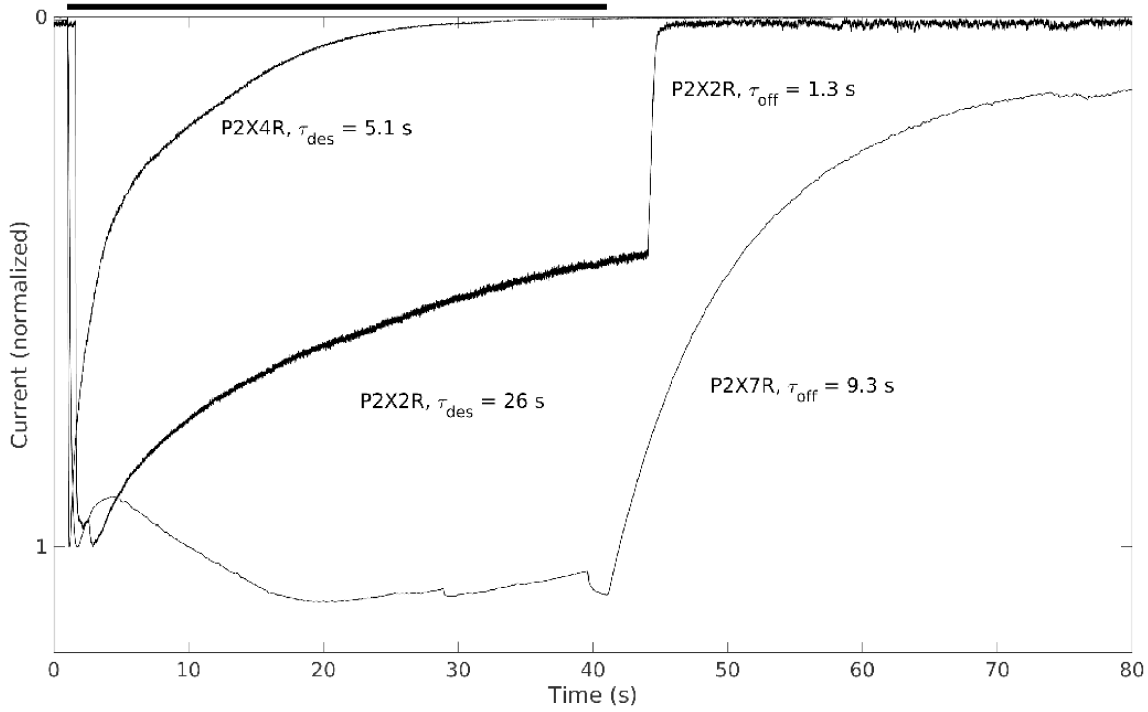


Figure 1–1: Profiles of P2XR currents generated during sustained agonist application (indicated by black bar). Recombinant rat receptors were expressed in HEK 293 cells and stimulated with supramaximal agonist (100 μM ATP for P2X4R and P2X2R, 320 μM BzATP for P2X7). $\tau_{\text{des}}$ and $\tau_{\text{off}}$ represent desensitization and deactivation time constants respectively. Both are a weighted sum ($\tau_w = (a_1\tau_1 + a_2\tau_2)/(a_1 + a_2)$) derived from biexponential fittings ($y(t) = a_1 \exp(-t/\tau_1)) + a_2 \exp(-t/\tau_2)$) of the current recordings.

Activation, desensitization, and deactivation were studied by Yan et. al. [42] using P2X4R as a model receptor. Both activation and desensitization times were found to be inversely related to agonist concentration and highly correlated to one another. On the other hand, the deactivation time of P2X4Rs was not agonist concentration-dependent. As far as we know this is a feature unique to P2X4Rs as both P2X2Rs and P2X7Rs display a strong concentration-dependence of their deactivation kinetics

[1, 43]. Interestingly, the deactivation kinetics of P2X4Rs are dependent on the concentration of the allosteric modulator Ivermectin (IVM) [41].

The gating properties of P2X2Rs and P2X4Rs seen in Fig. 1–1 seem very similar with some differences in kinetics. For many of the P2X receptors, this holds true. For example, it was determined experimentally that desensitization rates of some of these receptors satisfy (P2X2R<P2X4R<P2X1R<P2X3R). However, for P2X7Rs we see a delayed secondary growth of current which occurs at a time when other receptors would be in their desensitization phase, giving it a current profile which differs significantly from the other P2X receptors. Indicative of pore dilation, this secondary growth in current correlates temporally with a change in reversal potential when cells are bathed in a medium where sodium is replaced by $NMDG^+$[1]. $NMDG^+$is a large organic ($\sim$7.3 Å in mean diameter [44]) cation which has been used to detect pore dilation in P2X receptors [45, 41]. In such ionic conditions and prior to dilation, channels are only permeable to atomic cations such as sodium ($\sim$2.04 Å in diameter [46]). As the channel pore dilates, its selectivity for cations subsides and it becomes permeable to larger cations, including $NMDG^+$, which causes a change in reversal potential. P2X2Rs also exhibits this shift in reversal potential when stimulated with supramaximal ATP concentrations, while P2X4Rs requires the addition of IVM in order to dilate.

P2X receptors are considered allosteric systems, as opposed to simpler orthosteric systems, because they allow for the binding of modulator ligands at sites different from those of the orthosteric ligands (e.g., ATP) which can alter the properties of the

underlying orthosteric system. Receptors with allosteric properties will simultaneously interact with two ligands, an orthosteric ligand which can be directly correlated with receptor functioning and an allosteric ligand which modulates this functioning. Allosteric modulators can affect the potency of agonists in such a way that more or less agonist is required to produce half (or some other fraction) of the system's total effect. Alternatively, they can affect the efficacy of the system or the maximal response produced by the system. Both allosteric effects can be imparted on the system in agonist specific manners, which is of interest given its importance in the development of therapeutics that are both complex and of high specificity. The list of P2X allosteric modulators and effects is vast, and we refer the reader to the review of Couddou et. al. for comprehensive listings [9]. Ultimately, we will focus on the allosteric effects of IVM on P2X4Rs from a modeling perspective. However, the techniques we use here could be easily extended to model almost any allosteric modulator.

# Chapter 2
# Models of P2X Receptors

There has been a tremendous amount of experimental work done to try and characterize P2X receptor functioning both *in vitro* and *in vivo* [12, 9]. It is common in the field to use simple phenomenological models to quantitatively assess hypotheses and draw conclusions [37, 47]. However, more recently, biophysically detailed Markov models, which describe individual orthosteric binding sites and their allosteric modulation, have emerged and been very successful in reproducing varied experimental data from P2X homotrimers and succinctly explaining many phenomena [1, 38, 39, 40, 41]. We say that these are biophysically detailed because, rather than abstractly defining open and closed states, the conformational states of individual binding sites and other structural components of the receptor are considered. Open and closed states are associated with specific occupancy of the binding sites and other conformational changes which receptors can undergo.

These models have generally been constructed with the idea that all P2X receptors share the same underlying regulatory mechanism. This follows from one basic and experimentally supported assumption that a general model of all P2X receptors must exist, and that differences in its transition rate parameters yield the behaviour of specific receptors. New models have often inherited the properties of previously developed models of other receptors. For this reason, we will examine some features of the first P2X Markov model used to describe P2X7R kinetics.

## 2.1 P2X7R

According to the diagram in Fig. 2–1, we can see the P2X7R Markov model is organized into two horizontal rows of states, the top representing unsensitized states and the bottom the sensitized ones. The circles on top of the unsensitized states represent the individual agonist binding sites, where black circles indicate the number of bound agonist molecules. ATP is the primary agonist of P2X7Rs, but it has low potency. As a result, the majority of experiments have been carried out with the more potent agonist 2′,3′-O-(4-benzoylbenzoyl)-ATP (BzATP). In this and other models, we use the variable $A$ to represent agonist concentration such as ATP. We do not expect the action of these agonists to differ enough from ATP so as to require a different model for each agonist. Thus our models should remain valid for other agonists as long as they have the same number of binding sites and the appropriate changes to binding rates are made in order to reflect the different sensitivity to each agonist.

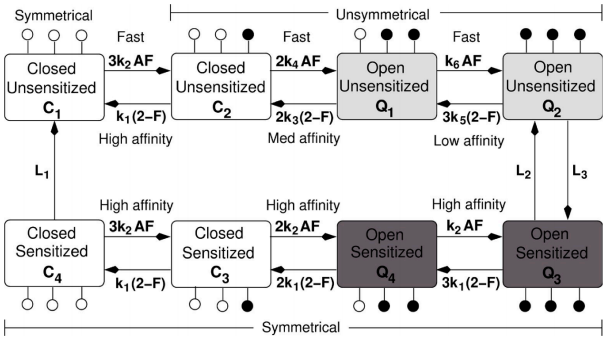


Figure 2–1: Markov model of P2X7R activation and dilation/sensitization by ATP. Copied with permission from Yan et. al. 2010. See publication for model equations.

The model in Fig. 2–1 assumes that unstimulated (naïve) receptors are all in the closed state $C_1$. Transitions to the right in the diagram requires the presence of ATP and represent binding of ATP to the receptor. Leftward transitions represent the unbinding of ATP. Once two ATP molecules bind to a receptor, the channel opens by transitioning from the state $C_2$ to $Q_1$. Subsequent binding of an ATP molecule puts the receptor in the state $Q_2$ where it spontaneously "sensitizes" to the state $Q_3$. From $Q_3$, ATP molecules may unbind until receptors reach the state $C_4$ which allows for a slow recovery to the naïve state $C_1$. The slow rate of return to $C_1$ gives sensitized receptors memory, such that repeated applications of agonist induces the patterns of activation and deactivation shown in Fig. 2–2.

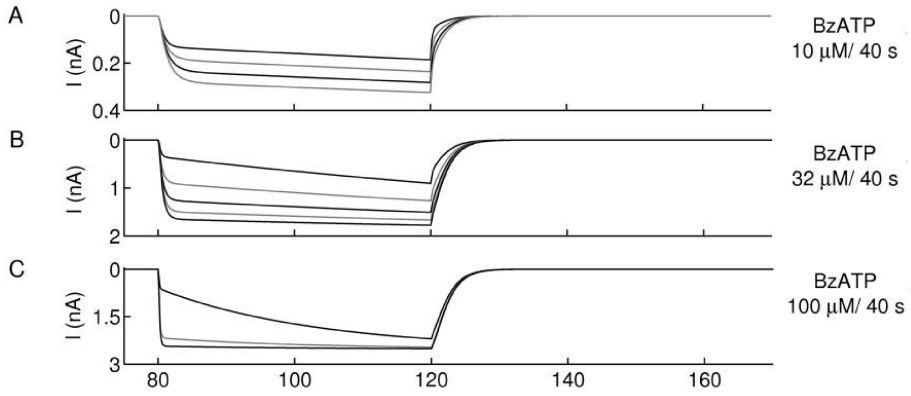


Figure 2–2: Model simulations of current responses to repeated application of agonist. A-C, Simulations of periodic stimulation with 10 μM/40 s BzATP (A), 32 μM/40 s BzATP (B) and 100 μM/40 s BzATP (C). Successive agonist application alternate between black and grey traces. Copied with permission from Yan et. al. 2010.

In the scheme of Fig. 2–1, each sequential binding of ATP in the top, unsensitized, row occurs with decreasing affinity. In contrast, in the sensitized row, each binding of ATP maintains a high affinity. If we considered each row as a separate Markov

model, the sensitized row would have a higher probability of opening at the same agonist concentration (hence the name sensitized). We view this sensitized row as an allosteric modification of the orthosteric system (unsensitized row) by an unspecified ligand (such as phosphorylation), leading to modification of agonist potency. This allosteric modification produces a copy of all the states in the orthosteric system in Fig. 2–1, as it is assumed that orthosteric ATP binding and sensitization are two processes that occur at distinct locations of the receptor. Transitions to the sensitized row are responsible for the secondary growth in current observed in Fig. 1–1. As such, they are to be associated with dilation and changes in permeability of the pore. Whereas the unsensitized open states ($Q_1$,$Q_2$) have the conductance $g_{12}$, the sensitized open states ($Q_3$,$Q_4$) have a larger conductance $g_{34}$. Therefore, the allosteric modification which transitions receptors to the sensitized row modifies both the potency and efficacy of agonists. This yields the following current equation for the Markov model

$$I = g_{12}\left(Q_1 + Q_2\right)\left(V - E\right) + g_{34}\left(Q_3 + Q_4\right)\left(V - E\right), \tag{2.1}$$

where E is the reversal potential of the cell (0 V for Krebs-Ringer-like bath medium) and V is the holding potential (-60 mV). Moreover, this model includes the allosteric effects of extracellular calcium (see Fig. 2–3 bottom row) described by the two multiplicative factors $F$ and $2 - F$ affecting all the ATP binding and unbinding rates, respectively, where

$$F = \alpha \frac{\beta^2}{\beta^2 + [DC]_e^2}, \tag{2.2}$$

is a monotonically decreasing Hill function of the extracellular divalent cation concentration $[DC]_e$ ($Ca^{2+}$), with kinetic constants$\alpha$ and $\beta$. Calcium is assumed to have the effect of reducing the potency of agonist binding by decreasing agonist binding rates while increasing unbinding rates (see Fig. 2–3). This is a second way in which receptor allostery is modeled. In contrast to sensitization, we do not expect the fraction of total receptors that are allosterically modified by calcium to change throughout experiments. This is because the allostery depends on extracellular calcium, a cation present in the bath medium long before experiments are carried out [48]. Assuming no cooperativity between ATP and calcium binding, the calcium binding process will have equilibrated before experiments are performed. In all of these experiments, calcium concentrations were kept the same, so we expect the ratio of calcium-bound and unbound receptors to stay the same. If the major effect of calcium is to alter ATP binding kinetics, there is no need to create a copy of each state to capture calcium allostery. Rather, we can approximate the calcium concentration dependent kinetics of the equilibrium mixture of calcium-bound and unbound receptors fairly well through the use of the multiplicative factor $F$. Evidence exists for partial agonist binding inducing many allosteric modifications of P2X receptors including increased agonist potency (sensitization) [49, 50]. We do not know the identity of the sensitizing allosteric modulator, but we do know that its action is subsequent to application of ATP. The timescales of these modifications, governed by the rate $L_3$ in the model of Fig. 2–1, is long compared to agonist binding. Thus a complete model of sensitization allostery will require modeling the modified (sensitized) and

unmodified (naïve) receptors independently of each other, as opposed to calcium allostery where it was not necessary to model the modified and unmodified receptors independently.

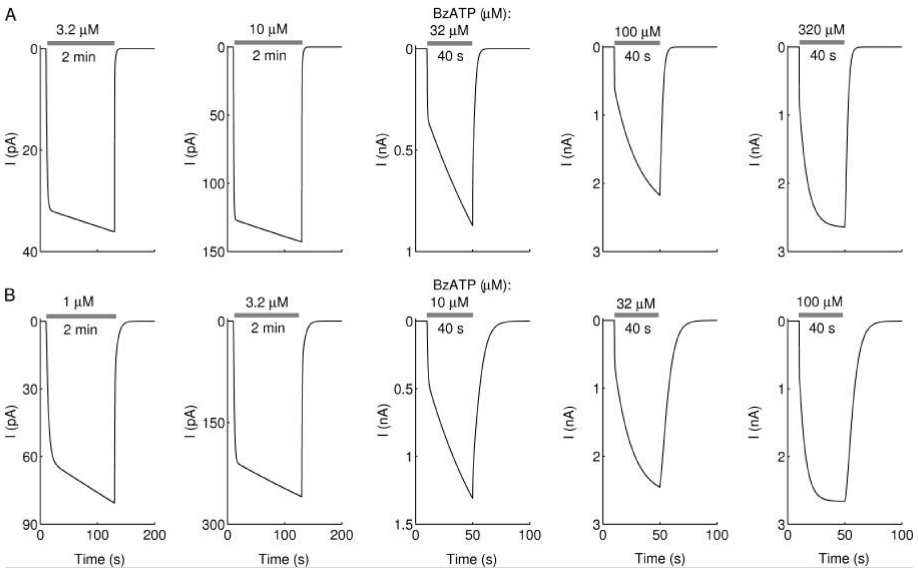


Figure 2–3: Dependence of current simulations on BzATP and extracellular calcium concentrations. (A and B) Current simulations generated by the kinetic model of Fig. 2–1 in the presence of 2 mM $Ca^{2+}$ (A) and in the absence (B) of bath $Ca^{2+}$ during 40-s and 2-min BzATP stimulation at increasing doses, as specified by the gray bars. Monophasic and biphasic responses are obtained at low and high concentrations of BzATP, respectively, in both cases, a result consistent with the experimental recordings. The loss of allosteric inhibition exerted by extracellular $Ca^{2+}$ in B led to an increase in the potency of BzATP, which manifested itself in requiring less agonist concentration (a reduction of ∼70% compared with A) to produce the same response. Copied with permission from Yan et. al., 2011. Originally published in Journal of General Physiology. doi:10.1085/jgp.201110647.

The model in Fig. 2–1 was further modified in order to include receptor desensitization which is necessary to explain current decay in the presence of agonist such as that observed in Fig. 1–1 [40]. These modifications can essentially be reduced to

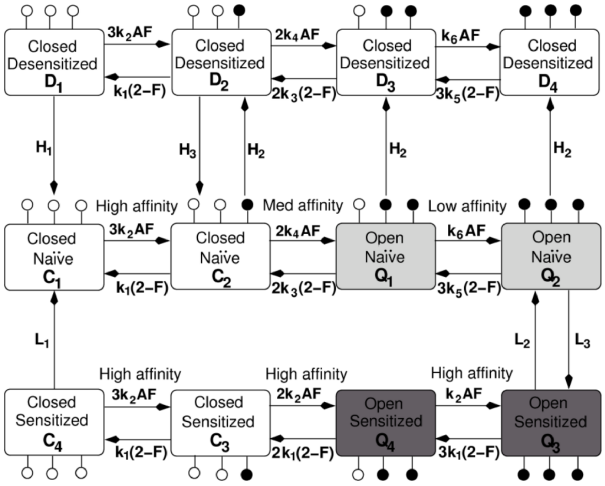


Figure 2–4: Markov model of P2X7R activation by ATP, desensitization, and dilation/sensitization. Copied with permission from Khadra et. al. 2013. See publication for model equations.

adding a third row connected to the naïve row as seen in Fig. 2–4. Receptor desensitization will be covered more thoroughly in the following section on the model of P2X2Rs.

## 2.2 P2X2R

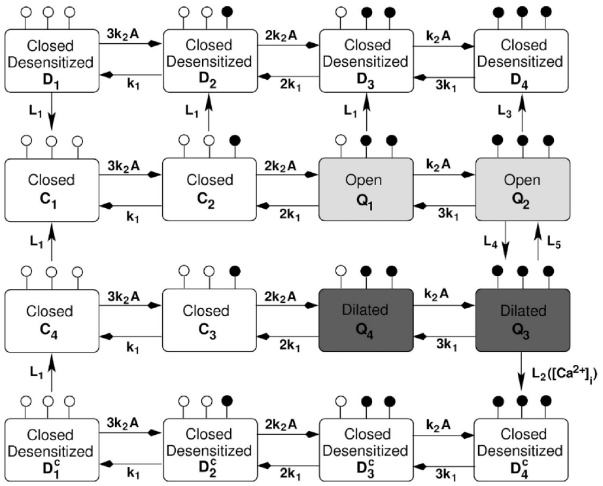


Figure 2–5: Markov model of P2X2R activation by ATP, desensitization, and dilation. Copied with permission from Khadra et. al., 2012. Originally published in Journal of General Physiology. doi:10.1085/jgp.201110716. See publication for model equations.

Following the success of the P2X7R Markov model, a closely related model was used to model P2X2Rs. At a glance, we can see that the P2X7R Markov model shown in Fig. 2–4 is represented by the top three rows of the P2X2R model of Fig. 2–5. In the model, the term sensitized has been replaced by dilated and unsensitized by naïve. One additional rows has been added to the model, representing calcium-dependent desensitized receptor states. The desensitized rows have ATP binding kinetics which are identical to naïve and dilated rows but lack states with conductivity. The model assumes that while ATP may still bind to desensitized receptors, this does not lead to the conformational changes associated with opening of the receptor pore. The inclusion of desensitized rows was necessary to model P2X2Rs, as desensitization is

more prominent and occurs significantly more quickly in P2X2Rs than in P2X7Rs and thus could not be omitted from the model. The naïve row desensitizes in a straightforward manner, once one or more ATP are bound to a naïve receptor. With one or two ATP molecules bound, desensitization is assumed to occur with a rate $L_1$, but when three are bound, desensitization is assumed to be faster with a rate $L_3$ ($L_3 > 10L_1$). This choice of rate parameters captures the dependence of the desensitization rate on agonist concentration (see Fig. 2–6). Once all ATP molecules become unbound from desensitized receptors, the model assumes these that receptors can return to the naïve state $C_1$ with a rate $L_1$, although this rate is taken to be 100 times larger than that of P2X7Rs because P2X2Rs do not exhibit the same memory upon activation. The model was used to capture the experimental behaviours of receptors from two different splice variants of P2X2 subunits, P2X2a the full-length protein and P2X2b a shorter splice form with 69 residues removed from the C-terminal domain.

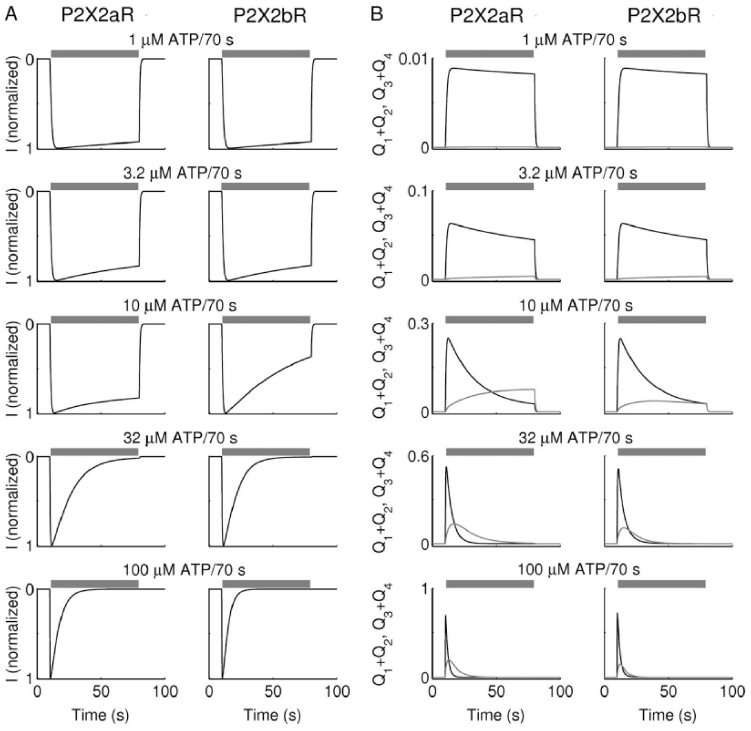


Figure 2–6: Dependence of current amplitude and desensitization of P2X2aR and P2X2bR on ATP concentration in the kinetic model. (A) Simulated currents normalized by their maximum amplitudes, and (B) receptor accumulation in $Q_1 + Q_2$ (black) and $Q_3 + Q_4$ (gray), obtained during the third stimulation of P2X2aR (left) and P2X2bR (right) with 1, 3.2, 10, 32, and 100 μM ATP for 70 s, as indicated by the gray bars (washout periods are 250 s). Increasing ATP dose increases receptor accumulation in $Q_3 + Q_4$, which in turn activates the Ca2+-dependent desensitization via the state $Q_3$, leading to an increase in current desensitization in both receptors. Because $L_5$ is smaller for P2X2aR, simulated P2X2b currents desensitize faster than P2X2a currents, in agreement with experimental observations (Fig. 9). Copied with permission from Khadra et. al., 2012. Originally published in Journal of General Physiology. doi:10.1085/jgp.201110716.

Calcium allostery is of a different nature in P2X2Rs than P2X7Rs. As such, it was not modeled using the multiplicative factors on agonist binding and unbinding rates.

Rather calcium is known to affect the rates of desensitization in a use dependent manner [39]. That is, with each stimulation by ATP in the presence of $Ca^{2+}$, the rate of desensitization tends towards a rate which is faster than that initially observed. It is even possible for the rate of desensitization to abruptly shift during a single ATP application. The increase in the rate of desensitization is achieved by making the transition to a desensitized row from dilated states a function of the intracellular calcium concentration. By including a differential equation for intracellular calcium ($[Ca^{2+}]_i$) in the model, it is possible to produce the use dependent increase in the desensitization rate. The following equation was used to describe calcium dynamics

$$\frac{d\left[Ca^{2+}\right]_i}{dt} = -f\left(\alpha SRI + k_c\left[Ca^{2+}\right]_i\right),$$

where $f$ is the fraction of free intracellular calcium, $\alpha$ is Faraday's constant, $S$ is the fraction of the current $I$ attributable to calcium, $R$ converts the units of I from A to pA, and $k_c$ is the calcium efflux rate. The model of Fig. 2–5 indicates that, when receptors open, they allow for an influx of calcium ions and thus the intracellular calcium concentration to increase throughout an experiment. The rate of calcium-dependent desensitization, $L_2$, is assumed to be a dynamic variable with the differential equation

$$\frac{dL_2}{dt} = \beta_1\left[Ca^{2+}\right]_i \frac{\beta_2 + \left(L_2^* - L_2\right)^2}{\beta_3^2 + \left(L_2^* - L_2\right)^2} - \beta_4 X L_2, \tag{2.3}$$

where $\beta_1$, $\beta_2$, $\beta_3$, $\beta_4$, and $L_2^*$ are constants. $X$ is the concentration of some unknown inhibitory agent which has been modeled as having a constant source $\sigma$ and a linear

decay with rate $k_x$, as follows

$$\frac{dX}{dt} = \sigma - k_x X.$$

Steady state analysis of $L_2$ reveals that it acts as a bistable switch with respect to calcium concentration (see Fig. 2–7). As calcium enters the cell the dynamic variable $L_2$ travels along the lower branch of the toggle switch. Eventually calcium influx brings $L_2$ to the right knee of the $L_2$ bistable switch. Any further increase in intracellular calcium causes $L_2$ to jump to the upper stable branch of its S-shaped bistable switch which leads to an abrupt increase in desensitization rate.

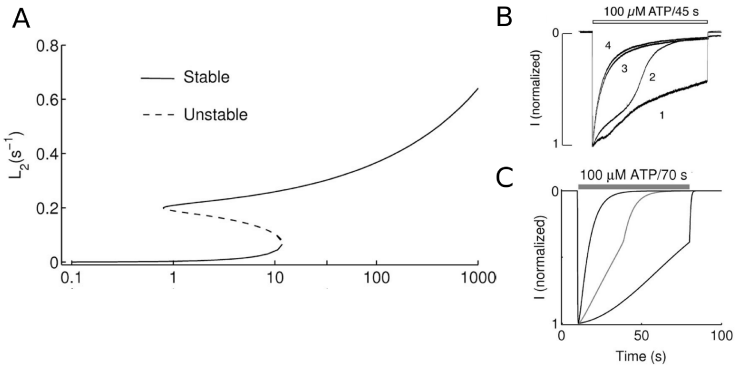


Figure 2–7: The bistable switch responsible for the change from slow to fast desensitization observed in P2X2Rs during repetitive ATP stimulation. (A) The nullcline of $L_2$ when $X$ is set to steady state and $[Ca^{2+}]_i$ is treated as a parameter. The S-shaped switch shows bistability and hysteresis. (B) Current profile of repeated stimulation of P2X2Rs with ATP. Notice the abrupt change in desensitization kinetics during the second application of ATP. (C) Model simulation of experimental protocol in B. During second stimulation (grey trace) there is a clear change in desensitization kinetics which corresponds to crossing the right knee in A. Copied with permission from Khadra et. al., 2012. Originally published in Journal of General Physiology. doi:10.1085/jgp.201110716.

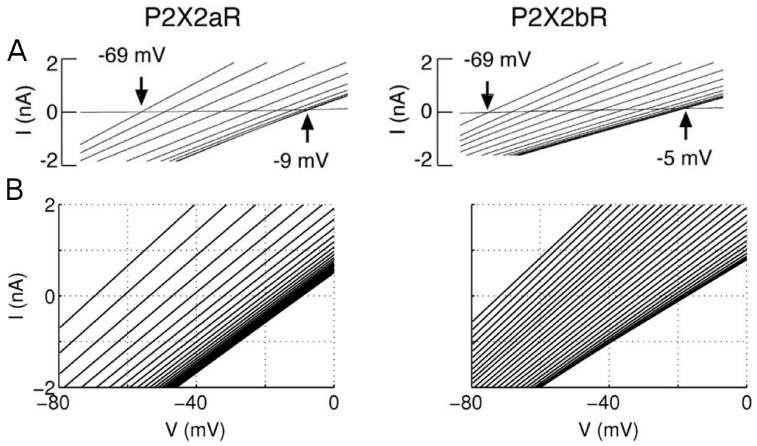


Figure 2–8: Pore dilation manifests itself as a reversal potential shift. The response of P2X2aR (left) and P2X2bR (right) to 100-μM ATP stimulation for 70 s in the presence of $NMDG^+$-only medium. (A) Positive shifts in reversal potential observed during the initial 25-s agonist application in HEK293 cells bathed in medium containing 155 mM $NMDG^+$, 10 mM HEPES, and 10 mM glucose only. Voltage ramps from -80 to +80 mV are delivered twice per second to P2X2aR and twice per 100 ms to P2X2bR (B) Model simulations also exhibit a shift in the reversal potential in the I-V curves during voltage ramps due to the very fast accumulation of receptors in open states $Q_1 + Q_2$ relative to the slow accumulation of receptors in the dilated states $Q_3 + Q_4$. The slopes of the I-V curves generated by the model decrease because $L_3 > L_4$ (i.e., desensitization masks dilation) For a more thorough treatment see section 4.1.1. Copied with permission from Khadra et. al., 2012. Originally published in Journal of General Physiology. doi:10.1085/jgp.201110716.

Furthermore, this model was used to simulate voltage ramp experiments, where the holding potential of the cell is rapidly varied from -80 mV to +80 mV twice per second. The current is then recorded as a function of the applied potential. Using graphical inspection we can find the reversal potential, which produces no net current. P2X2Rs display a shift in reversal potential over the first 25 seconds of agonist application, this is associated with dilation (see Fig. 2–8).

Due to the altered permeability of a dilated receptor, in $NMDG^+$ containing media, dilated states should have a different reversal potential from open states (see section 4.1.1). Thus the current equations were modified from that of the P2X7R model to

$$I = g_{12}\left(Q_1 + Q_2\right)\left(V - E_{12}\right) + g_2\left(Q_3 + Q_4\right)\left(V - E_{34}\right), \tag{2.4}$$

where $E_{12}$ and $E_{34}$ are the reversal potentials of the open and dilated states respectively. These reversal potentials are functions of the intracellular and extracellular environments as well as the cell's permeability to ions (see section 4.1.1). In the presence of $NMDG^+$ they differ, while in its absence with mixed ionic conditions they are taken to be the same. During the voltage ramp protocol, the reversal potential essentially shifts from $E_{12}$ to $E_{34}$. The current equation can be rewritten in a standard form to isolate the total conductance and reversal potential of the cell, as follows

$$I = g_{tot}\left(V - E_{tot}\right), \tag{2.5}$$

where $g_{tot}$ and $E_{tot}$ are the total conductance and reversal potential of the cell respectively. By equating Eqs. (2.4) and (2.5) and after some algebraic manipulations, we find that:

$$E_{tot} = \frac{g_{12}\left(Q_1 + Q_2\right)E_{12} + g_{34}\left(Q_3 + Q_4\right)E_{34}}{g_{12}\left(Q_1 + Q_2\right) + g_{34}\left(Q_3 + Q_4\right)}.$$

As receptors shift from being mostly in the open states $Q_1$ and $Q_2$ to the dilated states $Q_3$ and $Q_4$, this weighted sum shifts from $E_{12}$ to $E_{34}$. Provided that $E_{34}$ is chosen to be more positive than $E_{12}$, this will lead to a positive reversal potential shift.

## 2.3 P2X4R

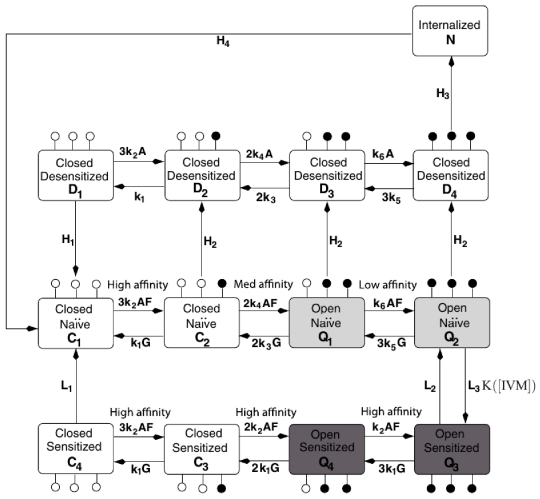


Figure 2–9: Markov model of P2X4R activation by ATP, desensitization, and dilation/sensitization by IVM. Copied with permission from Zemkova et. al. 2012. See publication for model equations.

For the P2X4R, the model is organized into three rows of states of sequential ATP binding and an auxiliary state N representing receptor internalization from the cell membrane into the cytoplasm. The three rows are essentially the top three rows of the P2X2R model. The middle row is deemed naïve as $C_1$ represents the state where population resides prior to any agonist exposure. From the state $Q_2$, the model assume that IVM binding (a pharmacological agent that causes P2X4R dilation) induces receptor sensitization and a shift to the bottom (sensitized) row. Because IVM is required to produce sensitization of P2X4Rs, the rate of sensitization/dilation $L_3K\left([IVM]\right)$ is assumed to be an increasing Hill function of IVM concentration

$$K\left([IVM]\right) = \frac{[IVM]}{\beta + [IVM]},$$

where $\beta$ is a constant. This is in contrast to the models of P2X7Rs and P2X2Rs which sensitize and dilate spontaneously once three ATP molecules are bound. For P2X4Rs, these processes are understood as allosteric modifications by IVM, while for P2X7Rs and P2X2Rs the nature of the receptor modulation which induces sensitization and dilation remains unclear [49, 50]. IVM is known to cause a leftward shift in $EC_{50}$ for ATP evoked current, an effect deemed to represent sensitization (see section 4.1.3). Similar to the sensitized row of the P2X7R model, this row has increased affinity to ATP binding as compared to the naïve row and thus produces the desired sensitization to ATP when IVM is applied. Transitions back to the naïve row are possible from the $Q_3$ and the ATP-unbound state $C_4$, but no transitions are allowed directly to the desensitized row. That is, in this model of P2X4R, sensitized receptors do not desensitize but must first revert back to the naive status (through IVM binding), before they can desensitize. It should be noted that sensitization (an increase in agonist potency) and desensitization (a decay in current amplitude during agonist application) are not reverse processes of one another as their names might suggest.

In the model of Fig. 2–9, the ATP binding and unbinding rates in both the naïve and sensitized rows are multiplied by factors which depend on the same Hill function of IVM concentration which governs the dilation process. The ATP binding rates are all multiplied by an increasing function of IVM concentration given by

$$F = 1 + \alpha K\left([IVM]\right),$$

where $\alpha$ is a constant. Conversely, the ATP unbinding rates are all multiplied by a decreasing function of IVM concentration, given by

$$G = 1 - K\left([IVM]\right).$$

These multiplicative factors are what produce the aforementioned sensitivity to ATP in the presence of IVM (see Fig. 2–10b). Moreover, the factor $G$ also has the experimentally justified effect of slowing down deactivation kinetics as the concentration of IVM is increased (see Fig. 2–10a). As will be discussed, according to this model, it is necessary to have such nonlinearities included in the naïve row in order to properly capture the experimentally observed data.

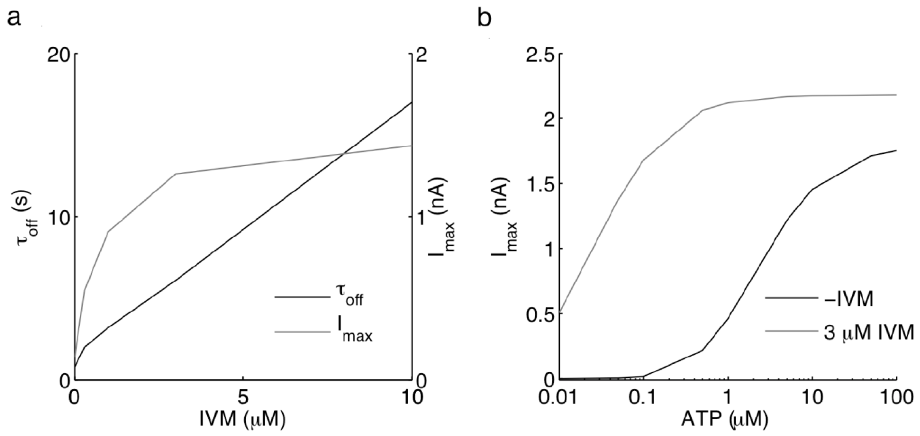


Figure 2–10: Effect of IVM on deactivation kinetics $\tau_{off}$ and maximum current amplitude $I_{max}$ on the P2X4R model. (a) $\tau_{off}$ (black) and $I_{max}$ (gray) as functions of IVM concentration. Both traces follow a Hill function with Hill coefficient n = 1; $EC_{50} \approx 9.2$ μM for $\tau_{off}$ and $EC_{50} \approx 0.6$ μM for $I_{max}$. (b) The maximum current amplitude produced by the model in the absence (black) and presence (gray) of 3 μM IVM in the medium when the model-cell is stimulated with various concentrations of ATP for 30 s. In both cases, Hill-like dose-response curves are generated with Hill coefficient n = 1; $EC_{50} \approx 4.2$ μM in the absence of IVM and $EC_{50} \approx 0.02$ μM in the presence of IVM. Copied with permission from Zemkova et. al. 2012.

We will subsequently refer to interactions between IVM and receptors as IVM binding, although we are not entirely sure that such interactions involve direct binding of IVM to the receptor. Moreover, we will maintain the naming convention for states used in this model for all subsequent models. Q is used to label open conducting states, C is used to label closed states, and finally D is used to label desensitized states. Adoption of this convention throughout all models will allow for simpler model descriptions.

# Chapter 3
# MCMC

## 3.1 Preliminaries

In all of the models presented in section §2 there are many parameters whose values must be chosen appropriately in order for the model to accurately capture experimental behaviour. Thus far, parameter values were chosen in an *ad hoc* manner to qualitatively reproduce the observed experimental currents. In this thesis, we prefer to take a more systematic approach where we quantitatively compare model predictions to experimental data for a given set of parameters. We have experimental data in the form of whole cell current recordings of various experimental protocols. Therefore, we will compare model generated currents to experimental data we have on hand. Ultimately, we will design error functions, which represent the error between model predictions and the experimental data, and try to minimize these functions. This can be done in a deterministic setting such as the least squares fitting. However, we have found this method to produce sub-optimal results, due to the existence of many local minimal, while its requirement to evaluate the gradient of the error function is prohibitively computationally expensive for models with large numbers of parameters. Moreover, we are interested in finding the probability distribution of parameters rather than a single parameter value which results from deterministic optimization techniques. Thus we employ Markov chain Monte Carlo (MCMC) algorithms which are capable of sampling from unknown probability distributions that

depend on an error function. Generally speaking, MCMC is a class of algorithms capable of approximating the integral

$$I = \int_{\mathbb{R}^d} f(\boldsymbol{x})\, \pi(\boldsymbol{x})\, d\boldsymbol{x}.$$

It does so not by finding the probability density function $\pi(\boldsymbol{x})$, but rather by sampling $\boldsymbol{x} \in \mathbb{R}^d$ in such a way to approximate the effect of $\pi(\boldsymbol{x})\, d\boldsymbol{x}$ on the integral. What is meant by this is that regions of $\mathbb{R}^d$ which have higher density, $\pi(\boldsymbol{x})$, will have more values drawn from them by an MCMC algorithm. The Markov chain of MCMC refers to the fact that the algorithm performs a Markovian random walk in $\mathbb{R}^d$. For our purposes, $\mathbb{R}^d$ represents the space of parameter values and $\pi(\boldsymbol{x})$ is thus the probability density function of the parameters.

Throughout the following sections we will deal with many probabilities and probability densities. For example, we will write the probability of some random variable $X$ having the value $A$ as $P(X = A)$ or more succinctly $P(A)$ to describe the probability of the event $A$. Moreover, we will denote the probability of $A$ given that $B$ is also true (i.e., the conditional probability of $A$) as $P(A|B)$. In order for it to be possible for both $A$ and $B$ to be true, they must be the value of two different random variables. Thus we can write the conditional probability more explicitly as $P(X = A|Y = B)$. Furthermore, for continuous random variables we say that they have probability density function (p.d.f.) $\pi(x)$ if the probability of being in some interval $[A, B]$ is given by the integral

$$P(A \leq X \leq B) = \int_A^B \pi(x)\, dx.$$

In order for probability distributions to be proper the total probability of all possible events must be 1. This means that the following probability conditions (or normalizing conditions) must be satisfied. If $A$ is an event (or set of events) in a countable state space and $\overline{A}$ is its complement, then the normalization condition is as follows

$$P(A) + P\left(\overline{A}\right) = 1.$$

Alternatively, if A and B are the limits of support of a p.d.f. on a continuous state space the normalization condition is

$$\int_A^B \pi(x)\, dx = 1.$$

### 3.2 Markov Chains

Ultimately, our goal is to use MCMC algorithms to produce Markov chains in parameter space. Since we will be working in a discrete digital representation of $\mathbb{R}^d$, we can consider it as a countable state space $E$. A Markov chain is a sequence of values $\{X_n, n \geq 0\}$ (possibly vector valued) produced by a discrete-time Markov process. A Markov process is a stochastic process that determines future values based only on its current state and not the past. As such, a Markov chain in parameter space will obey the following property for each set $i, j; i_0, i_1, \ldots, i_{n-1}$ of possible parameter values in the countable state space $E$:

$$P(X_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j | X_n = i), \qquad (3.1)$$

where $i_0$ is the initial value of the chain, $i$ is its $n^{th}$ value and $P(X_{n+1} = j|\cdot)$ is the probability of the chain moving to state $j$ at its $n^{th}$ step forward in discrete time.

Equation (3.1) is referred to as the Markov property. Moreover, the conditional probability $P(X_{n+1} = j|X_n = i)$ is referred to as the one-step transition probability. If the latter does not depend on $n$, for all states $i, j \in E$ then we say that the Markov chain has stationary transition probabilities and write

$$P_{ij} = P(X_{n+1} = j|X_n = i) = P(X_1 = j|X_0 = i)$$

We call $P_{ij}$ the one-step transitions probability, because it transforms the initial distribution vector $p(0)$ on the state space $E$ to the distribution $p(1)$ via one iteration of the Markov process defined by

$$p(1) = p(0) P.$$

Under some mild conditions, a Markov chain with stationary transition probabilities will also have an invariant distribution $\pi$ which satisfies the following property

$$\pi P = \pi.$$

That is, a single iteration of the Markov process leaves the distribution unchanged. The same will also hold true for any number of iterations.

### 3.3 Bayesian Inference: Bayes' Rule

Consider a hypothesis, $H$, about how some data, $D$, arises. Bayes' theorem states that

$$P(H|D) P(D) = P(D|H) P(H),$$

which can be rewritten to isolate the conditional probability of the hypothesis. In Bayesian statistics, this conditional probability of the hypothesis given data is referred to as the posterior probability. We will assume that the data is not subject to change, but there might be many different hypotheses we wish to test. Therefore, we can denote the posterior as a function of solely the hypothesis.

$$\pi(H) \stackrel{def}{=} P(H|D) = \frac{P(D|H)\,P(H)}{P(D)}.$$

Moreover, we can assume $P(D)$ to remain constant and thus acts as a normalizing factor. Because MCMC is insensitive to this normalizing factor (see Section 2.4), we will not include it in our analysis and define the posterior, up to a constant of proportionality, by

$$\pi(H) \propto P(D|H)\,P(H).$$

$P(D|H)$ is the probability of the observed data given our hypothesis, which is commonly referred to as the likelihood function of the hypothesis. $P(H)$ is the probability of the hypothesis regardless of the observed data. This probability reflects our prior knowledge of how probable different hypotheses are regardless of observed data. Thus it is referred to as the prior distribution. Using this formalism, we can now arrive at the formulation of Bayes' theorem known Bayes' rule, which lies at the heart of all Bayesian inference. This intuitive rule can be stated as follows: the posterior is proportional to the prior times the likelihood

$$posterior \propto likelihood \times prior.$$

### 3.4 MCMC: Metropolis-Hastings Algorithm

Having defined a posterior distribution, we do not have any viable means of sampling from this distribution. This is the task of the Metropolis-Hastings algorithm and other such MCMC algorithms, the former will be covered in this section as it forms the basis of almost all other algorithms we have seen in the literature. Firstly, consider an arbitrary Markov chain on the countable state space $E$ with stationary one-step transition probabilities given by $q_{ij}$, the specific stationary distribution of this Markov chain is unimportant. Secondly, consider another related Markov chain with transition probabilities given by

$$\begin{aligned} P_{ij} &= q_{ij}\alpha_{ij} \\ \alpha_{ij} &= \frac{s_{ij}}{1 + \frac{\pi_i}{\pi_j}\frac{q_{ij}}{q_{ji}}} \end{aligned} \tag{3.2}$$

where $\pi$ is a (possibly unnormalized) distribution on $E$ and $s_{ij}$ is a symmetric function of $i$ and $j$. For such a Markov chain, it can readily be shown that

$$\pi_i P_{ij} = \pi_j P_{ji}$$

which ensures that $\sum \pi_i P_{ij} = \pi_j$, for all $j$, and $\pi$ to be the stationary distribution of the transition matrix $P$. If we can produce a Markov chain whose transition probabilities obey equation (3.2), we can obtain samples from the distribution $\pi$. This is precisely what the Metropolis-Hastings algorithm is constructed to do. More

specifically, it uses

$$s_{ij}^{MH} = \begin{cases} 1 + \dfrac{\pi_i q_{ij}}{\pi_j q_{ji}} & if\ \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} \geq 1 \\ 1 + \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} & if\ \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} \leq 1 \end{cases},$$

such that

$$\alpha_{ij}^{MH} = \begin{cases} 1 & if\ \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} \geq 1 \\ \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} & if\ \dfrac{\pi_j q_{ji}}{\pi_i q_{ij}} < 1 \end{cases}.$$

It should be noted that all the transition probabilities depend on the distribution $\pi$ solely through ratios of two densities. Hence, the Metropolis-Hastings algorithm is capable of sampling from the distribution $\pi$ without knowledge of the appropriate normalization constant. For a more in-depth discussion of the details behind the Metropolis-Hastings algorithm and MCMC, we refer the reader to the seminal paper by Hastings [51].

The algorithm itself can thus be described as follows:

1. Starting with an initial value $x$, draw a new value $x'$ from a proposal density $q\left(x'|x\right)$.
2. Accept the new value $x = x'$ with the probability $\min\left(1, \dfrac{\pi\left(x'\right) q\left(x|x'\right)}{\pi\left(x\right) q\left(x'|x\right)}\right)$.
3. Restart at Step 1 until a specified number of iterations have been reached.

This simple algorithm allows one to sample from an unnormalized distribution on the countable state space $E$, and forms the basis of many Markov Chain Monte Carlo (MCMC) techniques. Despite the apparent simplicity of the algorithm there are many things to consider when using MCMC to explore multidimensional state spaces, which will be covered in the following sections.

## 3.5 Bayesian Inference: Likelihood Functions, Priors, and Experimental Data

Our goal is to use Bayesian inference in order to test Markov models against experimental data and to compare their probabilities given the data we have. In the context presented above, our hypothesis consists of a model, $M$, and specific numerical values for its various parameters which we denote as the vector $\boldsymbol{x}$. This yields

$$\pi\left(M, \boldsymbol{x}\right) \propto P\left(D|M, \boldsymbol{x}\right) P\left(M, \boldsymbol{x}\right). \tag{3.3}$$

$P\left(D|M, \boldsymbol{x}\right)$ is the probability that data would be observed given that our hypothesis, defined by the model $M$ and set of parameters $\boldsymbol{x}$, is true and is termed the likelihood function. We shall take this to be the negative exponential of the sum of squared errors (S.S.E.) between the mean value $\mu\left(D\right)$ of a quantity of interest measured from the experimental data and the same quantity produced by the differential equation model (e.g. Markov model), $\mu_M\left(y_M\left(\boldsymbol{x}\right)\right)$. i.e.,

$$P\left(D|M, \boldsymbol{x}\right) \propto \exp\left(-\sum_i {\left(\mu_{M,i}(y_M(\boldsymbol{x})) - \mu_i(D)\right)^2}/{2\sigma_i^2}\right),$$

where $\mu_i\left(D\right)$ is the mean experimental value of the $\mathrm{i}^{\mathrm{th}}$ quantity of interest and $\sigma_i^2$ is its variance (must first be estimated from experimental data). We have used $y_M\left(\boldsymbol{x}\right)$ to emphasize that the solutions to the ODE's will depend on not only on the model $M$ developed, but also the specific parameter values $\boldsymbol{x}$ associated with model $M$. Also, $\mu_{M,i}\left(\cdot\right)$ measures a quantity which is equivalent to $\mu_i\left(\cdot\right)$ but is dependent on the specification of the model $M$. For example, the quantities may be the current produced by a cell at given points in time, implying that the sum of squared errors

is the sum of squared distances between the current produced by the ODE system and the average of the current recordings produced experimentally, $\hat{I}$:

$$P\left(D|M,\boldsymbol{x}\right) \propto \exp\left(-\sum_i \left(I_M(y_i(\boldsymbol{x}))-\hat{I}_i\right)^2 / 2\sigma_i^2\right). \tag{3.4}$$

This type of likelihood function, which we refer to as a distance likelihood, works well when used to reproduce a single experimental recording. However, in this approach, the capacity for the sum of squared errors to determine parameter values is dependent on the contribution of each parameter in governing the solution of the ODE's. We have distinct phases in our experimental recordings and as a result, it would be very difficult to determine parameter values for the phases with short duration. For example, in P2XRs, the kinetics of the activation phase will depend largely on the ATP binding rates, whereas those of the desensitization phase will depend more on desensitization rates. In many of our experimental recordings, desensitization is apparent for only short periods of time and thus it is difficult to properly characterize the desensitization rates with this likelihood function. Moreover, the ODE solutions in these short phases seem physiologically unreasonable (i.e. the algorithm did not convergence to a physiologically reasonable distribution). This led us to try another approach using summary statistics to reduce the hundreds of thousands of data points recorded from cells to a handful of descriptive values. For example, we might fit an experimental current decay to an exponential function and compare the time constant, $\tau$, and amplitude, $A$, with those obtained from fittings of model currents as follows:

$$P\left(D|M,\boldsymbol{x}\right) \propto \exp\left(-(\tau_M(y(\boldsymbol{x}))-\hat{\tau})^2 / 2\sigma_\tau^2 - \left(A_M(y(\boldsymbol{x}))-\hat{A}\right)^2 / 2\sigma_A^2\right). \tag{3.5}$$

This allows us to easily give equal importance to different phases of evoked currents regardless of their duration. Moreover, because this type of likelihood function (referred to as the parametric likelihood hereafter) depends only on a few quantifiers which give a qualitative description of the solution, it is not nearly as sensitive to small changes in transient behavior as the distance likelihood function defined by equation (3.4). The success of this method hinges on how well the data is understood and an appropriate choice of quantifiers.

The expression $P(M, \boldsymbol{x})$ in Eq. (3.3) represents is the probability that the parameters of model $M$ have the given value $\boldsymbol{x}$. This represents our knowledge of the parameters prior to comparing them with the data and is termed the prior distribution. Due to their physical nature, we know that all parameters should be positive. We can also set an upper bound on state transition rates which seem appropriate given the timescales of experimental data obtained from P2X2Rs. From some preliminary testing, 10,000 $s^{-1}$ seems to be more than adequate for producing even the fastest of observed kinetics. This allows to assign a uniform probability over a range of parameter values for the prior. However, this is not particularly informative and produces chains which can have very poor convergence when used for our purposes [52, 53]. Ultimately, we have opted for the prior distribution of Siekmann et. al. which is exponentially distributed in the trace of the Markov model's infinitesimal generator $Q$ (see section §3.6) [54]. i.e., by taking

$$P(M, \boldsymbol{x}) \propto \exp\left(\frac{Tr\left(Q\left(\boldsymbol{x}\right)\right)}{\rho}\right),$$

where $\rho \in \mathbb{R}^+$. This expression is a well established prior distribution for ion channel parameter estimation [54]. Such a choice of prior prevents model rate parameters from becoming excessively large much as the uniform prior would. In fact, it will assign higher posterior density to parameter sets with smaller values that can reproduce the experimental data equally by skewing samples towards smaller values. This is because the prior distribution assumes that all the transition probability rates $q_{ij}$ are independent exponentially distributed random variables with parameter $\rho$ rather than being uniformly distributed within some interval. As was suggested by Siekmann et. al., we have used $\rho = 30$.

### 3.6 MCMC Estimation for an Ion Channel Model

Our goal is to use MCMC methods to infer parameters values for Markov models of an ion channel. Given a set of experimental data and an ion channel model, we wish to estimate parameter values for the model. In this thesis, we typically deal with models which are quite large and thus have a large number of parameters. The high dimensionality of the parameter space as well as the computational time required to evaluate the likelihood functions prohibits us from obtaining distributions of parameters. While MCMC does sample from the posterior distribution, there is a general rule of thumb that at least $2^k$ samples should be used for Bayesian analysis of a k-dimensional sample space. Unfortunately this was not feasible for us and so we have used MCMC as a stochastic optimization algorithm. Deterministic optimization was initially explored but when applied to complex models and sets of data, it converged to local minima that were not physiologically reasonable. MCMC has the capacity to escape these local minima (see section §3.7). Rather than obtaining the

parameter distribution, $\pi(x)$, we instead generated an estimate of the parameter set with the most posterior density $\hat{x} = arg\,\underset{x}{max}\ \pi(x)$. To illustrate how this works, we will explore a simple example of a Markov model and some constructed experimental data to see the effects of different likelihood functions on the sampled distributions. By calculating the mean and variance of currents as a function of time, this example will demonstrate the effectiveness of the likelihood function of equation (3.4) for estimation of parameters from experimental data and show that it consistently performs worse than that of equation (3.5).

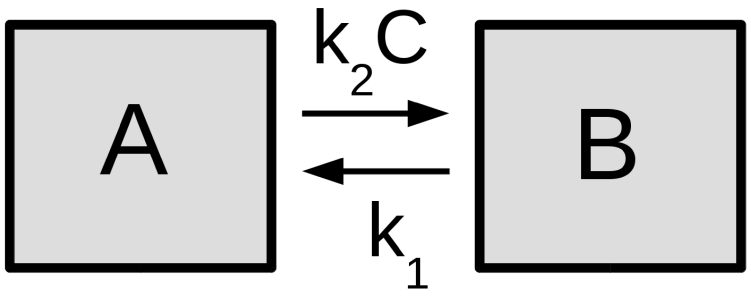


Figure 3–1: Simple Markov model of a ligand-gated ion channel.

We consider the simplest ligand-gated ion channel Markov model with two states: a closed state $A$ and an open state $B$ with conductance $g$ (see Fig. 3–1). In this model, the transition from $A$ to $B$ requires the presence of the a $C$. The model can be described by the following equations:

$$\begin{aligned} \frac{dA}{dt} &= k_1 B - k_2 C A \\ \frac{dB}{dt} &= k_2 C A - k_1 B \\ I(t) &= gB(t). \end{aligned}$$

Or in terms of the infinitesimal generator $Q$

$$\begin{aligned} \frac{dp(t)}{dt} &= p(t)\,Q \\ p(t) &= \begin{bmatrix} A(t) & B(t) \end{bmatrix} \\ Q &= \begin{bmatrix} -k_2 C & k_2 C \\ k_1 & -k_1 \end{bmatrix}. \end{aligned}$$

Because of the low-dimensionality of the system, it is possible to obtain analytical solutions and thus sample the posterior extensively. This will not be possible for larger models, where we will need to employ an ODE solver. However, the properties of the different posterior distributions can be extended to larger models.

The experimental characterization of any ion channel will require the use of multiple experimental protocols. When an ion channel model is under consideration, we seek to find the parameter set which best explains the behavior of the ion channel in all experimental conditions. For this reason, we will simulate two experimental protocols. The first is the application of the ligand $C$ at a concentration of 1 M for 4 seconds to a cell whose receptors are all in the closed state $A$. The second is of a cell whose receptors are all initially in the open state $B$, and the ligand $C$ is applied at a concentration of 1 M after 4 seconds of allowing receptors to close. Mimicking recording from a large number of cells, an ensemble of currents for both experimental protocols were generated (see Fig. 3–2) using normal distributions of the parameters $k_1$, $k_2$ and $g$. This allows us to use the distance likelihood function of equation (3.4) in order to redetermine these parameters. We can also fit the various phases of the

current to appropriate exponential functions, in order to determine parameters using the parametric likelihood function of equation (3.5).

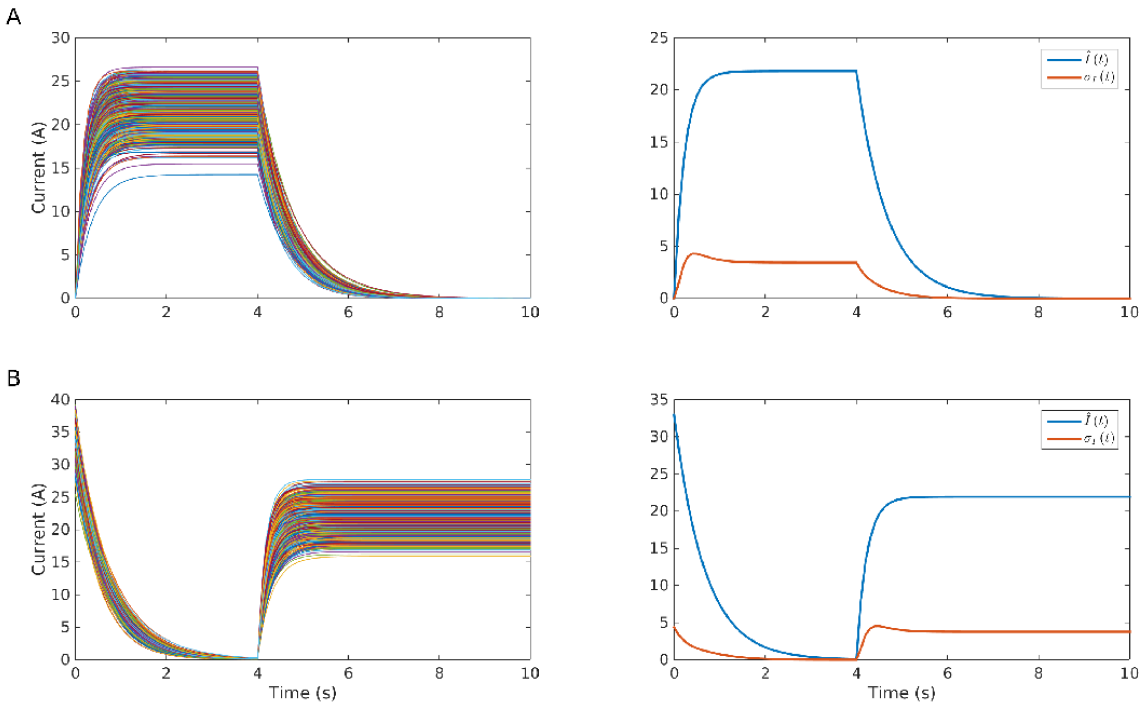


Figure 3–2: Simulated experimental datasets used for inference of parameters in the model of Fig. 3–1. On the left is the ensemble of generated currents and on the right are the estimated mean and variance of the current traces. Parameters were randomly chosen using $k_1 \sim \mathcal{N}(1.5, 0.1)$, $k_2 \sim \mathcal{N}(3, 0.5)$, $g \sim \mathcal{N}(33, 2)$. (A) Ligand $C$ is applied to a "closed cell" at a concentration of 1 M for 4 seconds. (B) An "open cell" is allowed to close for 4 seconds at which point the ligand $C$ is applied at a concentration of 1 M.

We found that when we used a distance likelihood, the parameter samples differed from one dataset to another. More specifically, the spread around the most likely value (the variance) differed significantly from the actual distribution (see Fig. 3–3A,B). We also found that varying the number of datapoints in each recording can

alter the variance significantly. Curiously, the samples of parameter $k_1$ were consistently confined to a very small region not representative of the actual distribution used to generate the data. This sensitivity to experimental protocol and sampling rate is certainly undesirable. *A priori*, we do not know how each experiment will skew the parameter samples, so it is impossible to compensate for this effect. Ultimately, this leads to the issue of our sampler rejecting parameter values which are perfectly reasonable. In a situation with a more complex set of experimental data, this can make it impossible to reproduce all experimental recordings very well with any sampled parameter set. On the other hand, using a parametric likelihood function seems to circumvent all of these issues. The sampler managed to sample the parameters $k_1$ and $k_2$ extremely well while more data would have been needed to determine the conductance $g$ to the same precision. As we have already stated, the overestimated variance of $g$ is not a concern for us since the sampler did not reject values which gave perfectly reasonable results. The excessive rejection of parameters leads to undersampling of the parameter space and thus limits the usefulness of our sampling methods.

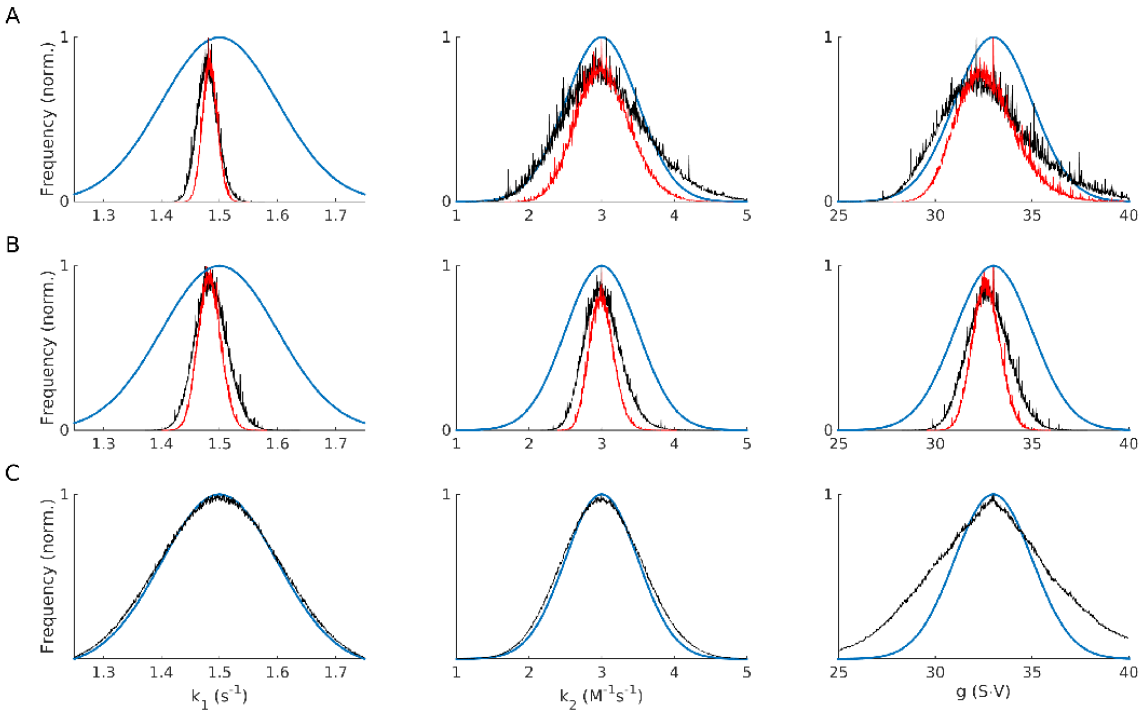


Figure 3–3: Effect of likelihood functions on parameter samples. From left to right the sampled parameters are $k_1$, $k_2$, and g respectively. In blue is the scaled frequency of parameters used to generate the data. (A,B) Parameters sampled using the distance likelihood function of equation (3.4). The currents used in A and B correspond to those in A and B of Fig. 3–2 respectively. Observe that the two datasets produce samples with very significantly different variances for $k_2$ and $g$. The data sampling rate also affects the variance of sampling. The black represents the sampling interval of 0.1 s whereas the red represents the sampling interval of 0.05 s. Increasing the sampling rate has the effect of increasing the confinement of parameter samples. (C) Parameters sampled using the parametric likelihood function of equation (3.5). Parameter samples of the $k_1$ and $k_2$ match very closely to their actual distributions. The variance of the conductance, $g$, is overestimated but the mean is well characterized. Changing the dataset used for fitting or the data sampling rate has little to no effect on the sampled distributions (data not shown).

Likelihood functions are a measure of the probability of the data given a model. It is unlikely that any model we design will ever explain all facets of experimental

data with a single parameter set. Aside from the problem of using an inadequate model, there are issues related to the way that experimental data is collected. Certain experimental protocols may have very low success rates and may, by the intrinsic properties of the system under study, effectively produce useful data for only certain regions of parameter space. Moreover, some experimental protocols may require so many attempts before obtaining useful data that only a single record of the experiment is available, which only yields information about a single point in parameter space. As an extreme example, consider the case where the experimental protocol of Fig. 3–2 A is such that the average ligand unbinding rate appears to be 1.5 $s^{-1}$ and that of B appears to be 4.5 $s^{-1}$. What will be the outcome of sampling parameters for the joint likelihood function of the two experiments? Obviously it will be impossible to fit both experiments perfectly well with a single parameter set. Alternatively, we can think of our simple two state model as an approximation of a more complicated model of multiple states with these two effective unbinding rates. As a result, we can still recover the average unbinding rate (3 $s^{-1}$) through sampling using the simplified model of Fig. 3–1.

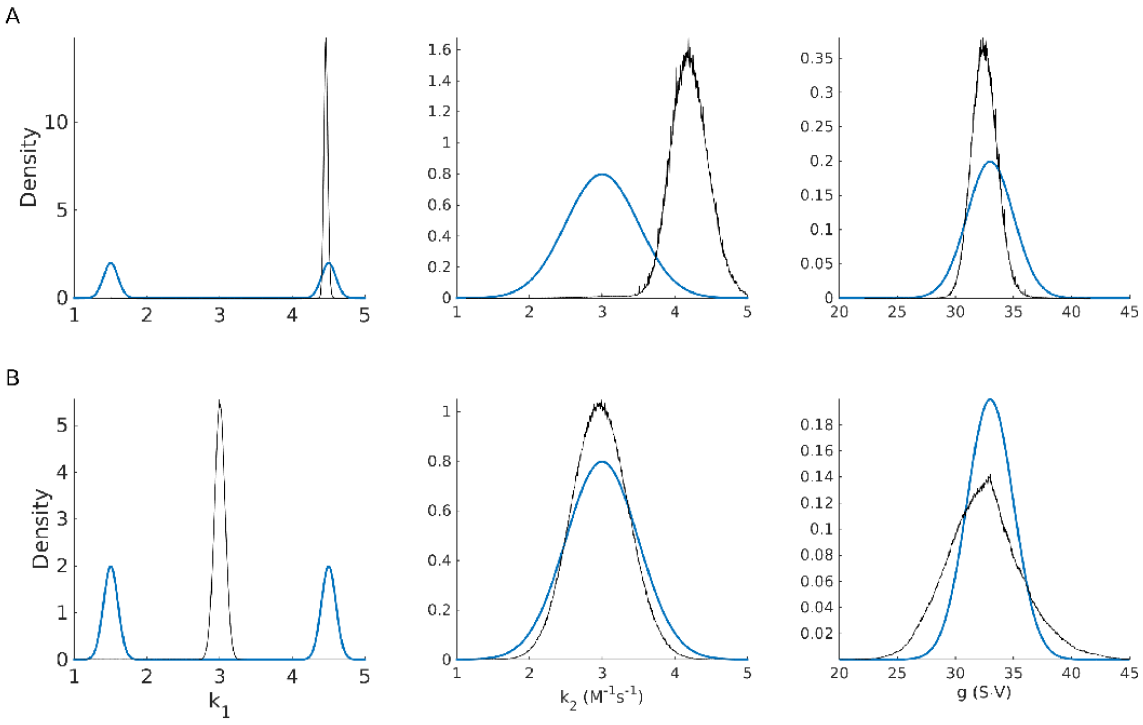


Figure 3–4: Effect of likelihood functions on parameter samples for data showing a strong heterogeneity in rate parameters. From left to right the sampled parameters are $k_1$, $k_2$, and $g$ respectively. In blue is the density function of parameters used to generate the data and in black is the sampled density. (A) Parameters sampled using the distance likelihood function of equation (3.4). Notice how samples of $k_1$ have been skewed towards the mode on the right and samples of $k_2$ do not have the correct mean value. (B) Parameters sampled using the parametric likelihood function of equation (3.5). Parameter samples of $k_1$ are indicative of the average ligand unbinding rate despite neither of the modes being well sampled. Parameter samples for $g$ and $k_2$ estimate the correct mean value and are less confined than in A.

By comparing the performance of the distance likelihood function with that of the parametric one to estimate the average, we found that the former heavily favors the faster unbinding rate ($k_1 = 4.5$ s$^{-1}$), despite all chains being started with $k_1 = 1.5$ s$^{-1}$

(see Fig. 3–4A). This means that the most likely outcome of sampling is a parameter set which reproduces one of the two experiments very well but not the other. Also, samples of $k_2$ incorrectly estimate the mean despite using the same parameter distribution for both datasets. We have not done extensive analysis as to why one dataset was favored over the other, but it is likely that it has something to do with the temporal structure of the variances in each dataset. When using the parametric likelihood, on the other hand, all average parameter samples correctly estimated the mean of the parameter distributions. Samples also had a larger variance compared to the distance likelihood. While the most likely parameter set for the parametric likelihood does not reproduce either dataset exceptionally well, it manages to make a compromise between the two rather than neglecting one of the modes (and its corresponding dataset). This kind of compromise in the sampler is greatly desired as it will allow for simple models (such as the one in Fig. 3–1) to exhibit more flexibility in reproducing experimental data, as opposed to favoring one set of data for unclear reasons. For these reasons, we believe that parametric likelihood functions are superior in their capacity for parameter estimation of ODE models.

### 3.7 MCMC: Trapping & Tempering

For distributions with multiple modes which do not have significant regions of overlap, the random walking of the Metropolis algorithm tends to get stuck in a single mode for long periods of time, leading to convergence issues and resulting in a severely incomplete sampling of the parameter space. To illustrate, we have produced an example density function which is the sum of three normal distributions ($P(x) = exp\left(-(x-\mu)^2/2\sigma^2\right)/\sqrt{2\pi\sigma^2}$, see Fig. 3–5A). Two modes ($\mu = 18, \mu = 20$) are

close enough that there is significant overlap of the distributions while a third mode ($\mu = 9$) is chosen to be well separated by a region of low density. To demonstrate the problem of local trapping, we ran 1000 MCMC simulations and analyzed the chains produced, all starting at the mode of one of the two closely packed distributions on the right ($x_0 = 20$). We found that only 6.1% of the chains ever manage to reach the more distant mode on the left ($\mu = 9$) and that in those rare cases it took around 500,000 iterations to reach the leftmost mode.

In all MCMC simulations a starting point, or initial set of parameters, is required. In the situation where we have managed to find one mode to the right ($x_0 = 20$ or $x_0 = 19$ ) *a priori,* we cannot expect to discover the leftmost mode or vice versa. This is quite problematic as we would like to use MCMC as an automated means of exploring parameter space in order to find all parameter sets of interest. Moreover, the evaluation of the likelihood function is significantly more computationally expensive when solving the ODE's of a Markov model. Even on specialized computers, the time required to produce 1,000,000 samples is prohibitively long and thus compels us to improve sampling efficiency.

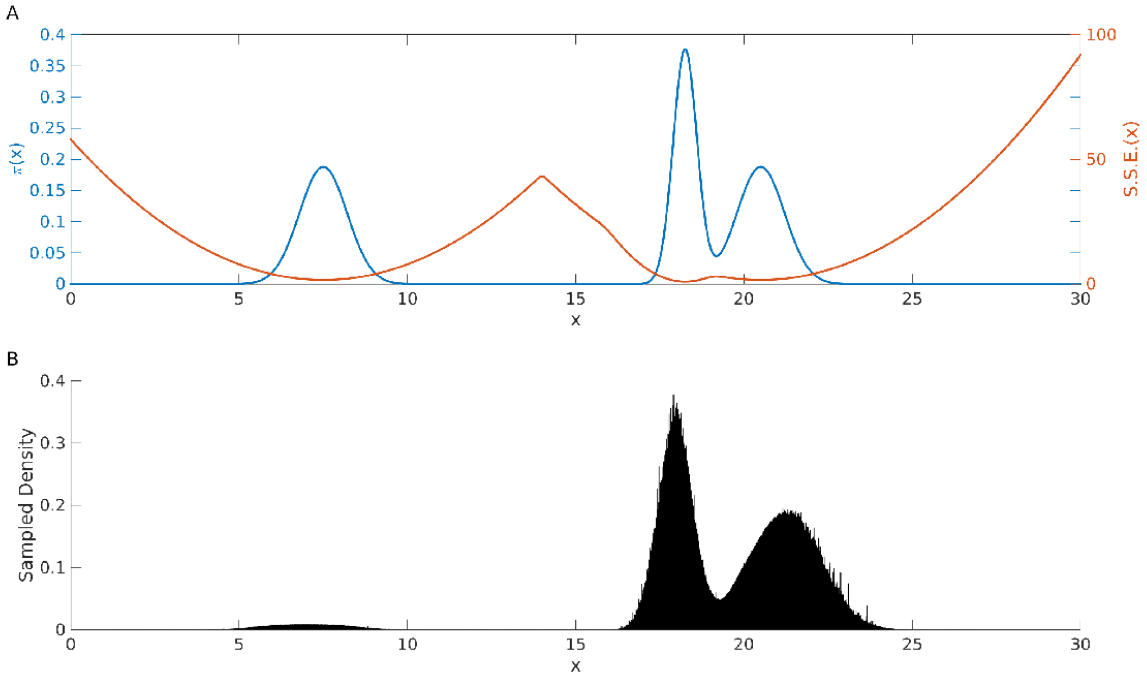


Figure 3–5: Incomplete sampling of the sample space due to being stuck in one mode. A) Posterior p.d.f. $\pi(x)$ (in blue) and sum of square errors $S.S.E(x)$ (in orange) as a function of a single parameter x. B) MCMC average sampled density of the p.d.f from A. All chains were started in the rightmost mode ($x_0$=20) and run for 1,000,000 iterations. Shown is the average density of 1000 such chains, of which only 6.1% ever reached the leftmost mode.

In order to overcome this local trapping of the sampler, we can employ a method known as tempering, where the sampled posterior is a "flattened" version of the posterior of interest.

$$\pi^{\beta}(\boldsymbol{x}),$$

where $\beta$ is referred to as the inverse temperature (confined to be between $(0, 1]$). The higher the temperature, the flatter the posterior (see Fig. 3–6). For our chosen likelihood function, this tempering is entirely equivalent to multiplying the S.S.E by

the scalar $\beta$. At high temperature (small $\beta$), it becomes much more likely to move through regions which are relatively too low in density prior to tempering.

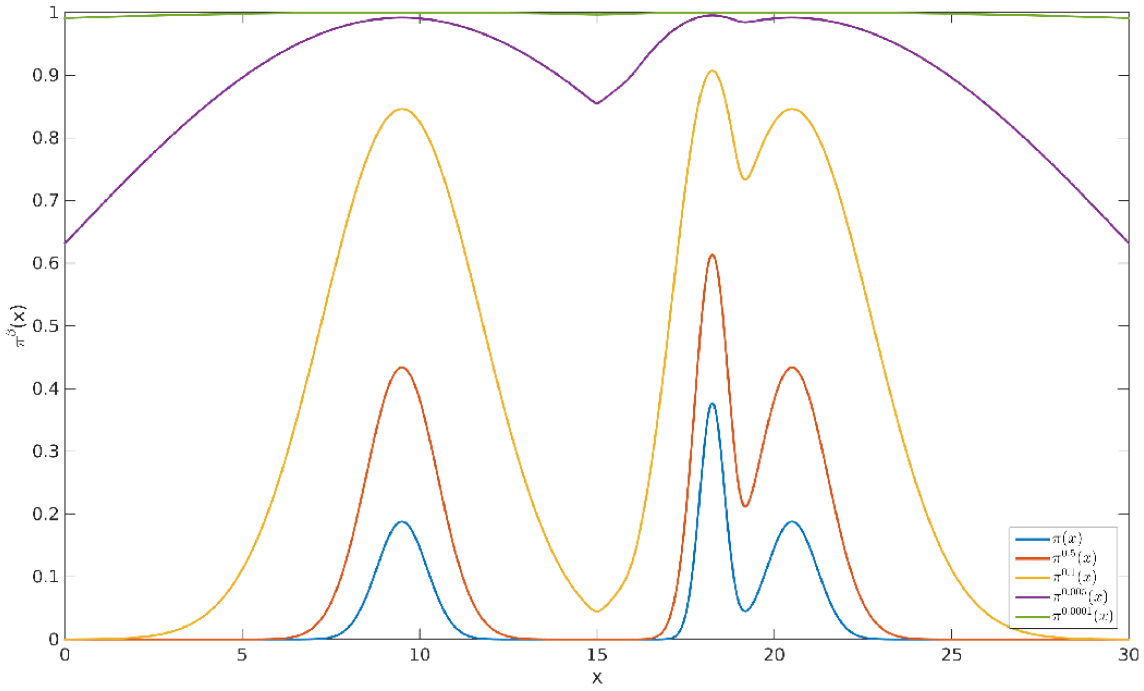


Figure 3–6: Tempering of the posterior probability density function. Notice how reducing the inverse temperature, $\beta$, leads to a flatter distribution.

We have sampled tempered versions of the posterior that locally trapped our sampler in Fig. 3–5 in order to illustrate the effect of tempering on a sampler. With a modest amount of tempering ($\beta = 0.5$, Fig. 3–7B), we were able to reach the leftmost mode with all chains. However, on average, it took 242,000 iterations to reach the leftmost mode. This is better than the average 500,000 iterations it took the untempered chains, but is still rather large. As tempering is increased ($\beta = 0.1$, Fig. 3–7C), we begin to see more significant spreading of the modes. Again, all chains reach the

leftmost mode but on average it took 762 iterations to reach it, which is a significant increase in sampling efficiency compared to the untempered case (see Table 3–1). Further increasing the temperature ($\beta = 0.005$, Fig. 3–7D), continued to increase the sampler's capacity to sample regions with very little density in such a way that it seemed not to be constrained within either of the modes (with the two modes on the right becoming indistinguishable). At this temperature level, chains reached the second mode on average after 580 iterations. In the limit $\beta \to 0$ the sampler is no longer informed by the posterior and the parameters are sampled uniformly, an outcome that we begin to observe in Fig. 3–7D.

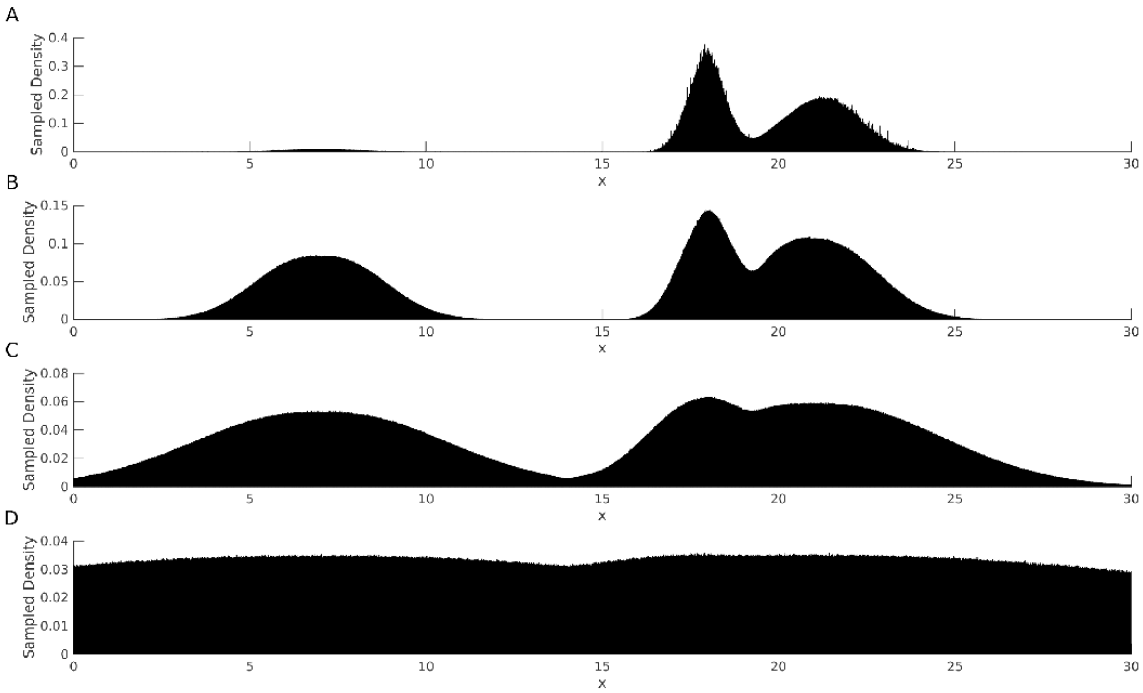


Figure 3–7: Effect of tempering on MCMC sampling. (A) MCMC of the original posterior. (B) Sampling of tempered posterior with $\beta = 0.5$. (C) Sampling of tempered posterior with $\beta = 0.1$ (D) Sampling of tempered posterior with $\beta = 0.005$. All chains had the same starting point at the rightmost mode ($x_0 = 20$) and were run for 1,000,000 iterations. Each plotted density was calculated as an average of 100 independent chains.

| $\beta$ | Chains reaching $x_i = 9$ | Average iterations required for $x_i = 9$ |
|---|---|---|
| 1 | 6.1% | 506,885 |
| 0.5 | 100% | 242,828 |
| 0.1 | 100% | 762 |
| 0.005 | 100% | 580 |

Table 3–1: Tempered MCMC sampling statistics from a multimodal distribution (n≥100).

With this simple example we can see both the strengths and weaknesses of sampling from tempered distributions. On one hand, it can increase the mixing of modes and even allow for exploration of regions of parameter space which might otherwise be missed by MCMC simulations of a finite length. On the other hand, it deforms the distribution of interest such that modes which are close by may become indistinguishable and in the limit of infinite tempering, all modes become essentially erased. None of the samples really gives us a perfect description of the posterior distribution, because either samples are too spread out around the modes or they tend to be trapped locally for very long periods of time. There is certainly some value of $\beta$ which is a good compromise between the two. How are we to choose it *a priori* when we begin our simulation? Or conversely, how can we determine whether or not it truly has explored the entire parameter space while giving importance to the appropriate regions? We do not know how to answer either question in a satisfactory manner. Instead, what we have found to be a common practice in the literature is the use of a more elaborate sampling procedure known as parallel tempering [55].

### 3.8 Parallel Tempering

Parallel Tempering (PT) is a technique which uses Markov chains generated in parallel from different tempered versions of the same distribution. Considering $L$ ordered temperature levels, $1 = \beta^{(1)} > \beta^{(2)} > \ldots > \beta^{(L)} > 0$, we can construct the following composite Markov chain

$$\boldsymbol{X}_k = \left( X_k^{(1)}, \ldots, X_k^{(L)} \right),$$

which has, as a stationary distribution, the product of posterior densities

$$\boldsymbol{\pi}_{\boldsymbol{\beta}}\left(x^{(1)},\ldots,x^{(L)}\right)=\frac{\pi^{\beta^{(1)}}\left(x^{(1)}\right)}{\int\pi^{\beta^{(1)}}\left(x\right)dx}\times\cdots\times\frac{\pi^{\beta^{(L)}}\left(x^{(L)}\right)}{\int\pi^{\beta^{(L)}}\left(x\right)dx}.$$

The Parallel Tempering algorithm can be reduced to successively applying the following two step scheme

$$\boldsymbol{X}_n\overset{swap}{\longrightarrow}\bar{\boldsymbol{X}}_n\overset{move}{\longrightarrow}\boldsymbol{X}_{n+1}.$$

Based on this, each Markov chain moves forward according to the Metropolis-Hastings algorithm for its tempered posterior, but this is also accompanied by swapping of parameters between chains at adjacent energy levels. We denote the parameters after swapping $\bar{\boldsymbol{X}}_k$. This swapping takes place in a very similar fashion to the metropolis step. The acceptance probability of a swap takes into account the energy difference between levels, such that adjacent energy levels which are closer in energy have a higher probability of exchange. This probability is defined by

$$P(\bar{X}_n^j=X_n^{j+1},\bar{X}_n^{j+1}=X_n^j)=min\left(1,\left(\frac{\pi\left(X_n^{j+1}\right)}{\pi\left(X_n^j\right)}\right)^{\beta^{(j)}-\beta^{(j+1)}}\right).$$

Repeating the same multimodal sampling task using the parallel tempering algorithm, we see a significant improvement in sampling efficiency. Firstly, on average, all chains manage to reach the second mode in fewer iterations than in the best case for serial tempering (see Table 3–2). Second, the untempered chain ($\beta=1$), by virtue of being run in parallel with tempered chains, reaches the second mode in all instances compared to only 6% for serial tempering (see Fig. 3–8A). One might

question if this is truly more efficient considering the necessity to evaluate the likelihood function L times for a single step of the PT algorithm. In this example, each iteration required the four likelihood evaluations (one for each temperature), but the speed of mode-mixing in the untempered chain was increased by 1000-fold. This is clearly an increase in sampling efficiency. Moreover, because the evaluations of the likelihood function at each temperature are independent of one another, they can be parallelized on a modern digital computer, allowing for L independent likelihood function evaluations in parallel (which requires far less time than L successive evaluations of the same likelihood function). In cases where the increase in mode mixing might not be as extreme, parallelization of the likelihood functions can still increase the number of samples obtained in a given amount of time [56]. This is particularly useful when the likelihood functions are computationally intensive such as when fitting large ODE systems to complex experimental data.

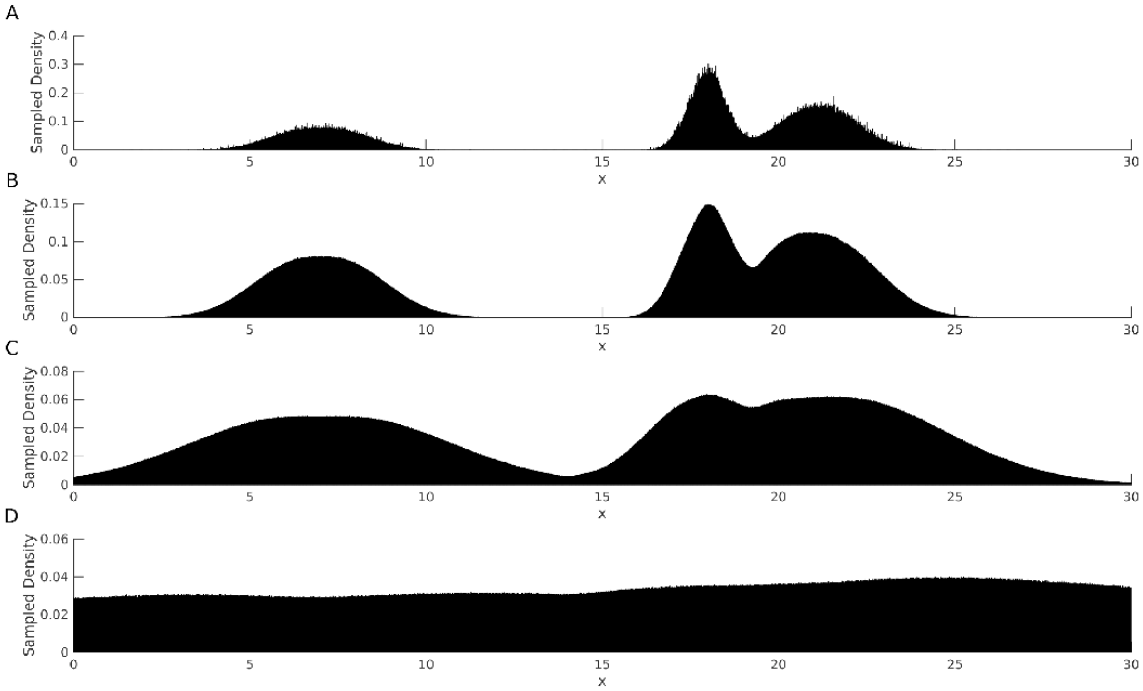


Figure 3–8: Effect of parallel tempering on MCMC sampling. (A) Sampling of the original posterior using the parallel tempering algorithm. (B) Parallel sampling of tempered posterior with $\beta = 0.5$. (C) Parallel sampling of tempered posterior with $\beta = 0.1$ (D) Parallel sampling of tempered posterior with $\beta = 0.005$. All chains had the same starting point at the rightmost mode ($x_0$=20) and were run for 10,000 iterations. Each plotted density was calculated as an average of 10,000 independent chains.

| $\beta$ | Chains reaching $x_i = 9$ | Average iterations required for $x_i = 9$ |
|---|---|---|
| 1 | 100% | 397 |
| 0.5 | 100% | 330 |
| 0.1 | 100% | 309 |
| 0.005 | 100% | 372 |

Table 3–2: Parallel Tempered MCMC sampling statistics from a multimodal distribution (n=100).

### 3.9 Adaptive Parallel Tempering (APT)

In the previous section, the choice of inverse temperatures ($\beta^{(\ell)}$) at which to run the parallel chains was arbitrarily chosen by trial and error. This choice worked well for the particular problem on hand, but may have been sub-optimal for others. The adaptive parallel tempering algorithm (APT) attempts to vary the temperature of the chains at each step in order to asymptotically obtain a specified swap acceptance rate, $\alpha^*$, between adjacent temperature chains [57]. Theoretical results suggest that an acceptance rate of 23.4% is optimal for maximizing the expected square jumping distance between individual swaps, and thus optimizing the exploration of parameter space. First, we introduce a parametrization, $\boldsymbol{\rho}$, of the inverse temperatures.

$$\{\boldsymbol{\rho}_n\}_{n\geq 0} \stackrel{def}{=} \left\{\rho_n^{(1:L-1)}\right\}_{n\geq 0}.$$

The parametrization is used to recursively determine the inverse temperatures as follows:

$$\beta^{(\ell+1)}\left(\rho^{(1:\ell)}\right) := \psi\left(\beta^{(\ell)}\left(\rho^{(1:\ell-1)}\right), \rho^{(\ell)}\right).$$

This parametrization provides a general formalism that can be used to produce the increasing sequence of temperatures. In particular, we use $\rho^{(\ell)} = log\left(\frac{1}{\beta^{(\ell+1)}} - \frac{1}{\beta^{(\ell)}}\right)$, as suggested by Miasojedow et. al., which in turn determines the inverse temperatures, as follows:

$$\beta^{(\ell+1)} = \frac{1}{\frac{1}{\beta^{(\ell)}} + \exp\left(\rho^{(l)}\right)}.$$

This parametrization is updated at each iteration of the algorithm by the stochastic approximation procedure

$$\rho_n^{(\ell)} = \left(\rho_{n-1}^{(\ell)} + \gamma_{n,1} H^{(\ell)}\left(\rho_{n-1}^{(1:\ell)}, \boldsymbol{x}_n\right)\right),$$

where the stepsize $\gamma_{n,1}$ is given by $\gamma_{n,1} = (n+1)^{\xi_1}, \xi_1 \in (1/2, 1)$ and

$$H^{(\ell)}\left(\rho_{n-1}^{(1:\ell)}, \boldsymbol{x}_n\right) = min\left(1, \left(\frac{\pi\left(x^{(\ell+1)}\right)}{\pi\left(x^{(\ell)}\right)}\right)^{\Delta\beta^{(\ell)}\left(\rho^{(1:\ell)}\right)}\right) - \alpha^*$$

$$\Delta\beta^{(\ell)}\left(\rho^{(1:\ell)}\right) = \beta^{(\ell)}\left(\rho^{(1:\ell-1)}\right) - \beta^{(\ell+1)}\left(\rho^{(1:\ell)}\right)$$

When the mean acceptance rate between an energy level $\ell$ and its adjacent level $\ell+1$ reaches $\alpha^*$, then $H \to 0$ and the parameter $\rho^{(\ell)}$ becomes quasi-stationary (in the sense that it will remain constant as long as the mean swap acceptance rate do not change). We refer the reader to the original paper by Miasojedow et. al. for more details [57].

One thing which we noticed when using the APT sampler was that it was not uncommon for the sampler to stray from parameter sets, which produced results relatively close to those we wished to see, to parameter sets which performed truly abysmally. Due to the Markov property, it is entirely possible for a random walk metropolis algorithm to move far from the mode in which it is initialized. This problem can be exacerbated by tempering. In particular, this can be observed in Fig. 3–7D, where the modes do not seem to have any real capacity to confine the chain to any part of the parameter space. Moreover, Miasojedow et. al. warn that in the case of a posterior with support of finite Lesbegue measure, it is possible that there exists no

solution $\hat{\boldsymbol{\rho}}$ for the parametrization of the inverse temperatures and their algorithm will assymptotically sample from a uniform distribution. The APT algorithm makes it such that once an energy level begins to sample from a uniform distribution, it will tend to swap parameters with an adjacent level and ultimately restart the process of parameter estimation without any of the knowledge accumulated by running the algorithm. Since it is the Markov property of a single chain that allows for excessive sampling of parameter space, we turn our attention to a class of samplers which are not Markovian in parameter space.

### 3.10 T-walk sampler

The t-walk is an MCMC sampler which is designed to be invariant to the structure of the target distribution. It is not an adaptive algorithm that performs tuning of parameters which govern the algorithm. Rather, it maintains two points $(x, x')$ in parameter space which stochastically perturb each other using proposal distributions. This effectively makes this method sample from the joint posterior $\pi(x)\,\pi(x')$. Instead of producing two independent chains in parameter space $\mathcal{X}$, the whole process lies in the product space $\mathcal{X} \times \mathcal{X}$ and thus the probability of a single point randomly walking to a region of low density will be greatly reduced. Only when a wandering point finds a region of higher density will it be likely to influence the position of the other point in parameter space. For this reason, its creators gave it the moniker "thoughtful walk" to distinguish it from random walk MCMC [58]. By maintaining two points in parameter space, the local shape of the distribution may be sampled efficiently because the two points carry information about the local structure of the distribution. The algorithm produces a new set of points in parameter space $(y, y')$

using the restricted proposal

$$(y, y') = \begin{cases} (x, h(x', x)) & with\ prob\ 0.5 \\ (h(x, x'), x') & with\ prob\ 0.5 \end{cases} \tag{3.6}$$

where $h(x, x')$ is itself a random variable whose p.d.f. is given by $g(\cdot|x, x')$. The Metropolis-Hastings acceptance ratio for the first case of equations is given by

$$\frac{\pi(y')\, g(x'|y', x)}{\pi(x')\, g(y'|x', x)},$$

whereas in the second case it is given by

$$\frac{\pi(y)\, g(x|y, x')}{\pi(x)\, g(y|x, x')}.$$

The t-walk uses four different proposal distributions: the walk move, the traverse move, and the hop and blow moves. Each proposal distribution is used with a certain probability. The details of the construction of each distribution are beyond the scope of this thesis but can be found in the seminal paper of Christen et. al. [58]. MATLAB implementations of the sampler are available online, but these codes required some modifications to make them suitable to work with high dimensional problems. Particularly, for high dimensional problems a random subset of the parameters are chosen to be updated rather than the whole set. For an n-dimensional set of parameters, this is done by sampling indicator variables, which are used to stochastically choose coordinates to vary according to a Bernouilli distribution, $I_j \sim Be(p), j = 1, 2, \ldots, n$. If $I_j = 0$, then the coordinate $x_j$ is not moved. The parameter p of the Bernouilli distribution is chosen to be $p = {}^{min(n, n_1)}/_{n}$, such that the expected number of moved

parameters at each iteration is $n_1$ for sample spaces of dimension greater than $n_1$ (chosen to be 4, as suggested by Christen et. al.).

### 3.11 Composite MCMC Sampler

Because certain aspects of both the APT and t-walk are useful, we developed our own sampler with properties of both. Here we describe briefly the sampler's algorithm. $L$ tempered t-walk samplers sample from the product space $\mathcal{X} \times \mathcal{X}$ at each iteration, producing the move step. After each move step, individual points in $\mathcal{X}$ are swapped between tempered chains according to the parallel tempering algorithm. The temperatures of each tempered chain are then updated according to the stochastic approximation defined by the APT sampler. Ultimately, this sampler samples from the product distribution given by

$$\boldsymbol{\pi}_{\boldsymbol{\beta}}(\mathbf{x})\,\boldsymbol{\pi}_{\boldsymbol{\beta}}(\mathbf{x}') = \frac{\pi^{\beta(1)}(x^{(1)})\pi^{\beta(1)}(x'^{(1)})}{\int \pi^{\beta(1)}(x)\pi^{\beta(1)}(x')dxdx'} \times \ldots \times \frac{\pi^{\beta(L)}(x^{(L)})\pi^{\beta(L)}(x'^{(L)})}{\int \pi^{\beta(L)}(x)\pi^{\beta(L)}(x')dxdx'}.$$

We found that this sampler was capable of quickly converging to regions of high density much like the t-walk but did not suffer as much from low move acceptance rates once the region of high density is reached (as the temperatures of the tempered chains could be adjusted to the properties of the local distribution using the APT temperature scheduling). This sampler was built from algorithms, rather than from mathematically analyzing proposal distributions. We diagnosed computational issues in each sampler that reduced their sampling efficiency and borrowed what we liked from each in order to overcome the limitations of the other. While a full suite of tests have been not performed to test the sampler's properties, we have been quite satisfied with its performance thus far. There does exist a generalization of the t-walk that

can handle any number of points in parameter space [59]. However, we found that the implementation of this sampler would have required too much time to implement based on our needs and would have produced significantly slower likelihood function evaluations. Thus it was never used, but it would be the next logical step in our efforts to find an efficient MCMC sampler for our problem.

Inclusion of the code within the appendices of this thesis is impossible due to constraints on length. Instead, we have made the code and some examples of how to use it publicly available at the following URL: https://github.com/metapfhor/MCION.

# Chapter 4
# Extension of P2X4R model

The model in Fig. 2–9 was successful in capturing many aspects of P2X4 gating except for one minor detail. The model IV curves transiently begin to increase in slope, indicative of an increase in conductance (see Fig. 4–1C). However, the slope of experimental IV curves decreases monotonically for as long as agonist is applied. There is also one experimental result with which this model is conceptually incompatible. IVM must be applied for some time before the application of ATP in order to have its full effect [41]. In the aforementioned model, applying IVM in the absence of ATP has no effect whatsoever on the dynamics of the model because the transition mediated by IVM from $Q_2$ to $Q_3$ can only occur when ATP is present. Thus, in order to generate dose response curves compatible with experiments (see Fig. 4–2 and section 4.1.3), it was necessary to have the IVM-dependent nonlinearities included in the naïve row. This incompatibility motivated us to further work in analyzing the interactions between receptors and IVM in more quantitative detail. here are three important questions that need to be addressed: (i) how do we model ATP-independent IVM interaction with P2X4R? (ii) which gating properties of P2X4R are affected by pretreatment with IVM; and (iii) What are the mechanisms underlying the decrease in the slopes of the IV-curves during IVM-induced sensitization? As we have mentioned, in the model of Fig. 2–9, there is only one ATP dependent IVM-receptor interaction; namely, the transition from $Q_2$ to $Q_3$. Moreover, the

model included only one such interaction although experimental evidence that there should be at least two distinct interactions [60, 41].

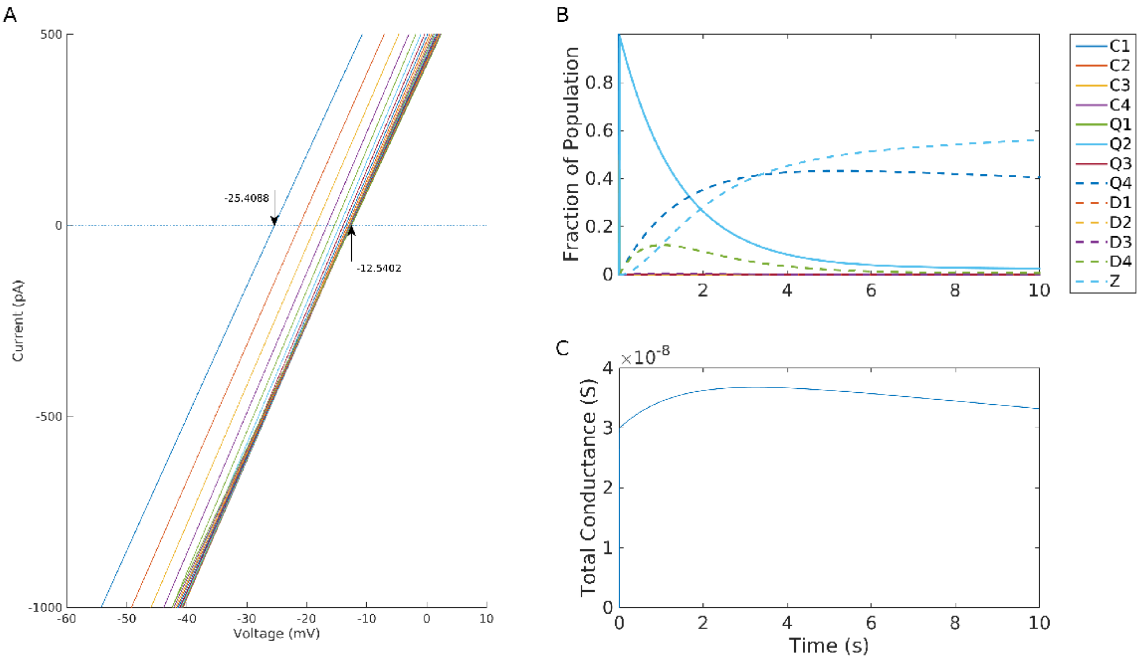


Figure 4–1: IV curves for the model in Fig. 2–9 produce a transient increase in conductance. (A) Model simulation of positive shift in reversal potential observed during initial 10 s of 100 μM ATP application in the presence of 3 μM IVM. (B) The time series of model states during the experimental protocol (C) The time series of the total conductance of a model cell. Notice the transient increase in conductance during the first three seconds of agonist application.

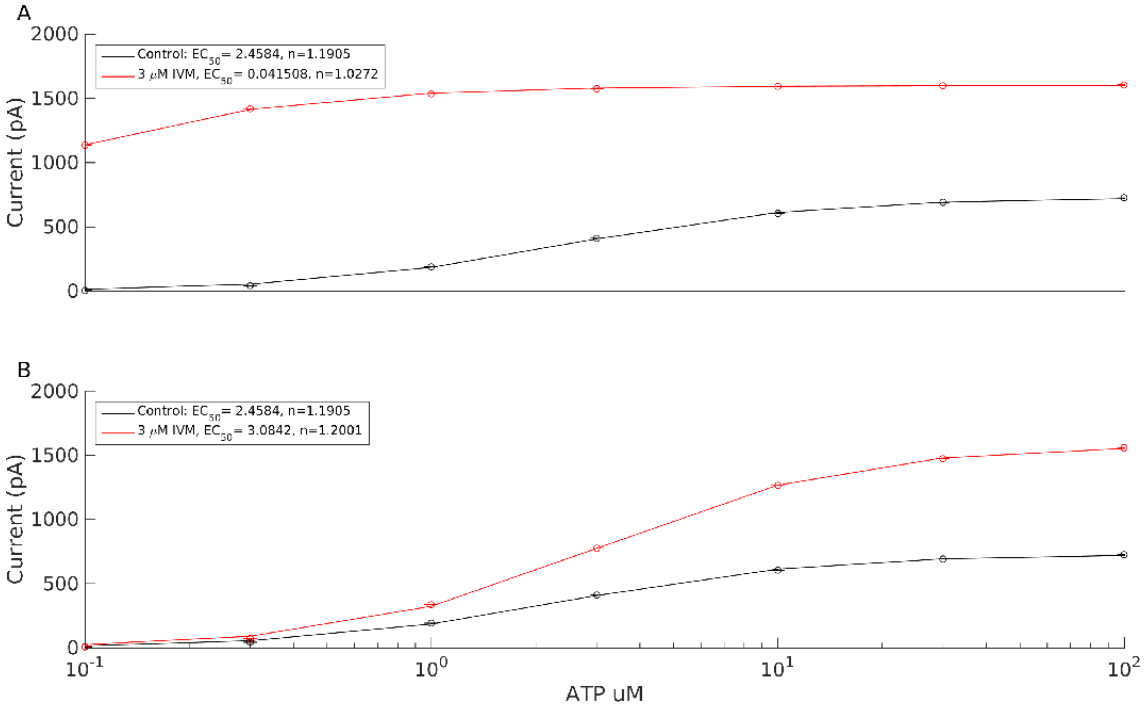


Figure 4–2: Nonlinearities in the naïve row of the model in Fig. 2–9 are necessary to capture IVM's sensitizing effect. (A) Model simulation of evoked current as a function of ATP concentration in the absence (black) and presence (red) of IVM. In the presence of IVM, the $EC_{50}$ is lower than in the control. This indicates that IVM sensitizes model receptors (see section 4.1.3). (B) Same as A but with the nonlinearities removed from the naïve row of the model. Notice that application of IVM leads to a higher $EC_{50}$ which is the opposite of the experimental effect.

## 4.1 The Effects of IVM on P2X4R

As indicated earlier, a Markov model describing P2X4 orthosteric activation by ATP as well as allosteric modulation by IVM had previously been developed. Our goal is to extend this model in order to capture some experimental characteristics which the previously developed simplified model did not account for. Before we proceed with our extension of P2X4 model, it is essential to go over all of the known effects of

IVM on P2X4 in order to elucidate the reasoning behind the choices we have made in the extension of the model.

### 4.1.1 IVM Causes an Increase in Unitary Conductance: Dilation

Some P2X receptors (more specifically P2X7R) exhibit biphasic currents at higher agonist concentrations. An initial fast current is observed followed by a secondary slower current growth (as can be seen in Fig. 1–1) . This secondary growth in current is associated with a slow dilation of the receptor pore which not only increases unitary conductance but also decreases cation selectivity. Such behaviour is deemed dilation because it has been shown that permeability to large cations (such as $NMDG^+$) slowly increases with time in the presence of high agonist concentration whereas at low agonist concentration the pore remains impermeable to such cations due to steric interactions between the receptor pore and cations. This was demonstrated largely through fluorescence studies where uptake of dye, initially contained only in the bath medium, is observed in cells overexpressing P2X receptors [45]. Interestingly, it was found that the rate of permeabilization is dependent on the cation size. This is highly suggestive that dilation might be a step-wise process whereby an initial "partial" dilation might allow for medium sized cations to pass before larger ones are allowed to do so.

In P2X4R, the primary agonist ATP is incapable of producing such behaviour. Nonetheless, it may be induced by application of the pharmacological agent IVM to the bath medium. Our group of collaborators has demonstrated this using the electrophysiological technique known as the ramp protocol (see section §2.2). In such a protocol the holding potential of the cell is rapidly swept across a range of voltages

which gives momentary snapshots of both the conductance of the cell and its reversal potential. As the permeability of the cell towards the large cations increases, the reversal potential of the cell also shifts according to the Nernst equation.

$$E_{rev} = \frac{RT}{F} ln \left( \frac{\sum_i P_{M_i^+} \left[M_i^+\right]_{out} + \sum_i P_{A_i^-} \left[A_i^-\right]_{in}}{\sum_i P_{M_i^+} \left[M_i^+\right]_{in} + \sum_i P_{A_i^-} \left[A_i^-\right]_{out}} \right),$$

where $P_x$ is the cell's permeability to the ion X and $[X]_{out}$ is the extracellular concentration of that ion, whereas $[X]_{in}$ is the intracellular concentration. Because P2X4Rs are cation channels, we do not need to consider the permeability of any of the anions in the medium.

In this experimental setup, cells are bathed in a medium where $Na^+$ has been replaced by equimolar $NMDG^+$. When ATP alone is applied, the cell does not become permeable to $NMDG^+$ but rather the lack of sodium in the medium causes it to escape from the cell through open P2X4Rs. This produces a fast transient, driven by the fact that the initial reversal potential is theoretically negative infinity. However, after the transient, the reversal potential becomes quasi-static as the intracellular and extracellular solutions equilibrate quickly. While a small shift in reversal potential does occur, it is negligible (see Fig. 4–3, left panel). Therefore, we may approximate the reversal potential in the absence of IVM to be that of sodium.

$$E_{ATP} \propto ln \left( \frac{[Na^+]_{out}}{[Na^+]_{in}} \right) \approx E_{Na^+}$$

On the other hand, when IVM is applied permeability to $NMDG^+$ will develop over time. Initially there is negligible permeability to $NMDG^+$, and so we essentially

return back to the same situation as in the absence of IVM.

$$\begin{aligned} E_{IVM,initial} &= \frac{RT}{F} ln\left(\frac{P_{NMDG}\left[NMDG^{+}\right]_{out} + P_{Na}[Na^{+}]_{out}}{P_{Na}\left[Na^{+}\right]_{in}}\right) \\ &\propto ln\left(\frac{[Na^{+}]_{out}}{[Na^{+}]_{in}}\right) \approx E_{Na^{+}}. \end{aligned}$$

As dilation takes place, the permeability of the cell towards $NMDG^{+}$ increases causing a shift in the cell's reversal potential closer to that of $NMDG^{+}$. It should be noted that final reversal potential is not exactly the same as that of $NMDG^{+}$ because the cell does not lose its permeability to $Na^{+}$, i.e.,

$$E_{IVM,final} \propto ln\left(\frac{P_{NMDG}\left[NMDG^{+}\right]_{out} + P_{Na}[Na^{+}]_{out}}{P_{NMDG}\left[NMDG^{+}\right]_{in} + P_{Na}\left[Na^{+}\right]_{in}}\right).$$

Experimentally it was found that the application of extracellular IVM to P2X4R greatly increases the shift in reversal potential when $NMDG^{+}$ is present in the bath medium [41]. This is strongly suggestive of pore dilation induced by the presence of extracellular IVM. The shift in reversal potential was consistently positive (see Fig. 4–3, right panel), indicating that $[NMDG^{+}]_{out}/[NMDG^{+}]_{in}$ is greater than $[Na^{+}]_{out}/[Na^{+}]_{in}$. Moreover, while we can use the x-axis intercept of the I-V curves generated by the voltage ramp protocol to determine reversal potential, we can also use the slope of the I-V curves to determine the cell's conductance. Under the same ramp protocol performed at supramaximal agonist concentrations, it was found that the conductance is a monotonically decreasing function of time [41].

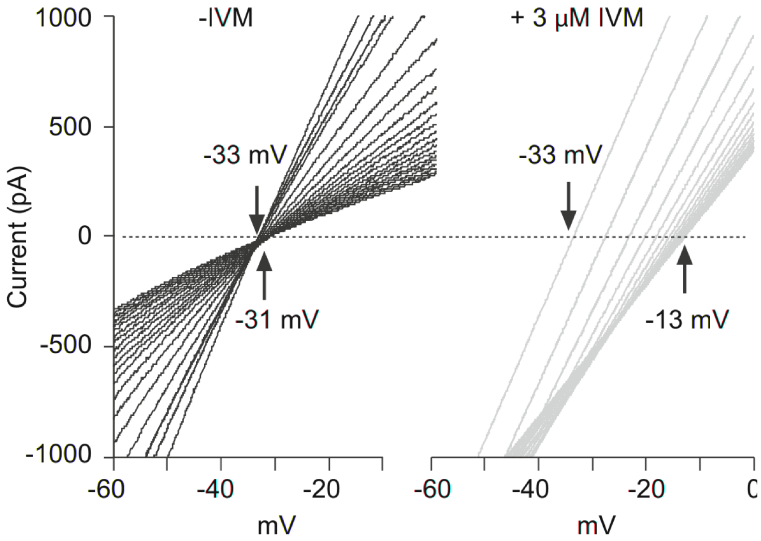


Figure 4–3: The application of 3 µM IVM causes a positive shift in the reversal potential in a cell stimulated with 100 µM ATP for 10 s. The bath medium has $Na^+$ replaced by $NMDG^+$. In the absence of IVM (left) a decrease in conductance with no change in reversal potential is observed (left). In the presence of IVM (right), the decrease in conductance is accompanied by a positive shift in reversal potential (right). The voltage was ramped from −80 mV to +80 mV twice per second from a holding potential of −60 mV, and cells were pretreated with IVM for 30-60 s. Traces shown are representative of 30 similar experiments.

#### 4.1.1.1 Modeling Decrease in Cell Conductance.

When IVM is applied, receptors gradually dilate which causes an increase in unitary conductance from an initial conductance of the open state to that of the dilated state. As stated above, the conductance of a cell decreases even when IVM is applied, due to desensitization. It is thus reasonable to assume that the probability of finding open receptors on the cell surface, $P(open)$, to be a strictly decreasing function of time ($\dot{P}\,(open) < 0$). On the other hand, we expect the slow process of dilation to make the probability of finding a dilated receptor, $P(dilated)$, an increasing function

of time ($\dot{P}\,(dilated) > 0$). Based on this description, the current equation for the cell is given by

$$I = g_{open} P\,(open)\,(V - E_{open}) + g_{dilated} P\,(dilated)\,(V - E_{dilated})\,,$$

where $g_{open}$ is the maximum conductance of the open receptor, $g_{dilated} > g_{open}$ is the maximum conductance of dilated receptors, and $E_{open}$ and $E_{dilated}$ are the reversal potentials of the cell in the fully open and dilated states respectively. We can rewrite the current equation (see section §2.2) as

$$I = g_{tot}\,(V - E_{tot})\,,$$

where

$$\begin{aligned} g_{tot} &= g_{open} P\,(open) + g_{dilated} P\,(dilated) \\ E_{tot} &= \frac{g_{open} P\,(open)\,E_{open} + g_{dilated} P\,(dilated)\,E_{dilated}}{g_{tot}}. \end{aligned}$$

Given that the slope of the I-V curves decrease during a ramp protocol, we expect the conductance to decrease over time, i.e., $\dot{g}_{tot} < 0$. After taking the time derivative of $g_{tot}$ and rearranging the terms, we obtain

$$\dot{g}_{tot} = g_{open}\left(\dot{P}\,(open) + \frac{g_{dilated}}{g_{open}}\dot{P}\,(dilated)\right),$$

which is strictly negative if we impose the condition

$$-\dot{P}\,(open) > \frac{g_{dilated}}{g_{open}}\dot{P}\,(dilated)\,. \tag{4.1}$$

It follows that

$$\left|\dot{P}\left(open\right)\right| > \frac{g_{dilated}}{g_{open}}\dot{P}\left(dilated\right).$$

That is, the total conductance of the cell will decrease if the open fraction of receptor population decreases at a rate which is faster than the ratio of dilated to open maximum conductances times the rate of increase of the dilated states. Thus, in order to capture the decrease in the slope of I-V curves in any model development, we can either increase desensitization of the open states or somehow reduce the rate of increase of the dilated states.

As a first approximation, we can attribute the decrease in the fraction of open states as coming from two processes, desensitization and dilation.

$$-\dot{P}\left(open\right) = \dot{P}\left(dilated\right) + \delta \tag{4.2}$$

where δ is the rate of change of open receptors associated with desensitization. We can substitute this back into equation (4.1) to obtain a new expression for the decrease in conduction

$$\delta > \left(\frac{g_{dilated}}{g_{open}} - 1\right)\dot{P}\left(dilated\right).$$

If we consider $g_{dilated}$ as $g_{open}$ added to some fractional increase in unitary conductance labeled f, as follows

$$g_{dilated} = g_{open}\left(1 + f\right),$$

then

$$\delta > f\dot{P}\left(dilated\right). \tag{4.3}$$

The desensitization rate of naïve receptors is well characterized by our data for experiments with prolonged application of ATP (see Fig. 1–1). It can be considered as a fixed parameter. The increase in unitary conductance has not been determined for rat P2X4 and neither has the rate of dilation. These are to be determined but will need to obey the above constraint in order to guarantee that cell's conductance to decrease in the ramp protocol experiment.

### 4.1.2 IVM Causes Slowing of Deactivation Kinetics

When a cell is stimulated by ATP, its receptors will bind with the agonist and eventually open leading to an increase in current. Once ATP is removed from the medium the reverse process begins to dominate and the current decays over time. This decay is referred to as receptor deactivation and is generally measured by the time constant, $\tau_{off}$, of a monoexponential function ($y=A\ \exp(-t/\tau_{off})$). In the experimental protocol, which we refer to as the "pulse protocol", cells are repeatedly stimulated by 1 μM ATP for 2 s twice per minute. In the absence of IVM, receptors deactivate quickly with a rate which remains constant throughout experiments (see Fig. 4–4A). Once IVM is applied to the medium, we begin to observe a decrease in receptor deactivation rates. The deactivation rates seem to gradually shift from that which is initially observed with ATP alone towards a slower value (see Fig. 4–4B). While our IVM experiments were generally not performed on long enough timescales to observe whether or not the deactivation rate asymptotically approach a final value, it has been shown that they do for human P2X4Rs [60].

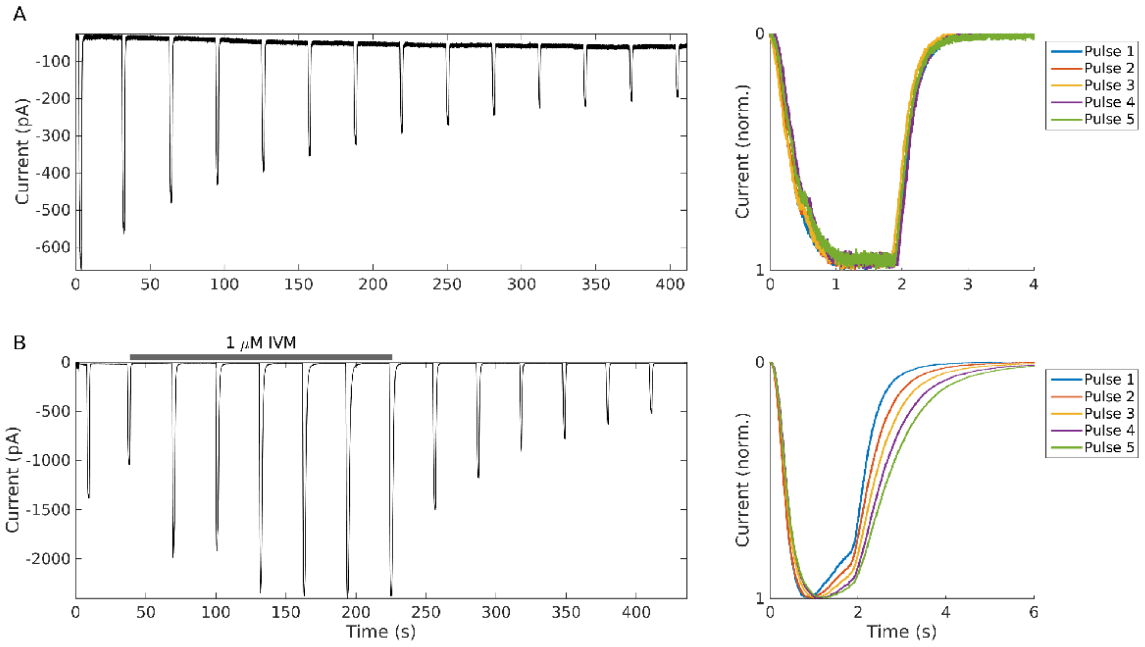


Figure 4–4: Effect of IVM on deactivation kinetics in rat P2X4R. 1 μM ATP is applied for 2 s twice per minute to HEK 293 cells transfected with rat P2X4R. On the left are current traces for the entire experimental protocol. On the right are normalized current traces, highlighting current kinetics. (A) Deactivation kinetics remain constant after repeated agonist applications. (B) Same protocol as A done in the presence of 1 μM IVM (application time is indicated by the grey bar). Deactivation kinetics progressively slow down in the presence of IVM. All pulse numbers are counted from the third application of ATP.

We can, to some extent, understand this gradual shift in deactivation kinetics by considering the population of receptors on the cell surface. Obviously IVM induces conformational changes in the receptors which results in a slower deactivation kinetic. However, it is known that IVM induces this effect rather slowly, with a time constant of 4.6 minutes in human P2X4R [60, 41]. Hence, we can interpret the slow shift in deactivation kinetics as the slow shift in receptor population between the

various conformational states. Furthermore, one can imagine that there are actually many different intermediate states between the naive receptors and those that have all of their IVM interaction sites modified by IVM. Such states might also have intermediate deactivation kinetics due to their incomplete degree of modification by IVM, leading to a nearly continuous spectrum of deactivation kinetics to choose from. At the end of a given pulse, the cell exhibits a dynamic mixture of all of these states which leads to the many deactivation time constants we observe. Ultimately, at a given IVM concentration, there will be some equilibrium mixture between all of these states that produce the asymptotically observed deactivation kinetics.

### 4.1.3 IVM Sensitizes Receptors to ATP

A common tool in the pharmacological characterization of drugs is the dose-response curve, whereby varying doses of a drug are applied to determine the response of the system. Typically dose response curves have the appearance of a saturating sigmoidal curve. For this reason, dose response curves are often fit to Hill functions of the form

$$y\left(x\right) = y_{max}\frac{x^n}{x^n + EC_{50}^n},$$

where x is the applied dosage, $y_{max}$ is the maximum predicted effect, n is the Hill coefficient, and $EC_{50}$ is half maximum of activation. For an experimentally obtained set of doses and effects, fittings are typically pursued to estimate the parameters $y_{max}$, n, and $EC_{50}$. The Hill coefficient is a measure of the steepness of the response, whereas the half maximum of activation is the dosage at which the effect reaches half of its maximum effect.

Dose response curves were constructed for rat P2X4R stimulated by ATP in the presence and absence of 3 μM IVM (see Fig. 4–5). ATP alone was found to produce a dose response curve with an $EC_{50}$ of 2.3±0.4 μM. On the other hand, in the presence of IVM, P2X4R produced responses with an $EC_{50}$ of 0.5±0.1 μM [41]. Moreover, similar studies were done with human P2X4R and the results were consistent with rat P2X4R [60]. While it is clear from Fig. 4–4B that IVM produces an increase in current amplitude, without constructing dose response curves, it is impossible to tell if this is because the receptors are more sensitive to ATP or because IVM simply increases the maximal observed effect. Because there is a nearly five-fold shift in the $EC_{50}$ when IVM is applied, we can conclude that IVM does indeed sensitize receptors towards ATP binding. The aforementioned slowing of deactivation kinetics is one possible way to achieve this, provided that we interpret the deactivation rate as being reflective of the ATP unbinding rate. Since an increase in ATP binding rate could also produce the same effect, it is imperative to test both possibilities.

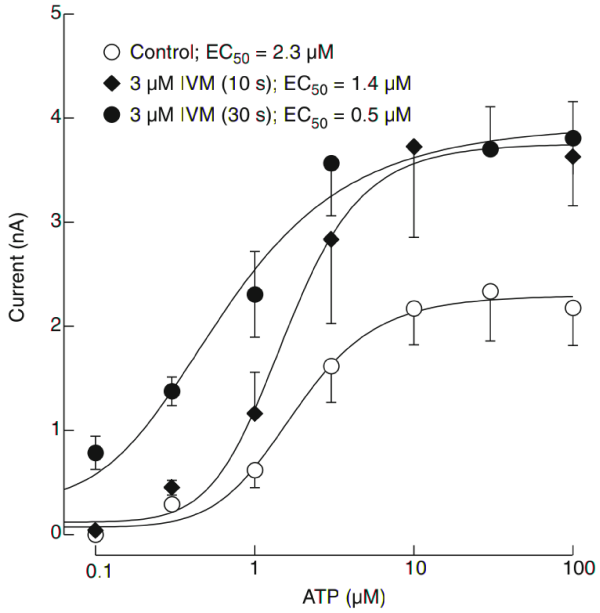


Figure 4–5: ATP concentration-dependent dose-response curve for model P2X4Rs in the absence (open circles) and presence (closed symbols) of 3 μM IVM. The cells were pretreated with IVM for 10 s (closed squares) or for 30 s (closed circles). Data shown are mean±SEM values from n= 4–20 cells per dose. The calculated EC50 values are indicated above traces. Copied with permission from Zemkova et. al. 2012.

In order to produce the sensitized $EC_{50}$ of 0.5 μM it was necessary to pretreat the cells with IVM for at least 30 s before applying ATP (see Fig. 4–5). A pretreatment period of 10 s was also considered and while it did produce the same maximal current response (see Fig. 4–5) an intermediate $EC_{50}$ of 1.6±0.3 μM was measured [41]. This suggests that there are two distinct effects which IVM has on receptors with separate time scales of action. Firstly, IVM increases the maximal current response, which is further discussed in section 4.1.4. Secondly, after prolonged exposures to IVM, receptors eventually become sensitized to ATP.

### 4.1.4 IVM Increases Maximal Open Probabilities

The dose response curves of P2X4R revealed that not only does IVM produce an increase in sensitivity to ATP, but that it also increases the maximum current amplitude evoked by ATP [60, 41]. There are two hypotheses that may explain this type of behaviour: (i) the unitary conductance of individual channels is increasing, which is precisely the dilating effect mentioned in section 4.1.1; (ii) the number of open receptors is rising. There is strong evidence that the first hypothesis does occur. However, this does not, by any means, confirm whether or not the number of open receptors changes with IVM application. In fact, there seems to be a consensus in the literature that the maximal open probability increases with the application of IVM [60, 37, 41]. It is reported that the maximal open probability in the absence of IVM is $\sim$0.2 compared to $\sim$0.8 in the presence of IVM [60, 37]. Previous modeling efforts were able to explain this by developing a model that effectively states that IVM modifies the ratio of the number of open to closed states which is somewhat contrived and cannot explain the pretreatment time required for IVM to take its effect (see section 4.1.3) [37]. Fortunately, we have been able to combine some experimental data with some basic modeling in order to shed some light onto another possible mechanism.

In the 2004 study of human P2X4R, Priel et. al. not only produced dose response curves exhibiting the increased sensitivity and maximal current response induced by IVM (such as in Fig. 4–5) but also managed to quantify its unitary conductance increase. In that study, it was found that IVM produces a roughly 5-fold increase in

maximal current amplitude while only inducing a 20% increase in unitary conductance [60]. Clearly the increase in current is not solely due to the increase in unitary conductance. They posit that, in the absence of IVM, desensitization plays a large role in reducing the current amplitude while when IVM is applied, desensitization is greatly reduced such that the observed current will accordingly be enhanced. While this is a very plausible theory, no mention is made of any receptors which are known to function in this manner. Moreover, no quantitative analysis is made to asses to what extent such a mechanism produces the observed effect.

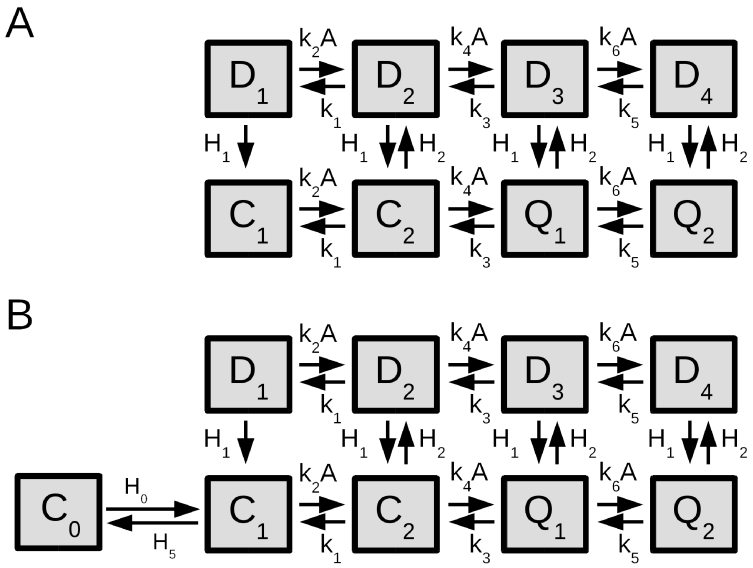


Figure 4–6: Desensitization models of ligand-gated ion channels with three agonist binding sites. (A) Model based on previously developed models of P2X receptors. (B) Model which includes inactivated receptors in the state $C_0$.

In order to test their hypothesis, we constructed a simple mathematical model of a desensitizing ligand activated receptor (see Fig. 4–6A). Dose response curves were

generated for model cells overexpressing these receptors, each with a progressively increasing rate of desensitization (see Fig. 4–7). It was found that although enhanced desensitization rates are very much capable of reducing the current amplitude at a given ATP concentration, such a mechanism is unable to significantly reduce the maximal current amplitude evoked by ATP. Rather it shifts the $EC_{50}$ of the dose response curves rightwards as well as decreases the hill coefficient such that the saturating phase of the dose response curves is shifted by many orders of magnitude. This is not the effect that is observed experimentally either by our group nor by Priel et. al. The presence of IVM consistently produces a dose response curve which has a different maximal current amplitude than when it is absent. The mechanism theorized by Priel et. al. simply does not suit its intended purpose.

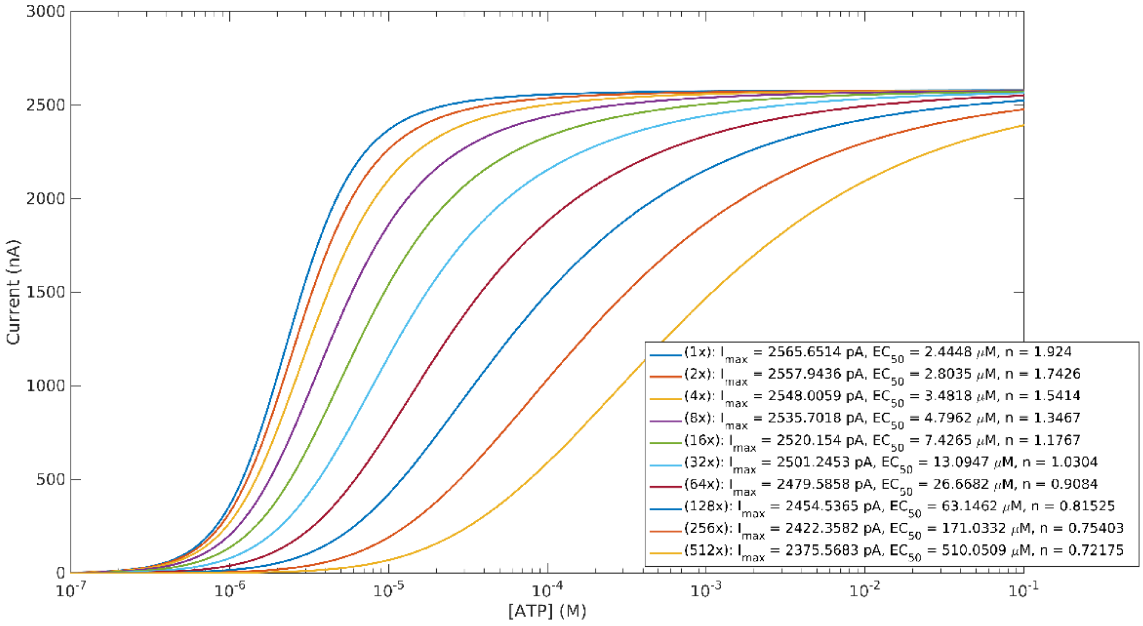


Figure 4–7: Enhanced desensitization does not explain the decrease in maximal current response of a receptor. Hill functions were fit to model simulations of dose response curves. The degree of enhanced desensitization is indicated in the legend along with the fitted parameters of the Hill function. Enhanced desensitization produces a rightward shift in $EC_{50}$ and a decrease in Hill coefficient. Note that dosages vary over 6 orders of magnitude in this figure as opposed to the experimentally relevant 2 orders of magnitude.

After having ruled out desensitization as a cause for the decreased maximal current amplitude in the absence of IVM, we decided to test an alternative hypothesis. Our idea is that the closed states exist in some equilibrium with states for which ATP is not an effective agonist, we deem such receptors as inactivated. This mechanism has already been established in Markov models of sodium channels [61]. Transitions linking the two states must be slow but the equilibrium mixture of the closed-inactivated subsystem makes the maximal open probability in the absence of IVM never attain

1, even at the highest of agonist concentrations. That is, initially when ATP is applied only some fraction of receptors are in closed states that can become activated by ATP. When IVM is applied, the transitions from inactivated to closed states are greatly enhanced such that the fraction of receptors in closed states increases in order to produce the higher current response and open probabilities at supramaximal agonist concentrations (see Fig. 4–8).

While this mechanism is somewhat more explicit about how the saturating response is reduced in the absence of IVM, it should still be tested quantitatively to see how effective it is at producing the observed effect. In order to do so, we extend the mathematical model, which we had used to asses the effect of desensitization on dose response curves, with an inactivated state whose transitions to a closed state are not mediated by interactions with an agonist (see Fig. 4–6B). The transition rate from closed to inactivated was then progressively enhanced in order to see the effect inactivation has on the dose response curves. It was found that this is highly effective at decreasing the maximal current amplitude without shifting the dose response curves or altering their Hill coefficients (see Fig. 4–8).

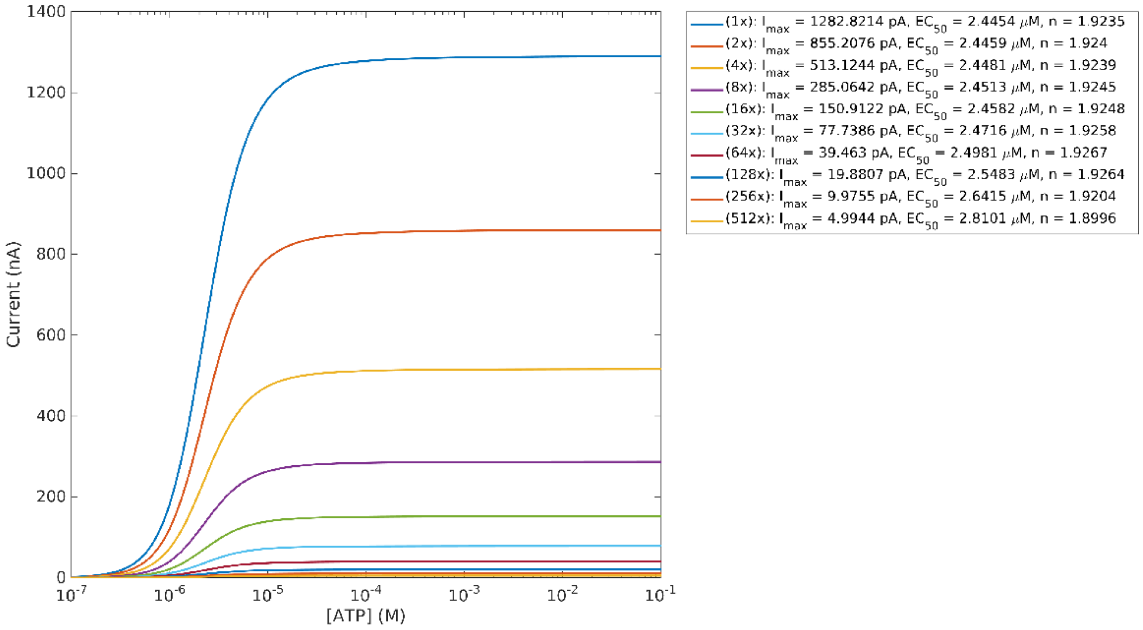


Figure 4–8: Enhanced inactivation does explain decreases in maximal current response of a receptor of an arbitrary degree. Hill functions were fit to model simulations of dose response curves. The degree of enhanced inactivation is indicated in the legend along with the fitted parameters of the Hill function. Enhanced inactivation produces a decrease in $I_{max}$. Note that while dosages span 6 orders of magnitude all responses reach saturation within 2 orders of magnitude as observed experimentally (compare to Figs. 4–7 and 4–5).

Experimentally, the dose response curves exhibit not only a decrease in maximal current but also a leftward shift in $EC_{50}$ as discussed in section 4.1.3, and an increase in Hill coefficients. It is most likely that both mechanisms are at play. However, what we wish to highlight in this section is that desensitization alone cannot explain the decrease in maximal current response. Rather, it would seem that IVM rescues receptors from some pool which is prevented from being activated by ATP. As opposed to the modeling done by Silberberg et. al., this mechanism explains why a

pretreatment period is necessary for IVM to take its full effect and does not postulate that IVM prevents receptor closing altogether.

### 4.1.5 IVM Causes a Transient Increase in Observed Receptor Desensitization Rate

Receptor desensitization is an inherent property of all P2X receptors [9]. It can be understood as a physiologically necessary braking mechanism which prevents receptors from allowing excessive amounts of extracellular calcium into the cell. In contrast to deactivation, desensitization is a decay in current amplitude which occurs during agonist application. Desensitization may not be immediately obvious in short term applications of ATP such as in Fig. 4–4, but when we look at smaller timescales, it becomes clear that the current does begin to decay well before agonist washout (see black trace in Fig. 4–9A). Following application of IVM, we can observe an initial increase in desensitization rate (blue trace in Fig. 4–9A), followed by a gradual decrease in desensitization rate with each subsequent pulse. By the fifth pulse with IVM (green trace in Fig. 4–9A), it seems that the desensitization rate reverts back to a rate which is comparable to that in the absence of IVM.

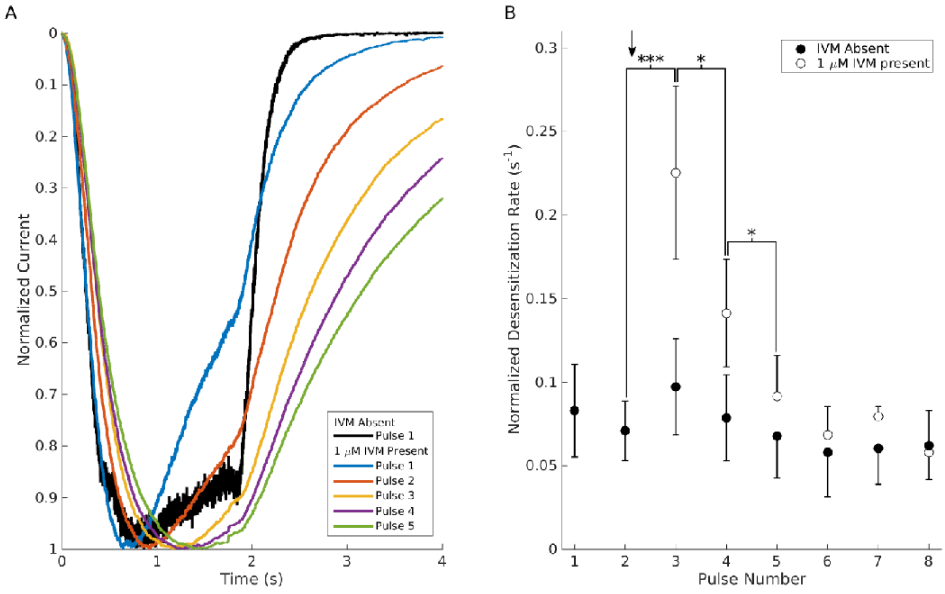


Figure 4–9: IVM induces a transient increase in receptor desensitization rates. (A) Cells are stimulated with 1 μM ATP for 2 s twice per minute. After two such pulses, 1 μM IVM is added to the bath medium. This results in an acute increase in desensitization rate which gradually subsides as desensitization reverts back to its initial rate. Normalized currents for each pulse are plotted in order to highlight the kinetics. (B) In the absence of IVM, there are no statistically significant variations in desensitization rate. Following IVM application (indicated by the arrow), the desensitization rate significantly increases followed by a gradual decline back to the desensitization rate in the absence of IVM. Statistical significance was calculated with the one sided Wilcoxon signed rank test. *$p<0.05$; ***$p<0.005$.

In order to test if this phenomena occurs consistently, we assessed the statistical significance of the increase in desensitization rate. To quantify the amount of desensitization seen in the recordings, we used linear fitting to measure the rate of receptor desensitization. As shown in Fig. 4–9 in the absence of IVM, the desensitization rate does not vary in a statistically significant way. However, when IVM is applied, the

first pulse following application exhibits a significant ($p<0.005$) increase in desensitization rates. Subsequent pulses then gradually drift back to the desensitization rate seen in the absence of IVM. As far as we know, this is novel data analysis and so we are unaware of experimental work which might elucidate the mechanism by which it happens.

### 4.2 One Layer Model

Following the success of the 9-state model of Fig. 2–9, we first set out to expand the detail of our model of IVM's action. This was accomplished in more or less the same way as IVM's action had been previously modeled. First, we know that IVM must be able to act in an ATP independent manner. Therefore, we will allow for IVM to induce transitions from all states in the naïve row as opposed to only when 3 ATP are bound. Second, since there may be multiple sites of interaction with IVM on a single receptor, we will also allow for IVM to transition to more than one row. These rows will have modified kinetics and conductances in order to reflect IVM's various effects. In some sense, this is like having multiple sensitized rows in the 9-state model. In the most general case, the multiple IVM interaction sites are independent of one another and so we should model them as independent transitions from the naïve row and further allow for transitions between each IVM modified row. However, the two effects associated with different IVM interaction sites are known to have distinct pharmacological profiles [60, 41]. This justifies modeling the numerous IVM interactions as a sequential binding process which in turn simplifies both the schematic drawing of the model as well as reduces the number of rate parameters. According to this model (see Fig. 4–10), from the naïve row, binding of IVM occurs

with maximal rates $L_{2i}$, with i=1,2,3,4 being the position along the naive row from which the transitions originate (see Fig. 4–10). Note that we continue to model all IVM forward (or binding) rates as Hill functions of IVM concentration, although we now employ fourth order hill functions (n=4) due to their increased steepness. The $EC_{50}$ of the first transition to IVM bound states is $\delta$. One binding of IVM brings receptors to the third row in the model (counting from top rows of Fig. 4–10) which has IVM unbinding rates $L_{1i}$, with i =1,2,3,4 as before. This IVM bound row has ATP binding and unbinding rates which are given by $\{k_8, k_{10}, k_{12}\}$ and $\{k_7, k_9, k_{11}\}$ respectively. Further binding of IVM transition receptors to a fourth binding row with a maximal rates $L_{4i}$, and $EC_{50}$ given by $\varepsilon$. The corresponding backward (or unbinding) rates is $L_{3i}$, while the ATP binding and unbinding rates are $\{k_{14}, k_{16}, k_{18}\}$ and $\{k_{13}, k_{15}, k_{17}\}$, respectively. Transitions to this fourth row represent receptor dilation, at least in the conductive states. We will refer to $L_{4i}$ as the dilation rates. Similarly this row transitions to a fifth binding row with maximal forward rate $L_{6i}$, $EC_{50}$ given by $\mu$ , backward rate $L_{5i}$ and ATP binding and unbinding rates $\{k_{20}, k_{22}, k_{24}\}$ and $\{k_{19}, k_{21}, k_{23}\}$, respectively. Conceptually, we wish for the steady state population to continuously shift towards the last IVM modified row as IVM concentration increases. Thus we will choose increasing $EC_{50}$ values for each IVM mediated transition (i.e. $\delta<\varepsilon<\mu$).

As was discussed in section 4.1.4, strong evidence suggests that IVM rescues receptors from an inactivated pool. This is incorporated into the model by inclusion of the inactivated state $C_0$ which exists in a (slow) equilibrium with the closed state $C_1$. The forward and backward rates between $C_0$ and $C_1$ are $H_0$ and $H_5$ respectively (see

Fig. 4–10). In the presence of IVM, transitions to the IVM modified closed state $C_3$ are possible with a maximal rate, $W_2$, which is higher than $H_0$ the transition rate between $C_0$ and $C_1$. Since the simultaneous binding of multiple IVM molecules is unlikely, we only allow for transitions from $C_0$ to $C_1$ and $C_3$. Also, the naïve row desensitizes when receptors are bound to at least one ATP molecule with rate $H_2$. The desensitized row has ATP binding and unbinding kinetics which are identical to the naïve row, but it lacks conductivity. Finally, receptors in the desensitized row can recover to the naïve row with rate $H_1$, whereas receptors with 3 bound ATP molecules are allowed to internalize at a rate $H_3$ and get recycled at a rate $H_4$.

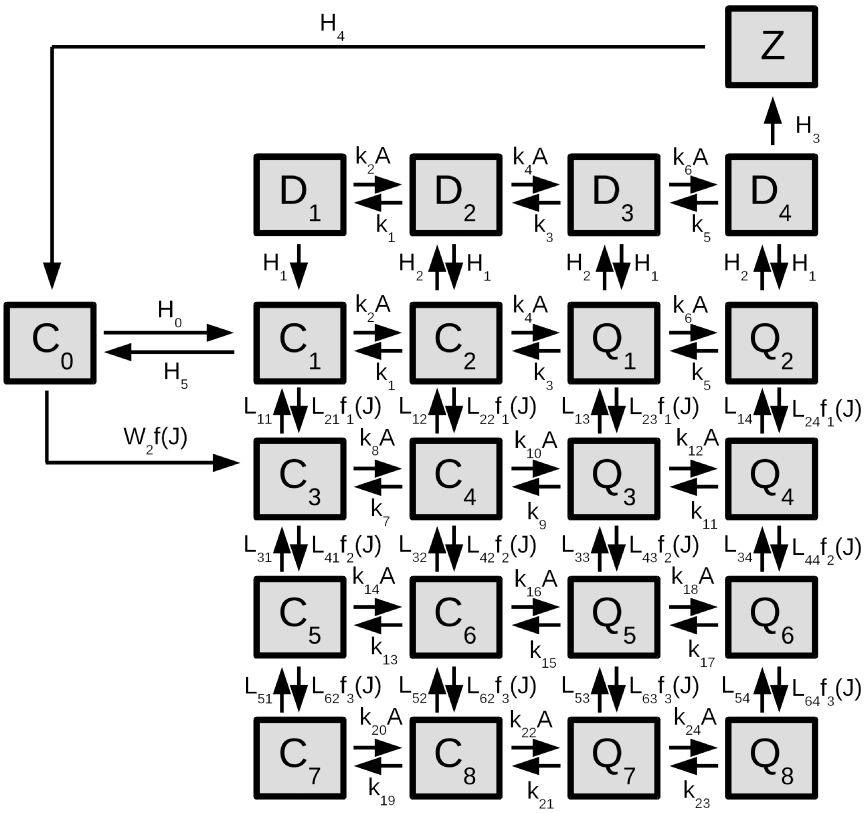


Figure 4–10: One layer model diagram. $C_i$ represents closed states, $D_i$ desensitized states, and $Q_i$ conducting states. $Q_1$, $Q_2$, $Q_3$, and $Q_4$ are open states whereas $Q_5$, $Q_6$, $Q_7$, and $Q_8$ are dilated states with conductance $g_2 > g_1$. $C_0$ is a deeply inactivated state. Model equations are displayed in Appendix B.

In order to assess the model's capacity to reproduce P2X4R gating we used MCMC techniques to determine parameter values which best explain the experimental data given the model of Fig. 4–10. The model was found to be quite capable of capturing many of the properties of P2X4R gating. In the absence of IVM, the model is essentially the same as that presented in Fig. 2–9 and as a result it still captures

naïve receptor activation, desensitization, and deactivation. The increase in current amplitude and deactivation time constant in the presence of IVM is also captured (see Fig. 4–11). This is not surprising, given that these features were captured by the model of Fig. 2–9 which essentially condenses the activation and deactivation kinetics of the three bottom rows into a single sensitized row.

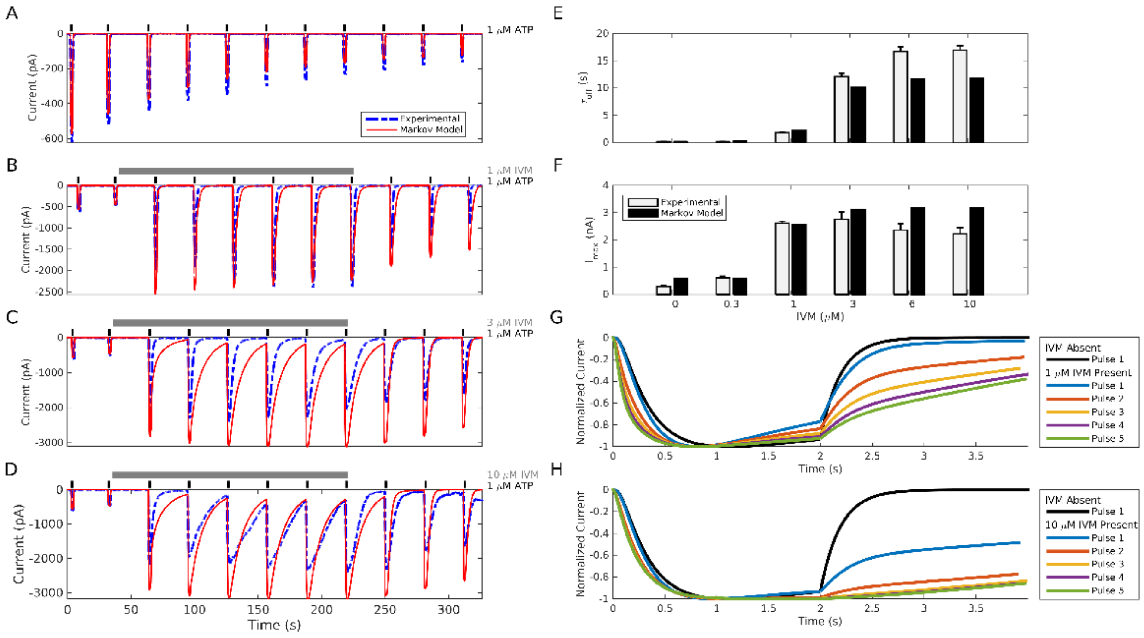


Figure 4–11: The one layer model of Fig 4–10 captures many aspects of the pulse protocol experiments. (A-D) Current time series produced by the model for the pulse protocol experiments performed with (A) 0, (B) 1, (C) 3, and (D) 10 μM IVM. Dashed blue lines are experimental recordings, whereas red lines are model simulations. (E, F) IVM-dependent dose–response curves of (E) peak current amplitude and (F) rate of receptor deactivation. Model deactivation kinetics are measured by a weighted sum of the time constants from a bi-exponential and are expressed as $\tau_{off}$. (G, H) Progression of activation, desensitization, and deactivation produced by the model during the pulse protocol performed with 1 μM and 10 μM IVM, respectively.

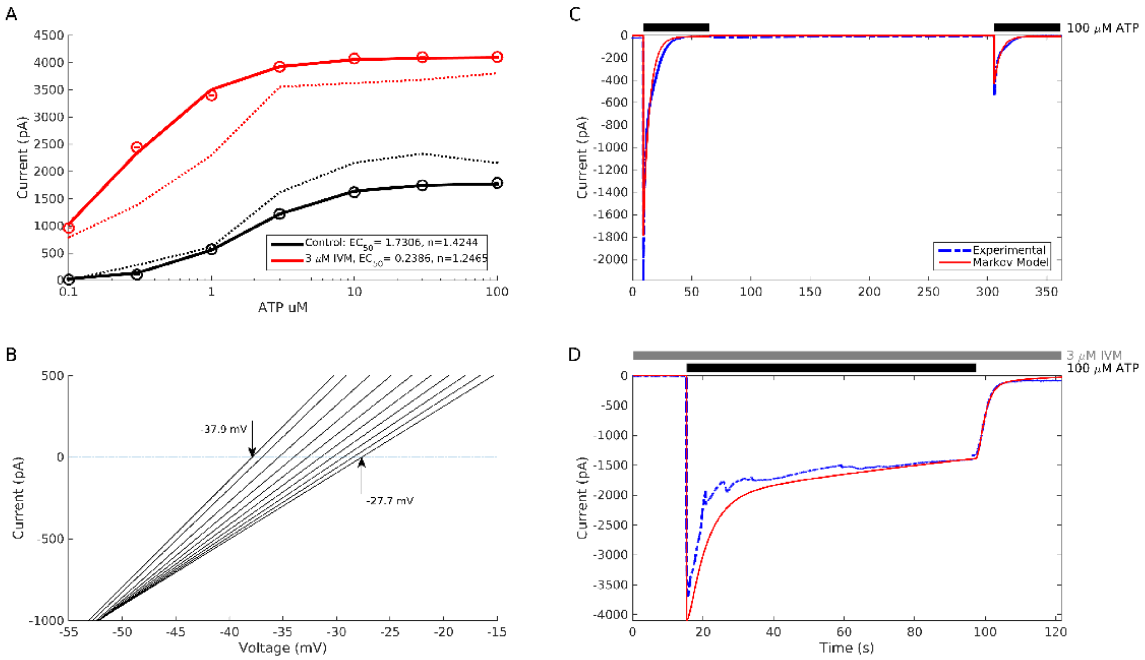


Figure 4–12: The one layer model of Fig 4–10 captures many aspects of the P2X4 gating with prolonged agonist application. (A) ATP concentration-dependent dose-response curve for model P2X4Rs in the absence (black) and presence (red) of 3 μM IVM. Model cells were pretreated with IVM for 30 s (red). Calculated from curve fittings, in the absence of IVM, the $EC_{50}$ was 1.73 μM and the hill coefficient was 1.42, whereas in the presence of IVM they were 0.24 μM and 1.25 respectively. Dashed lines are experimental data. (B) 3 μM IVM causes a positive shift in the reversal potential in a model cell stimulated with 100 μM ATP for 10 s. The bath medium has $Na^+$ replaced by $NMDG^+$ and as such we have set $E_1$ = -46.1 mV and $E_2$ = -21.9 mV. The voltage is ramped from −80 mV to +80 mV twice per second from a holding potential of −60 mV, and model cells were pretreated with IVM for 20 s. Compare with right panel in Fig. 4–3. (C) Two prolonged applications of 100 μM ATP to a model cell in the absence of IVM separated by a 3 min washout period. (D) Prolonged application of 100 μM ATP to a model cell in the presence of 3 μM IVM.

The continuously decreasing slope of the IV curves generated by a voltage ramp protocol are well captured by this model (see Fig. 4–12). We can understand this

by considering the conditions for decreasing cell conductance that were outlined in section 4.1.1.1. The model presented in section §2.3 is certainly capable of fulfilling the conditions. However, such a model requires a large increase (>150%) in unitary conductance in order to fulfill these conditions while simultaneously producing the increase in current amplitude induced by IVM. Experimentally, only a 20% increase in unitary conductance has been observed [60] so this large increase in unitary conductance is hard to justify. According to equation (4.3), the large conductance increase must be balanced by a slow dilation rate. To maintain such a delicate balance (by reducing the dilation rate) will ruin the model's ability to capture other experimental protocols. The model which we have proposed avoids the need for this balance and thus produces the decrease in conductance observed in the I-V curves more robustly. Because we allow for IVM to induce transitions from an inactivated state in our current model, we can easily achieve the current growth with IVM pretreatment using the experimentally validated 20% unitary conductance increase (or even less). This relaxes the need for a slow dilation rates to satisfy equation (4.3). Moreover, since the transitions to the fourth dilated row can occur with IVM alone, pretreatment with IVM will result in some non-negligible fraction of population in dilated states when the voltage ramp protocol begins. Conversely, this relaxes the need for a fast dilation rate to capture other experiments. By reducing both the conductance increase and the amount of dilation which must occur after agonist application, this model achieves the temporal decrease in conductance in a robust manner.

This model has a very large number of parameters. These allow for a very complicated pattern of cooperativity between ATP and IVM binding which manages to reproduce the time series of evoked currents very well. With such a large number of states and free parameters, it is very valid to question whether or not the model can actually be understood or is simply a machine which has been designed to produce the desired output by our fitting algorithm without having any relation to the system we are investigating. The idea behind having 24 different ATP forward and backward rates is that IVM binding might change the cooperativity between the ATP binding sites. The model of section §2.3 enforced negative cooperativity in the naïve row by including restrictions on the parameters ($3k_2/k_1 > k_4/k_3 > k_6/3k_5$), whereas its sensitized row was constructed with negative cooperativity built in ($3k_2/k_1 > k_2/k_1 > k_2/3k_1$). This was tested using statistical tests applied to the Markov chains of parameter values sampled from a posterior distribution reflecting the timeseries data which we have. Once samples were obtained, we compared the binding affinities of each binding site along each row. If there is indeed a specific cooperativity between ATP and IVM binding, we should see correlations between different binding affinities. We did not find any such correlations between ATP binding affinities. This led us to conclude that, while IVM does certainly change ATP binding kinetics, it does not change cooperativity in a manner which is readily discernible from our data. As such, we might be able to fit experiments with a smaller number of parameters using a model where all ATP binding and unbinding rates are proportional to those in the naïve row. This was implemented by thorough use of multiplicative factors as has been summarized in Table

4–1. After some preliminary testing, it became clear that such a simplification of model parameters has very little effect on the quality of the results which could be obtained while significantly reducing the number of parameters required for fitting. From an initial 61 parameters, these modifications brought the count down to 49.

This still leaves us with 24 IVM binding rates. These were initially included because we believed that cooperativity between ATP and IVM binding might help explain the effect of different IVM pretreatment times on the $EC_{50}$ of ATP dose response curves. From MCMC samples of the model parameters, we found that this model requires some degree of cooperativity. In fact, we have allowed for each ATP binding event to modify the IVM binding rate in this model, which made it unclear what the form of the cooperativity is. Along an ATP binding row the IVM binding and unbinding rates might increase with one ATP binding and then decrease with the next. Is such a pattern necessary or are there so many redundant IVM pathways that what seems to be cooperativity is nothing but a slow rate being compensated by a faster one? Initially we tried to apply some statistical tests to try and find any statistically validated patterns of cooperativity from parameter samples. This was not a fruitful endeavour, so we decided to see what occurs when all cooperativity is removed from IVM binding ($L_{21}$= $L_{22}$= $L_{23}$= $L_{24}$, etc..). The parameter changes we used have been summarized in Table 4–2. Overall, the loss of cooperativity seemed to only abolish the model's capacity to capture one effect of IVM application; namely, the increase in desensitization rate following the application of IVM of section 4.1.5 (Fig. 4–9). Recall that this model does not allow for IVM bound states to desensitize directly. They must first unbind all IVM molecules and return to the naïve row for

the receptors to desensitize. This model does not distinguish between desensitization and IVM unbinding in the three lower rows of the model, which seems unlikely. Thus, we tried various ways to rescue the behaviour by explicitly specifying different types of cooperativity between ATP binding and IVM unbinding in our model. The most effective of such models was one where only the first IVM row has progressively increasing IVM unbinding rates with each ATP binding (see Fig. 4–13). We have summarized the changes made to the model in Tables 4–1 and 4–2.

| Original Parameter | Modified Expression | Original Parameter | Modified Expression |
|---|---|---|---|
| $k_7$ | $v_1k_1$ | $k_{16}$ | $r_2k_4$ |
| $k_8$ | $r_1k_2$ | $k_{17}$ | $v_2k_5$ |
| $k_9$ | $v_1k_3$ | $k_{18}$ | $r_2k_6$ |
| $k_{10}$ | $r_1k_4$ | $k_{19}$ | $v_3k_1$ |
| $k_{11}$ | $v_1k_5$ | $k_{20}$ | $r_3k_2$ |
| $k_{12}$ | $r_1k_6$ | $k_{21}$ | $v_3k_3$ |
| $k_{13}$ | $v_2k_1$ | $k_{22}$ | $r_3k_4$ |
| $k_{14}$ | $r_2k_2$ | $k_{23}$ | $v_3k_5$ |
| $k_{15}$ | $v_2k_3$ | $k_{24}$ | $r_3k_6$ |

Table 4–1: ATP binding and unbinding transition rate changes in a reduced parameter one layer model.

| Original Parameter(s) | Modified Expression | Modified Expression (Unbinding Cooperativity) |
|---|---|---|
| $L_{11}$ | $L_1$ | $L_1$ |
| $L_{12}$ | $L_1$ | $\gamma_1 L_1$ |
| $L_{13}$ | $L_1$ | $\gamma_2 L_1$ |
| $L_{14}$ | $L_1$ | $\gamma_3 L_1$ |
| $L_{21}$, $L_{22}$, $L_{23}$, $L_{24}$ | $L_2$ | $L_2$ |
| $L_{31}$, $L_{32}$, $L_{33}$, $L_{34}$ | $L_3$ | $L_3$ |
| $L_{41}$, $L_{42}$, $L_{43}$, $L_{44}$ | $L_4$ | $L_4$ |
| $L_{51}$, $L_{52}$, $L_{53}$, $L_{44}$ | $L_5$ | $L_5$ |
| $L_{61}$, $L_{62}$, $L_{63}$, $L_{64}$ | $L_6$ | $L_6$ |

Table 4–2: IVM binding and unbinding transition rate changes in a reduced parameter one layer model.

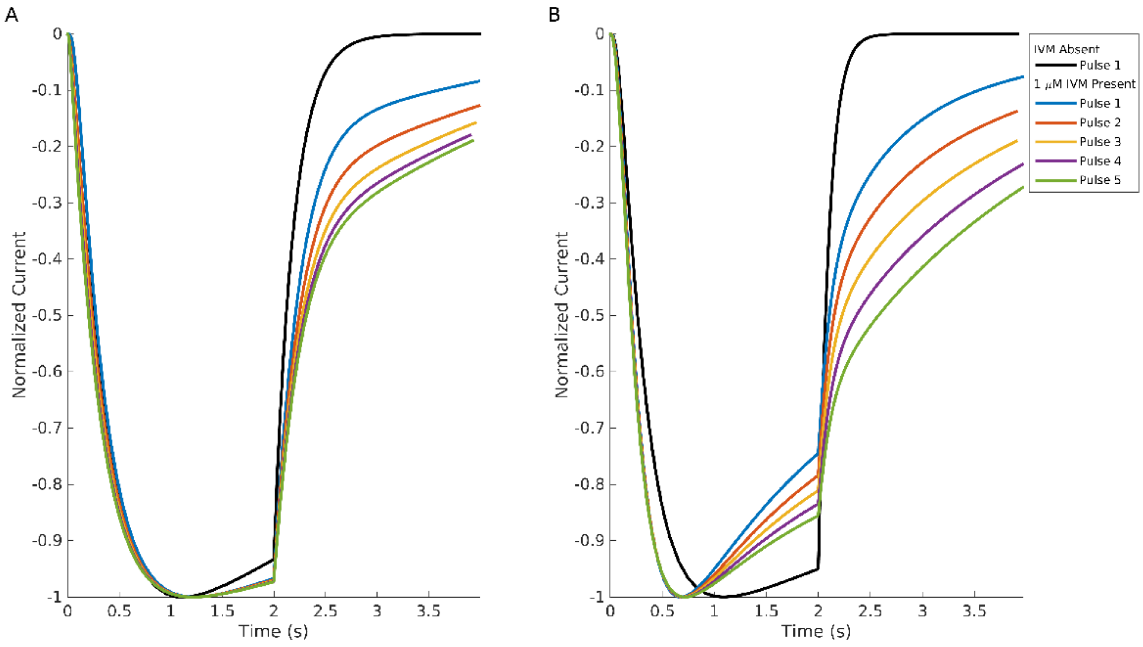


Figure 4–13: (A) Loss of cooperativity between the number of bound ATP molecules and IVM binding and unbinding rates (Table 4–1 and first column of Table 4–2) causes aberrations in the short timescale behaviour of model desensitization when 1 μM IVM is present in the bath medium (compare to Fig. 4–11G). (B) Coooperativity between ATP binding and IVM unbinding (Table 4–1 and second column of Table 4–2) in the first IVM row rescues the transient increase in desensitization. However, this model still remains unphysiological ($\gamma_1 = 3.26,\ \gamma_2 = 5.89,\ \gamma_3 = 33.5$).

It does not seem entirely reasonable physiologically reasonable to have IVM unbinding rates that increase in this manner. IVM concentration is kept constant in the bath medium, so ATP binding, in some sense, would be displacing the IVM molecules, which manifests itself in their increased unbinding rates (3-30 times faster than when no ATP is bound). Particularly, the model we found to have the best agreement with experiments has the IVM unbinding rates which increase with ATP binding only in the row where we assume one IVM to be bound to the receptor. When more IVM

molecules are bound, this effect vanishes altogether. Our experimental collaborators do not support such a claim about kinetics. Moreover, IVM has been found to bind preferentially to the ATP-bound open state of P2X4Rs [60]. Our one layer model is at odds with this finding. We suspect that our fitting algorithm has produced a Markov model which produces the right currents given the appropriate experimental stimuli without representing the underlying system. Because we cannot understand the model after judicious parameter simplifications have been made, we will explore an alternative kind of model. One where receptor sensitization/dilation does not abolish desensitization but rather modifies it.

### 4.3 Two Layer Model

While the one layer model had many of necessary ingredients to capture the gating properties of P2X4 and their allosteric modulation by IVM, there was a rather obvious issue with all one layer models tested. Generally speaking, they did not capture the short timescales of activation and desensitization, and when they did there was a need for a delicate balance between IVM binding rates which was somewhat unnecessary. Why must IVM unbinding rates increase as ATP binds? And why is it particularly crucial for this to happen along the first IVM bound row? We believe that this is because the one layer models does not allow for IVM bound receptors to desensitize. This is at odds with our findings about the transient increase in desensitization rate, which can be compensated for to some extent by having IVM unbinding rates increase with ATP binding. In one layer models, IVM bound states cannot desensitize; they do not have a pathway to directly transition to non-conducting states when ATP is present. This makes it hard to have receptors desensitize before ATP is

removed, particularly at high IVM concentrations where receptor population is concentrated in the bottom rows and must undergo multiple transitions in order to reach non-conducting states. In this section, we will investigate the effect of allowing for reversible transitions from IVM bound states directly to corresponding desensitized states.

Much of this work was motivated by inability of one layer models to faithfully reproduce the short timescale of receptor activation and desensitization. In the one layer models presented in section §4.2 (Fig. 4–10), IVM binding creates copies of an ATP binding system without copying its desensitized states. While *a priori* this may seem reasonable as there is a fair amount of evidence suggesting that IVM reduces desensitization, we feel this is somewhat of a bold claim about receptor dynamics. We believe it may be more reasonable to claim that IVM modifies receptor desensitization rather than preventing it from happening altogether. On top of the IVM-induced modified ATP kinetics which produce the slowed deactivation rates and increased sensitivity to ATP, we posit that IVM bound rows should have modified desensitization kinetics and should be able to readily recover from desensitized states. Consider the one layer model of Fig. 4–10 in the absence of IVM. This sub-model only allows for receptors to return from the desensitized row, with rate $H_1$, once all ATP molecules have unbound from the desensitized states. However, the rate $H_1$ is around 100 times smaller than the desensitization rate $H_2$ and so extending this transition back to the naïve row to all states in the desensitized row, as seen in Fig. 4–14, should have little to no effect in the absence of IVM. Our hypothesis is that this is the underlying binding system (see Fig. 4–14) which is always active, and we

will refer to as a binding configuration. What IVM does is, it induces changes in the rate parameters between different states in this binding configuration in order to produce the varied kinetics we see at different IVM and ATP concentrations as well as activation protocols.

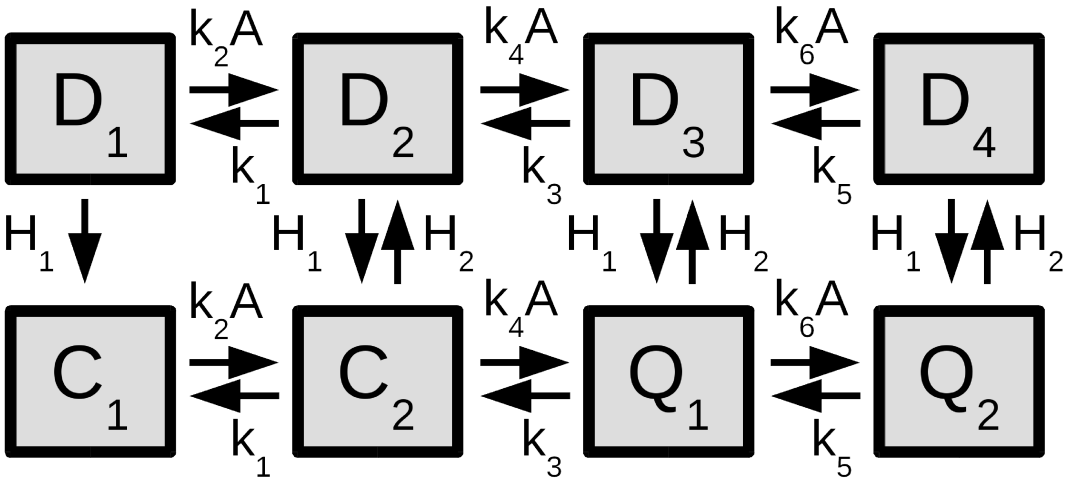


Figure 4–14: Putative underlying binding configuration of P2X4. Ultimately, this model will be used for simulating a few seconds of ATP application. Hence, in this reduced model, we are neglecting both activation/inactivation and internalization due to their long timescales of action relative to the kinetics considered here.

As has been previously discussed, there is strong evidence that there are at least two or three distinct IVM binding sites. At any given time, either of the two sites might be occupied by IVM; this gives us four possible permutations for the occupancy of the IVM binding site on one receptor. If IVM does transition receptors to new ATP binding configuration, a whole cell will have a mixture of receptors with various degrees of IVM modification(s). Therefore, we would expect that the pulses of current

we observe during the pulse protocol experiments to be produced by some mixture of these IVM modified configuration. During the pulse protocol experiments, we observe a gradual shift in the deactivation kinetics of the cell, suggesting that as time progresses the mixture of receptors on the cell's surface shifts towards those being progressively more modified by IVM. If we imagine these configurations as being linked by IVM dependent transition rates we should expect that at the lowest IVM concentrations, the mixture we observe will be mostly composed of naïve receptors and receptors which have only undergone a single modification by IVM. At the highest IVM concentrations the mixture will eventually saturate in those states which are most modified by IVM. We should be able to approximate the mixture of systems which produce a single pulse within the pulse protocol experiments by a single ATP binding configuration with appropriate choices for rate parameters.

In order to test our hypothesis, we chose two recordings of the pulse protocol performed at 1 μM and 10 μM IVM. Picking a single pulse from each recording, we fit model parameters for the putative underlying binding model in order to reproduce the two individual pulses. We found that this approach allowed for very good agreement between model simulations and recordings (see Fig. 4–15). Particularly, the most promising thing about this approach is that with the one layer models at 10 μM IVM, we had never managed to produce the fast activation, desensitization, and insensitivity to ATP removal with one set of parameters, while it seems to be readily reproducible once we assume IVM bound receptors to directly desensitize. Moreover, these fittings corroborate our idea that each pulse is actually a mixture of many different configurations. If one looks closely at Fig. 4–15B, one can see that although

the first pulse fits extremely well, by the last pulse, both the deactivation kinetics and current amplitude are off. Similarly, Fig. 4–15D shows good agreement between the last pulse and simulation while the first pulse of the recording clearly indicates that the mixture of configurations should lie more in favor of a configuration with faster deactivation kinetics. In both cases, we expect that building a model, where configurations of the type depicted in Fig. 4–14 are allowed to mix, will help better capture the complex experimental behaviour.

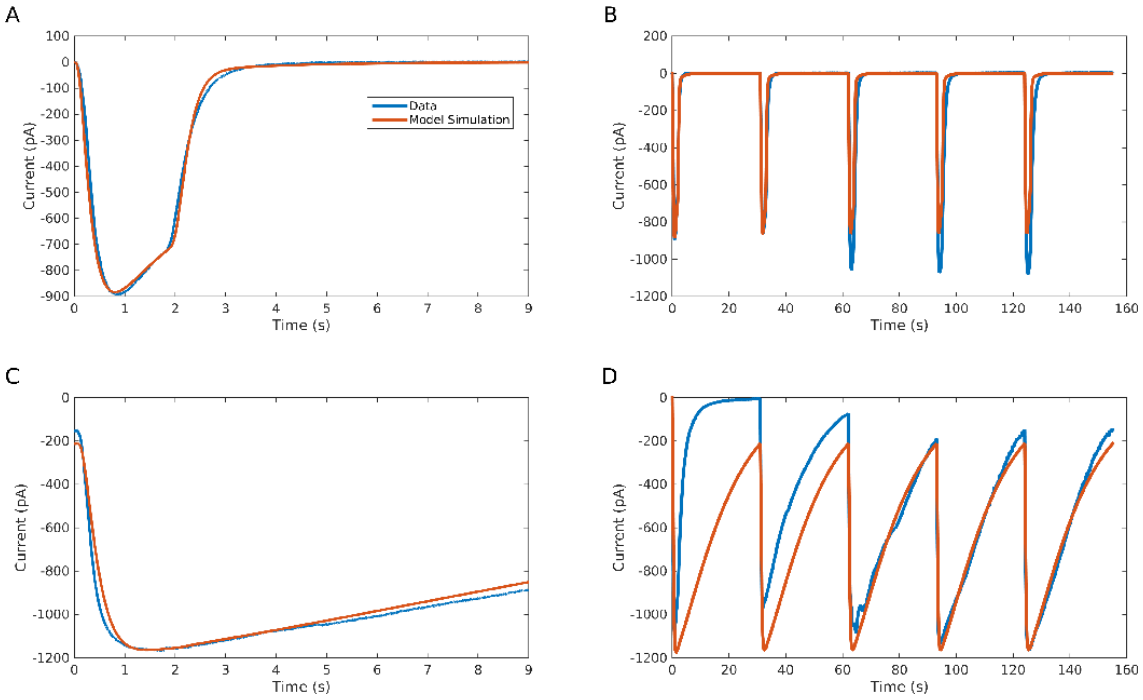


Figure 4–15: Putative binding configuration can easily capture the short timescale behaviour of a given single pulse of the pulse protocol experiments. (A) Single configuration fitting (orange line) to the first pulse of a 1 μM IVM pulse protocol recording (blue line). (B) Timeseries generated by configuration fit to the pulse in A when it is used for the remainder of the pulse protocol. (C) Single configuration fitting (orange line) to the last pulse of a 10 μM IVM pulse protocol recording (blue line). (B) Timeseries generated by configuration fit to the pulse in C when it is used to simulate the pulses leading up to that in C.

Motivated by the fact that the same underlying model could be used to reproduce single pulses of the pulse protocol with great fidelity at all IVM concentrations, we set out to develop a model which could capture not only the short timescale behaviour of receptors but also the progression in kinetics and amplitude which the single configuration fittings had failed to capture. When cooperativity between IVM and ATP had been included in the one layer model, we have seen many instances

where there was a clear progression in kinetics in amplitude. With this in mind, we have investigated models which bear a striking resemblance to those discussed in section §4.2. Because we will allow for all closed and open states to desensitize, we termed this model the two layered model, where one layer consists of closed and conductive states and another layer of desensitized states (see Fig. 4–16). While, this two layer model may seem to be a large departure from that of Fig. 2–9, it should be noted that the model for P2X2 also included a corresponding desensitized state for all of its closed and open states although the desensitization pathway for dilated states was calcium dependent [39].

Much in the same way that the one layer models had sequentially linked three ATP binding rows to a naïve row, we will link four distinct binding configurations together, again through fourth order Hill functions of IVM. In some sense there is a one-to-one correspondence between the binding configurations in the two layer model and the binding rows of the one layer model. As such, we will model ATP and IVM binding and unbinding kinetics in exactly the same manner as before, with each IVM binding event inducing changes in ATP kinetics. These binding processes, however, are accompanied by desensitization unlike the one layer model. The one layer model had a single desensitized row whose ATP kinetics were identical to the naïve row to which it was directly linked to. Continuing with this logic, the desensitized rows will have the same ATP-binding kinetics as the current producing rows to which they are directly linked to. This effectively states that desensitization is a process which does not modify agonist binding but rather affects the receptor pore such that it does not open despite agonist binding. Furthermore, we allowed for internalization from all

desensitized states which have 3 bound ATP molecules. The states with IVM bound were assumed to internalize at a rate, $H_{3D}$, which is different from naïve states, to reflect that IVM alters endocytosis of P2X4 receptors [62]. Altogether, this produced a large block of states where population largely resides in either desensitized states or closed and conducting states (see Fig. 4–16C). IVM's major effect, aside from modifying ATP-binding and activation kinetics, it also shifts receptor population into states where the transition to desensitized states is less favorable. This is achieved through the ratio of forward and backwards desensitization rates of the naive and IVM bound systems: $H_{21}/H_{11} \gg \{H_{22}/\ H_{12}, H_{23}/\ H_{13}, H_{24}/\ H_{14}\}$.

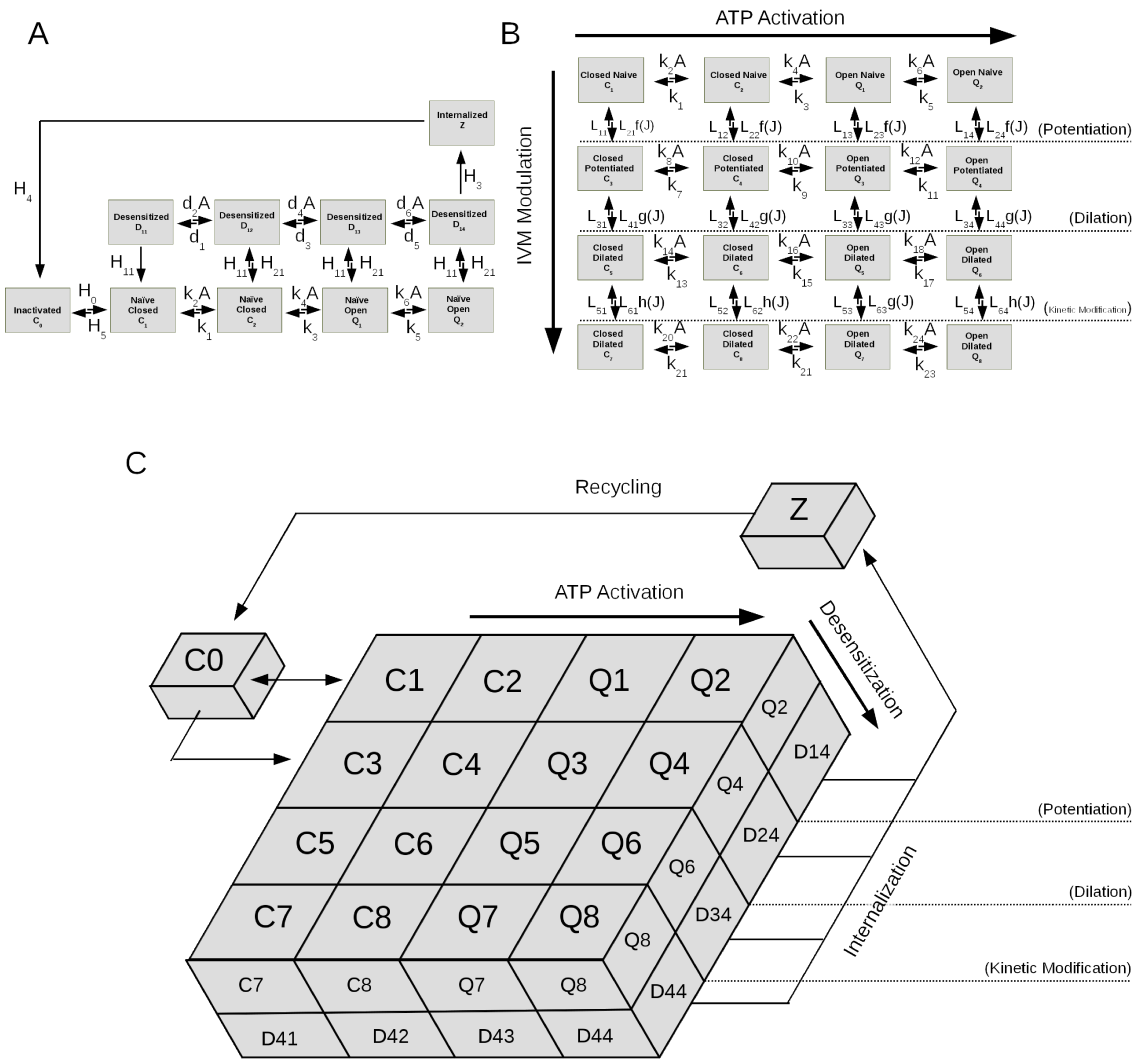


Figure 4–16: Diagram of the two layer model with all parameters independent of one another. (A) Sub-system in the absence of IVM. This illustrates the general motif of states and transitions used throughout each "slice" of the model. (B) IVM and ATP mediated transitions in the uppermost layer of the two layer model (open and closed states). These transitions are identical for each state's corresponding desensitized state. (C) Complete diagram of the model. The system in panel A forms the top slice of the large block of states, whereas that in B is the top face of the block. Model equations are in appendix B.

Given the success of the one layer model with independent binding rates, it is not surprising that this model also captures every aspect of P2X4R gating (see Figs. 4–17 and 4–18). Having decoupled desensitization from IVM unbinding, we expect this model to be able to produce the transient increase in desensitization under the ATP-related parameter simplifications detailed in Table 4–1 and the first column of Table 4–2. In fact, when this parameter simplification is applied to the two layer model, we did not lose the capacity to reproduce any of the features of experimentally observed data. Thus we can conclude that with this model, we understand all of the allosteric effects of IVM. Firstly, agonist potency is increased by IVM rescuing receptor population from a pool of inactivated receptors. The binding configuration with a single IVM bound has an increased desensitization rate which leads to a transient increase in observed desensitization rate. The increased sensitivity towards ATP in this and the other IVM-bound configurations is produced by their decrease in ATP unbinding rates and their increase in ATP binding kinetics. Finally, the return from desensitization rates of all the IVM-bound configurations are significantly larger than the naïve configuration while the internalization rate of these configurations is lower than for naïve receptors, which leads to the long-lasting currents during prolonged ATP applications in the presence of IVM (see Fig. 4–18).

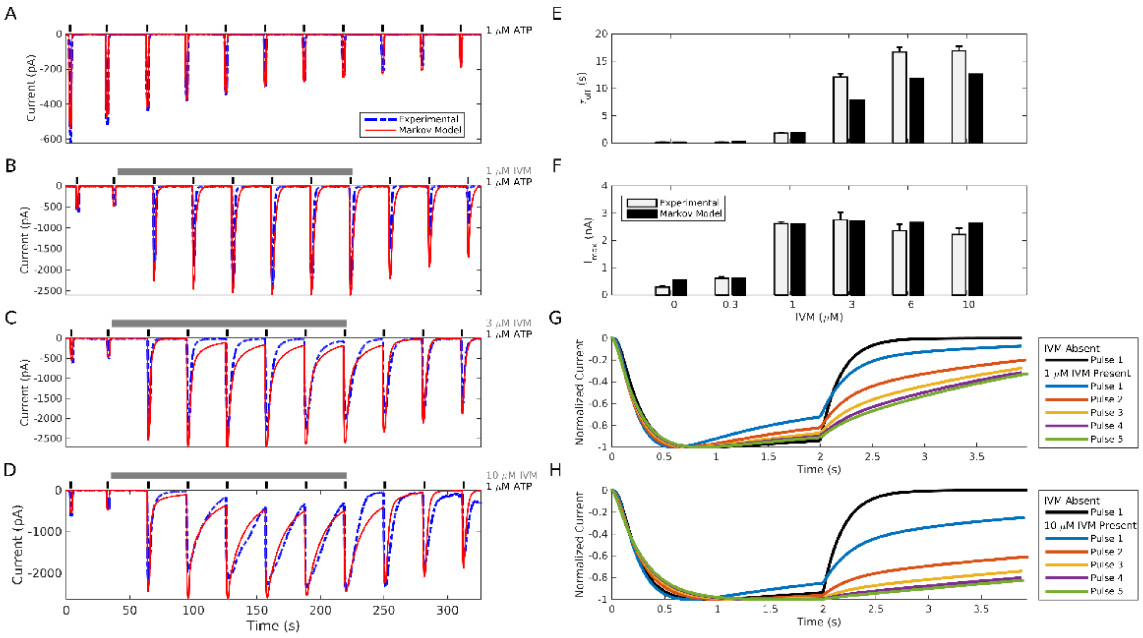


Figure 4–17: The two layer model of Fig 4–16 captures many aspects of the pulse protocol experiments. (A-D) Current time series produced by the model for the pulse protocol experiments performed with (A) 0, (B) 1, (C) 3, and (D) 10 μM IVM. Dashed blue lines are experimental recordings, whereas red lines are model simulations. (E, F) IVM-dependent dose–response curves of (E) peak current amplitude and (F) rate of receptor deactivation. Model deactivation kinetics are measured by a weighted sum of the time constants from a bi-exponential and are expressed as $\tau_{off}$. (G, H) Progression of activation, desensitization, and deactivation produced by the model during the pulse protocol performed with 1 μM and 10 μM IVM, respectively.

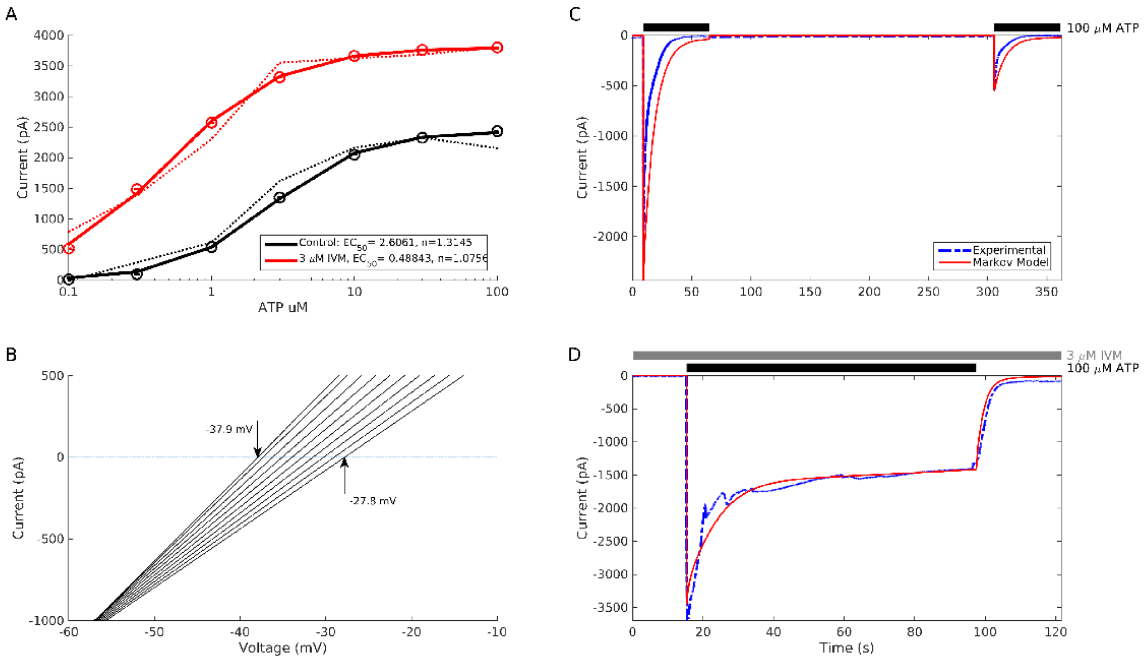


Figure 4–18: The two layer model of Fig 4–10 captures many aspects of the P2X4 gating with prolonged agonist application. (A) ATP concentration-dependent dose-response curve for model P2X4Rs in the absence (black) and presence (red) of 3 μM IVM. Model cells were pretreated with IVM for 30 s (red). Calculated from curve fittings, in the absence of IVM, the $EC_{50}$ was 2.60 μM and the hill coefficient was 1.31, whereas in the presence of IVM they were 0.49 μM and 1.07 respectively. Dashed lines are experimental data. (B) 3 μM IVM causes a positive shift in the reversal potential in a model cell stimulated with 100 μM ATP for 10 s. The bath medium has $Na^+$ replaced by $NMDG^+$ and as such we have set $E_1$ = -55.3 mV and $E_2$ = 12.8 mV. The voltage is ramped from −80 mV to +80 mV twice per second from a holding potential of −60 mV, and model cells were pretreated with IVM for 20 s. Compare with right panel in Fig. 4–3. (C) Two prolonged applications of 100 μM ATP to a model cell in the absence of IVM separated by a 3 min washout period. (D) Prolonged application of 100 μM ATP to a model cell in the presence of 3 μM IVM.

# Chapter 5
# Conclusions & Further Directions

In this thesis, we have analyzed and extended biophysically detailed Markov models of purinergic P2X receptors, with a focus on reproducing complex experimental data for P2X4Rs. Each aspect of P2X4R gating and its modulation by IVM has been analyzed from a modeling perspective in order to piece together the components required for a complete model of this receptor and thus tease out its gating behaviour. Ultimately, we believe that the experimental results are best captured by a two layered model with one layer representing desensitized receptors and another layer representing receptors that can be activated by agonist application. This type of model is preferred over a one layer model because it treats IVM unbinding and desensitization as two independent processes unlike the one layer model which conflates the two. While there may even be strong correlations between the two, we believe that it would be physiologically inaccurate not to treat them as separate. In fact, conflating the two processes in one, as in the one layer model, leads to questionable conclusions about the ligand cooperativity of P2X4Rs.

Previous models of P2XRs had placed significant emphasis on the opposing effects of sensitization/dilation and desensitization, leading to the formation of one layer models. However, P2X2Rs exhibit significant calcium-dependent desensitization even in the so-called sensitized state and thus it was necessary to include desensitization of sensitized receptors in the P2X2R Markov model. Although this model was not

presented as a "two layer model", it contains the essential idea of the two layer model that receptor desensitization is altered in the sensitized/dilated states. In P2X4Rs, sensitization and dilation are IVM dependent. Therefore, it is reasonable to assume that desensitization to be dependent IVM binding, as we have done in the two layer model.

We have presented a framework for generic modeling of allosteric modulation. Within this framework, any allosteric modulation which occurs on a timescale comparable to that of the experimental observation time produces a copy of all the states in the orthosteric system. This leads to the formation of Markov models with a large numbers of states. Although this may be physiologically required, it can be problematic as it leads to a large number of parameters which must be estimated and justified. In our case, it was not possible to collect more experimental data to unequivocally justify and estimate all of these parameters. There are a number of strategies that can be used to reduce the number of states and parameters in our model. We have alluded to the fact that calcium allostery in P2X2 could be modeled without duplication of states due to its fast timescales of action and the fact that its effects are likely fully induced before experiments are performed. Moreover, in the case of the distinct allosteric sites for IVM, we were able to reduce the number of transitions possible by noting that the effects of the two binding sites appear to have different concentration dependencies. Also, rather than assuming all rate parameters to be independent of each other, one can assume some form of correlation between these parameters. Not only does this reduce the number of parameters in a model, it can

also give insight into what relationships between transition rates are necessary to produce specific experimental phenomena.

Although, Markov models and MCMC have previously been used to analyze channel gating [63, 54], their use in analyzing P2XR kinetics is novel. Moreover, these studies were all conducted using single channel data whereas we have used whole-cell recordings for parameter estimation. We have adapted the Bayesian formalisms (including Adaptive Parallel Tempering and the T-Walk Method) for ion channel parameter estimation available in the literature the best we can to this setting. We have also shown some of the shortcomings of standard practices for Bayesian inference of time-series data for ODE models and presented an alternative formalism which manages to circumvent these issues. The methods for Bayesian inference we have presented are a key element in determining how orthosery and allostery affect the gating of these receptors and downstream signaling. Without the use of such quantitative approaches, it will be very difficult to systematically decipher the complex behaviours exhibited by these receptors and the kinetic differences observed between them.

The two layer model makes a number of claims about receptor desensitization kinetics and their IVM concentration-dependence. Unfortunately, the data used to fit the model to was not collected for the purpose of fully exhibiting desensitization kinetics at different concentrations of IVM. Thus, it was necessary to estimate desensitization rates from recordings which exhibit very short ($<$200 ms) desensitization phases. To complement this pulse protocol data, it would be extremely useful to have data on prolonged applications of 1 μM ATP at various IVM concentrations. This would allow for a more in depth analysis of the relationship between IVM concentration

and desensitization rates. Also, there is some data related to the concentration-dependence of activation and desensitization which we have only recently become aware of [42]. Namely, that both activation and desensitization rates are inversely related to agonist concentration. This data was not included in our fitting protocol, and may or may not be compatible with our model. More work will be needed to incorporate these features of P2X4R gating. Generally speaking, as more data becomes available we can use it to further test the model and improve its outcomes.

Our two layered model of P2X4R presents a specific model of allostery by IVM which produces sensitization and dilation. Particularly, it is modeled as being an orthosteric-agonist-independent process as opposed to previous models which required 3 ATP molecules to be bound for the receptor to sensitize/dilate. The key remaining question is: how can this model of P2X4Rs be adapted to yield a universal model for P2XRs? In order to answer such a question, we need firstly to identify the allosteric agonist and to investigate its effects, this can be done by using the P2X4-model where the IVM dependent transitions of our two layer P2X4R model are replaced by either constant rates or by rates that are agonist concentration-dependent. In the case where rates are agonist concentration-independent, we must ensure that allosteric modulation of the naive state is not incorporated into the model, because spontaneous sensitization/dilation does not occur prior to agonist application. Instead, we should assume that these transitions can occur when at least one ATP molecule is bound to the receptor. It remains to be tested if this type of allostery also produces all the characteristics of P2X2Rs and P2X7Rs. However, some preliminary testing of this model for allostery suggests that the agonist-dependent transition

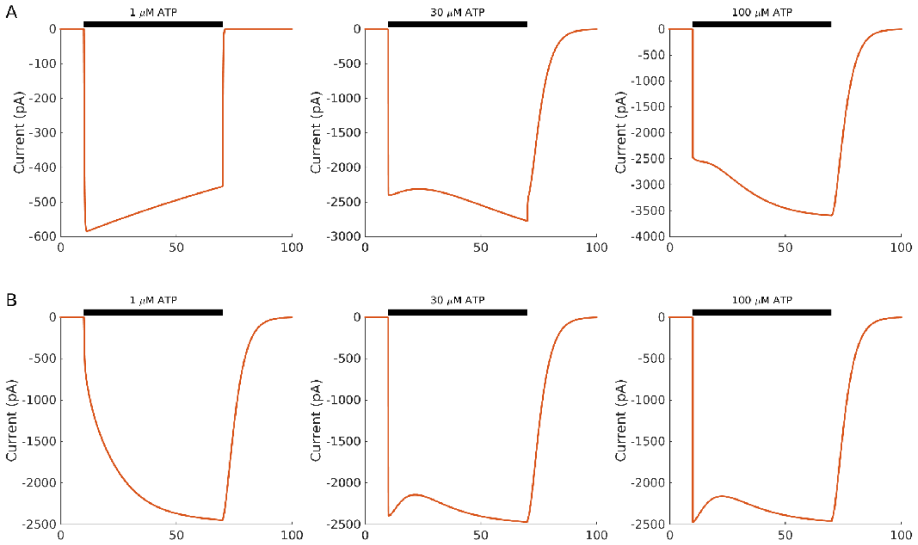


Figure 5–1: Adaption of P2X4R model of IVM-dependent allostery reproduces features of P2X7R concentration dependence of current amplitude and kinetics. (A) Currents generated by a model with allostery induced by the orthosteric agonist ATP. ATP application time is indicated by the black boxes above each recording while the concentrations used are 1 μM, 30 μM, and 100 μM from left to right. (B) Currents generated by a model with allostery induced by an unknown ligand whose concentration remains constant in all experiments. While a concentration-dependent switch in activation kinetics is obvious, this is not associated with an increase in current amplitude or a change in deactivation kinetics, unlike in A. Compare with Fig. 2–3 and recordings in Yan et. al. [1].

rates are better at producing the concentration-dependent switches in kinetics and ion-selectivity observed in P2X7Rs (see our preliminary results in Fig. 5–1). This should be tested more thoroughly in a systematic way using the parameter fitting algorithms we have used for P2X4Rs.

Bibliography

[1] Zonghe Yan, Anmar Khadra, Shuo Li, Melanija Tomic, Arthur Sherman, and Stanko S Stojilkovic. Experimental characterization and mathematical modeling of P2X7 receptor channel gating. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 30(42):14213–14224, October 2010.

[2] Annette Nicke, Hans G. Bäumert, Jürgen Rettinger, Annette Eichele, Günter Lambrecht, Ernst Mutschler, and Günther Schmalzing. P2X1 and P2X3 receptors form stable trimers: A novel structural motif of ligand-gated ion channels. *EMBO Journal*, 17(11):3016–3028, June 1998.

[3] Nelson P. Barrera, Susan J. Ormond, Robert M. Henderson, Ruth D. Murrell-Lagnado, and J. Michael Edwardson. Atomic force microscopy imaging demonstrates that P2X2 receptors are trimers but that P2X6 receptor subunits do not oligomerize. *Journal of Biological Chemistry*, 280(11):10759–10765, March 2005.

[4] Kazuhiro Mio, Yoshihiro Kubo, Toshihiko Ogura, Tomomi Yamamoto, and Chikara Sato. Visualization of the trimeric P2X2 receptor with a crown-capped extracellular domain. *Biochemical and Biophysical Research Communications*, 337(3):998–1005, November 2005.

[5] Toshimitsu Kawate, Jennifer Carlisle Michel, William T Birdsong, and Eric Gouaux. Crystal structure of the ATP-gated P2X(4) ion channel in the closed state. *Nature*, 460(7255):592–598, July 2009.

[6] Motoyuki Hattori and Eric Gouaux. Molecular mechanism of ATP binding and ion channel activation in P2X receptors. *Nature*, 485(7397):207–212, May 2012.

[7] Karina Kaczmarek-Hájek, Éva Lörinczi, Ralf Hausmann, and Annette Nicke. Molecular and functional properties of P2X receptors-recent progress and persisting challenges. *Purinergic Signalling*, 8(3):375–417, September 2012.

[8] Gonzalo E. Torres, Terrance M. Egan, and Mark M. Voigt. Hetero-oligomeric assembly of P2X receptor subunits. Specificities exist with regard to possible partners. *Journal of Biological Chemistry*, 274(10):6653–6659, March 1999.

[9] Claudio Coddou, Zonghe Yan, Tomas Obsil, J Pablo Huidobro-Toro, and Stanko S Stojilkovic. Activation and regulation of purinergic P2X receptor channels. *Pharmacological reviews*, 63(3):641–683, September 2011.

[10] Benjamin Marquez-Klaka, Jürgen Rettinger, Yogesh Bhargava, Thomas Eisele, and Annette Nicke. Identification of an intersubunit cross-link between substituted cysteine residues located in the putative ATP binding site of the P2X1 receptor. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 27(6):1456–1466, February 2007.

[11] Jonathan a. Roberts and Richard J. Evans. ATP binding at human P2X1 receptors: Contribution of aromatic and basic amino acids revealed using mutagenesis and partial agonists. *Journal of Biological Chemistry*, 279(10):9043–9055, March 2004.

[12] Geoffrey Burnstock and Gillian E. Knight. Cellular distribution and functions of P2 receptor subtypes in different systems. *International Review of Cytology*, 240(SPEC.ISS.):31–304, January 2004.

[13] a N Drury and a Szent-Györgyi. The physiological activity of adenine compounds with especial reference to their action upon the mammalian heart. *The Journal of physiology*, 68(3):213–237, November 1929.

[14] G. Burnstock. Purinergic nerves. *Pharmacological reviews*, 24(3):509–81, September 1972.

[15] G. Burnstock and a. Verkhratsky. Evolutionary origins of the purinergic signalling system. *Acta Physiologica*, 195(4):415–447, May 2009.

[16] Joan a Sim, Séverine Chaumont, Jihoon Jo, Lauriane Ulmann, Mark T Young, Kwangwook Cho, Gary Buell, R Alan North, and Francois Rassendren. Altered hippocampal synaptic potentiation in P2X4 knock-out mice. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 26(35):9006–9009, August 2006.

[17] Chaoying Li, Keming Xiong, and Forrest F. Weight. Ethanol inhibition of adenosine 5'-triphosphate-activated current in freshly isolated adult rat hippocampal CA1 neurons. *Neuroscience Letters*, 295(3):77–80, December 2000.

[18] Yi-Jen Huang, Yutaka Maruyama, Gennady Dvoryanchikov, Elizabeth Pereira, Nirupa Chaudhari, and Stephen D Roper. The role of pannexin 1 hemichannels in ATP release and cell-cell communication in mouse taste buds. *Proceedings of the National Academy of Sciences of the United States of America*, 104(15):6436–6441, April 2007.

[19] Roman a Romanov, Olga a Rogachevskaja, Marina F Bystrova, Peihua Jiang, Robert F Margolskee, and Stanislav S Kolesnikov. Afferent neurotransmission mediated by hemichannels in mammalian taste cells. *The EMBO journal*, 26(3):657–667, February 2007.

[20] Thomas E Finger, Vicktoria Danilova, Jennell Barrows, Dianna L Bartel, Alison J Vigers, Leslie Stone, Goran Hellekant, and Sue C Kinnamon. ATP signaling is crucial for communication from taste buds to gustatory nerves. *Science (New York, N.Y.)*, 310(5753):1495–1499, December 2005.

[21] Ryotaro Hayato, Yoshitaka Ohtubo, and Kiyonori Yoshii. Functional expression of ionotropic purinergic receptors on mouse taste bud cells. *The Journal of physiology*, 584(Pt 2):473–488, October 2007.

[22] Lucas T Woods, Jean M Camden, Josef M Batek, Michael J Petris, Laurie Erb, and Gary A Weisman. P2X7 receptor activation induces inflammatory responses in salivary gland epithelium. *American journal of physiology. Cell physiology*, 303(7):C790–801, October 2012.

[23] David J B Muñoz, Peter R. Thorne, and Gary D. Housley. P2X receptor-mediated changes in cochlear potentials arising from exogenous adenosine 5'-triphosphate in endolymph. *Hearing Research*, 138(1-2):56–64, December 1999.

[24] Yongbei Yu and William C de Groat. Sensitization of pelvic afferent nerves in the in vitro rat urinary bladder-pelvic nerve preparation by purinergic agonists and cyclophosphamide pretreatment. *American journal of physiology. Renal physiology*, 294(5):F1146–F1156, May 2008.

[25] M Vlaskovska, L Kasakov, W Rong, P Bodin, M Bardini, D a Cockayne, a P Ford, and G Burnstock. P2X3 knock-out mice reveal a major sensory role for

urothelially released ATP. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 21(15):5670–5677, August 2001.

[26] Michael F Jarvis, Edward C Burgard, Steve McGaraughty, Prisca Honore, Kevin Lynch, Timothy J Brennan, Alberto Subieta, Tim Van Biesen, Jayne Cartmell, Bruce Bianchi, Wende Niforatos, Karen Kage, Haixia Yu, Joe Mikusa, Carol T Wismer, Chang Z Zhu, Katharine Chu, Chih-Hung Lee, Andrew O Stewart, James Polakowski, Bryan F Cox, Elizabeth Kowaluk, Michael Williams, James Sullivan, and Connie Faltynek. A-317491, a novel potent and selective non-nucleotide antagonist of P2X3 and P2X2/3 receptors, reduces chronic inflammatory and neuropathic pain in the rat. *Proceedings of the National Academy of Sciences of the United States of America*, 99(26):17179–17184, December 2002.

[27] Tuan Trang, Simon Beggs, and Michael W. Salter. Purinoceptors in microglia and neuropathic pain. *Pflugers Archiv European Journal of Physiology*, 452(5):645–652, August 2006.

[28] Prisca Honore, Diana Donnelly-Roberts, Marian T Namovic, Gin Hsieh, Chang Z Zhu, Joe P Mikusa, Gricelda Hernandez, Chengmin Zhong, Donna M Gauvin, Prasant Chandran, Richard Harris, Arturo Perez Medrano, William Carroll, Kennan Marsh, James P Sullivan, Connie R Faltynek, and Michael F Jarvis. A-740003 [N-(1-{[(cyanoimino)(5-quinolinylamino) methyl]amino}-2,2-dimethylpropyl)-2-(3,4-dimethoxyphenyl)acetamide], a novel and selective P2X7 receptor antagonist, dose-dependently reduces neuropathic pain in the rat. *The Journal of pharmacology and experimental therapeutics*, 319(3):1376–1385, December 2006.

[29] Iain P. Chessell, Jonathan P. Hatcher, Chas Bountra, Anton D. Michel, Jane P. Hughes, Paula Green, Julie Egerton, Melanie Murfin, Jill Richardson, Wendy L. Peck, Caroline B a Grahames, Maria Anna Casula, Yiangos Yiangou, Rolfe Birch, Praveen Anand, and Gary N. Buell. Disruption of the P2X7 purinoceptor gene abolishes chronic inflammatory and neuropathic pain. *Pain*, 114(3):386–396, April 2005.

[30] W Ma, a Korngreen, N Uzlaner, Z Priel, and S D Silberberg. Extracellular sodium regulates airway ciliary motility by inhibiting a P2X receptor. *Nature*, 400(6747):894–897, August 1999.

[31] Weiyuan Ma, Alon Korngreen, Simy Weil, Enbal Ben-Tal Cohen, Avi Priel, Liubov Kuzin, and Shai D Silberberg. Pore properties and pharmacological

features of the P2X receptor channel in airway ciliated cells. *The Journal of physiology*, 571(Pt 3):503–517, March 2006.

[32] Alexander V Gourine, Enrique Llaudet, Nicholas Dale, and K Michael Spyer. Release of ATP in the ventral medulla during hypoxia in rats: role in hypoxic ventilatory response. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 25(5):1211–1218, February 2005.

[33] Davide Ferrari, Cinzia Pizzirani, Elena Adinolfi, Roberto M Lemoli, Antonio Curti, Marco Idzko, Elisabeth Panther, and Francesco Di Virgilio. The P2X7 receptor: a key player in IL-1 processing and release. *Journal of immunology (Baltimore, Md. : 1950)*, 176(7):3877–3883, April 2006.

[34] Francesco Di Virgilio. Liaisons dangereuses: P2X7 and the inflammasome. *Trends in Pharmacological Sciences*, 28(9):465–472, September 2007.

[35] R. Glas, N. S. Sauter, F. T. Schulthess, L. Shu, J. Oberholzer, and K. Maedler. Purinergic P2X7 receptors regulate secretion of interleukin-1 receptor antagonist and beta cell function and survival. *Diabetologia*, 52(8):1579–1588, August 2009.

[36] Baljit S Khakh and R Alan North. P2X receptors as cell-surface ATP sensors in health and disease. *Nature*, 442(7102):527–532, August 2006.

[37] Shai D. Silberberg, Mufeng Li, and Kenton J. Swartz. Ivermectin Interaction with Transmembrane Helices Reveals Widespread Rearrangements during Opening of P2X Receptor Channels. *Neuron*, 54(2):263–274, April 2007.

[38] Z. Yan, a. Khadra, a. Sherman, and S. S. Stojilkovic. Calcium-dependent block of P2X7 receptor channel function is allosteric. *The Journal of General Physiology*, 138(4):437–452, October 2011.

[39] Anmar. Khadra, Z. Yan, C. Coddou, M. Tomic, A. Sherman, and S. S. Stojilkovic. Gating properties of the P2X2a and P2X2b receptor channels: Experiments and mathematical modeling. *The Journal of General Physiology*, 139(5):333–348, May 2012.

[40] Anmar Khadra, Melanija Tomić, Zonghe Yan, Hana Zemkova, Arthur Sherman, and Stanko S. Stojilkovic. Dual gating mechanism and function of P2X7 receptor channels. *Biophysical Journal*, 104(12):2612–2621, June 2013.

[41] Hana Zemkova, Anmar Khadra, Milos B. Rokic, Vendula Tvrdonova, Arthur Sherman, and Stanko S. Stojilkovic. Allosteric regulation of the P2X4 receptor channel pore dilation. *Pflügers Archiv - European Journal of Physiology*, 467(4):713–726, April 2014.

[42] Zonghe Yan, Zhaodong Liang, Tomas Obsil, and Stanko S. Stojilkovic. Participation of the Lys313-Ile333 sequence of the purinergic P2X4 receptor in agonist binding and transduction of signals to the channel gate. *Journal of Biological Chemistry*, 281(43):32649–32659, October 2006.

[43] Hana Zemkova, Mu-Lan He, Taka-aki Koshimizu, and Stanko S Stojilkovic. Identification of ectodomain regions contributing to gating, deactivation, and resensitization of purinergic P2X receptors. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 24(31):6968–6978, August 2004.

[44] A Villarroel, N Burnashev, and B Sakmann. Dimensions of the narrow portion of a recombinant NMDA receptor channel. *Biophysical journal*, 68(3):866–75, March 1995.

[45] C. Virginio, a. Mackenzie, R. a. North, and a. Surprenant. Kinetics of cell lysis, dye uptake and permeability changes in cells expressing the rat P2X7 receptor. *Journal of Physiology*, 519(2):335–346, September 1999.

[46] Kenneth Barbalace. Periodic Table of Elements - Sorted by Ionic Radius, 2015.

[47] T Riedel, I Lozinsky, G Schmalzing, and F Markwardt. Kinetics of P2X7 receptor-operated single channels currents. *Biophysical journal*, 92(7):2377–2391, April 2007.

[48] Laurent MacKay, Arthur Sherman, Stanko S. Stojilkovic, and Anmar Khadra. Deciphering the Kinetic and Gating Properties of Purinergic P2X7 Receptor Channels. *Athens Journal of Natural & Formal Sciences*, 1(1):43–55, 2014.

[49] Y. Bhargava, J. Rettinger, and a. Mourot. Allosteric nature of P2X receptor activation probed by photoaffinity labelling. *British Journal of Pharmacology*, 167(6):1301–1310, November 2012.

[50] L. E. Browne and R. a. North. P2X Receptor Intermediate Activation States Have Altered Nucleotide Selectivity. *Journal of Neuroscience*, 33(37):14801–14808, September 2013.

[51] W. K. Hastings. Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109, April 1970.

[52] Nicole E. Radde and Jonas Offtermatt. Convergence of posteriors for structurally non-identified problems using results from the theory of inverse problems. *Journal of Inverse and Ill-posed Problems*, 22(2):251–276, January 2014.

[53] Patrick Weber and Jan Hasenauer. Parameter Estimation and Identifiability of Biological Networks Using Relative Data. In *Proc. of 18th IFAC World*, volume 18, pages 11648–11653, August 2011.

[54] Ivo Siekmann, James Sneyd, and Edmund J. Crampin. MCMC can detect non-identifiable models. *Biophysical Journal*, 103(11):2275–2286, December 2012.

[55] David J Earl and Michael W Deem. Parallel tempering: theory, applications, and new perspectives. *Physical chemistry chemical physics : PCCP*, 7(23):3910–3916, November 2005.

[56] Cristian Predescu, Mihaela Predescu, and Cristian V. Ciobanu. On the efficiency of exchange in parallel tempering Monte Carlo simulations. *Journal of Physical Chemistry B*, 109(9):4189–4196, March 2005.

[57] Blazej Miasojedow, Eric Moulines, and Matti Vihola. An Adaptive Parallel Tempering Algorithm. *Journal of Computational and Graphical Statistics*, 22(3):649–664, May 2013.

[58] J. Andrés Christen and Colin Fox. A general purpose sampling algorithm for continuous distributions (the t -walk). *Bayesian Analysis*, 5(2):263–281, June 2010.

[59] Jonathan Goodman and Jonathan Weare. Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, 2010.

[60] Avi Priel and Shai D Silberberg. Mechanism of ivermectin facilitation of human P2X4 receptor channels. *The Journal of general physiology*, 123(3):281–293, March 2004.

[61] C a Vandenberg and F Bezanilla. A sodium channel gating model based on single channel, macroscopic ionic, and gating currents in the squid giant axon. *Biophysical journal*, 60(6):1511–1533, December 1991.

[62] Estelle Toulmé, Florentina Soto, Maurice Garret, and Eric Boué-Grabot. Functional properties of internalization-deficient P2X4 receptors reveal a novel mechanism of ligand-gated channel facilitation by ivermectin. *Molecular pharmacology*, 69(2):576–587, February 2006.

[63] Elan Gin, Martin Falcke, Larry E Wagner, David I Yule, and James Sneyd. Markov chain Monte Carlo fitting of single-channel data from inositol trisphosphate receptors. *Journal of theoretical biology*, 257(3):460–74, April 2009.

# Appendix A
# Parameter Tables

| Label | Value | Transitions |
|---|---|---|
| $k_1$ | 3.806e+00 | $\{C_2, D_{12}\} \rightarrow \{C_1, D_{11}\}$ |
| $k_2$ | 2.684e+06 | $\{C_1, D_{11}\} \rightarrow \{C_2, D_{12}\}$ |
| $k_3$ | 5.044e+00 | $\{Q_1, D_{13}\} \rightarrow \{C_2, D_{12}\}$ |
| $k_4$ | 7.055e+06 | $\{C_2, D_{12}\} \rightarrow \{Q_1, D_{13}\}$ |
| $k_5$ | 4.428e+01 | $\{Q_2, D_{14}\} \rightarrow \{Q_1, D_{13}\}$ |
| $k_6$ | 2.765e+05 | $\{Q_1, D_{13}\} \rightarrow \{Q_2, D_{14}\}$ |
| $k_7$ | 2.213e-03 | $C_4 \rightarrow C_3$ |
| $k_8$ | 4.138e+06 | $C_3 \rightarrow C_4$ |
| $k_9$ | 5.153e+00 | $Q_3 \rightarrow C_4$ |
| $k_{10}$ | 1.326e+07 | $C_4 \rightarrow Q_3$ |
| $k_{11}$ | 5.547e+00 | $Q_4 \rightarrow Q_3$ |
| $k_{12}$ | 1.404e+07 | $Q_3 \rightarrow Q_4$ |
| $k_{13}$ | 2.179e-03 | $C_6 \rightarrow C_5$ |
| $k_{14}$ | 3.539e+06 | $C_5 \rightarrow C_6$ |
| $k_{15}$ | 8.228e-01 | $Q_5 \rightarrow C_6$ |
| $k_{16}$ | 1.761e+07 | $C_6 \rightarrow Q_5$ |
| $k_{17}$ | 6.793e-01 | $Q_6 \rightarrow Q_5$ |
| $k_{18}$ | 8.690e+06 | $Q_5 \rightarrow Q_6$ |
| $k_{19}$ | 7.997e-03 | $C_8 \rightarrow C_7$ |
| $k_{20}$ | 1.258e+07 | $C_7 \rightarrow C_8$ |
| $k_{21}$ | 0.000e+00 | $Q_7 \rightarrow C_8$ |
| $k_{22}$ | 4.781e+06 | $C_8 \rightarrow Q_7$ |
| $k_{23}$ | 0.000e+00 | $Q_8 \rightarrow Q_7$ |
| $k_{24}$ | 5.820e+06 | $Q_7 \rightarrow Q_8$ |

Table A–1: ATP binding and unbinding rates in current producing rows. Values used in Fig. 4–11 & 4–12.

| Label | Value | Transitions |
|---|---|---|
| $L_{11}$ | 1.598e-03 | $C_3 \rightarrow C_1$ |
| $L_{12}$ | 1.055e-02 | $C_4 \rightarrow C_2$ |
| $L_{13}$ | 7.496e-01 | $Q_3 \rightarrow Q_1$ |
| $L_{14}$ | 7.139e-01 | $Q_4 \rightarrow Q_2$ |
| $L_{21}$ | 4.617e-01 | $C_1 \rightarrow C_3$ |
| $L_{22}$ | 2.937e-01 | $C_2 \rightarrow C_4$ |
| $L_{23}$ | 2.259e-01 | $Q_1 \rightarrow Q_3$ |
| $L_{24}$ | 2.536e-01 | $Q_2 \rightarrow Q_4$ |
| $L_{31}$ | 4.561e-02 | $C_5 \rightarrow C_3$ |
| $L_{32}$ | 2.376e-03 | $C_6 \rightarrow C_4$ |
| $L_{33}$ | 4.437e-03 | $Q_5 \rightarrow Q_3$ |
| $L_{34}$ | 8.184e-03 | $Q_6 \rightarrow Q_4$ |
| $L_{41}$ | 1.725e-02 | $C_3 \rightarrow C_5$ |
| $L_{42}$ | 4.472e-01 | $C_4 \rightarrow C_6$ |
| $L_{43}$ | 8.250e-03 | $Q_3 \rightarrow Q_5$ |
| $L_{44}$ | 2.033e-01 | $Q_4 \rightarrow Q_6$ |
| $L_{51}$ | 5.122e-03 | $C_7 \rightarrow C_5$ |
| $L_{52}$ | 2.742e-02 | $C_8 \rightarrow C_6$ |
| $L_{53}$ | 2.490e-02 | $Q_7 \rightarrow Q_5$ |
| $L_{54}$ | 9.690e-02 | $Q_8 \rightarrow Q_6$ |
| $L_{61}$ | 4.074e-01 | $C_5 \rightarrow C_7$ |
| $L_{62}$ | 4.688e-01 | $C_6 \rightarrow C_8$ |
| $L_{63}$ | 2.724e-02 | $Q_5 \rightarrow Q_7$ |
| $L_{64}$ | 3.188e-02 | $Q_6 \rightarrow Q_8$ |

Table A–2: IVM binding and unbinding rates in current producing rows. Values used in Fig. 4–11 & 4–12.

| Label | Value | Transitions |
|---|---|---|
| $\delta$ | 1.300e-06 | $\{C_0, C_1, C_2, Q_1, Q_2\} \rightarrow \{C_3, C_3, C_4, Q_3, Q_4\}$ |
| $\epsilon$ | 1.775e-06 | $\{C_3, C_4, Q_3, Q_4\} \rightarrow \{C_5, C_6, Q_5, Q_6\}$ |
| $\mu$ | 3.686e-06 | $\{C_5, C_6, Q_5, Q_6\} \rightarrow \{C_7, C_8, Q_7, Q_8\}$ |

Table A–3: $EC_{50}$ values for fourth order hill functions which govern IVM binding rates. Values used in Fig. 4–11 & 4–12.

| Label | Value | Transitions |
|---|---|---|
| $H_0$ | 9.511e-04 | $C_0 \rightarrow C_1$ |
| $W_2$ | 8.981e-01 | $C_0 \rightarrow C_3$ |
| $H_5$ | 1.196e-03 | $C_1 \rightarrow C_0$ |

Table A–4: Rates of activation and inactivation. Values used in Fig. 4–11 & 4–12.

| Label | Value | Transitions |
|---|---|---|
| $H_3$ | 6.811e-03 | $D_{14} \rightarrow Z$ |
| $H_4$ | 0.000e+00 | $Z \rightarrow C_0$ |

Table A–5: Rates of internalization and recycling. Values used in Fig. 4–11 & 4–12.

| Label | Value | Transitions |
|---|---|---|
| $k_1$ | 6.010e+00 | $\{C_2, D_{12}\} \rightarrow \{C_1, D_{11}\}$ |
| $k_2$ | 6.886e+06 | $\{C_1, D_{11}\} \rightarrow \{C_2, D_{12}\}$ |
| $k_3$ | 5.184e+00 | $\{Q_1, D_{13}\} \rightarrow \{C_2, D_{12}\}$ |
| $k_4$ | 2.946e+06 | $\{C_2, D_{12}\} \rightarrow \{Q_1, D_{13}\}$ |
| $k_5$ | 9.293e+01 | $\{Q_2, D_{14}\} \rightarrow \{Q_1, D_{13}\}$ |
| $k_6$ | 2.596e+04 | $\{Q_1, D_{13}\} \rightarrow \{Q_2, D_{14}\}$ |
| $k_7$ | 1.900e-03 | $\{C_4, D_{22}\} \rightarrow \{C_3, D_{21}\}$ |
| $k_8$ | 7.113e+06 | $\{C_3, D_{21}\} \rightarrow \{C_4, D_{22}\}$ |
| $k_9$ | 4.089e+00 | $\{Q_3, D_{23}\} \rightarrow \{C_4, D_{22}\}$ |
| $k_{10}$ | 1.301e+07 | $\{C_4, D_{22}\} \rightarrow \{Q_3, D_{23}\}$ |
| $k_{11}$ | 9.057e+00 | $\{Q_4, D_{24}\} \rightarrow \{Q_3, D_{23}\}$ |
| $k_{12}$ | 2.864e+06 | $\{Q_3, D_{23}\} \rightarrow \{Q_4, D_{24}\}$ |
| $k_{13}$ | 1.858e-02 | $\{C_6, D_{32}\} \rightarrow \{C_5, D_{31}\}$ |
| $k_{14}$ | 5.608e+06 | $\{C_5, D_{31}\} \rightarrow \{C_6, D_{32}\}$ |
| $k_{15}$ | 1.806e+00 | $\{Q_5, D_{33}\} \rightarrow \{C_6, D_{32}\}$ |
| $k_{16}$ | 6.385e+06 | $\{C_6, D_{32}\} \rightarrow \{Q_5, D_{33}\}$ |
| $k_{17}$ | 5.045e-01 | $\{Q_6, D_{34}\} \rightarrow \{Q_5, D_{33}\}$ |
| $k_{18}$ | 8.787e+06 | $\{Q_5, D_{33}\} \rightarrow \{Q_6, D_{34}\}$ |
| $k_{19}$ | 3.247e-04 | $\{C_8, D_{42}\} \rightarrow \{C_7, D_{41}\}$ |
| $k_{20}$ | 1.398e+06 | $\{C_7, D_{41}\} \rightarrow \{C_8, D_{42}\}$ |
| $k_{21}$ | 8.307e-04 | $\{Q_7, D_{43}\} \rightarrow \{C_8, D_{42}\}$ |
| $k_{22}$ | 3.573e+06 | $\{C_8, D_{42}\} \rightarrow \{Q_7, D_{43}\}$ |

| Label | Value | Transitions |
|---|---|---|
| $k_{23}$ | 2.801e-03 | $\{Q_8, D_{44}\} \rightarrow \{Q_7, D_{43}\}$ |
| $k_{24}$ | 1.608e+06 | $\{Q_7, D_{43}\} \rightarrow \{Q_8, D_{44}\}$ |

Table A–6: ATP binding and unbinding rates in current producing rows. Values used in Fig. 4–17 & 4–18.

| Label | Value | Transitions |
|---|---|---|
| $L_{11}$ | 2.259e-02 | $\{C_3, D_{21}\} \rightarrow \{C_1, D_{11}\}$ |
| $L_{12}$ | 1.156e-03 | $\{C_4, D_{22}\} \rightarrow \{C_2, D_{12}\}$ |
| $L_{13}$ | 9.528e-01 | $\{Q_3, D_{23}\} \rightarrow \{Q_1, D_{13}\}$ |
| $L_{14}$ | 1.425e-01 | $\{Q_4, D_{24}\} \rightarrow \{Q_2, D_{14}\}$ |
| $L_{21}$ | 2.087e-01 | $\{C_1, D_{11}\} \rightarrow \{C_3, D_{21}\}$ |
| $L_{22}$ | 3.133e-01 | $\{C_2, D_{12}\} \rightarrow \{C_4, D_{22}\}$ |
| $L_{23}$ | 2.739e-02 | $\{Q_1, D_{13}\} \rightarrow \{Q_3, D_{23}\}$ |
| $L_{24}$ | 4.483e-01 | $\{Q_2, D_{14}\} \rightarrow \{Q_4, D_{24}\}$ |
| $L_{31}$ | 4.843e-03 | $\{C_5, D_{32}\} \rightarrow \{C_3, D_{21}\}$ |
| $L_{32}$ | 4.876e-03 | $\{C_6, D_{31}\} \rightarrow \{C_4, D_{22}\}$ |
| $L_{33}$ | 4.015e-03 | $\{Q_5, D_{33}\} \rightarrow \{Q_3, D_{23}\}$ |
| $L_{34}$ | 4.597e-03 | $\{Q_6, D_{34}\} \rightarrow \{Q_4, D_{24}\}$ |
| $L_{41}$ | 1.769e-02 | $\{C_3, D_{21}\} \rightarrow \{C_5, D_{32}\}$ |
| $L_{42}$ | 1.646e-01 | $\{C_4, D_{22}\} \rightarrow \{C_6, D_{31}\}$ |
| $L_{43}$ | 3.595e-02 | $\{Q_3, D_{23}\} \rightarrow \{Q_5, D_{33}\}$ |
| $L_{44}$ | 5.639e-02 | $\{Q_4, D_{24}\} \rightarrow \{Q_6, D_{34}\}$ |
| $L_{51}$ | 7.812e-03 | $\{C_7, D_{41}\} \rightarrow \{C_5, D_{31}\}$ |
| $L_{52}$ | 7.034e-03 | $\{C_8, D_{42}\} \rightarrow \{C_6, D_{32}\}$ |
| $L_{53}$ | 1.516e-04 | $\{Q_7, D_{43}\} \rightarrow \{Q_5, D_{33}\}$ |
| $L_{54}$ | 7.741e-02 | $\{Q_8, D_{44}\} \rightarrow \{Q_6, D_{34}\}$ |
| $L_{61}$ | 2.909e-02 | $\{C_5, D_{31}\} \rightarrow \{C_7, D_{41}\}$ |
| $L_{62}$ | 3.590e-01 | $\{C_6, D_{32}\} \rightarrow \{C_8, D_{42}\}$ |
| $L_{63}$ | 4.780e-02 | $\{Q_5, D_{33}\} \rightarrow \{Q_7, D_{43}\}$ |
| $L_{64}$ | 4.246e-02 | $\{Q_6, D_{34}\} \rightarrow \{Q_8, D_{44}\}$ |

Table A–7: IVM binding and unbinding rates in current producing rows. Values used in Fig. 4–17 & 4–18.

| Label | Value | Transitions |
|---|---|---|
| $\delta$ | 7.546e-07 | $\{C_0, C_1, C_2, Q_1, Q_2, D_{11}, D_{12}, D_{13}, D_{14}\} \rightarrow \{C_3, C_3, C_4, Q_3, Q_4, D_{21}, D_{22}, D_{23}, D_{24}\}$ |
| $\epsilon$ | 1.162e-06 | $\{C_3, C_4, Q_3, Q_4, D_{21}, D_{22}, D_{23}, D_{24}\} \rightarrow \{C_5, C_6, Q_5, Q_6, D_{32}, D_{31}, D_{33}, D_{34}\}$ |
| $\mu$ | 5.646e-06 | $\{C_5, C_6, Q_5, Q_6, D_{31}, D_{32}, D_{33}, D_{34}\} \rightarrow \{C_7, C_8, Q_7, Q_8, D_{41}, D_{42}, D_{43}, D_{44}\}$ |

Table A–8: $EC_{50}$ values for fourth order hill functions which govern IVM binding rates. Values used in Fig. 4–17 & 4–18.

| Label | Value | Transitions |
|---|---|---|
| $kd_1$ | 1.071e-01 | $\{C_2, Q_1, Q_2\} \rightarrow \{D_{12}, D_{13}, D_{14}\}$ |
| $ks_1$ | 2.141e-06 | $\{D_{11}, D_{12}, D_{13}, D_{14}\} \rightarrow \{C_1, C_2, Q_1, Q_2\}$ |
| $kd_2$ | 1.959e-01 | $\{C_4, Q_3, Q_4\} \rightarrow \{D_{22}, D_{23}, D_{24}\}$ |
| $ks_2$ | 9.077e-01 | $\{D_{21}, D_{22}, D_{23}, D_{24}\} \rightarrow \{C_3, C_4, Q_3, Q_4\}$ |
| $kd_3$ | 1.780e-01 | $\{C_6, Q_5, Q_6\} \rightarrow \{D_{32}, D_{33}, D_{34}\}$ |
| $ks_3$ | 2.318e-01 | $\{D_{31}, D_{32}, D_{33}, D_{34}\} \rightarrow \{C_5, C_6, Q_5, Q_6\}$ |
| $kd_4$ | 1.597e-01 | $\{C_8, Q_7, Q_8\} \rightarrow \{D_{42}, D_{43}, D_{44}\}$ |
| $ks_4$ | 2.798e-01 | $\{D_{41}, D_{42}, D_{43}, D_{44}\} \rightarrow \{C_7, C_8, Q_7, Q_8\}$ |

Table A–9: Desensitization and return from desensitization rates. Values used in Fig. 4–17 & 4–18.

| Label | Value | Transitions |
|---|---|---|
| $H_0$ | 2.879e-03 | $C_0 \rightarrow C_1$ |
| $W_2$ | 7.985e-02 | $C_0 \rightarrow C_3$ |
| $H_5$ | 1.464e-03 | $C_1 \rightarrow C_0$ |

Table A–10: Rates of activation and inactivation. Values used in Fig. 4–17 & 4–18.

| Label | Value | Transitions |
|---|---|---|
| $H_3$ | 2.450e-01 | $D_{14} \rightarrow Z$ |
| $H3D$ | 8.961e-03 | $\{D_{24}, D_{34}, D_{44}\} \rightarrow \{Z, Z, Z\}$ |
| $H_4$ | 4.970e-04 | $Z \rightarrow C_1$ |

Table A–11: Rates of internalization and recycling. Values used in Fig. 4–17 & 4–18.

## Appendix B
## Model Equations

The differential equations associated with the one layer model are as follows:

$$\frac{dC_0}{dt} = H_5C_1 + H_4Z - (H_0 + W_2f(J))C_0$$
$$\frac{dC_1}{dt} = H_0C_0 + k_1C_2 + L_{11}C_3 + H_1D_{11} - (H_5 + k_2A + L_{21}f(J))C_1$$
$$\frac{dC_2}{dt} = k_2AC_1 + k_3Q_1 + H_1D_{12} + L_{12}C_4 - (k_1 + k_4A + H_2 + L_{22}f(J))C_2$$
$$\frac{dC_3}{dt} = k_7C_4 + W_2f(J)C_0 + L_{21}f(J)C_1 + L_{31}C_5 - (k_8A + L_{11} + L_{41}g(J))C_3$$
$$\frac{dC_4}{dt} = k_8AC_3 + k_9Q_3 + L_{22}f(J)C_2 + L_{32}C_6 - (k_7 + k_{10}A + L_{12} + L_{42}g(J))C_4$$
$$\frac{dC_5}{dt} = k_{13}C_6 + L_{41}g(J)C_3 + L_{51}C_7 - (k_{14}A + L_{31} + L_{61}h(J))C_5$$
$$\frac{dC_6}{dt} = k_{14}AC_5 + k_{15}Q_5 + L_{42}g(J)C_4 + L_{52}C_8 - (k_{13} + k_{16}A + L_{32} + L_{62}h(J))C_6$$
$$\frac{dC_7}{dt} = k_{19}C_8 + L_{61}h(J)C_5 - (k_{20}A + L_{51})C_7$$
$$\frac{dC_8}{dt} = k_{20}AC_7 + k_{21}Q_7 + L_{62}h(J)C_6 - (k_{19} + k_{22}A + L_{52})C_8$$
$$\frac{dQ_1}{dt} = k_4AC_2 + k_5Q_2 + H_1D_{13} + L_{13}Q_3 - (k_3 + k_6A + H_2 + L_{23}f(J))Q_1$$
$$\frac{dQ_2}{dt} = k_6AQ_1 + H_1D_{14} + L_{14}Q_4 - (k_5 + H_2 + L_{24}f(J))Q_2$$
$$\frac{dQ_3}{dt} = k_{10}AC_4 + k_{11}Q_4 + L_{23}f(J)Q_1 + L_{33}Q_5 - (k_9 + k_{12}A + L_{13} + L_{43}g(J))Q_3$$
$$\frac{dQ_4}{dt} = k_{12}AQ_3 + L_{24}f(J)Q_2 + L_{34}Q_6 - (k_{11} + L_{14} + L_{44}g(J))Q_4$$
$$\frac{dQ_5}{dt} = k_{16}AC_6 + k_{17}Q_6 + L_{43}g(J)Q_3 + L_{53}Q_7 - (k_{15} + k_{18}A + L_{33} + L_{63}h(J))Q_5$$

$$\frac{dQ_6}{dt} = k_{18}AQ_5 + L_{44}g(J)Q_4 + L_{54}Q_8 - (k_{17} + L_{34} + L_{64}h(J))Q_6$$
$$\frac{dQ_7}{dt} = k_{22}AC_8 + k_{23}Q_8 + L_{63}h(J)Q_5 - (k_{21} + k_{24}A + L_{53})Q_7$$
$$\frac{dQ_8}{dt} = k_{24}AQ_7 + L_{64}h(J)Q_6 - (k_{23} + L_{54})Q_8$$
$$\frac{dD_{11}}{dt} = k_1D_{12} - (k_2A + H_1)D_{11}$$
$$\frac{dD_{12}}{dt} = k_2AD_{11} + k_3D_{13} + H_2C_2 - (k_1 + k_4A + H_1)D_{12}$$
$$\frac{dD_{13}}{dt} = k_4AD_{12} + k_5D_{14} + H_2Q_1 - (k_3 + k_6A + H_1)D_{13}$$
$$\frac{dD_{14}}{dt} = k_6AD_{13} + H_2Q_2 - (k_5 + H_1 + H_3)D_{14}$$
$$\frac{dZ}{dt} = H_3D_{14} - H_4Z$$

The differential equations associated with the two layer model are as follows:

$$\frac{dC_0}{dt} = H_5C_1 - (H_0 + W_2f(J))C_0$$
$$\frac{dC_1}{dt} = H_0C_0 + k_1C_2 + L_{11}C_3 + k_{s,1}D_{11} + H_4Z - (H_5 + k_2A + L_{21}f(J))C_1$$
$$\frac{dC_2}{dt} = k_2AC_1 + k_3Q_1 + k_{s,1}D_{12} + L_{12}C_4 - (k_1 + k_4A + k_{d,1} + L_{22}f(J))C_2$$
$$\frac{dC_3}{dt} = k_7C_4 + W_2f(J)C_0 + L_{21}f(J)C_1 + L_{31}C_5 + k_{s,2}D_{21} - (k_8A + L_{11} + L_{41}g(J))C_3$$
$$\frac{dC_4}{dt} = k_8AC_3 + k_9Q_3 + k_{s,2}D_{22} + L_{22}f(J)C_2 + L_{32}C_6 - (k_7 + k_{10}A + k_{d,2} + L_{12} + L_{42}g(J))C_4$$
$$\frac{dC_5}{dt} = k_{13}C_6 + L_{41}g(J)C_3 + L_{51}C_7 + k_{s,3}D_{31} - (k_{14}A + L_{31} + L_{61}h(J))C_5$$
$$\frac{dC_6}{dt} = k_{14}AC_5 + k_{15}Q_5 + k_{s,3}D_{32} + L_{42}g(J)C_4 + L_{52}C_8 - (k_{13} + k_{16}A + k_{d,3} + L_{32} + L_{62}h(J))C_6$$
$$\frac{dC_7}{dt} = k_{19}C_8 + L_{61}h(J)C_5 + k_{s,4}D_{41} - (k_{20}A + L_{51})C_7$$
$$\frac{dC_8}{dt} = k_{20}AC_7 + k_{21}Q_7 + k_{s,4}D_{42} + L_{62}h(J)C_6 - (k_{19} + k_{22}A + k_{d,4} + L_{52})C_8$$
$$\frac{dQ_1}{dt} = k_4AC_2 + k_5Q_2 + k_{s,1}D_{13} + L_{13}Q_3 - (k_3 + k_6A + k_{d,1} + L_{23}f(J))Q_1$$

$$\frac{dQ_2}{dt} = k_6 A Q_1 + k_{s,1} D_{14} + L_{14} Q_4 - \left(k_5 + k_{d,1} + L_{24} f\left(J\right)\right) Q_2$$
$$\frac{dQ_3}{dt} = k_{10} A C_4 + k_{11} Q_4 + k_{s,2} D_{23} + L_{23} f\left(J\right) Q_1 + L_{33} Q_5 - \left(k_9 + k_{12} A + k_{d,2} + L_{13} + L_{43} g\left(J\right)\right) Q_3$$
$$\frac{dQ_4}{dt} = k_{12} A Q_3 + k_{s,2} D_{24} + L_{24} f\left(J\right) Q_2 + L_{34} Q_6 - \left(k_{11} + k_{d,2} + L_{14} + L_{44} g\left(J\right)\right) Q_4$$
$$\frac{dQ_5}{dt} = k_{16} A C_6 + k_{17} Q_6 + k_{s,3} D_{33} + L_{43} g\left(J\right) Q_3 + L_{53} Q_7 - \left(k_{15} + k_{18} A + k_{d,3} + L_{33} + L_{63} h\left(J\right)\right) Q_5$$
$$\frac{dQ_6}{dt} = k_{18} A Q_5 + k_{s,3} D_{34} + L_{44} g\left(J\right) Q_4 + L_{54} Q_8 - \left(k_{17} + k_{d,3} + L_{34} + L_{64} h\left(J\right)\right) Q_6$$
$$\frac{dQ_7}{dt} = k_{22} A C_8 + k_{23} Q_8 + k_{s,4} D_{43} + L_{63} h\left(J\right) Q_5 - \left(k_{21} + k_{24} A + k_{d,4} + L_{53}\right) Q_7$$
$$\frac{dQ_8}{dt} = k_{24} A Q_7 + k_{s,4} D_{44} + L_{64} h\left(J\right) Q_6 - \left(k_{23} + k_{d,4} + L_{54}\right) Q_8$$
$$\frac{dD_{11}}{dt} = k_1 D_{12} + L_{11} D_{21} - \left(k_2 A + L_{21} f\left(J\right) + k_{s,1}\right) D_{11}$$
$$\frac{dD_{12}}{dt} = k_2 A D_{11} + k_3 D_{13} + k_{d,1} C_2 + L_{12} D_{22} - \left(k_1 + k_4 A + k_{s,1} + L_{22} f\left(J\right)\right) D_{12}$$
$$\frac{dD_{13}}{dt} = k_4 A D_{12} + k_5 D_{14} + k_{d,1} Q_1 + L_{13} D_{23} - \left(k_3 + k_6 A + k_{s,1} + L_{23} f\left(J\right)\right) D_{13}$$
$$\frac{dD_{14}}{dt} = k_6 A D_{13} + k_{d,1} Q_2 + L_{14} D_{24} - \left(k_5 + k_{s,1} + L_{24} f\left(J\right) + H_3\right) D_{14}$$
$$\frac{dD_{21}}{dt} = k_7 D_{22} + L_{21} f\left(J\right) D_{11} + L_{31} D_{32} - \left(k_8 A + L_{11} + L_{41} g\left(J\right) + k_{s,2}\right) D_{21}$$
$$\frac{dD_{22}}{dt} = k_8 A D_{21} + k_9 D_{23} + k_{d,2} C_4 + L_{22} f\left(J\right) D_{12} + L_{32} D_{31} - \left(k_7 + k_{10} A + k_{s,2} + L_{12} + L_{42} g\left(J\right)\right) D_{22}$$
$$\frac{dD_{23}}{dt} = k_{10} A D_{22} + k_{11} D_{24} + k_{d,2} Q_3 + L_{23} f\left(J\right) D_{13} + L_{33} D_{33} - \left(k_9 + k_{12} A + k_{s,2} + L_{13} + L_{43} g\left(J\right)\right) D_{23}$$
$$\frac{dD_{24}}{dt} = k_{12} A D_{23} + k_{d,2} Q_4 + L_{24} f\left(J\right) D_{14} + L_{34} D_{34} - \left(k_{11} + k_{s,2} + L_{14} + L_{44} g\left(J\right) + H_{3,D}\right) D_{24}$$
$$\frac{dD_{31}}{dt} = k_{13} D_{32} + L_{42} g\left(J\right) D_{22} + L_{51} D_{41} - \left(k_{14} A + L_{32} + L_{61} h\left(J\right) + k_{s,3}\right) D_{31}$$
$$\frac{dD_{32}}{dt} = k_{14} A D_{31} + k_{15} D_{33} + k_{d,3} C_6 + L_{41} g\left(J\right) D_{21} + L_{52} D_{42} - \left(k_{13} + k_{16} A + k_{s,3} + L_{31} + L_{62} h\left(J\right)\right) D_{32}$$
$$\frac{dD_{33}}{dt} = k_{16} A D_{32} + k_{17} D_{34} + k_{d,3} Q_5 + L_{43} g\left(J\right) D_{23} + L_{53} D_{43} - \left(k_{15} + k_{18} A + k_{s,3} + L_{33} + L_{63} h\left(J\right)\right) D_{33}$$
$$\frac{dD_{34}}{dt} = k_{18} A D_{33} + k_{d,3} Q_6 + L_{44} g\left(J\right) D_{24} + L_{54} D_{44} - \left(k_{17} + k_{s,3} + L_{34} + L_{64} h\left(J\right) + H_{3,D}\right) D_{34}$$

$$\frac{dD_{41}}{dt} = k_{19}D_{42} + L_{61}h\left(J\right)D_{31} - \left(k_{20}A + L_{51} + k_{s,4}\right)D_{41}$$
$$\frac{dD_{42}}{dt} = k_{20}AD_{41} + k_{21}D_{43} + k_{d,4}C_8 + L_{62}h\left(J\right)D_{32} - \left(k_{19} + k_{22}A + k_{s,4} + L_{52}\right)D_{42}$$
$$\frac{dD_{43}}{dt} = k_{22}AD_{42} + k_{23}D_{44} + k_{d,4}Q_7 + L_{63}h\left(J\right)D_{33} - \left(k_{21} + k_{24}A + k_{s,4} + L_{53}\right)D_{43}$$
$$\frac{dD_{44}}{dt} = k_{24}AD_{43} + k_{d,4}Q_8 + L_{64}h\left(J\right)D_{34} - \left(k_{23} + k_{s,4} + L_{54} + H_{3,D}\right)D_{44}$$
$$\frac{dZ}{dt} = H_3D_{14} + H_{3,D}D_{24} + H_{3,D}D_{34} + H_{3,D}D_{44} - H_4Z$$