Mathematical modelling of GPCR-mediated calcium signalling

Wodu Majin, BSc.

Thesis submitted to The University of Nottingham
for the degree of Doctor of Philosophy

January 2012

To you mummy. For supporting, sacrificing, believing and cheering, yet telling the truth.

Abstract

Ca^{2+} is an important messenger which mediates several physiological functions, including muscle contraction, fertilisation, heart regulation and gene transcription. One major way its cytosolic level is raised is via a G-protein coupled receptor (GPCR)–mediated release from intracellular stores. GPCR's are the target of approximately 50% of all drugs in clinical use. Hence, understanding the underlying mechanisms of signalling in this pathway could lead to improved therapy in disease conditions associated with abnormmal Ca^{2+} signalling, and to the identification of new drug targets. To gain such insight, this thesis builds and analyses a detailed mathematical model of key processes leading to Ca^{2+} mobilisation.

Ca^{2+} signalling is considered in the particular context of the M3 muscarinic receptor system. Guided by available data, the Ca^{2+} mobilisation model is assembled, first by analysing a base G-protein activation model, and subsequently extending it with downstream details. Computationally efficient designs of a global parameter sensitivity analysis method are used to identify the key controlling parameters with respect to the main features of the Ca^{2+} data. The underlying mechanism behind the experimentally observed, rapid, amplified Ca^{2+} response is shown to be a rapid rate of inositol trisphosphate (IP_3) formation from Phosphatidylinositol 4,5-bisphosphate (PIP_2) hydrolysis. Using the same results, potential drug targets (apart from the GPCR) are identified, including the sarco/endoplasmic reticulum Ca^{2+}-ATPase (SERCA) and PIP_2. Moreover, possible explanations for therapeutic failures were found when some parameters exerted a biphasic effect on the relative Ca^{2+} increase.

The sensitivity analysis results are used to simplify the process of parameter estimation by a significant reduction of the parameter space of interest. An evolutionary algorithm is used to successfully fit the model to a significant portion of the Ca^{2+} data. Subsequent sensitivity analyses of the best-fitting parameter sets suggest that mechanistic modelling of kinase-mediated GPCR desensitisation, and SERCA dynamics may be required for a comprehensive representation of the data.

Acknowledgements

Thanks to you Lord, it wasn't just mathematics I learnt during this PhD.

Thanks to my mathematics supervisors, Markus Owen and John King for their expertise and guidance. Markus, thank you for easing me into mathematical biology, and keeping such an open door; also, to my pharmacology supervisor, Stephen Hill: your explanations were super-clarifying, and your enthusiasm for your field is inspiring. Thanks to Stephen Coombes, and Nick Monk for giving crucial advice. Thanks, Lloyd, for being an unofficial mentor, and for a million helpful conversations; it was incredibly helpful to have someone who understood my challenges. Thanks to B84 and B86 for being amazingly helpful, especially Martin, Anna, and Sarah. The worlds of MATLAB and LaTeXwould have been much stranger without your help. Thanks to all the other students and postdocs, especially those friends I made more recently. The administrative staff, especially Dave and Helen, were so helpful *throughout* my studies, and for that I am grateful. Also, thanks to the authors of [15, 78] for making example MATLAB codes available.

Erin, thank you so much for (amongst other things) enthusiastically proofreading all my chapters! Thanks to all my old and new friends, and extended family that have supported me; uncle Sola, thanks for convincing me to study mathematics in the first place. Ebun, Nneka, and Ndeshi, thanks for trading stories. ;)

Finally, to you who are blood *and* water: you all are not afraid to tell me the truth, but are a hundred percent on my side. Your support was (and always will be) crucial!

Contents

CHAPTER 1

Introduction

1.1 Cell signalling

The proper functioning of any organism depends on the ability of its individual cells to correctly interact with their environment and each other. For example, a normal immune response to harmful compounds (like insect bites) is the release of Histamine, which then interacts with cells in the affected area to induce protective mechanisms like swelling and muscle contraction. The processing of a variety of such signals by various cells is what enables the proper functioning of organisms, and the underlying mechanisms can collectively be referred to as cell signalling. Understanding those mechanisms can help explain and rectify failures in physiological functioning. For instance, in allergy sufferers, the protective release of Histamine occurs in response to innocuous substances, and therapies (such as antihistamines) were developed using knowledge of the mechanisms of Histamine's interaction with cells.

One of the ways that cells respond to signals is via proteins called *receptors*, which act as sensors. They are activated by ligands, which are molecules that induce a response by binding to receptors and forming a complex; they often alter the conformation, and hence functionality of the receptors. Receptors which reside in the cell membrane are called *cell surface receptors*; these possess an extracellular portion to which ligands can bind, and a portion on the membrane interior where other molecules can interact (see Figure 1.1). In this way the extracellular signal can be detected and relayed into the cell. Molecules which can act as ligands, include hormones, neurotransmitters and drugs [106]. Hormones are chemical messengers which, from organs where they are produced, travel through the bloodstream to other organs [56], where they can induce a cellular response by interacting with cell surface receptors; physiological functions as varied as growth, metabolism, digestion and heart rate regulation are initiated by hormones. Neurotransmitters are chemical messengers which transmit impulses between nerves, or between nerves and muscles. Many existing drugs are synthetic

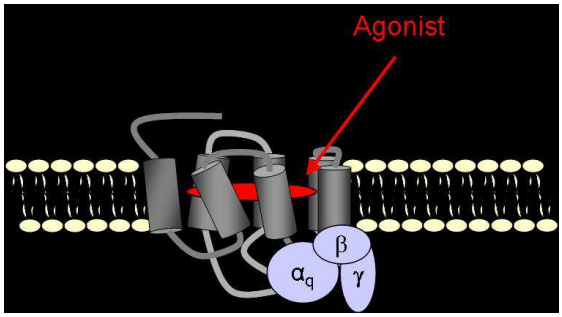


Figure 1.1: Schematic of a G-protein coupled receptor, which is a cell surface receptor.

compounds, either designed to induce similar responses to endogenous ligands or to inhibit their action. For instance, antihistamines work by blocking the binding of Histamine to its receptor.

1.1.1 G-protein coupled receptor signalling

The interest of this work is confined to cell signalling pathways initiated by *G-protein coupled receptors* (GPCR's), a family of cell surface receptors. GPCR's are the largest family of membrane-bound receptors, through which the majority of extracellular signals are processed [1]. They span the membrane by winding back and forth seven times through it, and G proteins, which are located on the inside of the cell membrane, can bind to their intracellular portion. It is reported that more than 50% of all drugs in clinical practice are targeted toward GPCR's [18]; this is because they mediate numerous physiological functions, including metabolism, reproduction, development, hormonal homeostasis, and behaviour [98].

G-proteins are so called because of the role that guanine nucleotides play in regulating their activity [79]. The G-protein is a heterotrimer, i.e., it is composed of three subunits: the α, β, γ units. There is a guanine nucleotide binding site on the α subunit, and β and γ are very closely associated, and thought to operate as one subunit [27]. In the G-protein's inactive state, the α subunit's binding site is occupied by the guanine nucleotide, guanosine diphosphate (GDP). The receptor can be activated with or without a bound ligand, and when (in its active state) it binds to the G-protein, GDP dissociates and leaves room for the guanine nucleotide, guanosine triphosphate (GTP), to bind. GTP activates the G-protein, causing a change in conformation, and causing the G-protein to dissociate from the ligand-receptor complex, which can subsequently bind to other inactive G-proteins (see Figure 1.2).

A widely held hypothesis is that the G-protein's subunits dissociate on activation. However, there are objections to this hypothesis which state that while it is certain that

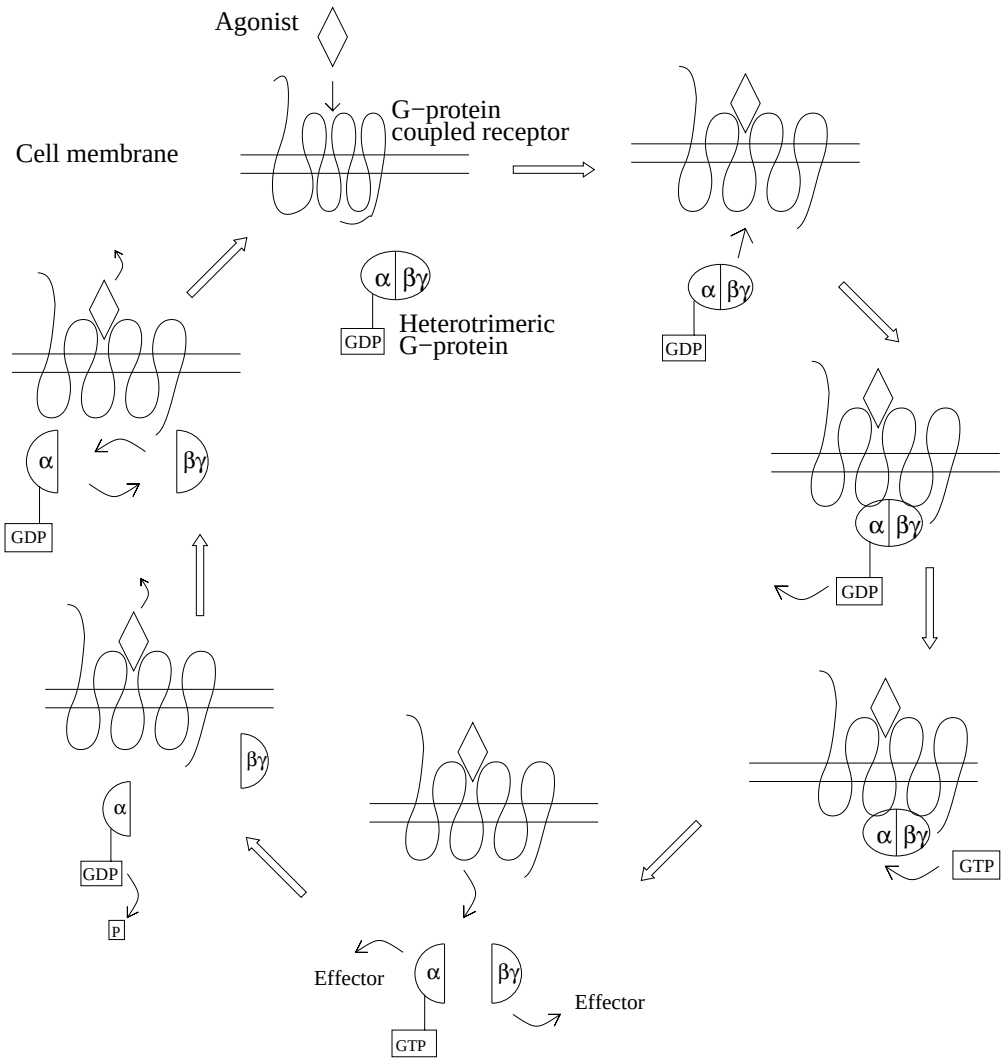


Figure 1.2: Schematic of the G-protein activation cycle (reproduced with permission from [106]), based on the hypothesis that the G-protein's subunits dissociate; an alternative hypothesis states that they do not dissociate.

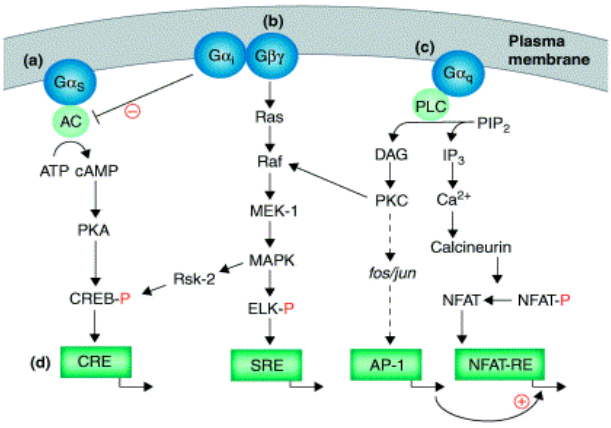


Figure 1.3: Pathways activated by different G-proteins, and interactions between them. Ca^{2+} mobilisation is mediated via the G_q protein. Picture reproduced with permission from [32].

the G-protein can dissociate under certain experimental conditions (for example, in detergent solution, in the presence of high concentrations of magnesium ions), it might not necessarily do so in cell membranes [27]. Consequently, an alternative, experimentally supported hypothesis is that the active α subunit does not need to dissociate in order to mediate downstream signalling [27, 79]. Most mathematical models are based on the former hypothesis [37, 44, 48, 86, 106], but in this chapter both will be considered.

The α subunit is inactivated when GTP is hydrolysed back to GDP, through an intrinsic ability of the α subunit, referred to as *GTPase activity*. There are other proteins, including Regulators of G protein signalling (RGS) and the enzyme Phospholipase C (PLC), which have the ability to bind to the α subunit and accelerate its intrinsic GTPase activity.

There are four main G-protein classes: G_s, G_i, G_q, and $G_{12/13}$, to which all the aforementioned G-protein activation mechanisms apply, but each type is capable of mediating different pathways. G_s-coupled receptors stimulate adenylyl cyclase, which synthesises cyclic adenosine monophosphate (cAMP), which in turn mediates gene transcription; G_i-coupled receptors inhibit adenylyl cyclase and so reduce cAMP formation (see Figure 1.3). The $\beta\gamma$ subunits from G_i and other G-proteins are also able to mediate downstream signalling [32]. The $G_{12/13}$ class of G-proteins is the most recently discovered of the four [89], and mediates signalling to the cytoskeleton via the GTPases Rac and RhoA [27]. This thesis focuses on the pathway activated by the G_q protein, which leads to calcium release from intracellular stores.

The active G_q-protein activates the β isoform of PLC ($PLC\beta$) by binding, after which

its membrane-bound substrate phosphatidylinositoldiphosphate (PIP_2) is cleaved by $PLC\beta$ into inositol triphosphate (IP_3) and diacylglycerol (DAG). Subsequently, IP_3 and DAG trigger divergent pathways. DAG remains in the membrane and activates Protein Kinase C (PKC), which contributes to receptor desensitisation, and continues a pathway that eventually reaches the nucleus of the cell [55]. IP_3 diffuses into the cytosol and goes on to stimulate calcium (Ca^{2+}) release from intracellular stores into the cytosol, ultimately leading to cellular responses such as smooth muscle contraction and changes in gene expression via the activation of transcription factors. This thesis aims to develop a mathematical model that encompasses signalling events from ligand binding to Ca^{2+} mobilisation, but because of the complexity involved, this will be done in stages, of which the early signalling event, G-protein activation is the first.

1.1.2 Motivation

Pharmacology has been described as "a way of thinking and designing experiments [which] compare the activities of natural and synthetic compounds on a biological target (usually a GPCR) and use quantitative measurements to gain insight into mechanisms of action" [33]. Exploring the relationship between drugs (synthetic compounds) and the nature and extent of the cellular response they induce is an important part of the process of *drug discovery*, where drugs are tested for potential use in medicine. The work in this thesis was done in collaboration with pharmacological experimentalists from the Institute of Cell Signalling at the University of Nottingham, who have studied in particular the M3 muscarinic GPCR system, for which there are synthetic drugs that successfully induce a similar Ca^{2+} response as the receptor's natural ligand, Acetylcholine.

The model which will be assembled in this thesis will be tailored to match Ca^{2+} mobilisation data from the M3 muscarinic pathway [52, 54], from experiments carried out by Dr Lauren May under the supervision of Professor Stephen Hill. Guided by these data, the aim is to use a detailed model to provide further insight into mechanisms of drug action in this pathway, and where possible, other GPCR-mediated Ca^{2+} pathways.

Pharmacological analysis frequently uses simplified functional forms which phenomenologically describe the data, but, by lumping several processes together, might obscure mechanistic detail. Hence, a detailed model can be used to shed light on important mechanisms in the pathway. It is hoped that this will inspire new ways of thinking about, and designing, experiments. The M3 receptor is already a drug target in the treatment of lung-related and urological disorders [5], but understanding the pathway mechanisms might highlight other drug targets. Also, this pathway is in-

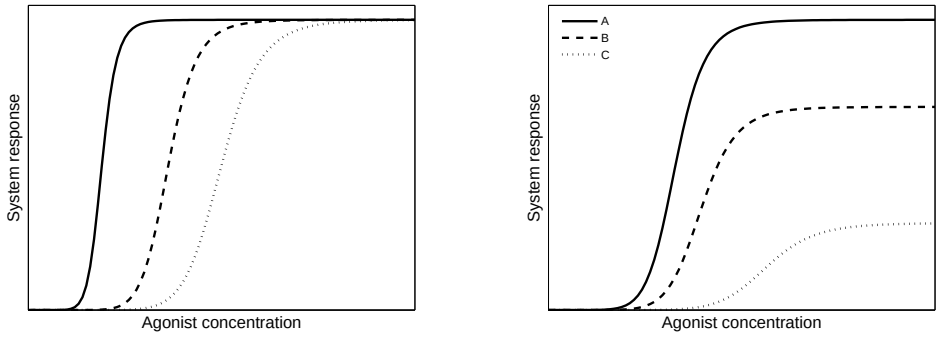
volved in controlling cell-cycle progression and cell death in activated T cells [49], so important insights for regenerative medicine might be gained by better understanding this pathway.

1.2 Experimental studies of GPCR signalling

For obvious reasons, testing the effects of ligands on receptor systems cannot routinely be carried out in the human host. In previous years, the lack of access to human receptors necessitated the use of animal receptor systems, as many of the ligands mediating response via these receptors were the same in animals as in humans [38]. Most of the known therapeutic drugs were discovered using these surrogate animal tissues. However, these experiments have largely given way to experiments carried out in engineered human receptor systems known as *recombinant systems* [40]. These are set-ups in which human DNA is used to make non-human cells express human receptors. This process of introducing nucleic acid into cells to express particular proteins is what is referred to as *transfection*. Another important term is *overexpression*, which refers to the process by which these surrogate cells are exposed to high DNA concentrations in order to yield increased expression of particular proteins [40]. Current technology allows for both receptor and G-protein overexpression in appropriate GPCR systems [39, 85].

Recombinant systems are often used in drug discovery. Two terms often used to describe the properties of a drug are *affinity* and *efficacy* (For a glossary of these, and other, pharmacological terms, see appendix A). Affinity is a ligand-specific property which refers to the ability of a drug to bind (or stay bound) to its receptor, and so is purely a measure of the strength of interaction, while efficacy is a measure of the drug's ability to induce a response. The latter would depend on a ligand's *intrinsic* efficacy, and system-specific mechanisms [40].

An *agonist* is a ligand which acts to positively increase cellular response, so it possesses positive efficacy; different agonists can have different efficacies. An *inverse agonist* acts to decrease basal activity, hence it possesses negative efficacy, while an *antagonist* merely occupies the binding site and in itself does not act to influence the response either way. Figure 1.4 shows examples of dose-response curves which are often used in pharmacology to quantify drug activity. The steepness of a curve depends on both the affinity and efficacy of the agonist it represents, while its maximum value is dependent on its efficacy alone. So in Figure 1.4(a), agonists A, B and C all produce the maximum response, i.e. they are *full agonists*, but A and B are more *left-shifted* than C because the agonists have varying affinites. In Figure 1.4(b), only A is a full agonist. B and C are *partial* agonists, but B has the greater efficacy because the highest response it can



(a) (b)

Figure 1.4: Dose-response curves illustrating (a) the action of full agonists; some curves are more left-shifted, but all produce the maximal response. (b) the action of a full agonist (A) and partial agonists (B and C), which do not induce the maximal response.

produce is higher than C's.

In the study of receptor systems, cellular response cannot always be measured at every level of the pathway. For GPCR's there are currently no experimental methods for tracking α_{GTP} or $\beta\gamma$ over time, hence there are no data with which to directly compare time course simulations. But there are available procedures to measure the binding of a GTP analogue $GTP\gamma S$ to the α subunit [58], which cannot be hydrolysed, so that the whole G-protein cycle is not completed; hence, the G-protein response measured in such experiments is likely higher than it would normally be. Also, the accumulation of ligand-receptor complexes and time courses for second messengers like Ca^{2+} are easily measured. Experimental methods for measuring Ca^{2+} will be discussed in Chapter 3.

1.3 Previous mathematical models of GPCR signalling

Over the years, many models have been developed to describe the behaviour of receptor signalling systems, which are largely a series of biochemical reactions. The majority of models utilise the *law of mass action*, which states that the rate of a reaction is proportional to the product of the concentrations of its reactants. Given the following reversible reaction

$$A + B \underset{k_-}{\overset{k_+}{\rightleftharpoons}} AB, \tag{1.1}$$

the rate of production of the product AB, is given by

$$\frac{d[AB]}{dt} = -\frac{d[A]}{dt} = -\frac{d[B]}{dt} = k_+[A][B] - k_-[AB], \tag{1.2}$$

where (and throughout this thesis), $[*]$ refers to the concentration of $*$. If the reaction is irreversible, the rates are given by substituting $k_- = 0$.

The majority of existing GPCR signalling models are *equilibrium models*. These are models which describe the equilibrium states of the receptors with algebraic equations derived from applying steady state conditions to the mass action equations can be used to describe how the addition of ligand changes the steady state distribution of receptor states [46].

The simplest of these equilibrium models is Clark's classical model [16], proposed in 1933, which describes the interaction of two components only, the receptor, R and ligand, L, and assume that they bind following the law of mass action, so that the receptor can exist in two states: unbound (R) and ligand-bound (LR) [103]. The only possible reactions are:

$$L + R \underset{k_-}{\overset{k_+}{\rightleftharpoons}} LR, \tag{1.3}$$

Using the law of mass action and the assumption of equilibrium (that the forward rate is equal to the backward rate):

$$K_a = \frac{k_+}{k_-} = \frac{[LR]}{[L][R]}, \tag{1.4}$$

where K_a is called the equilibrium association constant. Its reciprocal, K_D, is called the dissociation constant, and is measured in the same unit as the ligand. Both constants are measures of affinity. A high K_a (or equivalently, low K_D) indicates a strong affinity for the receptor.

In this model, the efficacy of a drug is taken to be the ratio of ligand-bound to total receptor concentration, and is quantified by the following saturation function:

$$\text{Response} = \frac{[\text{L}]}{[\text{L}] + K_D}, \tag{1.5}$$

using the fact that the total number of receptors is conserved and given by $R_{TOT} = [R] + [LR]$. No provision was made in this model for different drugs with similar affinities but varying efficacies. In 1954, E. J. Ariens introduced a proportionality factor κ to address this limitation as follows [4]:

$$\text{Response} = \frac{\kappa \,.\, [\text{L}]}{[\text{L}] + K_D}. \tag{1.6}$$

κ values exist in the range of 0 to 1, representing a spectrum of agonists that produce no response to those that produce the maximum response. Since then many more models with differing levels of complexity [103] have been developed to account for further experimental observations. After it came to light that receptors can bind to other membrane proteins (the G-protein in the case of GPCR's), the ternary complex model (TCM) [19] was proposed, which incorporates G-protein binding into the classical model. It allows the receptor to bind to a third component, the G-protein (G). The reactions for this model are:

$$L + R \rightleftharpoons LR \tag{1.7}$$

$$LR + G \rightleftharpoons LRG \tag{1.8}$$

In this model the receptor can only bind to the G-protein when it has itself been bound by the ligand.

New experimental evidence revealed that unbound receptors can spontaneously become active and in turn activate G-proteins in the absence of any ligand; such a response is referred to as *constitutive activity*. Increased constitutive activity has been observed in systems where receptors have been overexpressed; for instance, it was observed in Sf9 insect cells that the second messenger cAMP can be generated at a rate which is directly proportional to the level of receptor expression [27]. Constitutive activity has been detected by overexpressing many other GPCR's, including the Ca^{2+}-activating G_q-coupled Histamine H1 receptor [84]. Another source of constitutive activity is G-protein overexpression; in [85], GTPase activity was reported to increase with G-protein expression levels. The TCM could not account for constitutive activity, so an extension was proposed: the extended ternary complex model (eTCM). This model assumes that the receptor can exist in two states: one active (R), one inactive (R^*), with the ligand able to bind to both states with different affinities. The active receptor, whether ligand-bound or not, can bind and activate the G-protein [103], although the ligand can alter its affinity for the G-protein. Consequently, there are six receptor species: R, R^*, LR, LR^*, R^*G, LR^*G (see Figure 1.5).

The discovery of promiscuous receptors, which can bind to more than one G-protein type, prompted the proposal of a three-state receptor model. This model assumes the simplest case, in which the receptor can exist in two active conformations, R^* and R^{**}, each able to bind to a unique G-protein type [71]. An explanation for the ability of a drug to be an agonist in one system and an inverse agonist in another was offered by the model.

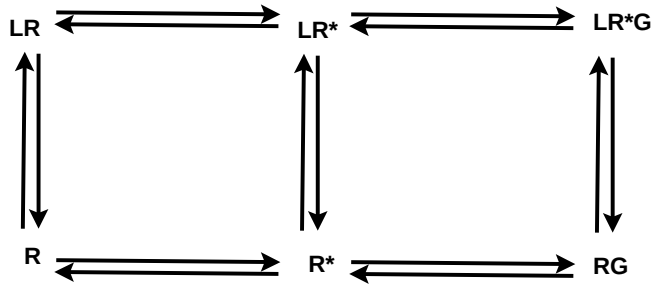


Figure 1.5: The extended cubic ternary complex model (eCTM) accounts for the fact that receptor can become spontaneously become active (R^*).

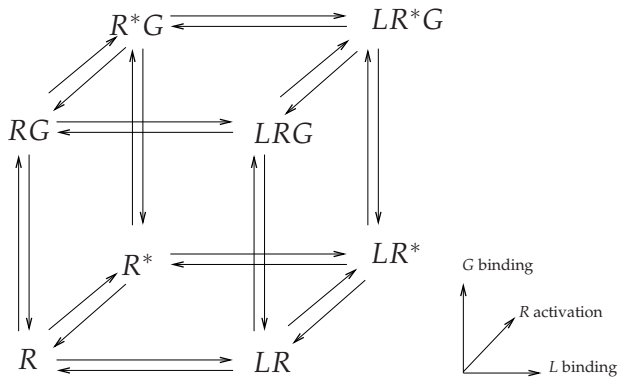


Figure 1.6: The cubic ternary complex (CTC) model accounts for all possible states.

All the previously mentioned models (except the three state model) are subsets of the cubic ternary complex model (CTC), a comprehensive model which includes all possible combinations of receptor, ligand and one G-protein type, based on the assumption that the ligand and G-protein never directly encounter each other [103]. In this model the receptor can exist in eight states, as shown in Figure 1.6, and agonist efficacy is the ability of a ligand to increase the number of receptors in the active, signalling state (R^*G and LR^*G) at equilibrium.

1.3.1 Equilibrium vs kinetic models

The assumption of equilibrium in the previously discussed models limit the ability to characterise a drug's properties because important time-dependent (e.g. transient) responses might go uncharacterised. As a result, kinetic models are now more commonly used to model GPCR signalling. The cubic ternary complex activation model (cTCAM), for example, integrates the kinetics of G protein activation and deactivation in the CTC model with three additional irreversible reactions that describe the G-protein activation cycle [86]. Without the assumption of equilibrium, the model consists of a system

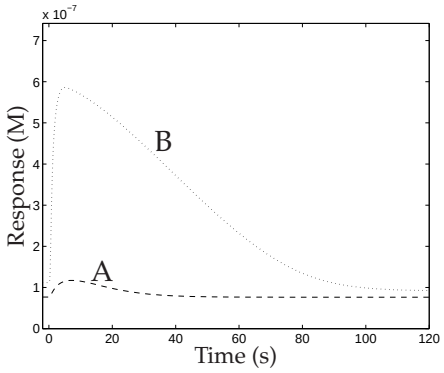


Figure 1.7: Transient responses induced by two different drugs. If only the equilibrium response is measured, both drugs appear to have little efficacy. However, if the peak response is used, drug B is efficacious.

of ODEs, allowing, for instance, a transient α_{GTP} peak to be predicted [41]. Figure 1.7 is used to illustrate why it may be important to use transient responses as measures of a drug's action. If only the equilibrium response is measured, both drugs would appear to be similar, having little efficacy. However, if the peak response is used, only drug A has little efficacy, and such behaviour can be captured by a kinetic model.

1.3.2 A kinetic model: The cubic ternary complex activation model

The cTCAM, which will be used to analyse the dynamics of G-protein activation in this thesis, is now described in detail.

Receptor-Ligand binding

The model assumes that the receptor switches between an active and inactive state, as follows:

$$R \underset{k_{deact}}{\overset{k_{act}}{\rightleftharpoons}} R^*, \tag{1.9}$$

The ligand can bind to both receptor states with different affinites, modelled by the thermodynamic constant $\zeta = \frac{\zeta_+}{\zeta_-}$, as follows:

$$L + R \underset{k_{lb-}}{\overset{k_{lb+}}{\rightleftharpoons}} LR, \tag{1.10}$$

$$L + R^* \underset{\zeta_- k_{lb-}}{\overset{\zeta_+ k_{lb+}}{\rightleftharpoons}} LR^*. \tag{1.11}$$

$\zeta = 1$ indicates an *antagonist*, a drug which binds indiscriminately to both receptor states; $\zeta > 1$ indicates an agonist, which favours the receptor's active state; $\zeta < 1$

models an *inverse agonist*, which favours the inactive state of the receptor. The ligand concentration is assumed constant, since it is usually supplied in excess.

The inactive receptor-ligand complex can also spontaneously become active, as follows:

$$LR \underset{\zeta_- k_{deact}}{\overset{\zeta_+ k_{act}}{\rightleftharpoons}} LR^*. \tag{1.12}$$

The four receptor states can also bind to the G-protein with different affinities. This includes the possibility of *precoupling*, a situation in which a G protein binds to a receptor before the ligand:

$$R + G \underset{k_{g-}}{\overset{k_{g+}}{\rightleftharpoons}} RG, \tag{1.13}$$

$$R^* + G \underset{\mu_- k_{g-}}{\overset{\mu_+ k_{g+}}{\rightleftharpoons}} R^*G, \tag{1.14}$$

$$LR + G \underset{\nu_- k_{g-}}{\overset{\nu_+ k_{g+}}{\rightleftharpoons}} LRG, \tag{1.15}$$

$$LR^* + G \underset{\theta_{\nu\mu}\mu_-\nu_- k_{g-}}{\overset{\theta_{\nu\mu}\mu_+\nu_+ k_{g+}}{\rightleftharpoons}} LR^*G, \tag{1.16}$$

where μ_+ and μ_- make up the equilibrium constant, $\mu = \frac{\mu_+}{\mu_-}$, which models the extent to which the active receptor binds more or less preferably to the G-protein than the inactive receptor. ν_+ and ν_- make up the equilibrium constant $\nu = \frac{\nu_+}{\nu_-}$, which models the extent to which the ligand binds more preferably to the precoupled receptors (RG, R^*G) than to the uncoupled ones. $\theta_{\nu\mu}$ quantifies the extent to which ligand binding and receptor activation jointly facilitate the binding of G-protein.

The inactive receptor-G-protein complexes can also spontaneously become active:

$$RG \underset{\mu_- k_{deact}}{\overset{\mu_+ k_{act}}{\rightleftharpoons}} R^*G, \tag{1.17}$$

$$LRG \underset{\theta_{\zeta\mu}\zeta_-\mu_- k_{deact}}{\overset{\theta_{\zeta\mu}\zeta_+\mu_+ k_{act}}{\rightleftharpoons}} , LR^*G. \tag{1.18}$$

where $\theta_{\nu\mu}$ quantifies the extent to which ligand binding and G-protein coupling jointly facilitate receptor activation.

The ligand can also bind to precoupled receptors, as follows:

$$L + RG \underset{\nu_- k_{lb-}}{\overset{\nu_+ k_{lb+}}{\rightleftharpoons}} LRG, \tag{1.19}$$

$$L + R^*G \underset{\theta_{\zeta\nu}\zeta_-\nu_-k_{lb-}}{\overset{\theta_{\zeta\nu}\zeta+\nu_+k_{lb+}}{\rightleftharpoons}} LR^*G, \tag{1.20}$$

where $\theta_{\zeta\nu}$ quantifies the extent to which receptor activation and G-protein coupling jointly facilitate ligand binding.

It is assumed that only the active receptor states, R^*G and LR^*G, can carry on the signalling pathway, by activating the G-protein.

The G-protein activation cycle

As noted previously, there are divided views on the exact mechanisms of G-protein activation and deactivation. The cTCAM is built on the hypothesis that GTP binding causes the α and the $\beta\gamma$ subunits of the G-protein to dissociate into two activated components, α_{GTP} and $\beta\gamma$, which can interact with effector molecules and propagate signalling [1], as given by the following reactions:

$$\begin{aligned} R^*G + GTP &\longrightarrow R^* + \alpha_{GTP} + GDP + \beta\gamma \\ LR^*G + GTP &\longrightarrow LR^* + \alpha_{GTP} + GDP + \beta\gamma. \end{aligned}$$

It is assumed that GTP and GDP are present at constant levels. This means that their concentrations can be incorporated into the rate constants, so that the activation reactions are given by:

$$R^*G \xrightarrow{k_{GTP+}} R^* + \alpha_{GTP} + \beta\gamma, \tag{1.21}$$

$$LR^*G \xrightarrow{k_{GTP+}} LR^* + \alpha_{GTP} + \beta\gamma. \tag{1.22}$$

As previously mentioned, α_{GTP} has the ability to deactivate itself by hydrolysing GTP to GDP; its bound GTP loses one phosphate molecule, leaving the inactive form, α_{GDP},; the inactive G-protein (G) is recovered when α_{GDP} reunites with $\beta\gamma$:

$$\alpha_{GTP} \xrightarrow{k_{gd+}} k_{gd-} \quad \alpha_{GDP}, \tag{1.23}$$

$$\alpha_{GDP} + \beta\gamma \xrightarrow{k_{RA+}} k_{RA-} \quad G. \tag{1.24}$$

The law of mass action is used to convert all reactions to ODEs which can then be integrated and used to describe time-dependent behaviour. A variant model, with a G-protein activation and deactivation cycle built on the hypothesis that the G-protein subunits do not dissociate will also be described in the next chapter.

There are other relevant kinetic mathematical models of G-protein signalling that describe events beyond G-protein activation. Lemon *et al.* incorporated the dynamics of IP_3 formation and Ca^{2+} mobilisation into the G-protein cascade [44], while Kang

and Othmer incorporated the dynamics of IP_3 formation, PKC feedback and Ca^{2+} mobilisation [37]. In another study, agonist-induced Ca^{2+} mobilisation was coupled to downstream phosphorylation events [48]. These models will be discussed when events downstream of G-protein activation are considered, later in Chapter 3.

1.4 Thesis Outline

The rest of the thesis is outlined below.

In Chapter 2, the cTCAM is analysed by considering the effects of varying cell-specific parameters (total receptor and G-protein concentration) on the active G-protein response. The effects of ligand-specific parameters are also considered, when particular drugs are modelled. In addition, the cTCAM is compared with a variant model, in which the G-protein subunits do not dissociate.

In Chapter 3, Ca^{2+} signalling is discussed in more detail. The cTCAM is then extended by adapting downstream details from an existing Ca^{2+} model [37]. Guided by experimental results, the parameters of the extended model are then fine-tuned using manual variation and numerical simulation. A more systematic parameter analysis is then carried out in Chapter 4, by using a global sensitivity analysis method to identify parameters which most influence key features of the Ca^{2+} response seen in the data. Such parameters are then adjusted to enable the model match the data more closely, while providing insight about pathway dynamics.

In Chapter 5, possible adjustments to the model (suggested by the results of Chapter 4) are considered, resulting in variant Ca^{2+} mobilisation models, which are subsequently analysed.

In Chapter 6, a global parameter optimisation is carried out to improve the quantitative fit to the data, and to provide further insight about the model, while a further sensitivity analysis points to possible model extensions.

Finally, in Chapter 7, the main results of the thesis are summarised, and suggestions are made about future work that can build on these results.

CHAPTER 2

Modelling G-protein Activation

2.1 Introduction

The pathway from ligand binding to Ca^{2+} release consists of multiple processes, many of which are complex in nature. This thesis seeks to assemble a model of these processes which is detailed enough to provide insight into important pathway mechanisms; however, this results in a complex model; it therefore makes sense to start by analysing a simpler, existing model of early signalling events (such as the cubic ternary complex activation model (cTCAM)), which can subsequently be developed into a more extensive model. This allows the model to be better understood before it is used as a base model in the extension process. In addition, insight can be gained about the earlier signalling events.

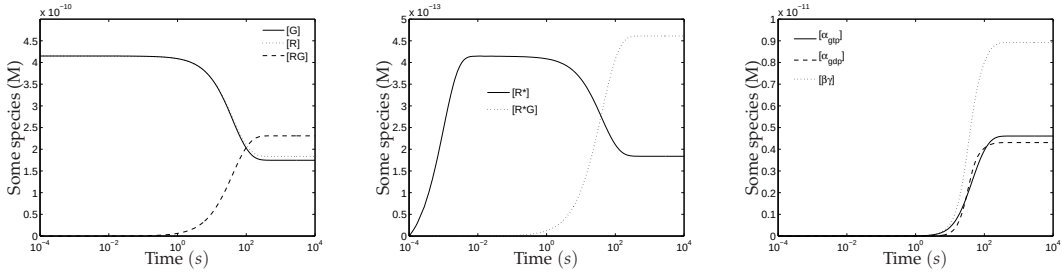
In this chapter, the cTCAM (see section 1.3), is used to study the responses of the mediators of downstream signalling, α_{GTP} and $\beta\gamma$; in particular, the effects of changing model parameters such as the total receptor, total G-protein, and ligand concentrations are explored; these parameters can be experimentally manipulated, and so have direct practical application. Also, ligand-specific parameters such as affinity and efficacy are modelled for actual drugs which were used in Dr Lauren May's M3 muscarinic experiments [52, 54].

As noted in Chapter 1, traditionally, GPCR modelling has made use of equilibrium models, so time courses for active G-protein have only recently been simulated; α_{GTP}, but not $\beta\gamma$, has been simulated in some relatively recent studies [41, 44, 106]. Although Ca^{2+} release in the M3 muscarinic pathway is mediated through α_{GTP}, the $\beta\gamma$ response will also be discussed in this chapter, for comparison, since both subunits partake in the G-protein activation cycle.

The cTCAM model equations, derived by applying the law of mass action to reactions (1.9) - (1.24) in Chapter 1, are given by:

$$\begin{aligned}
\frac{d[R]}{dt} &= -k_{act}[R] + k_{deact}[R^*] - k_{lb+}[L][R] + k_{lb-}[LR] \\
&\quad -k_{g+}[R][G] + k_{g-}[RG] && (2.1)\\
\frac{d[R^*]}{dt} &= k_{act}[R] - k_{deact}[R^*] - \zeta_+ k_{lb+}[L][R^*] + \zeta_- k_{lb-}[LR^*] \\
&\quad -\mu_+ k_{g+}[R^*][G] + \mu_- k_{g-}[R^*G] + k_{GTP+}[R^*G] && (2.2)\\
\frac{d[LR]}{dt} &= k_{lb+}[L][R] - k_{lb-}[LR] - \zeta_+ k_{act}[LR] + \zeta_- k_{deact}[LR^*] \\
&\quad -\nu_+ k_{g+}[LR][G] + \nu_- k_{g-}[LRG] && (2.3)\\
\frac{d[LR^*]}{dt} &= \zeta_+ k_{lb+}[L][R^*] - \zeta_- k_{lb-}[LR^*] + \zeta_+ k_{act}[LR] - \zeta_- k_{deact}[LR^*] \\
&\quad -\theta_{\nu\mu}\mu_+\nu_+ k_{g+}[LR^*][G] + \theta_{\nu\mu}\mu_-\nu_- k_{g-}[LR^*G] \\
&\quad +k_{GTP+}[LR^*G] && (2.4)\\
\frac{d[RG]}{dt} &= k_{g+}[R][G] - k_{g-}[RG] - \mu_+ k_{act}[RG] + \mu_- k_{deact}[R^*G] \\
&\quad -\nu_+ k_{lb+}[L][RG] + \nu_- k_{lb-}[LRG] && (2.5)\\
\frac{d[R^*G]}{dt} &= \mu_+ k_{g+}[R^*][G] - \mu_- k_{g-}[R^*G] + \mu_+ k_{act}[RG] - \mu_- k_{deact}[R^*G] \\
&\quad -\theta_{\zeta\nu}\zeta_+\nu_+ k_{lb+}[L][R^*G] + \theta_{\zeta\nu}\zeta_-\nu_- k_{lb-}[LR^*G] - k_{GTP+}[R^*G] && (2.6)\\
\frac{d[LRG]}{dt} &= \nu_+ k_{g+}[LR][G] - \nu_- k_{g-}[LRG] - \theta_{\zeta\mu}\zeta_+\mu_+ k_{act}[LRG] \\
&\quad +\theta_{\zeta\mu}\zeta_-\mu_- k_{deact}[LR^*G] + \nu_+ k_{lb+}[L][RG] - \nu_- k_{lb-}[LRG] && (2.7)\\
\frac{d[LR^*G]}{dt} &= \theta_{\nu\mu}\mu_+\nu_+ k_{g+}[LR^*][G] - \theta_{\nu\mu}\mu_-\nu_- k_{g-}[LR^*G] \\
&\quad +\theta_{\zeta\mu}\zeta_+\mu_+ k_{act}[LRG] - \theta_{\zeta\mu}\zeta_-\mu_- k_{deact}[LR^*G] \\
&\quad +\theta_{\zeta\nu}\zeta_+\nu_+ k_{lb+}[L][R^*G] - \theta_{\zeta\nu}\zeta_-\nu_- k_{lb-}[LR^*G] \\
&\quad -k_{GTP+}[LR^*G] && (2.8)\\
\frac{d[G]}{dt} &= -\nu_+ k_{g+}[LR][G] + \nu_- k_{g-}[LRG] \\
&\quad -\theta_{\nu\mu}\mu_+\nu_+ k_{g+}[LR^*][G] + \theta_{\nu\mu}\mu_-\nu_- k_{g-}[LR^*G] \\
&\quad -\mu_+ k_{g+}[R^*][G] + \mu_- k_{g-}[R^*G] \\
&\quad -k_{g+}[R][G] + k_{g-}[RG] + k_{RA+}[\alpha_{GDP}][\beta\gamma] && (2.9)\\
\frac{d[\alpha_{GTP}]}{dt} &= k_{GTP+}[R^*G] - k_{gd+}[\alpha_{GTP}] + k_{GTP+}[LR^*G] && (2.10)\\
\frac{d[\beta\gamma]}{dt} &= k_{GTP+}[R^*G] + k_{GTP+}[LR^*G] - k_{RA+}[\alpha_{GDP}][\beta\gamma] && (2.11)\\
\frac{d[\alpha_{GDP}]}{dt} &= k_{gd+}[\alpha_{GTP}] - k_{RA+}[\alpha_{GDP}][\beta\gamma] && (2.12)
\end{aligned}$$

The variables are the concentrations of the different signalling molecules (measured in moles per litre, M) and time, measured in seconds. Because of the number of reactions involved in the signalling pathway, the model is a high order system of coupled



(a) (b) (c)

Figure 2.1: The unstimulated phase of signalling. (Note the log scale of the x axes.) Initially, only the unbound receptor (R) and G-protein states (G) exist. (a) Receptor pre-coupling occurs when the receptor and G-protein bind; (b) spontaneous receptor activation also occurs, allowing the active pre-coupled state (R^*G) to form; (c) some G-protein activation also occurs. The steady state levels of these species then serve as initial conditions for ligand stimulation.

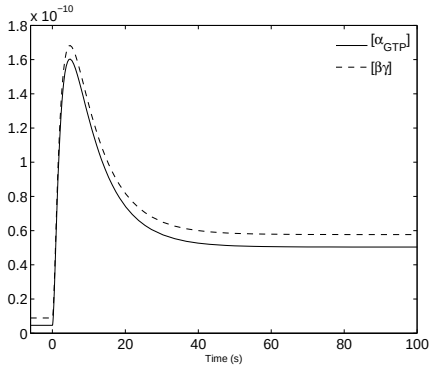
ODEs which cannot be solved analytically, hence computational methods are heavily relied on for the analysis of the models in this thesis. Model simulations are carried out in MATLAB using the ode15s solver for stiff systems, because the existence of various time scales makes the model stiff [106]. Parameter values for this chapter are listed in Appendix B and were taken from a previous study [106].

In order to compare the hypotheses of G-protein subunit dynamics discussed in section 1.1.1, this chapter finally compares the cTCAM with an adapted version, in which the G-protein does not dissociate into α_{GTP} and $\beta\gamma$.

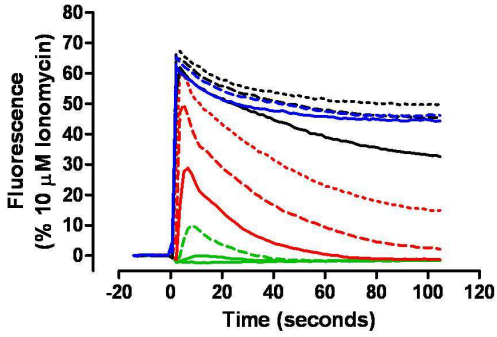
2.2 Agonist-induced G-protein activation

This section uses the cTCAM to mimic experiments in which cells are in a basal state before stimulation by a drug (specifically, an agonist) [54]. Hence, the ligand concentration is initially set as $[L] = 0$. In addition, the total (unbound) G-protein concentration and total (unbound) receptor concentration are given by $G_{TOT} = 4.15 \times 10^{-10} M$ and $R_{TOT} = 4.15 \times 10^{-10} M$ respectively (following a previous study [106]), while all other species are at zero concentration. During this unstimulated phase, formation of other species then occurs until a steady state is reached, as illustrated in Figure 2.1; precoupled receptors, RG and R^*G form, (Figure 2.1(a) (b)), and from the latter, basal α_{GTP} and $\beta\gamma$ form (Figure 2.1(c)). The equilibrium values of all species then serve as the initial conditions for ligand stimulation.

Unless otherwise indicated, a stimulation of $[L] = 10^{-5} M$ of a strong agonist will



(a)



(b)

Figure 2.2: (a) α_{GTP} and $\beta\gamma$ responses in the cTCAM to an agonist, added after 10000s (when a steady state has been reached). Both subunits exhibit the same transient peak response when $G_{TOT} = 4.15 \times 10^{-10}M$, $R_{TOT} = 4.15 \times 10^{-10}M$; (b) Experimental results reveal that Ca^{2+} also exhibits a transient peak response over various agonist concentrations [54].

be used in the simulations; the strength of the agonist is modelled by the values

$$\begin{aligned} k_{lb+} &= 8.4 \times 10^7 M^{-1}s^{-1}, \\ k_{lb-} &= 0.37s^{-1}, \\ \zeta &= \frac{\zeta_+}{\zeta_-} = 1000. \end{aligned}$$

This is an agonist that quickly binds to the receptor and strongly favours the receptor's active state, since ζ is much greater than 1 (see section 1.3.2). As mentioned, the cellular response will be defined as the concentration of activated G-protein.

2.2.1 The transient peak response

Previous studies illustrate that in response to agonist stimulation, α_{GTP} can exhibit a rapid, transient peak response [41, 106], brought on by rapid agonist binding which increases α_{GTP} production via reaction (1.22); consequent GTPase action via reaction 1.24) then makes the peak transient [106]. With the parameter set in Table B.2.1, the cTCAM model predicts that rapid ($\approx$ 5 seconds), transient α_{GTP} peak, as illustrated in Figure 2.2(a); $\beta\gamma$ also exhibits an early peak due to rapid agonist binding, but unlike α_{GTP}, $\beta\gamma$ does not undergo GTPase action, hence its descent to steady state level (also referred to as the *plateau* in this chapter) is due to its subsequent re-association with α_{GDP} (see reaction (1.24)).

Figure 2.2(b) shows Ca^{2+} time courses from the M3 muscarinic receptor pathway [54], from which it can be seen that Ca^{2+} also exhibits a rapid, transient peak. Because

α_{GTP} mediates the production of IP_3, which is the trigger for Ca^{2+} mobilisation, the lifetime of the α_{GTP} response can be expected to have an impact on the lifetime of cytosolic Ca^{2+}. Even though the pathway has downstream mechanisms for IP_3 and Ca^{2+} removal [74, 94], which would contribute to the transience of the Ca^{2+} response, their efficiency would depend on the strength of α_{GTP} activation. Hence, this section explores the hypothesis that the transience of the Ca^{2+} peak depends on the existence of a transient α_{GTP} peak (referred to as the *transient peak hypothesis*). In that case, parameters that control α_{GTP}'s peak could also be used to manipulate Ca^{2+} release in a similar fashion.

2.2.2 The influence of receptor and G-protein expression levels

In a previous study [41], a sensitivity analysis identified the key parameters that influence the dynamics of α_{GTP} production in the cTCAM. 1000 parameter sets were sampled (using latin hypercube sampling) from a parameter space constrained by defining uncertainty ranges for each parameter, based on values from the literature, and it was shown that these parameter sets predict varying α_{GTP} time courses. The variations included disparate basal levels (which could vary by factors of up to 350) and either transient or sustained responses to agonist stimulation. Total G-protein G_{TOT} and total receptor (R_{TOT}) were implicated as the 1st and 3rd most important drivers of the dynamics of α_{GTP}, so they are key to understanding G-protein activation, which could then have implications for the Ca^{2+} response. Hence this section considers the effects of varying G_{TOT} and R_{TOT}, particularly on the transient α_{GTP} (and $\beta\gamma$) peak.

R_{TOT} and G_{TOT} are sometimes quantified according to how many molecules are present per cell. Ranges for R_{TOT} and G_{TOT} have been given as $10^2 - 10^5$ per cell in the literature [41]; conversion to moles per litre can be carried out by multiplying by the number of cells per litre, and dividing by Avogadro's number, $N_A = 6.022 \times 10^{23}$. Assuming a typical value of 5×10^9 cells per litre [86], $10^2 - 10^5$ converts to $8.3 \times 10^{-13} - 8.3 \times 10^{-10} M$. Obviously this conversion depends on the number of cells per litre, which might vary across experiments, hence receptor and G-protein concentrations up to $10^{-6} M$ have also been reported [37]. An uncertainty range of $10^{-12} - 10^{-6} M$ can thus be defined for R_{TOT} and G_{TOT}.

The influence of G-protein concentration

The total G-protein concentration is now varied, while R_{TOT} is fixed at $4.15 \times 10^{-10} M$.

Figures 2.3(a) and 2.3(b) show the $[\alpha_{GTP}]$ and $[\beta\gamma]$ time courses for various G-protein concentrations; the peak concentrations of both subunits increase with G_{TOT},

which can be seen more clearly in Figure 2.3(c), where the peak values are plotted against $\log(G_{TOT}/R_{TOT})$. In Figure 2.3(d), the size of the transient peak, quantified by subtracting the plateau from the peak (Peak $-$ Plateau), is plotted against $\log(G_{TOT}/R_{TOT})$. When the peak is sustained, Peak $-$ Plateau$= 0$, but when it is transient, Peak $-$ Plateau> 0. The significant peak increases occur when $\log(G_{TOT}/R_{TOT})$ lies between 1 and 3, while transience is lost after $\log(G_{TOT}/R_{TOT}) = 1$ (which is when G_{TOT}:R_{TOT} is 10:1). Also, around this value, the peaks become significantly slower, as illustrated by Figure 2.3(e). Hence when G-proteins significantly outnumber the receptors, the response increases but its qualitative nature changes. Hence, if the transience of the Ca^{2+} peak depends on the existence of a transient α_{GTP} peak, it can potentially be removed by expressing more G-protein than receptor.

In general, the dynamics of α_{GTP} and $\beta\gamma$ are strikingly similar; the only differences occur at the lowest G-protein concentrations, when the peaks are relatively insignificant; in these cases, $\beta\gamma$ takes longer to peak than α_{GTP} (see Figure 2.3(e)).

The influence of receptor concentration

The total receptor concentration is now varied, while G_{TOT} is fixed at $4.15 \times 10^{-10}M$.

Figures 2.4(a) and 2.4(b) show the $[\alpha_{GTP}]$ and $[\beta\gamma]$ time courses for various receptor concentrations; the peak values of both subunits increase with $[R_{TOT}]$, which can be seen more clearly in Figure 2.4(c), where the peak values are plotted against $\log(R_{TOT}/G_{TOT})$. In Figure 2.4(d), (Peak $-$ Plateau) is plotted against $\log(R_{TOT}/G_{TOT})$. The significant peak increases occur when $\log(R_{TOT}/G_{TOT})$ lies between -1 and 2, while transience is lost shortly before $\log(R_{TOT}/G_{TOT}) = 2$ (which is when the R_{TOT}:G_{TOT} ratio is 100:1). At this point the peaks also become significantly slower, as can be seen in Figure 2.4(e). Hence, the qualitative nature of the response also changes when the receptors significantly outnumber the G-proteins (and not just the other way round), and if the peak hypothesis is true, these results suggest that the Ca^{2+} transient peak observed in the M3 muscarinic G_q pathway mobilisation experiments may also be removable by increasing receptor expression. Hence, such experiments are worth designing.

An important observation is that, in general, the agonist generates less of a response through receptor overexpression than it does through G-protein overexpression (compare Figures 2.3(c) and 2.4(c)). This is perhaps because of reaction (1.22) ($LR^*G \xrightarrow{k_{GTP+}} LR^* + \alpha_{GTP} + \beta\gamma$) which allows for instantaneous receptor 'recycling'. After α_{GTP} and $\beta\gamma$ are released, the dissociated LR^* complex is free to bind to a different unbound G-protein, and when the total G-protein concentration is high, many more cycles of G-protein activation are then possible. However, when R_{TOT} is varied the effect is not the same; at high receptor concentrations, the number of possible cycles is limited

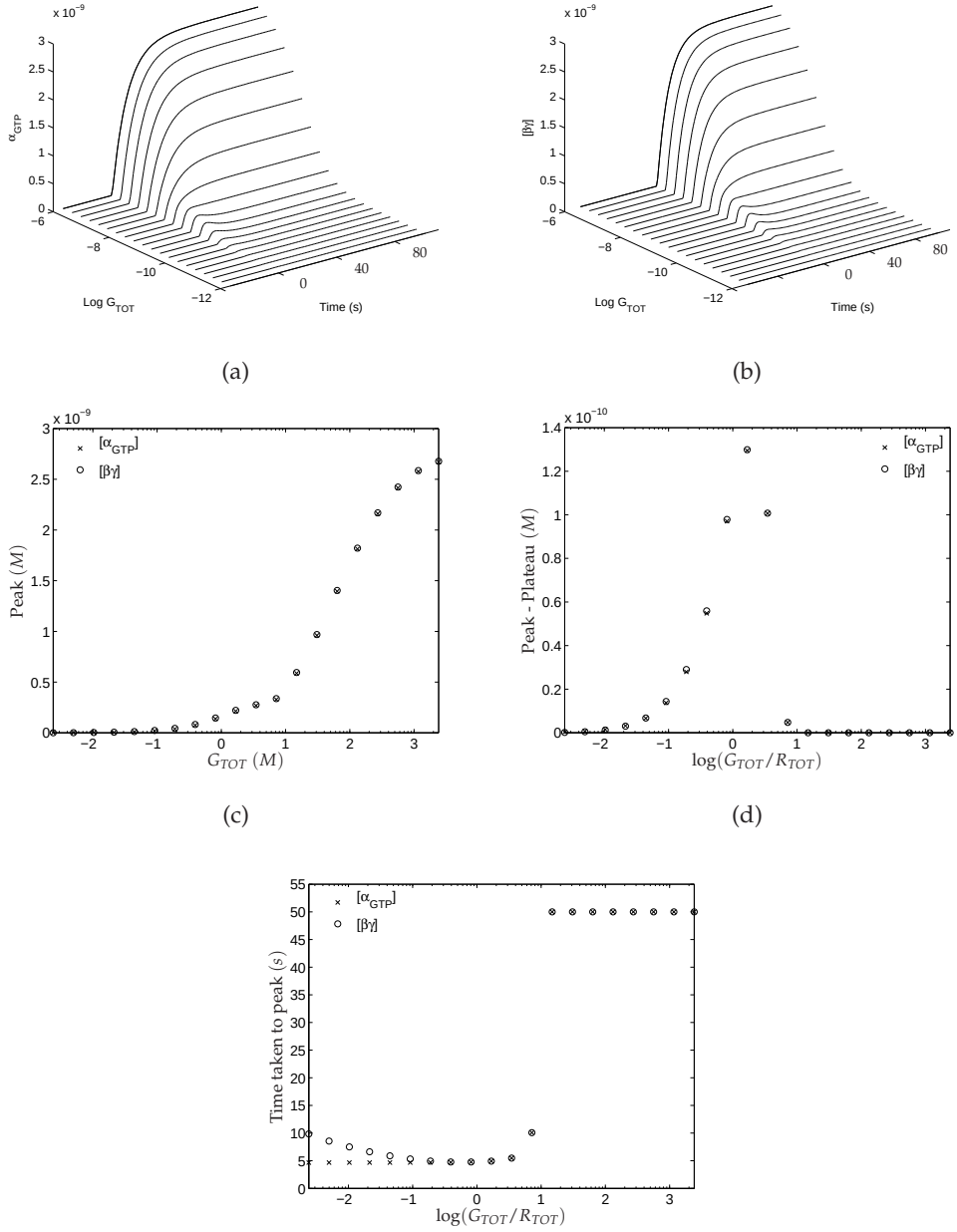


Figure 2.3: The effects of varying the total G-protein concentration ($R_{TOT} = 4.15 \times 10^{-10}M$) on α_{GTP} and $\beta\gamma$ after the addition of an agonist at $t = 10000$s, when the system is at steady state. (a),(b) α_{GTP} and $\beta\gamma$ time courses; (c) α_{GTP} and $\beta\gamma$ peak values, which increase with G_{TOT}; (d) Peak$-$Plateau for α_{GTP} and $\beta\gamma$; (e) the times taken to reach α_{GTP} and $\beta\gamma$ peaks, which are very similar for the significant peaks.

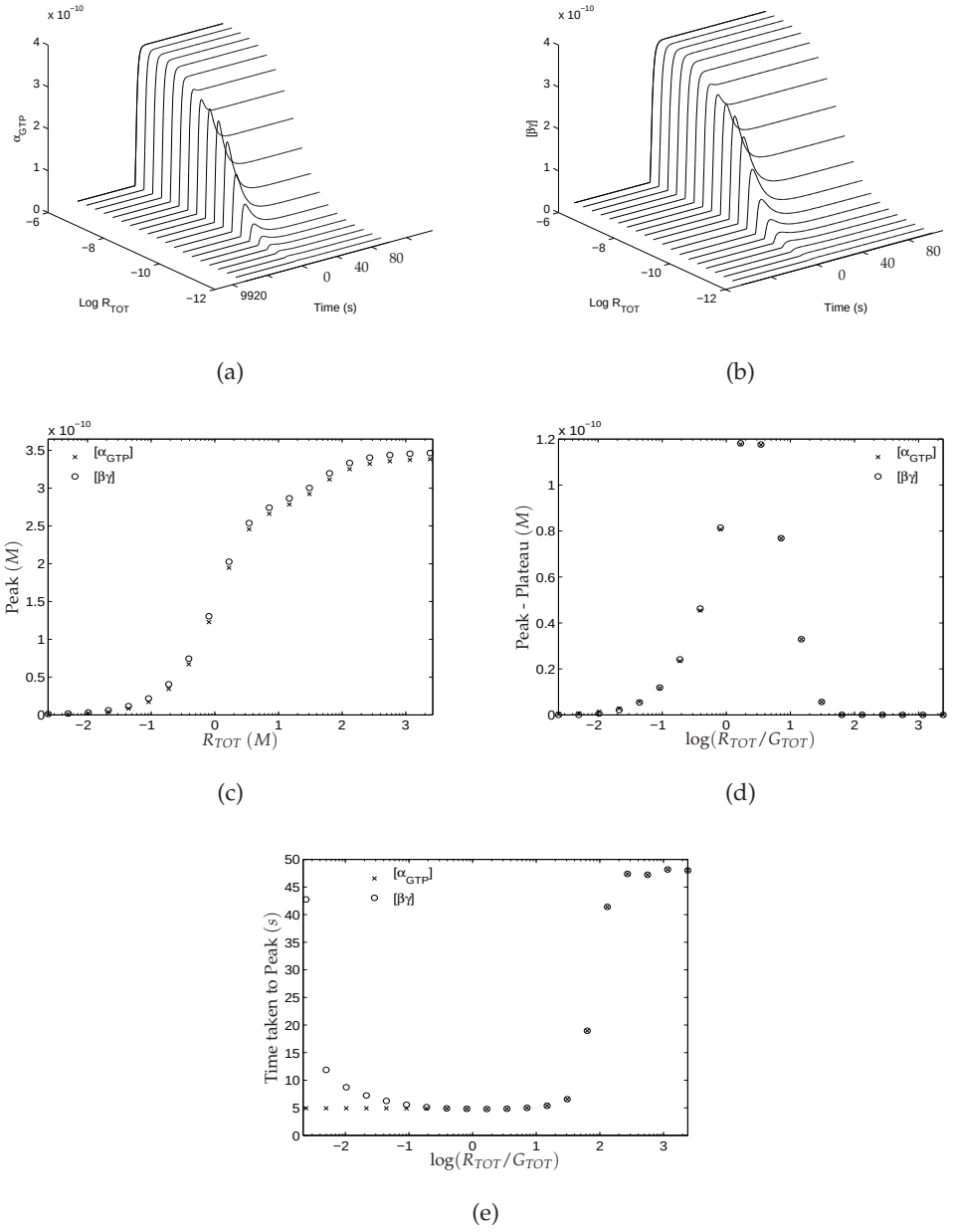


Figure 2.4: The effects of varying the total receptor concentration ($G_{TOT} = 4.15 \times 10^{-10}M$) on α_{GTP} and $\beta\gamma$ after the addition of an agonist. (a),(b) α_{GTP} and $\beta\gamma$ time courses; (c) α_{GTP} and $\beta\gamma$ peak values, which increase with R_{TOT}; (d) Peak$-$Plateau for α_{GTP} and $\beta\gamma$; (e) the times taken to reach α_{GTP} and $\beta\gamma$ peaks, which are very similar for the significant peaks.

by the relatively small G-protein concentration ($G_{TOT} = 4.15 \times 10^{-10} M$). The G-protein does undergo recycling when α_{GDP}, the product of α_{GTP} hydrolysis ($\alpha_{GTP} \xrightarrow{k_{gd+}} \alpha_{GDP}$), reassociates with $\beta\gamma$ to reform the whole G-protein ($\alpha_{GDP} + \beta\gamma \xrightarrow{k_{RA+}} G$), which can thus bind again to another free receptor; however, because it requires two extra reactions, this is a comparatively delayed recycling process. As the G-protein cannot bind as effectively to multiple receptors, as the receptor can to multiple G-proteins, G-protein overexpression is a more effective measure. The receptor concentration is more routinely manipulated in experimental studies [41], but these results suggest that more effort should also go into manipulating the G-protein concentration.

It has been shown that variations in the ratio of the total G-protein to receptor concentration can alter the transience of the peak response. The ratios were carried out either with G_{TOT} fixed at $4.15 \times 10^{-10} M$ or R_{TOT} fixed at $4.15 \times 10^{-10} M$. A more thorough analysis may be carried out by simulating with various other $G_{TOT} : R_{TOT}$ ratios within the defined ranges. This kind of analysis will be covered by a more global parameter analysis on an extension of the cTCAM in Chapter 4.

2.2.3 The influence of ligand properties

The effects of ligand related properties, such as affinity and efficacy (and concentration), on the dynamics of G-protein activation are now considered. Pharmacological GPCR experiments are often carried out by stimulating cells with various concentrations of various agonists and measuring the response at a selected point in the pathway. The ability of these ligands to increase the cellular response can then be observed. The levels of response tend to depend on where response is measured in the pathway. For instance, in the previous section it was seen that one active ligand receptor complex (LR^*G) can activate multiple G-protein cycles (see reaction (1.22)), so the α_{GTP} response is considerably higher than the LR^*G one, as illustrated in Figure 2.5. Such mechanisms, which take small signals and amplify their effect, occur along the pathway, and the phenomenon is referred to as *signal amplification*.

At the University of Nottingham's Institute of Cell Signalling, Dr Lauren May's experiments on the M3 Muscarinic receptor pathway measured Ca^{2+} levels induced by a range of concentrations of four different agonists; these data are presented as Ca^{2+} time courses (see Figures 2.2(b)), and corresponding dose response curves, constructed using the Ca^{2+} peak values as measures of the system's response (Figure 2.8(c)). As mentioned in Chapter 1, α_{GTP} cannot be measured in such an experiment. The cTCAM can be used to mimic these experiments, but with output at the α_{GTP} level. Although the cTCAM has previously been analysed in different ways [41, 86, 106], it does not appear that it has been used to model specific agonists. Here, each agonist will be

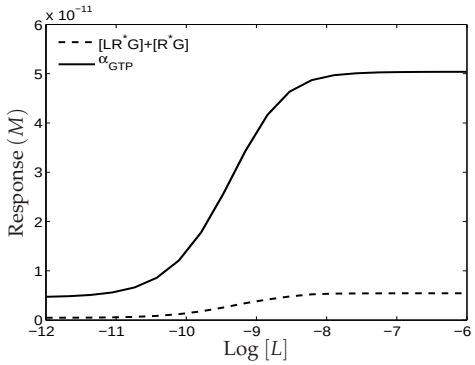


Figure 2.5: The α_{GTP} response is considerably greater than the $[LR^*G]$ response, due to signal amplification ($G_{TOT} = 4.15 \times 10^{-10}M$, $R_{TOT} = 4.15 \times 10^{-10}M$, $k_{lb+} = 8.4 \times 10^7 M^{-1}s^{-1}$, $k_{lb-} = 0.37s^{-1}$, $\zeta = \frac{\zeta_+}{\zeta_-} = 1000$, $[L] = 10^{-7}M$.

modelled by adjusting the parameters that pertain to ligand binding (affinity) and efficacy. Comparing the simulated α_{GTP} response with the Ca^{2+} data then allows the agonist's effects at both levels of the pathway to be observed. In addition, since the hypothesis of the previous sections is that the transience of the Ca^{2+} peak depends (in some measure) on that of the α_{GTP} peak, dose-response curves for the α_{GTP} peak values are compared to the Ca^{2+} dose-response curves from the experiments. In this section, only α_{GTP}, the mediator of Ca^{2+} release in this receptor system, is considered.

The initial values of $G_{TOT} = R_{TOT} = 4.15 \times 10^{-10}M$ will be used in all the simulations.

Modelling ligand-specific properties

The four agonists used in Dr May's experiments are Oxotremorine-M, Carbachol, Pilocarpine and Bethanachol. Binding studies have been carried out on these agonists which reveal that they have different affinities for the M3 muscarinic receptor, illustrated by the fact that their (experimentally estimated) binding and unbinding rate constants, k_{lb+} and k_{lb-} (which are shown in Table 2.2.1)) differ from each other [92]. Modelling affinity is then simply a matter of using these values in the simulations. (Note that the experimental estimation method made use of a simplified model, which excludes the active receptor state (R^*), hence the rate constants may not be completely fine-tuned for the more detailed cTCAM. However, at the very least, they are useful for comparing the different agonist affinities.)

Modelling efficacy is a less direct procedure. The parameter that controls ligand efficacy in the cTCAM is $\zeta = \frac{\zeta_+}{\zeta_-}$, so each of the four agonists can have a different ζ value. This is modelled by fixing $\zeta_+ = 1$ and assigning different ζ_- values to each agonist; ζ_-, and not ζ_+, is varied because experimental studies suggest that an agonist's

Agonist	$k_{lb+}(M^{-1}s^{-1})$	$k_{lb-}(s^{-1})$	ζ
Oxotremorine-M	$k_{lb+} = 2885.6$	$k_{lb-} = 0.14$	543.48
Carbachol	$k_{lb+} = 691.2$	$k_{lb-} = 0.14$	294.11
Pilocarpine	$k_{lb+} = 10222.2$	$k_{lb-} = 0.35$	29.94
Bethanachol	$k_{lb+} = 796.8$	$k_{lb-} = 0.26$	49.5

Table 2.2.1: Rate constants for the association (k_{lb+}) and dissociation (k_{lb-}) of some agonists, taken from the experimental studies of Sykes *et al.* [92], as well as estimates of their efficacy, ζ.

efficacy might be governed by how *slowly* it dissociates [92]. ζ values for each agonist can be estimated by converting experimentally estimated measures of relative efficacy, τ, from the aforementioned study [92]; the τ values were obtained by fitting to Black and Leff's [9] operational model of agonism, given by

$$\frac{[L](E_{max}\frac{\tau}{\tau+1})}{\frac{K_D}{\tau+1}+[L]}, \tag{2.13}$$

where E_{max} is the maximum possible response and K_D is the dissociation constant of ligand-receptor binding. It can be seen that, the higher its τ value ($\frac{\tau}{\tau+1} \to 1$, $\frac{K_D}{\tau+1} \to 0$), the closer an agonist comes to producing the maximum response. There is no direct way to make the conversions from τ to ζ, but since τ is a measure of relative efficacy, a reasonable approach is to obtain corresponding ζ values which preserve the same ratios as respective τ values. It is assumed that the agonist with the maximum efficacy is represented by $\zeta = 1000$, corresponding to the highest experimentally estimated τ value, $\tau_{max} = 42701$ and then for the i^{th} agonist:

$$\frac{\tau_{max}}{\tau^i} = \frac{1000}{\zeta^i}. \tag{2.14}$$

Then,

$$\zeta^i_- = \frac{\zeta_+}{\zeta^i} = \frac{1}{\zeta^i}. \tag{2.15}$$

The ζ values obtained for each agonist are given in Table 2.2.1. Oxotremorine-M is the most efficacious, with $\zeta = 543.48$, while Pilocarpine, with $\zeta = 29.94$, is the least.

Simulations

Figures 2.6 and 2.7 show that, for all the agonists, the simulated α_{GTP} and experimental Ca^{2+} time courses are similar, and the α_{GTP} and Ca^{2+} peaks remain transient as agonist concentration increases (although, as the previous section showed, receptor or G-protein overexpression can remove the transience). Another similarity is that the

simulated α_{GTP} time-to-peak and the experimental Ca^{2+} time-to-peak exhibit a similar dependence on the agonist concentration; Figures 2.8(c) and 2.8(d) show that both times-to-peak decrease as each agonist's concentration increases. It should be noted that the times taken to reach the α_{GTP} peak are generally longer than the times taken to reach the Ca^{2+} peak in the experiments (Figure 2.8), which would seem to contradict the transient peak hypothesis. However, the discrepancy might be explained by the fact that α_{GTP}'s interaction with PLC which occurs further down in the pathway is not included in the model. PLC considerably accelerates α_{GTP} hydrolysis (see section 1.1.1) [27], and so would cause the α_{GTP} peak to occur much earlier.

The main difference between the experimental Ca^{2+} peak and the simulated α_{GTP} one is that all agonists are able to produce the same maximal Ca^{2+} response (through signal amplification), while they produce different maximal α_{GTP} responses. In other words, at the Ca^{2+} level, all four drugs are full agonists, but not at the α_{GTP} level; this is especially indicated by the fact that the Ca^{2+} curves are left-shifted, with smaller concentrations of agonist required to produce the maximal response. For instance, each agonist, supplied at $10^{-5}M$, produces a small α_{GTP} response, but the same concentration produces the maximal Ca^{2+} response. This indicates that beyond the G-protein activation cycle, the pathway has other mechanisms which enable considerable signal amplification. These mechanisms will be explored later in Chapters 3 and 4 when the cTCAM is extended to include Ca^{2+} mobilisation.

Figure 2.9 shows that the maximal α_{GTP} response increases with ζ; it is less influenced by the ligand binding rate, k_{lb+}. For instance, even though Pilocarpine is by far the fastest binding of the agonists (Table 2.2.1), it produces the smallest maximal α_{GTP} peak, because its ζ value does not favour the active receptor state as strongly as the other agonists, making it the least efficacious agonist for the α_{GTP} response. This is one advantage of studying the G-protein model, the intrinsic pharmacological properties of the drugs are more evident, before the signal is completely amplified.

2.3 The influence of subunit dissociation

The cTCAM and other mathematical models of GPCR signalling assume the widely embraced hypothesis that the G-protein dissociates into α_{GTP} and $\beta\gamma$ on activation [37, 86, 106]. However, as discussed in section 1.1.1, another experimentally supported hypothesis is that the activated subunits do not need to dissociate in order to mediate downstream signalling [27, 79]; a variant of the cTCAM can be derived which reflects the latter hypothesis. It is interesting and important to compare the models, looking for parameter regimes under which they might or might not be interchangeable. In

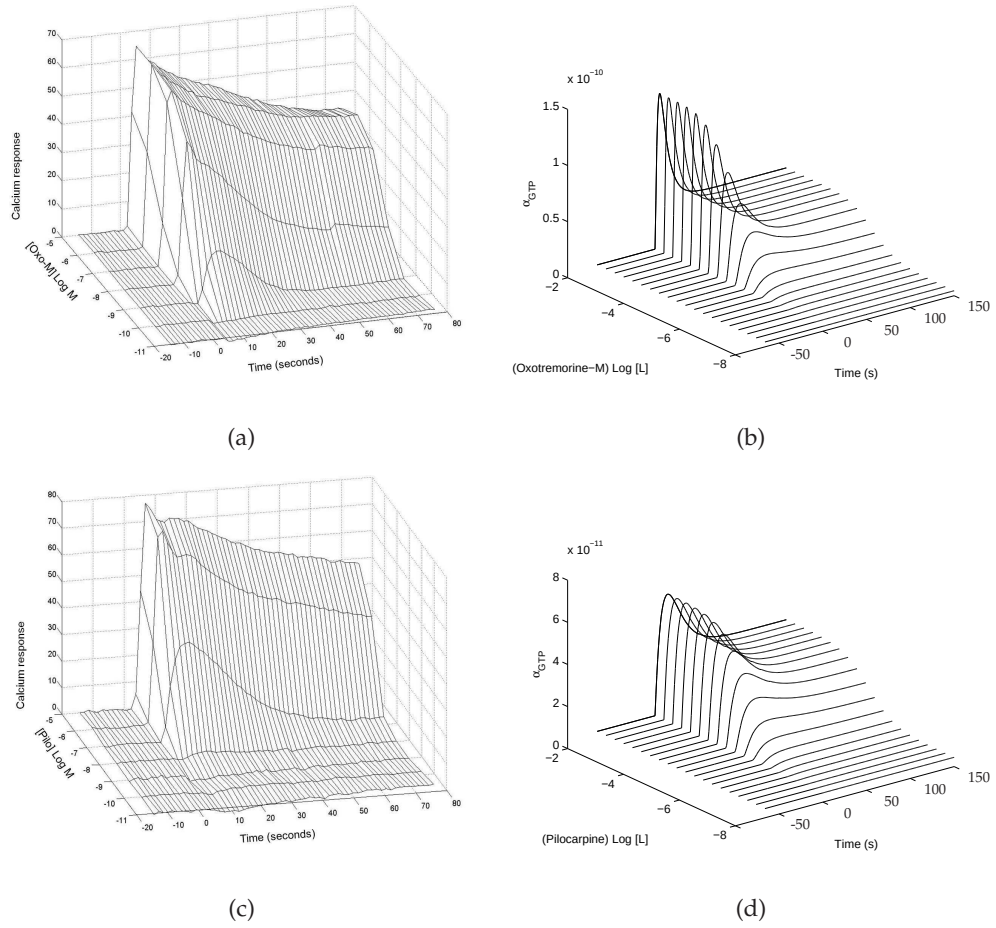


Figure 2.6: Ca^{2+} time courses from the experiments and simulated α_{GTP} time courses are qualitatively similar. (a) Ca^{2+} time courses for Oxotremorine-M; (b) Simulated α_{GTP} time courses for Oxotremorine-M ($k_{lb+} = 2885.6; k_{lb-} = 0.14; \zeta = 543.48$); (c) Ca^{2+} time courses for Pilocarpine; (d) Simulated α_{GTP} time courses for Pilocarpine ($k_{lb+} = 10222.2; k_{lb-} = 0.35; \zeta = 29.94$).

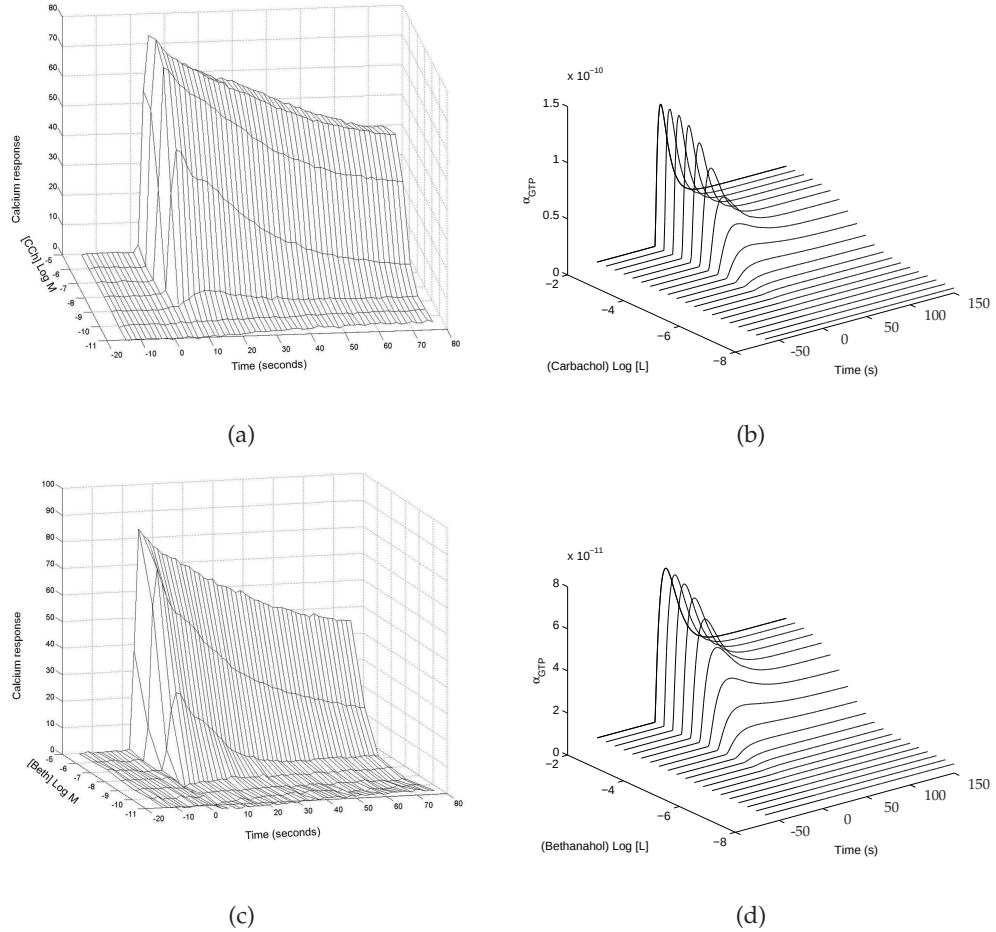
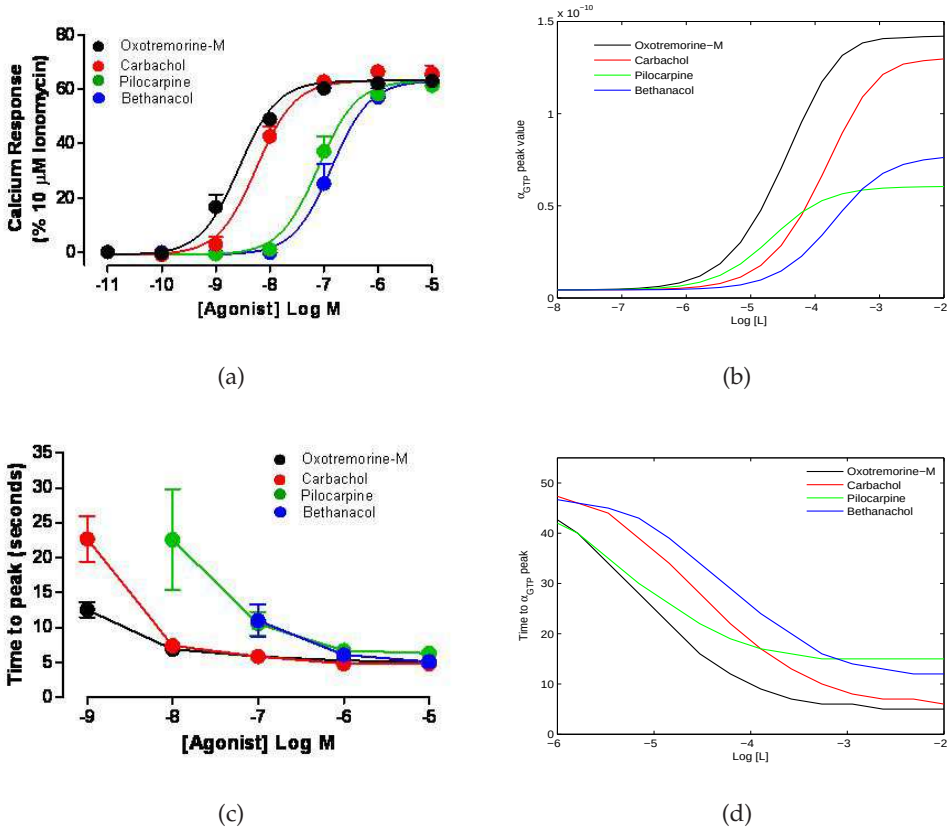


Figure 2.7: Ca^{2+} time courses from the experiments and simulated α_{GTP} time courses are qualitatively similar. (a) Ca^{2+} time courses for Carbachol; (b) Simulated α_{GTP} time courses for Carbachol ($k_{lb+} = 691.2, k_{lb-} = 0.14, \zeta = 294.11$;) (c) Ca^{2+} time courses for Bethanachol; (d) Simulated α_{GTP} time courses for Bethanachol ($k_{lb+} = 796.8, k_{lb-} = 0.26, \zeta = 49.5$).



(a) (b)

(c) (d)

Figure 2.8: Comparing agonist effects seen in the Ca^{2+} data [54] with effects on simulated α_{GTP} (predicted by the cTCAM) (a) All four drugs are full agonists for the Ca^{2+} peak response due to signal amplification; (b) partial agonism occurs for the α_{GTP} peak response; (c) Ca^{2+} times-to-peak decrease as the agonist concentrations increase ; (d) α_{GTP} times-to-peak do the same.

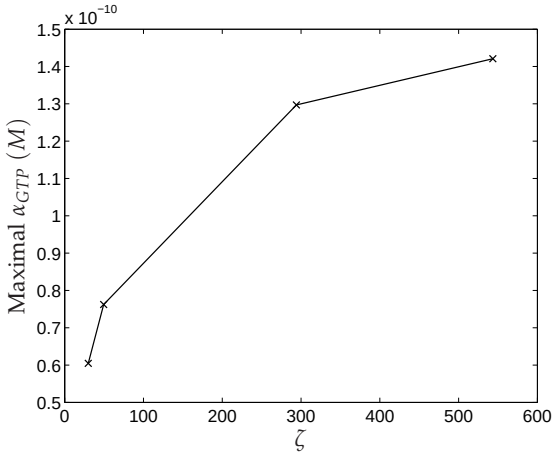


Figure 2.9: The maximal α_{GTP} peak increases with ζ.

essence, what is being explored is if one model of G-protein activation can be flexible enough to account for both hypotheses.

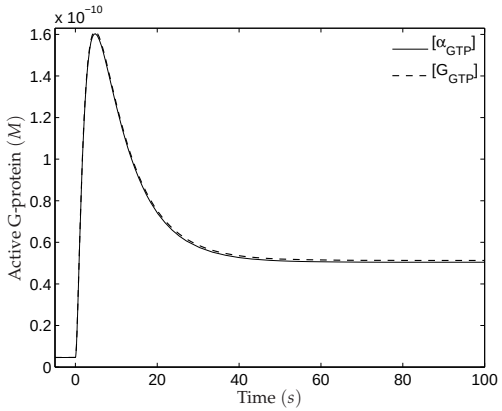
An adapted version of the cTCAM is derived in which the G-protein does not dissociate (called the *No-dissociation cTCAM*). Both models differ only in reactions pertaining to the G-protein activation and deactivation cycle, outlined and named as follows:

cTCAM	**No-dissociation cTCAM**	
$R^*G \xrightarrow{k_{GTP+}} R^* + \alpha_{GTP} + \beta\gamma$	$R^*G \xrightarrow{k_{GTP+}} R^* + G_{GTP}$	(Activation step)
$LR^*G \xrightarrow{k_{GTP+}} LR^* + \alpha_{GTP} + \beta\gamma$	$LR^*G \xrightarrow{k_{GTP+}} LR^* + G_{GTP}$	(Activation step)
$\alpha_{GTP} \xrightarrow{k_{gd+}} \alpha_{GDP}$	$G_{GTP} \xrightarrow{\psi k_{gd+}} G$	(Deactivation step)
$\alpha_{GDP} + \beta\gamma \xrightarrow{k_{RA+}} G$		(Re-association step)

where G_{GTP} is the activated, undissociated G-protein. ψ is set to 1, based on the assumption that the rate constant for the deactivation step is the same in the cTCAM and No-dissociation cTCAM, as being bound to the $\beta\gamma$ subunit is not expected to influence the α subunit's inherent GTPase ability.

The same agonist will be used to simulate both models ($k_{lb+} = 8.4 \times 10^7 M^{-1}s^{-1}$, $k_{lb-} = 0.37s^{-1}$, $\zeta = \frac{\zeta_-}{\zeta_+} = 1000$). The initial values are $R_{TOT} = G_{TOT} = 4.15 \times 10^{-10}M$.

Figure 2.10(a) shows that the α_{GTP} and G_{GTP} time courses are virtually indistinguishable, hence the cTCAM and the No-dissociation cTCAM make almost identical predictions of the activated G-protein response. This is due to the quick re-association of subunits, built into the cTCAM via a high value of k_{RA+}, which at $1.2 \times 10^{10} M^{-1}s^{-1}$, is the largest parameter in the model among those with units of $M^{-1}s^{-1}$. This is illustrated in Figure 2.11(a), where $[\alpha_{GTP}]$ decreases with the value of k_{RA+} in the range $1.2 \times 10^5 - 1.2 \times 10^{10} M^{-1}s^{-1}$. Between $k_{RA+} = 1.2 \times 10^8 - 1.2 \times 10^{10} M^{-1}s^{-1}$, $[\alpha_{GTP}]$



(a)

Figure 2.10: Comparison of the activated G-protein responses in the cTCAM and No-dissociation cTCAM simulated with $G_{TOT} = 4.15 \times 10^{-10}M$, $R_{TOT} = 4.15 \times 10^{-10}M$ and $k_{RA+} = 10^{10}M^{-1}s^{-1}$ in the cTCAM. The responses in both models are similar for this parameter set.

does not vary much, but below $k_{RA+} = 1.2 \times 10^{8}M^{-1}s^{-1}$, begins to noticeably decrease, to the point where the response is insignificant at $k_{RA+} = 1.2 \times 10^{5}M^{-1}s^{-1}$. The decrease in $[\alpha_{GTP}]$ can be explained as follows. As re-association slows down, the G-protein stays increasingly in the $\beta\gamma$ and inactive α_{GDP} forms, instead of re-forming the intact G-protein G; this is evidenced by the fact that $\beta\gamma$ and α_{GDP} increase as k_{RA+} decreases (see Figures 2.11(b) and 2.11(c)); as a result, fewer and fewer G-proteins (G) are available to re-associate with free receptors to form the active receptor-G-protein complexes ($[R^*G]$,$[LR^*G]$) from which α_{GTP} is formed; this is illustrated in Figure 2.11(d), where $[R^*G] + [LR^*G]$ decreases with k_{RA+}.

The cTCAM, under a parameter regime of fast subunit re-association, is thus equivalent to the No-dissociation cTCAM. Outside of that regime, the α_{GTP} response to an efficacious agonist is much suppressed or trivial (see Figure 2.11(a)), which does not seem likely. Therefore, it seems that if the subunits dissociate in cell membranes, they re-associate quickly enough to ensure that the active G-protein response is similar to a system in which the subunits do not dissociate. However, even though slow re-association is less likely, it can be easily accounted for by the cTCAM, but not by the No-dissociation cTCAM; so the cTCAM is the more flexible model.

The aim of this section was to determine which model is more suitable for modelling G-protein activation. However, the results also suggest that agonist stimulation experiments are worth designing, in which compounds that competitively and neutrally bind to either subunit are present; this would delay the rate of re-association, and according to the cTCAM, potentially reduce the cellular response. Conclusions

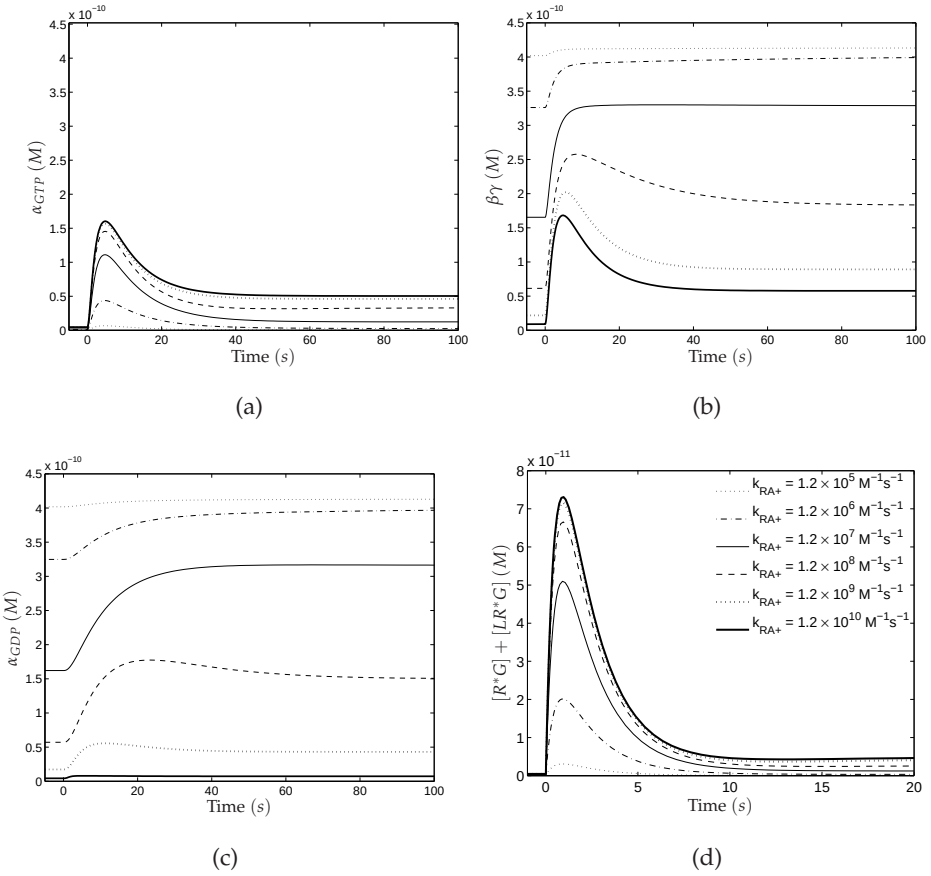


Figure 2.11: When the subunit re-association constant, k_{RA+}, is decreased, $[\alpha_{GTP}]$ and the active receptor concentration decrease ((a), (d)), while $[\alpha_{GDP}]$ and $[\beta\gamma]$ increase and accumulate ((c), (b)).

can then be drawn about the dependence of agonist potency on G-protein dissociation dynamics.

2.4 Discussion

As a first step in the construction of a comprehensive model of GPCR-mediated mobilisation of Ca^{2+} from intracellular stores, this chapter analysed models of G-protein activation alone.

The cTCAM was used to analyse the α_{GTP} and $\beta\gamma$ response to agonist stimulation. In virtually indistinguishable manners, $[\alpha_{GTP}]$ and $[\beta\gamma]$ both increased with the G-protein and receptor expression levels (the previous section suggests however, that the responses may not be so similar if subunit re-association is slow, see Figures 2.11(a)–(b)). It was also seen that the effect of agonist stimulation is more enhanced by G-protein overexpression than by receptor overexpression. This is because the receptor

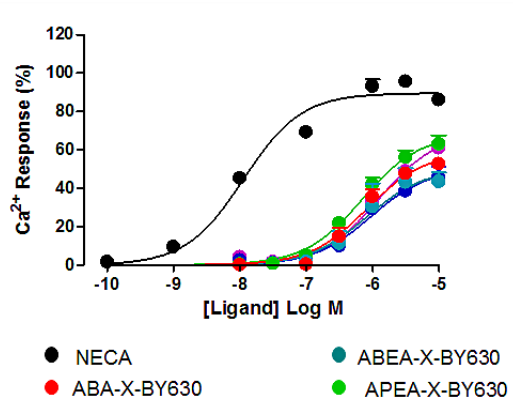


Figure 2.12: Partial agonism occurs for the Ca^{2+} response in the $\beta\gamma$-mediated A1 system [53], unlike the α_{GTP}-mediated M3 receptor system, where full agonism occurs.

has a more effective way of (sequentially) activating multiple G-proteins, and so more G-protein activation cycles are made possible as the total G-protein concentration increases. It thus seems that the manipulation of G-protein expression levels should be the focus of more experimental work.

The effects of ligand-specific properties such as the binding rate and efficacy were also explored, by comparing the responses induced by four different agonists. It was seen that partial agonism occurs for the G-protein-level response, in contrast to the full agonism which has been observed in experimental data obtained at the Ca^{2+} level, suggesting that the M3 pathway is *strongly* amplifying, which is not always the case with other Ca^{2+} mobilising pathways; for example, partial agonism occurs in the $\beta\gamma$-mediated Adenosine A1 receptor system for the Ca^2 response, as illustrated by the dose response curves in Figure 2.12, which resemble the previously simulated α_{GTP} curves (Figure 2.8(b)).

The hypothesis that the transience of the Ca^{2+} peak is due to a transient α_{GTP} peak was explored. The simulated results suggest that G-protein and receptor overexpression, by removing the transient α_{GTP} peak, may also remove the transient Ca^{2+} peak, replacing it with a sustained maximum response. The transient peak was more sensitive to changes in the total G-protein concentration than the total receptor concentration, suggesting that the same holds for the transient Ca^{2+} peak, if the hypothesis is true. It should be noted that receptor desensitisation of the M3 muscarinic receptor (which is not included in the cTCAM and would contribute to the transience of the peak) is known to occur; therefore the peak might be more robust to changes in the G-protein and receptor concentrations than the cTCAM suggests. The effects of receptor desensitisation will be explored when the model is extended in Chapter 3.The simulated α_{GTP} and experimental Ca^{2+} transient peaks also responded similarly to in-

creasing agonist concentrations, thus supporting the hypothesis, which will be further explored in Chapter 3.

It was also shown that the cTCAM, a model of G-protein subunit dissociation, is equivalent to one in which the subunits do not dissociate, under a parameter regime of fast re-association. Outside of that regime, the α_{GTP} response was either much repressed or trivial (see Figure 2.11(a)), which is unlikely to be the case in the cell membrane, as agonist efficacy would be greatly inhibited. This suggests that experiments in which subunit re-association is delayed or inhibited could be carried out to determine the dependence of agonist efficacy on subunit dynamics.

The cTCAM with fast subunit reassociation will be used to assemble the Ca^{2+} model, as either fast subunit re-association or no dissociation at all are what likely occurs in the experimental system. The cTCAM is also more versatile than the No-dissociation model, as it is able to account for the G-protein dynamics caused by slow re-association.

CHAPTER 3

GPCR-mediated Ca^{2+} Mobilisation Model

3.1 Introduction

As mentioned in Chapter 2, the activation of α_{GTP} in the M3 pathway leads to the mobilisation of Ca^{2+} from intracellular stores into the cytosol. Having carried out some analysis on the G-protein activation model (cTCAM), that model may now be extended to include details of the events leading to Ca^{2+} mobilisation.

Ca^{2+} is an essential and versatile second messenger with roles in numerous physiological processes (and consequently, pathologies), including muscle contraction, fertilisation, gene transcription and cell division [27]. In heart muscle cells for instance, activation of the GPCR's, Ang-2 receptor-Type 1 and Type 2, elevates cytosolic Ca^{2+}, which activates the phosphatase, calcineurin, [17]; this can cause massive cardiac enlargement and lead to heart failure [24]. In the M3 muscarinic receptor pathway, agonist-stimulated Ca^{2+} mobilisation influences function in bladder and intestinal smooth muscle by mediating contraction; in disease conditions, such as diabetes, smooth muscle is hypersensitive to Ca^{2+} [96]. The same pathway is also implicated in enhanced gene transcription via the activation of the transcription factors, the early growth factor, Egr-1, and activator protein, AP-1, both of which can mediate cell division [69]. Hence, understanding the mechanisms behind Ca^{2+} mobilisation has implications for developing or improving therapies for many disease conditions.

There are two major ways by which the level of cytosolic Ca^{2+} is raised in non-excitable cells, either via a GPCR-activated release from intracellular stores or an influx from the extracellular environment across plasma membrane channels, activated when the stores near depletion [27]. Organelles that can act as intracellular stores include the endoplasmic reticulum (ER), the sarcoplasmic reticulum (SR), and the mitochon-

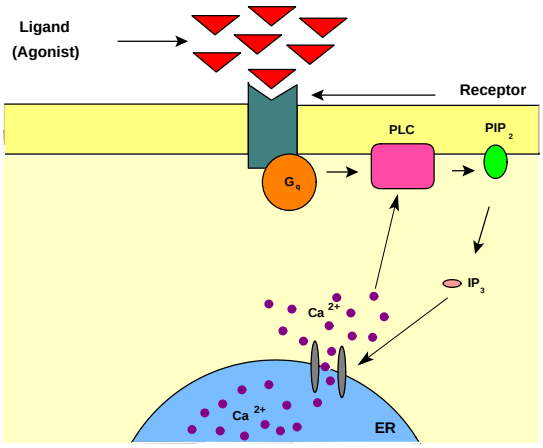


Figure 3.1: GPCR activation leads to the production of IP_3, which mediates Ca^{2+} release.

dria. Agonist-stimulated Ca^{2+} release from the ER is a central process [27, 82] and so, for simplicity, will be the focus of this work. There are mechanisms for removing excess Ca^{2+} from the cytosol, which could otherwise cause cell death or damage [94]. Excess Ca^{2+} can be transported back into the ER by the sarcoplasmic/endoplasmic reticulum calcium ATPase (SERCA) pumps (located in the organelle membrane), or removed from the cytosol into the extracellular environment by pumps located in the plasma membrane, called plasma membrane Ca^{2+} ATPases (PMCA) [27]. Another way in which the cytosolic Ca^{2+} concentration is regulated is via binding proteins, called buffers, which thus decrease the concentration of free Ca^{2+} and inhibit the activation of Ca^{2+}-mediated processes [10].

In the M3 (G_q–coupled) pathway, α_{GTP} binds to the enzyme Phospholipase C (PLC), which is located in the plasma membrane, to form $\alpha_{GTP}PLC$, which is then activated by Ca^{2+}. The signal generated by $\alpha_{GTP}PLC$ is terminated by the α subunit's intrinsic GTPase activity, which is accelerated by PLC, resulting in the dissociation of the complex into α_{GDP} and PLC. There are three known PLC isoforms: $PLC\gamma$, $PLC\delta$ and $PLC\beta$, the last of which is activated by the G_q protein. The $\alpha_{GTP}PLC$ complex hydrolyses its substrate, phosphatidylinositol bisphosphate (PIP_2), to produce inositol triphosphate (IP_3) and diacylglycerol (DAG). IP_3 diffuses into the cytosol, eventually inducing Ca^{2+} release from intracellular stores, and is subject to a multi-step degradation into an inositol phosphate pool from which PIP_2 is re-synthesised [74]. DAG remains at the membrane and causes the activation of Protein kinase C (PKC), which activates a pathway that can involve the phosphorylation and consequent desensitisation of the GPCR, and can extend to the nucleus of the cell [55]. Other kinases might also mediate desensitisation [105]. For simplicity, only the IP_3 pathway (see Figure 3.1) will be explicitly modelled, as it involves the main processes that lead to Ca^{2+} release.

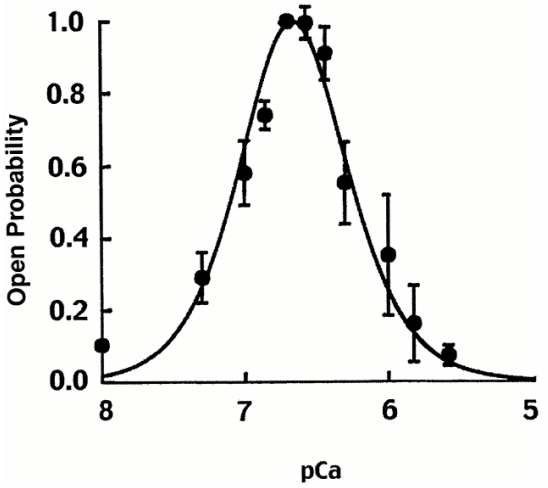


Figure 3.2: Ca^{2+} dependence of the IP_3R open probability. Reproduced from [62]. Initially the open probability increases with $[Ca^{2+}]$ (pCa means $|\log[Ca^{2+}]|$), but past a threshold value, begins to decrease.

IP_3 diffuses through the cytosol to bind to receptors (IP_3R) on the membrane of the ER, which are ion channels that open on activation to release Ca^{2+} into the cytosol. The IP_3R consists of four subunits, each possessing binding sites for IP_3 and Ca^{2+}, both of which co-operatively regulate the channels. There are at least three IP_3R subtypes, designated type 1, type 2, type 3 – which differ in aspects of their regulation, and more than one type may be expressed in the same cell [27]. All subtypes are believed to be regulated in a biphasic manner by IP_3-induced Ca^{2+} [91]. The accepted theory is that IP_3R possesses three binding sites, one for IP_3 and two for Ca^{2+} – one stimulatory, the other inhibitory. On binding, IP_3 increases the Ca^{2+} sensitivity of the IP_3R's, causing them to bind to Ca^{2+} and open; this leads to an increase in cytosolic Ca^{2+}, which when high enough, causes the channels to begin to close again. In other words, Ca^{2+} first exerts positive feedback on its own mobilisation, and then negative feedback. This phenomenon is normally referred to as calcium-induced calcium release (CICR) [27, 82], illustrated in Figure 3.2.

3.2 Experimental studies of Ca^{2+} mobilisation

Agonist-stimulated Ca^{2+}-mobilisation experiments are often carried out in either human or surrogate cells, which can be engineered to express a variety of Ca^{2+}-activating GPCR's. Ca^{2+} is easily measured when the cells are loaded with high affinity indicators that emit fluorescence when bound to Ca^{2+}. These experiments can be carried out in the presence or absence of extracellular Ca^{2+}; the data that will be discussed here are from the latter experimental setting, allowing for a simpler model which does not need

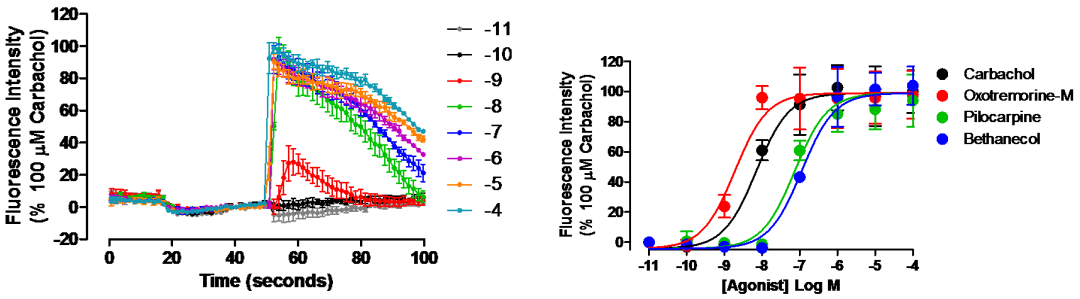


Figure 3.3: Data from Ca^{2+}-mobilisation (in the absence of extracellular Ca^{2+}) experiments performed at the University of Nottingham's Institute of Cell Signalling (ICS) [52]. Each point is the mean $\pm$ standard error (S.E.M) of the cell population, and is normalised to the fluorescence due to stimulation by $10^{-4}M$ of Carbachol (a) Time courses for Ca^{2+}-mobilisation induced by various concentrations of the agonist Oxotremorine-M. (b) Dose response curves of various agonists for the Ca^{2+} peak response.

to include the influx of extracellular Ca^{2+}.

Figure 3.3 shows data from the Ca^{2+} mobilisation experiments carried out at the Institute of Cell Signalling [54], in which Chinese hamster ovary cells (CHO) are transfected with the human muscarinic $M3$ receptor and loaded with a fluorescent indicator that has a high affinity for Ca^{2+}; the system is then left to settle down for 30 minutes, and basal fluorescence is measured. Agonist treatment increases the fluorescence, indicating the release of Ca^{2+} ions into the cytosol, to which the indicators rapidly bind. At several time points, fluorescence is measured, and data for the average of the cell population indicates that a transient Ca^{2+} peak occurring in $5 - 20s$ is the response induced by virtually all concentrations of the agonist Oxotremorine-M (Figure 3.3(a)), and Ca^{2+} tends to return to its basal concentration. Dose-response curves for Oxotremorine-M and other agonists, Carbachol, Pilocarpine and Bethanecol (modelled in Chapter2) indicate that they all act as full agonists, producing the same maximal Ca^{2+} response, although at different concentrations (Figure 3.3(b)). The time courses for these agonists are also similar (results not shown).

Studies have indicated that many fluorescent indicators, in their Ca^{2+}-free forms, are competitive antagonists for IP_3 binding to IP_3R, causing a concentration-dependent inhibition of IP_3 binding. Antagonism is only significant at unusually high concentrations of the fluorescence indicators [64, 77], indicating that their binding dynamics do not need to be included in the model.

3.3 Existing mathematical models of Ca^{2+} dynamics

There are many models of Ca^{2+} signalling which have shown good agreement with available experimental results. Many of the existing models were developed to investigate Ca^{2+} oscillations, and it should be noted that the timecourses in Figure 3.3(a) are non-oscillatory, so oscillations are not the focus here. However, in other systems, oscillations have been experimentally observed [94].

Early models of Ca^{2+} dynamics did not include explicit details of IP_3R's mechanisms, because they had not yet emerged from experimental studies. The early models were based on the knowledge that Ca^{2+} oscillations rely on the phenomenon of CICR, and generally focused on the cytosolic aspects.

The most influential of these early phenomenological models was by Goldbeter and Dupont [26]; it assumes the existence of two intracellular Ca^{2+} stores, one sensitive to IP_3, the other sensitive to Ca^{2+}. IP_3 causes the release of Ca^{2+} from the IP_3-sensitive store; this increase in cytosolic Ca^{2+} then triggers release from the Ca^{2+}-sensitive store. Influx and efflux of Ca^{2+} across membranes is included in this model as well. Influx is modelled as constant while efflux depends linearly on cytosolic Ca^{2+} [82]. A later version of this model allowed for the existence of just one Ca^{2+} store. Both versions predicted Ca^{2+} oscillations similar to those observed experimentally.

3.3.1 Models of the IP_3 receptor

Later experimental studies uncovered the role of the IP_3 receptor (IP_3R) in CICR (that it is activated by low Ca^{2+} concentrations and subsequently inactivated by high Ca^{2+} concentrations); hence, the relationship between the open probability of the channel and $[Ca^{2+}]$ is bell-shaped (see Figure 3.2). Models subsequently incorporated mechanistic details of IP_3R interactions with IP_3 and Ca^{2+}.

The De Young-Keizer [20] model was the earliest model to include details of IP_3R kinetics. It assumes that there are three IP_3R subunits, each with three binding sites for IP_3, activating Ca^{2+}, and inhibiting Ca^{2+}, which are allowed to bind in any order. Hence there are eight possible receptor states. The notation S_{ijk}, $i, j, k = 0$ or 1 may be used to describe fractions of IP_3R in particular states, with 1 corresponding to an occupied site, and 0, to an unoccupied one; so the possible fractional states are S_{000}, S_{010}, S_{001}, S_{011}, S_{100}, S_{110}, S_{101} and S_{111}. The receptor is activated when in the state S_{110}. It is assumed that the channels open only when three subunits are activated, hence S_{110}^3 is the open probability. The model is a system of eight ODEs, seven representative of the receptor states and derived from mass action kinetics; its Ca^{2+} ODE has a source term proportional both to the Ca^{2+} gradient between the cytosol and ER and to the

open probability of the channels, and its sink term is a Hill function modelling SERCA flux, given by:

$$\frac{d[Ca^{2+}]}{dt} = v_r(v_1 S_{110}^3 + v_2)([Ca_{ER}^{2+}] - [Ca^{2+}]) - \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2}, \tag{3.1}$$

where Ca_{ER}^{2+} is the concentration of Ca^{2+} in the ER, v_r is the ER to cytosol volume ratio, and v_1 and v_2 set the maximal Ca^{2+} fluxes. The factor v_r accounts for the fact that the flux is between the cytoplasm and ER, which have different volumes. Hence, fluxes between these compartments do not cause the same concentration change. $\frac{d[Ca_{ER}^{2+}]}{dt}$ is then given by $(v_1 S_{110}^3 + v_2)([Ca_{ER}^{2+}] - [Ca^{2+}]) - \frac{p_1[Ca^{2+}]^2}{v_r[Ca^{2+}]^2 + p_2^2}$. Note that p_1 is a composite term including v_r. The Hill function represents more complex mechanisms, in which, after Ca^{2+} binds to the pump, it undergoes several transformations, including conformational changes which send Ca^{2+} back into the ER [94]. The Hill coefficient is 2, representing two binding sites for Ca^{2+}. The model was fit to steady state data to obtain the rate constants, and was able to reproduce Ca^{2+} oscillations via a Hopf bifurcation from a stable steady state [95]. No extracellular Ca^{2+} fluxes were modelled, so $[Ca_{ER}^{2+}]$ could be eliminated, [45].

The Li-Rinzel model is a simplification of the De Young-Keizer model, derived by making assumptions of rapid kinetics for the binding of IP_3 and stimulatory Ca^{2+}, resulting in a two-dimensional system. This simplified model is able to reproduce most of the qualitative predictions of the eight-state Keizer-De Young model [45].

Another simpler alternative to the Keizer-De Young model is the model by Othmer and Tang [70], derived using the assumption that IP_3, activating Ca^{2+} and inhibitory Ca^{2+} bind to the IP_3R sequentially (an experimentally supported theory [91]) so that the only possible IP_3R states are S_{000}, S_{100}, S_{110}, S_{111}. As in the Keizer-De Young model, the active state is S_{110}, but it is assumed that only one subunit needs to be active for the channel to open. The Ca^{2+} rate is given by:

$$\frac{d[Ca^{2+}]}{dt} = (1 + v_r)(\gamma_0 + \gamma_1^{'} S_{110})(Ca_{AV}^{2+} - [Ca^{2+}]) - \frac{p_1[Ca^{2+}]^4}{[Ca^{2+}]^4 + p_2^4}, \tag{3.2}$$

where γ_0 quantifies a basal leak, and $\gamma_1^{'}$, the maximal Ca^{2+} flux. Ca_{AV}^{2+} is the volume average concentration of the cell, the level cytosolic Ca^{2+} approaches if the ER network is ruptured [36]; it is given by $\frac{[Ca^{2+}] + v_r[Ca_{ER}^{2+}]}{1 + v_r}$, and is assumed to be constant due to no flux across the plasma membrane. Like the previous models, this one successfully reproduces Ca^{2+} oscillations [70].

The above models do not include the details of ligand-receptor-G-protein dynamics but merely utilise $[IP_3]$ as the input parameter. The literature also includes more comprehensive models of GPCR-mediated Ca^{2+} mobilisation, which are discussed in

the next section.

3.3.2 Comprehensive models of Ca^{2+} mobilisation

There are at least four relevant comprehensive Ca^{2+}-mobilisation models in the literature, which incorporate details of signalling all the way from GPCR activation. Two of these models cover events from agonist stimulation through to Ca^{2+} mobilisation [37, 44], while the others go beyond to cover downstream processes: Ca^{2+}-mediated initiation of smooth muscle contraction via the phosphorylation of myosin II [48] and the secretion of luteinizing hormone (LH) [6, 30].

Lemon *et al.* presented an ODE model of GPCR-mediated IP_3 formation and consequent Ca^{2+} mobilisation [44] in the PTY2 receptor system. This model took into account the desensitisation, and subsequent recycling of the receptor, as well as PIP_2 depletion and replenishment; assumptions of rapid kinetics were used to significantly simplify the model equations. A rapid transient peak ($\approx 25s$) was predicted for α_{GTP}, IP_3 and Ca^{2+} over different rates of PIP_2 replenishment. It was predicted that the desensitisation of the ligand-receptor complex causes a decrease in α_{GTP} on a time scale of minutes.

Using the Virtual cell simulation and analysis suite software, Lukas [48] modelled a Ca^{2+} signalling process of agonist binding through to Ca^{2+}-mediated protein phosphorylation of the myosin light chain (MLC), which initiates contraction in smooth muscle. G-protein activation was modelled similarly to the cTCAM (see Chapters 1 and 2) except that constitutively active receptors (R^*) were not included. Receptor desensitisation is also included as an irreversible phosphorylation reaction by a G-protein receptor kinase (GRK), hence there are five possible receptor states compared to eight in Chapter 2's G-protein activation model [48]. The model was simulated both with generic parameters, and with parameters specific to the Bradykinin receptor. In both cases, the Ca^{2+} response was a rapid, transient peak (see Figure 3.4 for the. When G-protein-receptor pre-coupling was removed in the generic case, the peak was slower. The model also predicted that the timing and magnitude of the downstream phosphorylation event is dependent on the agonist concentration.

Washington *et al.* [102] presented a model of GPCR-mediated Ca^{2+} mobilisation leading to the secretion of luteinizing hormone (LH) (which triggers ovulation) in the Gonadotropin-releasing hormone receptor (GnRH) system. In this model, the receptors could bind to each other to form *dimers*, as has been experimentally observed in that receptor system [30]. In addition to sustained hormone stimulation, the model also considered the effects of stimulation applied in brief pulses, to mirror the way that GnRH is secreted physiologically. Transient Ca^{2+} peaks were predicted, which could

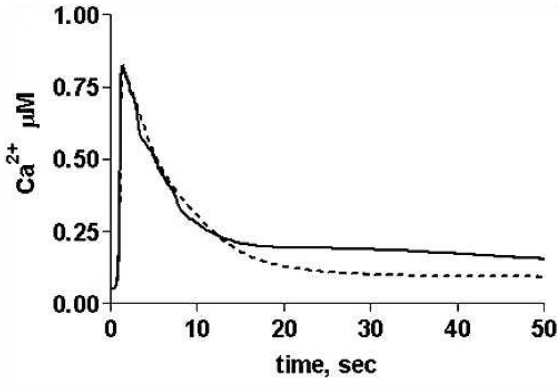


Figure 3.4: GPCR-mediated Ca^{2+} models in the literature predict the experimentally observed transient Ca^{2+} peak. Shown are transients predicted by the Lukas model for the bradykin receptor system [48]. The dotted line is the response to 3×10^{-6} M of agonist, while the solid one is for 30×10^{-6} M.

re-occur at each point of stimulation, provided the dose and frequency of stimulation were low. Armstrong *et al.* [6] used the same model to show significant reduction of the Ca^{2+} response at high GnRH receptor concentration, with the peaks much lower after the initial stimulation.

Kang and Othmer [37] modelled agonist-induced Ca^{2+} mobilisation in the Glutamate receptor (GPCR) system, incorporating details of the pathways activated not only by IP_3, but also by DAG. In this model, the DAG pathway activates PKC, which phosphorylates PLC and the active ligand-receptor complex (LR^*), causing temporary desensitisation. The receptor can only become activated when bound by the ligand, so the R^* state is not included in the model. All reaction rates are governed by mass action, and the $[Ca^{2+}]$ ODE is based on the Othmer and Tang model of sequential binding [70]:

$$\frac{d[Ca^{2+}]}{dt} = (1 + v_r)(\gamma_0 + \gamma_1[IP_3IP_3RCa^{2+}])(Ca^{2+}_{AV} - [Ca^{2+}]) - \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2}, \quad (3.3)$$

The full model is thus a system of 25 coupled ODEs, which was reduced to 19, using equations of conservation.

The influence of PKC on the nature of the Ca^{2+} response was explored. The model was able to reproduce oscillations and transient peak responses. A bifurcation analysis showed that the transient peak occurs for high IP_3 production rates and zero or low PKC regulation, while oscillations occurred for intermediate IP_3 production rates and zero or low PKC regulation. Hence, oscillations are possible without the inclusion of PKC dynamics, which is in accordance with the previously mentioned IP_3R models. The range of IP_3 production rates for which oscillations occurred were widened by

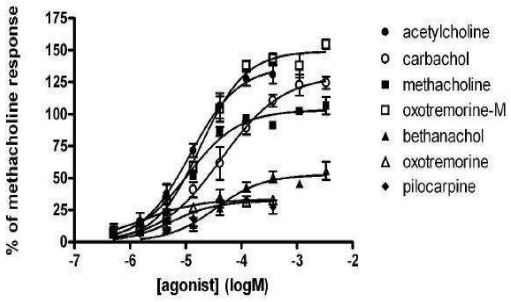


Figure 3.5: Dose-response curves from the experiments of Sykes *et al.* measuring $\alpha_{GTP\gamma S}$ in the M_3 Muscarinic receptor system [93] reveal that partial agonism occurs for the G-protein response. The system thus has amplification machinery through which partial agonists at that level of the pathway are full agonists at the Ca^{2+} level (compare Figure 3.3(b)).

increasing PKC activity. Hence, the model predicts that PKC, in conjunction with PLC, can act as a switch between different Ca^{2+} response patterns, because of the ability to modulate $[IP_3]$, which is the main determinant of response patterns.

The Kang and Othmer model is comprehensive, like the Ca^{2+} model to be constructed in this chapter, and models IP_3R dynamics based on an experimentally supported (and model-simplifying) theory of sequential binding. Hence, out of existing Ca^{2+} models in the literature, Chapter 2's G-protein activation model (cTCAM) is extended using downstream details from the Kang and Othmer model.

3.4 Signal amplification

Chapter 2's G-protein activation model predicted that partial agonism occurs for the α_{GTP} response, a result which is apparently supported by the experimental results in Figure 3.5, which show partial agonism for $\alpha_{GTP\gamma S}$, which is a non-hydrolysable analogue of α_{GTP}. The experimental study included the four agonists studied in Chapter 2, and three others, including the receptor's endogenous ligand, Acetylcholine [93].

A typical pharmacological measure of agonist efficacy is the EC_{50}, which is the concentration of agonist that induces the half-maximal response [40]. A quantitative measure of signal amplification is found by comparing the EC_{50} values (or variants of it) of the $\alpha_{GTP\gamma S}$ and Ca^{2+} responses. It is convenient to use the absolute value of the logarithm of the EC_{50}, referred to as pEC_{50} in pharmacology, and denoted here by $|\log EC_{50}|$. The higher the value of an agonist's $|\log EC_{50}|$, the more potent it is; for instance, $|\log EC_{50}| = 9$ means that only $10^{-9}M$ of an agonist is required to induce a half-maximal response. The values that pertain to $\alpha_{GTP\gamma S}$ and Ca^{2+} are shown in Table

Agonist	$\lvert \log EC_{50} \rvert$ of response	
	$\alpha_{GTP\gamma S}$	Ca^{2+}
Oxotremorine-M	5.59	8.76
Carbachol	4.64	7.98
Pilocarpine	5.05	7.72
Bethanechol	4.62	6.77

Table 3.4.1: $|\log EC_{50}|$ values of the $\alpha_{GTP\gamma S}$ response and that of Ca^{2+} [93]. The values that pertain to Ca^{2+} and the corresponding ones for $\alpha_{GTP\gamma S}$ always differ by at least 2.15, indicating that signal amplification has occurred between these stages in the pathway.

3.4.1 [93], and it can be seen that, for each agonist, they differ by at least 2.15 (thus the Ca^{2+} curves are more left shifted). This shows that the pathway has mechanisms for amplifying the signal between the levels of α_{GTP} production and Ca^{2+} release, as was concluded in Chapter 2.

None of the aforementioned comprehensive models attempt to model the signal amplification that occurs between the α_{GTP} and Ca^{2+} levels of the pathway. A major aim of the Ca^{2+}-mobilisation model of this chapter will be to reproduce that pathway feature, which is so evident in the data. This continues the pharmacologically flavoured modelling from Chapter 2.

3.5 Extension of the G-protein activation model

Here, the full Ca^{2+} mobilisation model is assembled by extending Chapter 2's G-protein activation model to include the downstream processes which culminate in Ca^{2+} mobilisation. As previously mentioned, this is accomplished using relevant components of the Kang and Othmer model.

The output of the cTCAM (see Chapter 2) is α_{GTP}, and its role in activating the cascade of processes that lead to Ca^{2+} release has been explained in section 3.1. These processes are represented by the following reactions, after Kang and Othmer [37]:

$$\alpha_{GTP} + PLC \underset{k_{PLCb-}}{\overset{k_{PLCb+}}{\rightleftharpoons}} \alpha_{GTP}PLC \tag{3.4}$$

$$Ca^{2+} + \alpha_{GTP}PLC \underset{k_{PLCdeact}}{\overset{k_{PLCact}}{\rightleftharpoons}} \alpha_{GTP}PLC^{*} \tag{3.5}$$

$$\alpha_{GTP}PLC^{*} + PIP2 \underset{k_{PIP2b-}}{\overset{k_{PIP2b+}}{\rightleftharpoons}} \alpha_{GTP}PLC^{*}PIP2 \tag{3.6}$$

$$\alpha_{GTP}PLC^{*}PIP2 \xrightarrow{k_{IP_3}} \alpha_{GTP}PLC^{*} + IP_3 + DAG \tag{3.7}$$

$$\alpha_{GTP}PLC^* \xrightarrow{k_{PLCdis}} \alpha_{GDP} + PLC + Ca^{2+} \tag{3.8}$$

$$IP_3 \xrightarrow{k_{IP_3deg}} 0 \tag{3.9}$$

$$IP_3 + IP_3R \underset{k_{IP_3Rb-}}{\overset{k_{IP_3Rb+}}{\rightleftharpoons}} IP_3IP_3R \tag{3.10}$$

$$IP_3IP_3R + Ca^{2+} \underset{k_{Ca^{2+}act-}}{\overset{k_{Ca^{2+}act+}}{\rightleftharpoons}} IP_3IP_3RCa^{2+} \tag{3.11}$$

$$IP_3IP_3RCa^{2+} + Ca^{2+} \underset{k_{Ca^{2+}inh-}}{\overset{k_{Ca^{2+}inh+}}{\rightleftharpoons}} IP_3IP_3RCa^{2+}Ca^{2+} \tag{3.12}$$

$$\frac{d[Ca^{2+}]}{dt} = (1+\nu_r)(\gamma_0 + \gamma_1[IP_3IP_3RCa^{2+}])(Ca^{2+}_{AV} - [Ca^{2+}]) - \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2}. \tag{3.13}$$

Reactions (3.4) - (3.12) represent the processes from α_{GTP}-PLC binding to IP_3 production and consequent binding to its receptor (IP_3R), while equation (3.13) describes the rate of change of cytosolic Ca^{2+}, which is not modelled by mass action, but is based on the Tang and Othmer model described in the previous section. Ca^{2+}'s rate of release from the ER is proportional to the concentration gradient between the cytosol and ER, and the concentration of activated IP_3R's ($[IP_3IP_3RCa^{2+}]$). ν_r is the ER/cytosol ratio, γ_0 represents the basal Ca^{2+} permeability of the ER, while γ_1 is a measure of the sensitivity of IP_3R to IP_3. The rate of Ca^{2+} efflux into the ER is modelled by a Hill function with exponent two, in which p_1 represents the maximal rate of the SERCA pumps, and p_2 is the Michaelis constant.

As in Chapter 2, the law of mass action is used to convert the reactions into ODEs, which when coupled to the G-protein activation model constitute the Ca^{2+} mobilisation model, a system of 22 ODEs, which can be found in Appendix B.3; the whole pathway is illustrated in Figure 3.6. The G-protein activation model of Chapter 2 was a model of 10 governing ODEs, so this represents a considerable increase in model complexity. Complex models can sometimes be simplified based on timescale considerations (using asymptotic analysis) [106], but this requires confidence in the accuracy of the model parameters, which is rarely the case with biological models as many parameters cannot be measured experimentally, while some vary from cell to cell [41], hence such model simplification is not attempted here.

The data being used to guide this modelling effort pertain to experiments carried out in the absence of extracellular Ca^{2+}; as a result, there is no need to model the membrane channels which would normally facilitate influx to the cytosol. As previously mentioned, efflux occurs via membrane pumps, but initially this is not modelled, as cytosolic Ca^{2+} removal is already modelled via the SERCA pumps which may only

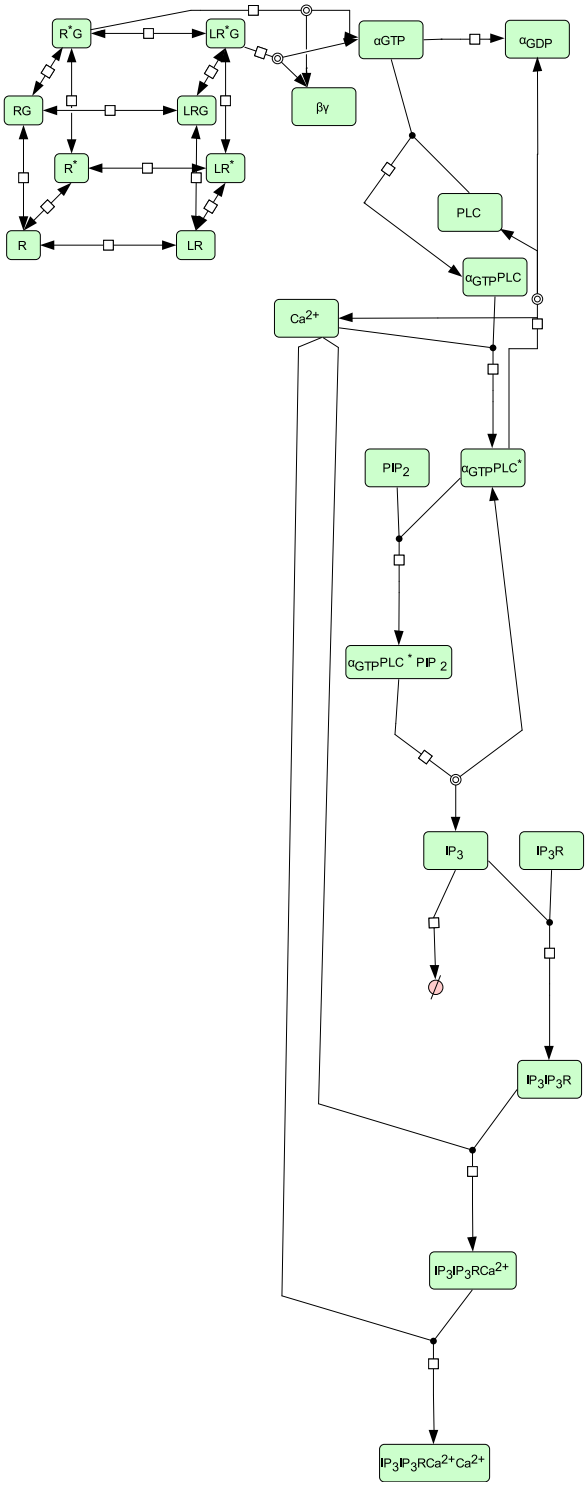


Figure 3.6: Schematic representation of the Ca^{2+} model, which is derived by extending the cTCAM [86] (see Chapter 2) using downstream details from the Kang and Othmer model [37].

require careful parameter choices to account for the action of the membrane pumps. In addition, Ca^{2+} buffers are not modelled, for the sake of simplicity; however the model can easily be extended by including the relevant binding reactions to buffers.

3.5.1 Parameter values

The parameter values used in Chapter 2's G-protein activation model are retained, while the remaining relevant parameter values are taken (or modified) from the Kang-Othmer model [37]. The resultant set of parameter values is referred to as the *preliminary parameter set*; it can be found in Table B.2.2, and will subsequently be subjected to some data-motivated analysis. Stimulation will be by $10^{-7}M$ of Oxotremorine-M, unless otherwise specified.

3.6 Data-driven modelling aims

The aim of this model is to predict the key features of the experimentally observed Ca^{2+} response to varying agonist concentrations, and so to understand the pathway processes which drive these features. Figure 3.3(a) shows that, in response to the agonist Oxotremorine-M (added after 50s) Ca^{2+} rapidly peaks (5-20s, depending on agonist concentration), and then more slowly returns to a plateau level; hence Ca^{2+} release is rapid *and* transient (three other agonists similarly tested in these experiments produced qualitatively similar time courses, which are not shown here). Figure 3.3(b) also shows that all agonists produce the same maximal Ca^{2+} response, with high EC_{50} values, indicating that signal amplification (previously discussed in section 3.4) occurs. Hence, there are three key features of the Ca^{2+} response that the model needs to reflect, and which are summarised below:

1. Rapid time-to-peak
2. Transient peak
3. Signal amplification

3.7 Data-motivated parameter variation

The Ca^{2+} model is initially simulated with the preliminary parameter set, and subsequent adjustments motivated by the data are made to relevant parameter values in an initial attempt to reflect the key features of the data.

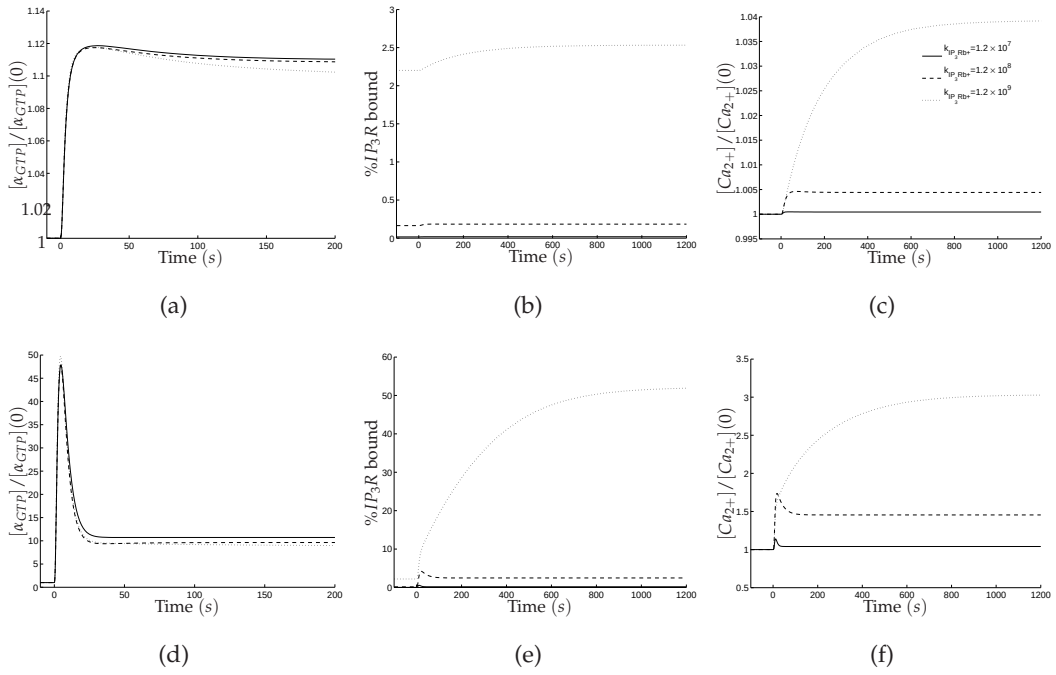


Figure 3.7: Time courses of some species in the Ca^{2+}-mobilisation model, simulated with the preliminary parameter set in Table B.2.2 (in the appendices) and varying k_{IP_3Rb+}. (a)-(c) show the response to 10^{-7} (M) of Oxotremorine-M, while (d)-(f) show the response to 10^{-4} (M). α_{GTP} and Ca^{2+} are normalised to their basal levels.

3.7.1 Preliminary parameter set

The previously discussed experimental results pertain to the α_{GTP} and Ca^{2+} levels of the pathway (see Figure 3.5), hence, for the sake of brevity, only model simulations of α_{GTP} and relevant downstream species will be discussed here.

With the preliminary parameter set, the model predicts that there is a small α_{GTP} response (Figure 3.7(a), solid line), and an insignificant Ca^{2+} response to $10^{-7}M$ of Oxotremorine-M, as Ca^{2+} hardly rises above its basal concentration (Figure 3.7(c), solid line). This is contrary to what is seen in the data, in which this concentration of the drug elicits the maximum Ca^{2+} response (see Figure 3.3). When a high concentration ($10^{-4}(M)$) is used, there is a significant α_{GTP} response (Figure 3.7(d), solid line), and a Ca^{2+} response which is small compared to the data, but qualitatively similar, being rapid and transient (see Figure 3.7(f), solid line).

The simulated Ca^{2+} dose-response curves can now be directly compared to the experimentally generated ones, something that could not be done with just the G-protein activation model in Chapter 2. Figure 3.8 shows the simulated curves (using the ligand-specific parameters from Chapter 2); signal amplification is not predicted, since the agonists do not all produce the same maximal response, acting as partial agonists in-

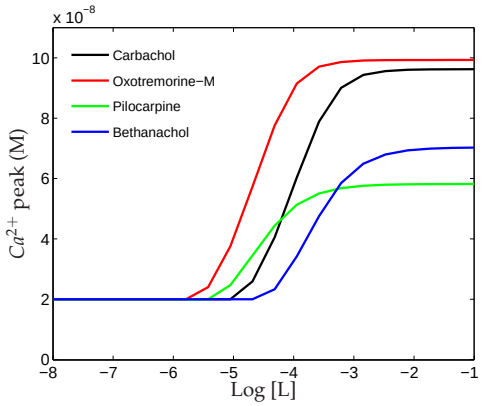


Figure 3.8: Experimental dose response curves for the Ca^{2+} peak response show that all four drugs are full agonists Dose-response curves simulated with the preliminary parameter set in Table B.2.2 ($k_{IP_3Rb+} = 1.2 \times 10^7 M^{-1}s^{-1}$) do not predict full agonism for all the agonists, unlike the experimental results (see Figure 3.3(b)).

stead. In contrast, agonist concentrations as low as $10^{-7}M$ can produce the maximum Ca^{2+} response in the data (see Figure 3.3(b)). Hence, with the preliminary parameter set, the model does not reflect the experimentally observed signal amplification, and so under-predicts agonist efficacy; therefore, parameters which influence the extent of Ca^{2+} release need to be identified and manipulated in order to predict the desired level of response.

3.7.2 Varying the binding rate of IP_3 and its receptor

In the model simulations with the preliminary parameter set, the majority of the IP_3 receptors (IP_3R) are left unbound by IP_3 (see Figures 3.7(b) and 3.7(e), solid lines), which means that only a very small percentage of receptors can be activated by Ca^{2+}, even at high agonist concentrations; it is therefore unsurprising that the predicted Ca^{2+} response is insignificant, since only a very low concentration of IP_3 receptors ever open up to release Ca^{2+} into the cytosol. This suggests that, in order for the model to predict a response comparable to the data, parameters in the preliminary parameter set that can promote the binding of IP_3 to its receptor should be varied appropriately.

The rate constants for the association and dissociation of IP_3 and IP_3R (k_{IP_3Rb+} and k_{IP_3Rb-} respectively) in the preliminary parameter set give a K_D of $\approx 6.7 \times 10^{-7}M$. Experimental studies in which Ca^{2+}-mobilisation was induced by directly transporting IP_3 to IP_3 receptors purified from rat cerebellar membranes give the K_D value of the binding as $6.25 \pm 0.4 \times 10^{-9}M$ [64], which is roughly 100 times smaller than in the preliminary parameter set. Recall that the K_D of a reversible reaction is given by the

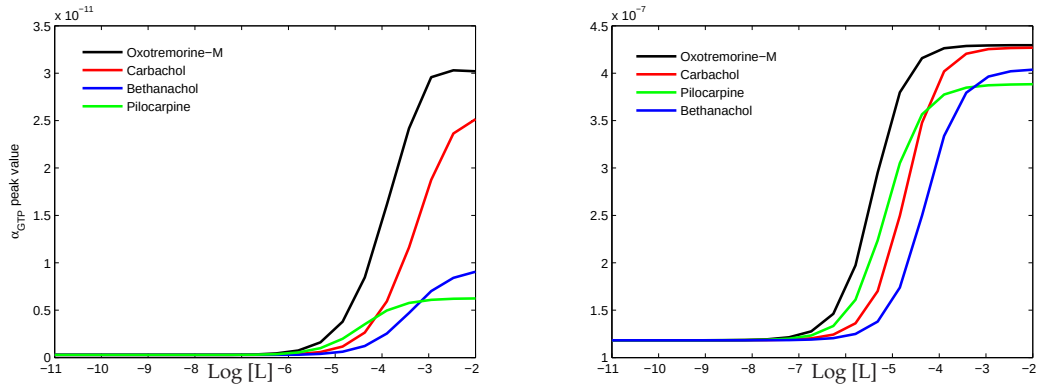
ratio of the forward rate constant to the backward one; the K_D of the IP_3-IP_3R reaction is then given by $\frac{k_{IP_3Rb-}}{k_{IP_3Rb+}}$, and to reduce its value in the preliminary parameter set, k_{IP_3Rb+} can be fixed at its current value while k_{IP_3Rb-} is decreased, or k_{IP_3Rb+} can be increased while k_{IP_3Rb-} is fixed. The aim here is to promote the binding of IP_3 to its receptor, and the most direct way to do this is to increase the rate constant k_{IP_3Rb+} alone (this should also promote the binding of activating Ca^{2+}, and consequently, Ca^{2+} mobilisation). Using the K_D value of $6.25 \pm 0.4 \times 10^{-9}M$ [64] as a boundary, the value of k_{IP_3Rb+} can be increased up to a hundred-fold.

A ten-fold increase in k_{IP_3Rb+} ($k_{IP_3Rb+} = 1.2 \times 10^8 M^{-1}s^{-1}$) leads to a small increase in the percentage of channels bound by IP_3 when $[L] = 10^{-7}M$ (Figure 3.7(b), dashed line), and so gives a trivial increase in Ca^{2+} release (Figure 3.7(c), dashed line). When $[L] = 10^{-4}M$, though the increase in IP_3-bound receptors appears to be insignificant (only $\approx 4\%$, see Figure 3.7(e), dashed line), the increase in the Ca^{2+} response is significant: approximately 1.75 times more than basal (Figure 3.7(f)).

When k_{IP_3Rb+} is increased a hundred-fold to $1.2 \times 10^9 M^{-1}s^{-1}$ (giving a K_D value of $6.67 \times 10^{-9}M$, which fits the previously discussed published value of $6.25 \pm 0.4 \times 10^{-9}M$ [64]) and $[L] = 10^{-7}M$, there is still no significant agonist-induced Ca^{2+} increase (Figure 3.7(c), dotted line). However, the dynamics are significantly altered when $[L] = 10^{-4}M$ (Figures 3.7(d)-3.7(e), dotted lines). The transient peak is no longer exhibited by Ca^{2+}, but still seen in α_{GTP}. 50% of the IP_3 receptors are now bound by IP_3, leading to the opening of more channels, and at three times over basal, a more significant Ca^{2+} release. Also the time-to-peak is no longer rapid, as it is in the data. Dose-response curves for this parameter set reveal that very high concentrations of each agonist induce an almost-maximal Ca^{2+} peak (Figure 3.9(b)), but this is still dissimilar to the the data, where the maximal peak occurs at lower concentrations.

In summary, by increasing the rate constant of the $IP_3 - IP_3R$ binding reaction (k_{IP_3Rb+}) a hundred-fold from its original value in the preliminary parameter set, as guided by a published K_D value of the reaction [64], a parameter set was obtained under which the model predicts an approximately maximal Ca^{2+} response to all four agonists, but at concentrations higher than in the data (i.e., without the left shift, compare Figures 3.9(a) and (b)), and without the rapid, transient peak. The challenge is to attempt to find, initially through this method of data-motivated parameter variation, a parameter set that replicates *all* three key features at once: signal amplification, and the rapid, transient peak.

At this point, the value of k_{IP_3Rb+} in the preliminary parameter set (see Table B.2.2) is replaced by $1.2 \times 10^9 M^{-1}s^{-1}$, because with this value the agonists induce non-trivial Ca^{2+} responses, indicating that the G-protein activation model is better coupled to the



(a) (b)

Figure 3.9: Dose-response curves (a) α_{GTP} peak response (b) Ca^{2+} peak response , simulated with the preliminary parameter set in Table B.2.2 but with k_{IP_3Rb+} increased a hundred-fold to $1.2 \times 10^9 M^{-1}s^{-1}$

downstream modules.

3.7.3 A theory of IP_3R occupancy

With the various values of k_{IP_3Rb+} utilised up until this point, the model has only been able to predict significant agonist-induced Ca^{2+} peak responses at very high agonist concentrations. However if the model is to agree with the data, an Oxotremorine-M concentration of $[L] = 10^{-7}M$ should predict a maximal response (see Figures 3.8 and 3.9(b)). Hence the change made to the value of k_{IP_3Rb+} is not sufficient; other parameters need to be considered for variation.

Marchant *et al.* carried out experimental studies to compare the kinetics of Ca^{2+} mobilisation evoked by endogenous IP_3, with those evoked by a stable analogue. In independent experiments Ca^{2+}-release was induced by transporting different concentrations of both analogues directly to IP_3 receptors in rat hepatocytes [50]. The maximally effective concentrations of IP_3 for Ca^{2+} mobilisation were between $10^{-6}M$ and $10^{-5}M$. It was reported that 99.8% of the receptors were bound when $[IP_3]$ was $10^{-5}M$, suggesting that high occupancy occurs when Ca^{2+} release is maximal. Hence, for the Ca^{2+} model it can be assumed that a high percentage of the IP_3 receptors should be occupied for a maximal Ca^{2+} release.

Based on this observation it becomes clear that a maximal Ca^{2+} release cannot be expected at $[L] = 10^{-7}M$ with the adjusted value of $k_{IP_3Rb+} = 1.2 \times 10^9 M^{-1}s^{-1}$, since the concentration of IP_3 produced is enough to bind only 2.7% of the total IP_3R concentration (see Figure 3.7(b)); therefore, the concentration of IP_3 produced needs to be

higher. There is a need to identify and adjust appropriate parameters that can promote the production of more IP_3.

One way to produce more IP_3 would be to simply produce more α_{GTP}, which would increase the concentration of $\alpha_{GTP}PLC$ available to hydrolyse PIP_2, but this would shift the α_{GTP} dose-response curves unnecessarily to the left, predicting an overly amplified signal at the wrong level of the pathway. In order to produce the desired concentration of IP_3 therefore, some amplification is required somewhere in the steps between α_{GTP} production and IP_3 production, that is, between reactions (3.4) and (3.9). The maximum possible concentrations of the products of reactions (3.4) - (3.6) are limited by the concentration of α_{GTP} produced in reactions that occur earlier in the pathway; that is, the production of the species $\alpha_{GTP}PLC$, $\alpha_{GTP}PLC^*$ and $\alpha_{GTP}PLC^*PIP_2$ cannot be amplified. Hence, a higher IP_3 concentration must be achieved directly; one way is to increase the rate at which it is hydrolysed, via k_{IP_3} in reaction (3.7); another way is to increase the concentration of available $\alpha_{GTP}PLC^*PIP_2$ by decreasing k_{PIP2b-} or increasing k_{PIP2b+} in reaction (3.6); alternatively, it might be achieved by making k_{PLCdis} small in reaction (3.8) so that $\alpha_{GTP}PLC$ can re-associate with PIP_2, keeping $\alpha_{GTP}PLC^*PIP_2$ available to produce IP_3; reducing the rate of IP_3 degradation via k_{IP_3deg} in reaction (3.9) might also be another way.

All the previously mentioned measures to increase IP_3 production, and consequently, IP_3R occupancy, lead to excessively high basal activity, as illustrated in Figure 3.10; this is unsupported by the experimental data, in which the basal level is negligible relative to the agonist-induced response (see Figure 3.3). This raises an interesting question: what processes in Ca^{2+}-activating pathways are responsible for ensuring that signal amplification only occurs *after* the addition of agonist?

3.7.4 Data-driven analysis of G-protein parameters

The basal level of IP_3 is driven by basal α_{GTP}, which is produced from the R^*G state of the GPCR. This means that any parameters of the G-protein activation model which control the extent of basal activity relative to the agonist-induced response need to be identified, and appropriately adjusted. Hence, the parameters of the G-protein activation model will be re-assessed.

There is more than one way of lowering the basal production of α_{GTP} in the G-protein activation model. To facilitate the discussion of these possibilities, the rate equations for some of the GPCR species and α_{GTP} are reproduced below.

$$\frac{d[R^*]}{dt} = k_{act}[R] - k_{deact}[R^*] - \theta_\zeta \zeta_+ k_{lb+}[L][R^*] + \theta_\zeta \zeta_- k_{lb-}[LR^*]$$

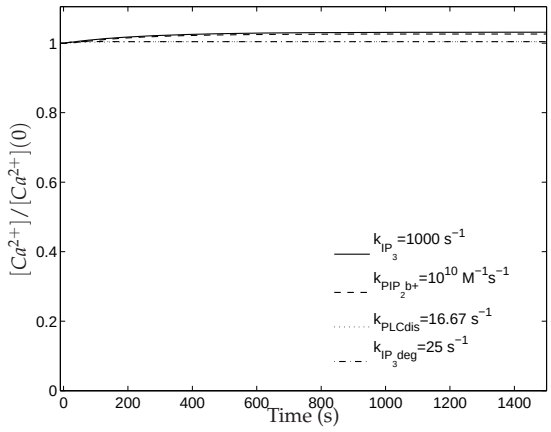


Figure 3.10: Discussed changes in certain parameter values cause the basal response to be excessively high, suggesting that some parameter fine-tuning needs to occur, in order to maintain a significant peak-to-basal ratio.

$$-\theta_\mu\mu_+k_{g+}[R^*][G] + \theta_\mu\mu_-k_{g-}[R^*G] + k_{GTP+}[R^*G] \tag{3.14}$$

$$\begin{aligned}\frac{d[LR^*]}{dt} &= \theta_\zeta\zeta_+k_{lb+}[L][R^*] - \theta_\zeta\zeta_-k_{lb-}[LR^*] + \zeta_+k_{act}[LR] - \zeta_-k_{deact}[LR^*] \\ &\quad -\theta_{\nu\mu}\mu_+\nu_+k_{g+}[LR^*][G] + \theta_{\nu\mu}\mu_-\nu_-k_{g-}[LR^*G] \\ &\quad +k_{GTP+}[LR^*G]\end{aligned} \tag{3.15}$$

$$\begin{aligned}\frac{d[R^*G]}{dt} &= \theta_\mu\mu_+k_{g+}[R^*][G] - \theta_\mu\mu_-k_{g-}[R^*G] + \mu_+k_{act}[RG] - \mu_-k_{deact}[R^*G] \\ &\quad -\theta_{\zeta\nu}\zeta_+\nu_+k_{lb+}[L][R^*G] + \theta_{\zeta\nu}\zeta_-\nu_-k_{lb-}[LR^*G] \\ &\quad -k_{GTP+}[R^*G]\end{aligned} \tag{3.16}$$

$$\begin{aligned}\frac{d[LR^*G]}{dt} &= \theta_{\nu\mu}\mu_+\nu_+k_{g+}[LR^*][G] - \theta_{\nu\mu}\mu_-\nu_-k_{g-}[LR^*G] \\ &\quad +\theta_{\zeta\mu}\zeta_+\mu_+k_{act}[LRG] - \theta_{\zeta\mu}\zeta_-\mu_-k_{deact}[LR^*G] \\ &\quad +\theta_{\zeta\nu}\zeta_+\nu_+k_{lb+}[L][R^*G] - \theta_{\zeta\nu}\zeta_-\nu_-k_{lb-}[LR^*G] \\ &\quad -k_{GTP+}[LR^*G]\end{aligned} \tag{3.17}$$

$$\frac{d[\alpha_{GTP}]}{dt} = k_{GTP+}[R^*G] - k_{gd+}[\alpha_{GTP}] + k_{gd-}[\alpha_{GDP}] + k_{GTP+}[LR^*G] \tag{3.18}$$

The rate of α_{GTP} formation from R^*G could be slowed directly by reducing the value of k_{GTP+} (see equation (3.18)); less directly, either the R^*G formation rate could be slowed by reducing the ratio $\frac{\mu_+}{\mu_-}$ (see equation (3.16)), or the rate of R^* formation slowed by lowering the ratio $\frac{k_{act}}{k_{deact}}$ (see equation (3.14)). However, all of these measures would have a similar effect on the corresponding ligand-bound states, as can be seen from equations (3.15), (3.17), (3.18), and consequently, the level of agonist-induced α_{GTP} production would also drop; that is, the peak-to-basal ratio would not necessarily be increased. What is desired is a parameter change that would result in an increase of the peak-to-basal ratio for $[\alpha_{GTP}]$, resulting in the production of enough IP_3 to cause full

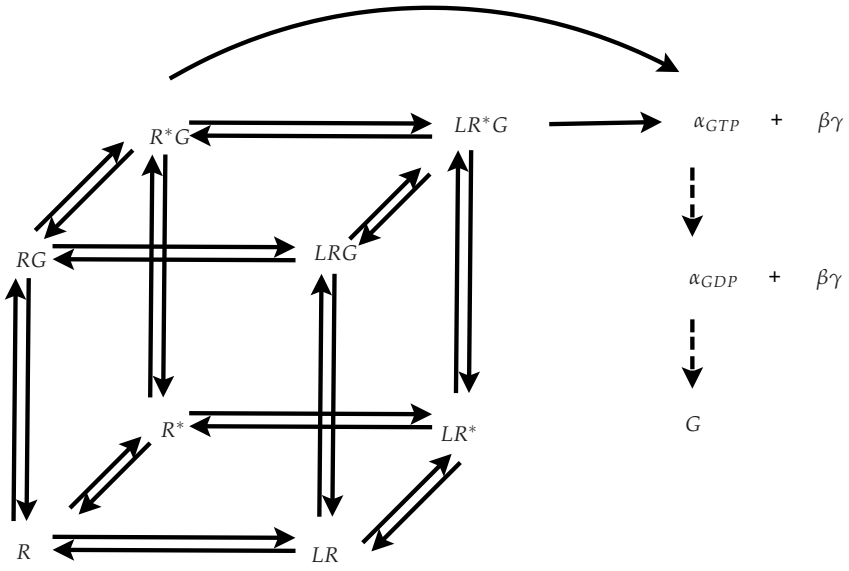


Figure 3.11: The $\alpha_{GTP\gamma S}$ model (represented by the figure without the dotted lines) is derived, in order to mimic data from an $\alpha_{GTP\gamma S}$ assay [93]; it is a subset of the (cubic ternary complex activation model) cTCAM, which is represented by the whole figure (see Chapter 2).

occupancy of its receptors only after agonist stimulation.

The parameters in the preliminary parameter set which pertain to the G-protein activation model can be re-assessed with the help of the α_{GTP}-level data discussed in section 3.4, which were obtained from a $GTP\gamma S$ binding assay [93]. The experiment measures the level of agonist-induced G-protein activation by measuring the concentration of the complex formed from the binding of the non-hydrolysable GTP analogue, $GTP\gamma S$ to the α subunit of the G-protein (forming $\alpha_{GTP\gamma S}$). Since $GTP\gamma S$ is non-hydrolysable, there is no deactivation of $\alpha_{GTP\gamma S}$ into α_{GDP} or subsequent reunion with $\beta\gamma$ to reconstitute the inactive G-protein. These experiments were carried out in CHO-cells expressing the M_3 Muscarinic receptor, the same system used to produce the Ca^{2+} data of Dr Lauren May here at the University of Nottingham (see section 2.2.3). The $\alpha_{GTP\gamma S}$ response to several agonists, measured after one hour, is shown in dose-response curves in Figure 3.5.

A model of the $GTP\gamma S$ binding experiment is derived by setting $k_{gd+} = k_{RA+} = 0$ in the G-protein activation model (see section 2.1). The relationship between the two models is illustrated in Figure 3.11. Comparing the predictions of the $\alpha_{GTP\gamma S}$ model with the $\alpha_{GTP\gamma S}$ data serves as a way to analyse some parameters in the preliminary parameter set that pertain to the G-protein activation model.

Figure 3.12(a) shows the dose response curves predicted by the $\alpha_{GTP\gamma S}$ model with

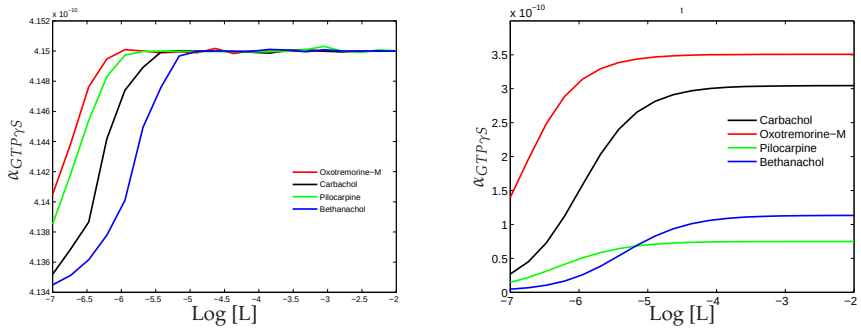


Figure 3.12: Dose response curves for $\alpha_{GTP\gamma S}$ predicted by the $\alpha_{GTP\gamma S}$ accumulation model ((a)) simulated with the parameter set in Appendix B.2.1 but with $k_{gd+} = k_{RA+} = 0$, (b) simulated with the parameter set in Appendix B.2.1 but with $k_{gd+} = k_{RA+} = 0$ and adjusted values $k_{GTP+} = 0.001s^{-1}$, $\nu = 100$, which allow the curves resemble the data more closely (compare Figure 3.5.

relevant parameters taken from the preliminary parameter set. In contrast to the data, the model predicts that all of the agonists are full agonists, capable of producing the maximum possible $[\alpha_{GTP\gamma S}]$, which is approximately equal to the total G-protein concentration ($G_{TOT} = 4.15 \times 10^{-10}M$). It is also clear from Figure 3.12(a) that over the different agonist concentrations, the model response range is very small, $4.135 - 4.15 \times 10^{-10}M$. In fact, in the absence of agonist, the system, via the R^*G state, is able to produce almost the maximum possible $\alpha_{GTP\gamma S}$ response (result not shown); this is a significant overestimation of the basal activity when compared to the data. These results then imply that one or more of the G-protein parameters require adjustment in order to predict a more realistic basal response. Since all the parameters in the $\alpha_{GTP\gamma S}$ model are also in the G-protein activation model and in the Ca^{2+} model, the consequent parameter changes informed by the $\alpha_{GTP\gamma S}$ data can be applied to those models as well.

The fact that the basal $\alpha_{GTP\gamma S}$ concentration is too high suggests that the rate of the reaction $R^*G \xrightarrow{k_{GTP+}} R^* + \alpha_{GTP} + \beta\gamma$ is too fast, hence an obvious parameter to lower is k_{GTP+}. An examination of the $\alpha_{GTP\gamma S}$ rate equation

$$\frac{d[\alpha_{GTP\gamma S}]}{dt} = k_{GTP+}([R^*G] + [LR^*G]), \tag{3.19}$$

indicates that lowering k_{GTP+} would lower the basal concentration of $\alpha_{GTP\gamma S}$, produced from R^*G, but also its agonist-induced peak concentration, produced from LR^*G. Therefore lowering the value of k_{GTP+} on its own would not necessarily increase the peak-to-basal ratio. From equations (3.16) and (3.17), it can be seen that if the value of ν_+ is also increased, the effect would be to promote the LR^*G state of the GPCR while

inhibiting the R^*G state; this counteracts the negative effect on the peak-to-basal ratio of decreasing k_{GTP+}.

With the adjusted values $k_{GTP+} = 0.001s^{-1}$ (previously $k_{GTP+} = 1s^{-1}$) and $\nu_+ = 100$ (previously $\nu_+ = 1$) the model makes qualitatively similar predictions to the $\alpha_{GTP\gamma S}$ data, predicting different maximal responses for each agonist (Figure 3.12(b)). The model's dose-response curves are steeper than those of the data, which means that the predictions of efficacy for the agonists are overestimated. This may be accounted for by the fact that mechanisms of receptor desensitisation are not included in the model.

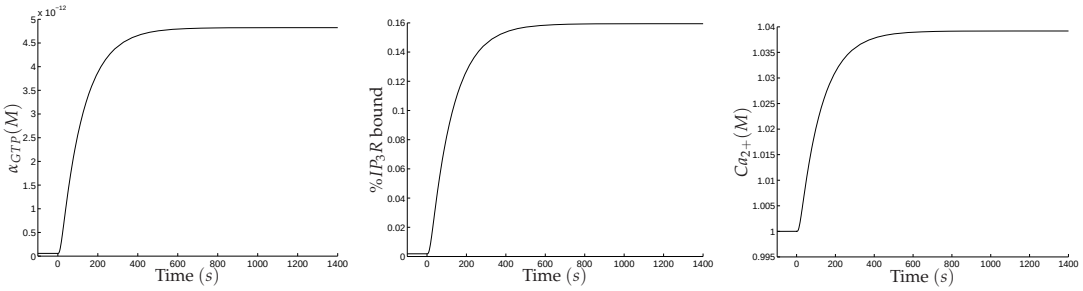
Following previous modelling efforts, the model had so far assumed that the agonist binds indiscriminately to the pre-coupled (RG and R^*G) and un-coupled states of the GPCR (R and R^*); that is, $\nu = \frac{\nu_+}{\nu_-} = 1$ [86, 106]. The need to increase the value of ν_+ suggests that this might not be the case. The model suggests that in reality, the agonist might favour the pre-coupled state of the receptor.

The values $k_{GTP+} = 0.001s^{-1}$, $\nu_+ = 100$ have improved the qualitative predictions of the $\alpha_{GTP\gamma S}$ model and as such replace their former values in the Ca^{2+} model.

3.7.5 Final variation of the Ca^{2+} model parameters

As a reminder, the current parameter set has been derived by replacing relevant parameters in the preliminary parameter set with the new values: $k_{IP_3Rb+} = 1.2 \times 10^9 M^{-1}s^{-1}$, $k_{GTP+} = 0.001s^{-1}$ and $\nu_+ = 100$. The last two parameter choices have the effect of keeping the basal $[\alpha_{GTP}]$ low, as illustrated by Figure 3.13(a), but (as would be expected) do not necessarily increase the number of IP_3-bound receptors (see Figure 3.13(b)), or consequently, the Ca release (see Figure 3.13(b)). To maximally bind the IP_3 receptors, parameter adjustments that increase IP_3 production need to be made.

In Chapter 2, it was shown that increasing the total G-protein concentration (G_{TOT}) up to $\approx 10^{-6}M$ in the G-protein activation model also increases the maximum $[\alpha_{GTP}]$ (see Figure 2.3), and that increasing the total receptor concentration (R_{TOT}) up to $10^{-8}M$ has a similar effect (see Figure 2.4), all without significantly elevating the basal concentration of α_{GTP}. Hence, the new values, $G_{TOT} = 10^{-6}M$ and $R_{TOT} = 10^{-8}M$ are now used instead of their former values ($G_{TOT} = 4.15 \times 10^{-10}M$, $R_{TOT} = 4.15 \times 10^{-10}M$) in the preliminary set; by increasing α_{GTP} production, IP_3 production can also be increased, without unduly increasing the basal level. Any of the parameter changes discussed in section 3.7.3 can also be implemented to increase the production of IP_3, since the peak-to-basal ratio will be controlled by the values of new values of k_{GTP+}, ν_+, G_{TOT} and R_{TOT}. The most direct parameter change is to increase the rate constant of $\alpha_{GTP}PLC$ and PIP_2 binding, k_{PIP_2b+} (see reaction 3.6), and consequently, the available $\alpha_{GTP}PLCPIP_2$ concentration, from which IP_3 is hydrolysed.



(a) (b) (c)

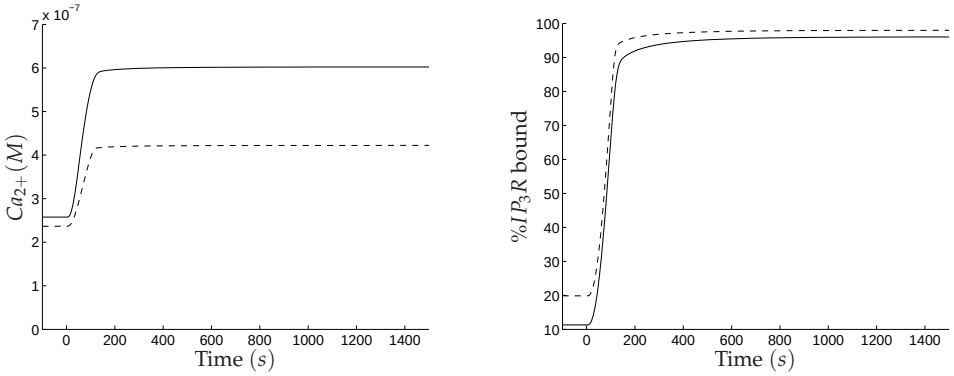
Figure 3.13: Time courses predicted by the Ca^{2+} model, simulated with the preliminary parameter set in Table B.2.2 but with the adjusted values: $k_{IP_3Rb+} = 1.2 \times 10^9 M^{-1}s^{-1}$, $k_{GTP+} = 0.001s^{-1}, \nu = 100$. The changes in k_{GTP+} and ν only impact the α_{GTP} peak-to-basal ratio.

Increasing k_{PIP_2b+} a hundred-fold, from $10^9 M^{-1}s^{-1}$ to $10^{11} M^{-1}s^{-1}$ does cause production of enough IP_3 to occupy nearly all receptors *only after* stimulation by $10^{-7}M$ of Oxotremorine-M (see Figure 3.14(b), dashed line), and Ca^{2+} release is increased (see Figure 3.14(a), dashed line), but it is not maximal (results not shown). This is because of the way that the IP_3-bound receptors are distributed between states; Figure 3.15(a) shows that at any time after stimulation, more receptors are in the inhibitory state ($[IP_3IP_3RCa^{2+}Ca^{2+}]$) than the active state ($[IP_3IP_3RCa^{2+}]$), which limits Ca^{2+} release. This indicates that the value of the binding rate constant of inhibitory Ca^{2+}, $k_{Ca^{2+}inh+}$ must be too large (or the unbinding rate constant too small). Figure 3.15(b) indicates that lowering $k_{Ca^{2+}inh+}$ ten-fold, from $1.8 \times 10^6 M^{-1}s^{-1}$ to $1.8 \times 10^5 M^{-1}s^{-1}$ causes more receptors to stay in the active, than the inhibitory state, without significantly changing the total percentage bound to IP_3R (see Figure 3.14(b), solid line); Ca^{2+} release is also increased (Figure 3.14(a), solid line). Figure 3.16 indicates that this Ca^{2+} release is maximal, and all other agonists are able to produce it. The simulated dose-response curves thus resemble the data in predicting that all the drugs are full agonists.

All the parameter changes made to get to this point form a new parameter set found in Table B.2.2, referred to as the *derived* parameter set. The model predicts signal amplification with these new parameters, but the rapid, transient peak has been lost.

3.8 Model extension: Effects of receptor desensitisation

So far, the model has not simultaneously predicted all the features of the Ca^{2+} data with any of the parameter sets that have been tested. In each case (apart from the cases where the response was trivial) either the rapid, transient peak was predicted without



(a) (b)

Figure 3.14: Time courses predicted by the Ca^{2+} model, simulated with the preliminary parameter set in Table B.2.2 but with the adjusted values: $k_{GTP+} = 0.001s^{-1}$, $G_{TOT} = 4.15 \times 10^{-6}M$, $R_{TOT} = 10^{-8}M$, $k_{PIP_2b+} = 10^{11}M$.

signal amplification, or vice versa. With the preliminary parameter set (see Table B.2.2), the Ca^{2+} removal process (carried out by SERCA pumps) was influential enough to make the peak transient, probably because the Ca^{2+} response was not amplified (see Figures 3.7 and 3.8); with the derived parameter set (see Table B.2.2) came the model's ability to predict signal amplification (see Figure 3.16), but at the cost of the transient peak, with the action of the SERCA pumps no longer influential enough to bring the peak down (see Figure 3.14).

The Ca^{2+} removal process is modelled by a Hill function, as similarly used in some existing Ca^{2+} models [44, 48, 70]. It might be that varying the parameters p_1 and p_2 of the Hill function could enable the prediction of a transient peak, which will occur when $\frac{d[Ca^{2+}}{dt}]$ is negative; that is, when,

$$(1 + v_r)(\gamma_0 + \gamma_1[IP_3IP_3RCa^{2+}])(Ca^{2+}_{AV} - [Ca^{2+}]) < \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2}. \tag{3.20}$$

Other upstream parameters might also modulate equation 3.20 by influencing the production of $IP_3IP_3RCa^{2+}$, thus influencing the lifetime of the peak. The next chapter will attempt to identify such parameters.

However, there is an indirect process of Ca^{2+} removal which has not yet been included in the model, but which may have a role in making the peak transient. In the data, Ca^{2+} returns to its basal level, indicating that the signal from the agonist becomes ineffective through desensitisation. Hence, it might be necessary to incorporate receptor phosphorylation into the Ca^{2+} model, in order to predict the return from peak to basal level. As previously mentioned, PKC, which is activated by DAG (a co-product

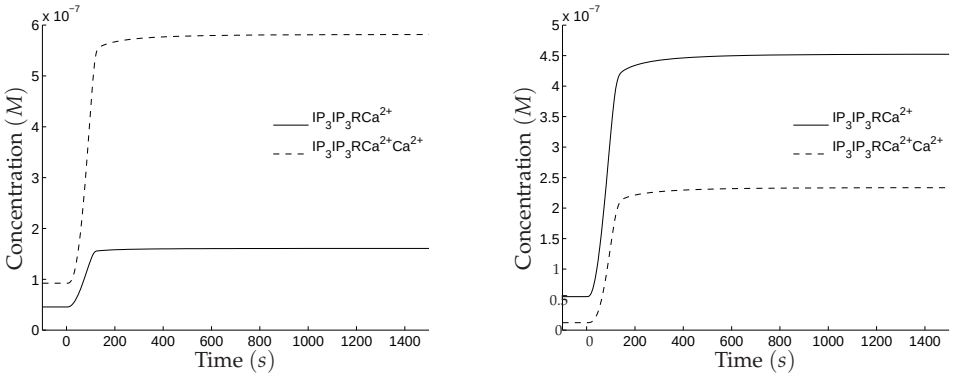


Figure 3.15: When the model is simulated with the adjusted values: $k_{GTP+} = 0.001s^{-1}$, $G_{TOT} = 4.15 \times 10^{-6}M$, $R_{TOT} = 10^{-8}M$, $k_{PIP_2b+} = 10^{11}M$, the response is not maximal because $k_{Ca^{2+}inh+} = 1.8 \times 10^{6}M^{-1}s^{-1}$ is too high, favouring the inhibited IP_3R state (a), but when $k_{Ca^{2+}inh+}$ is lowered to $1.8 \times 10^{6}M^{-1}s^{-1}$, more active states are formed.

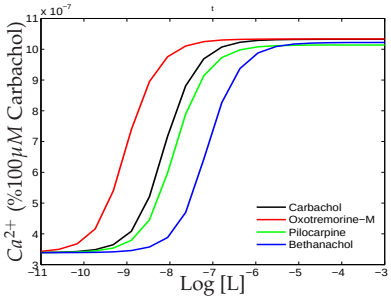


Figure 3.16: Dose response curves for the peak $[Ca^{2+}]$ predicted by the Ca^{2+} model, with the parameter set in Table B.2.2, but with the adjusted values: $k_{GTP+} = 0.001s^{-1}$, $G_{TOT} = 4.15 \times 10^{-6}M$, $R_{TOT} = 10^{-8}M$, $k_{PIP_2b+} = 10^{11}M$, $k_{Ca^{2+}inh+} = 1.8 \times 10^{5}M^{-1}s^{-1}$.

with IP_3 of PIP_2 hydrolysis) causes GPCR desensitisation by phosphorylating GPCR's in the G_q pathway, and causing termination of the signal. In the Kang and Othmer model, where the active, unbound receptor (R^*) is incapable of signalling, the phosphorylation reaction (in the terminology of this work) is given by

$$LR^*G + PKC \underset{k_{-PKLR}}{\overset{k_{PKLR}}{\rightleftharpoons}} LR^*_{des}G. \tag{3.21}$$

The DAG-PKC pathway will not be included in this chapter's Ca^{2+} model, as that would significantly increase model complexity. Instead, a simple desensitisation scheme, which accounts for the fact that the lifetime of active GPCR's is shortened by PKC-mediated phosphorylation is incorporated by adding the following phosphoryla-

tion reactions:

$$R^*G \underset{k_{des-}}{\overset{k_{des+}}{\rightleftharpoons}} R^*_{des}G \tag{3.22}$$

$$LR^*G \underset{\rho_- k_{des-}}{\overset{\rho_+ k_{des+}}{\rightleftharpoons}} LR^*_{des}G, \tag{3.23}$$

where ρ_+, ρ_- are factors that potentially differentiate the rate constants of the desensitisation and re-sensitisation of R^*G from LR^*G. Values for k_{des+} are available from a GPCR-mediated IP_3 production model of Cooling *et al.* [17], which incorporates receptor phosphorylation via the forward reaction in (3.23); in that model $k_{des+} = 6.22 \times 10^{-6} s^{-1}$ and $k_{des+} = 4 \times 10^{-4} s^{-1}$ were used to model phosphorylation in two different GPCR systems. The forward rate constant in the Kang and Othmer model (reaction 3.21) is not directly applicable because it has units of $M^{-1}s^{-1}$; however, the reverse reaction, which is similar to those in reactions (3.22) and (3.23), makes use of the value, $k_{-PKLR} = 0.1 s^{-1}$ [37]. There are thus values in the literature that can be used as guide for setting k_{des+} and k_{des-}. Without evidence to the contrary, it is assumed that the agonist-bound receptor is phosphorylated at the same rate as the unbound receptor, so that $\rho_+ = \rho_- = 1$.

Suitable values of k_{des+} and k_{des-} need to be found which permit the model to still satisfy the IP_3R maximal occupancy theory (see section 3.7.3). Hence, one criterion that will be used to fine-tune these parameter choices is the percentage of bound IP_3R,

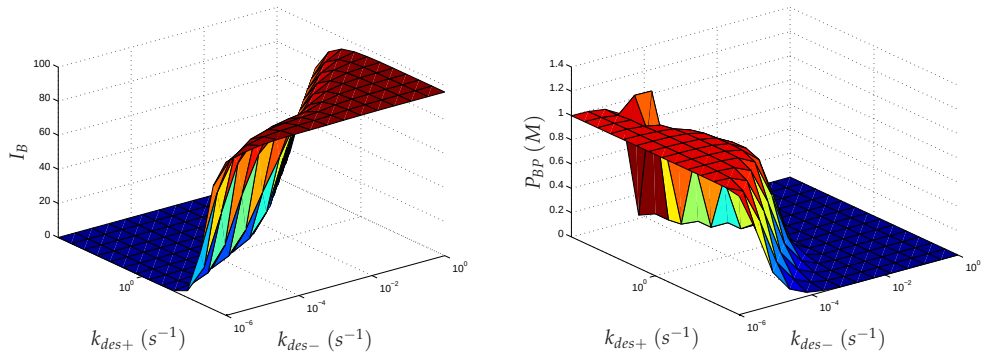
$$I_B = \frac{100(IP_3R_{TOT} - IP_3R^P)}{IP_3R_{TOT}}, \tag{3.24}$$

where IP_3R^P is the peak IP_3R concentration. Another measure is the ratio,

$$P_{BP} = \frac{\text{Peak} - \text{Plateau}}{\text{Peak} - \text{Basal}}, \tag{3.25}$$

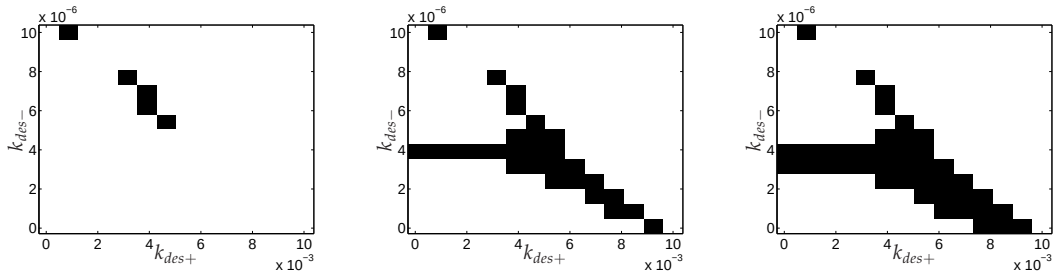
which should be 1 if the plateau returns to the basal value. (Note that Peak $-$ Basal $\neq 0$ in the parameter range explored).

Figure 3.17 shows three-dimensional plots of P_{BP} and I_B on the z axes with k_{des+} and k_{des-} on the x and y axes. In Figure 3.17(a) there is a region in which I_B is high, while in Figure 3.17(b) there is one in which $P_{BP} \approx 1$. However, only small portions of these regions intersect, meaning that only pairs in a constricted region of (k_{des+},k_{des-}) parameter space cause a post-peak return to basal *and* high IP_3R occupancy. Figure 3.18 shows areas in space for which $0.95 \leq P_{BP} \leq 1.04$ *and* IP_3R occupancy is high. Only a very small area causes $\geq 90\%$ IP_3R occupancy or more (Figure 3.18(a)), and the



(a) (b)

Figure 3.17: The dependence of (a) the percentage of bound IP_3 receptors and of (b) Ca^{2+}'s plateau relative to its basal (b) on the values of k_{des+} and k_{des-}. Only a few pairs k_{des+} and k_{des-} cause a return to basal *and* high IP_3R occupancy.



(a) (b) (c)

Figure 3.18: Shown in black are areas of the k_{des+}-k_{des-} plane where $P_{BP} \approx 1$ *and* (a) IP_3R occupancy $\geq$ 90%, (b) IP_3R occupancy $\geq$ 70%, (c) IP_3R occupancy $\geq$ 50%.

area increases for $\geq$ 70% IP_3R occupancy (Figure 3.18(b)) and $\geq$ 50% IP_3R occupancy (Figure 3.18(c)).

A pair of values, $k_{des+} = 5 \times 10^{-3} s^{-1}$ and $k_{des-} = 5 \times 10^{-6} s^{-1}$, from the $\geq$ 90% IP_3R occupancy area (Figure 3.18(a)) are used to simulate the model. To illustrate the role that the desensitisation reactions have on the model, Figure 3.19 compares the GPCR states in the presence and absence of desensitisation. Figure 3.19(a) shows that, in the absence of desensitisation, the majority of receptors are in the inactive RG state before agonist stimulation. On stimulation, RG and R^*G get converted into the LRG and LR^*G states, whose maximum concentrations are sustained as the steady state (the sustained Ca^{2+} response in Figure 3.14(a) might be explained by the fact that the active state, LR^*G is sustained). However, when desensitisation is incorporated, the majority of receptors are distributed between the inactive (RG) and desensitised ($R^*_{des}G$) states

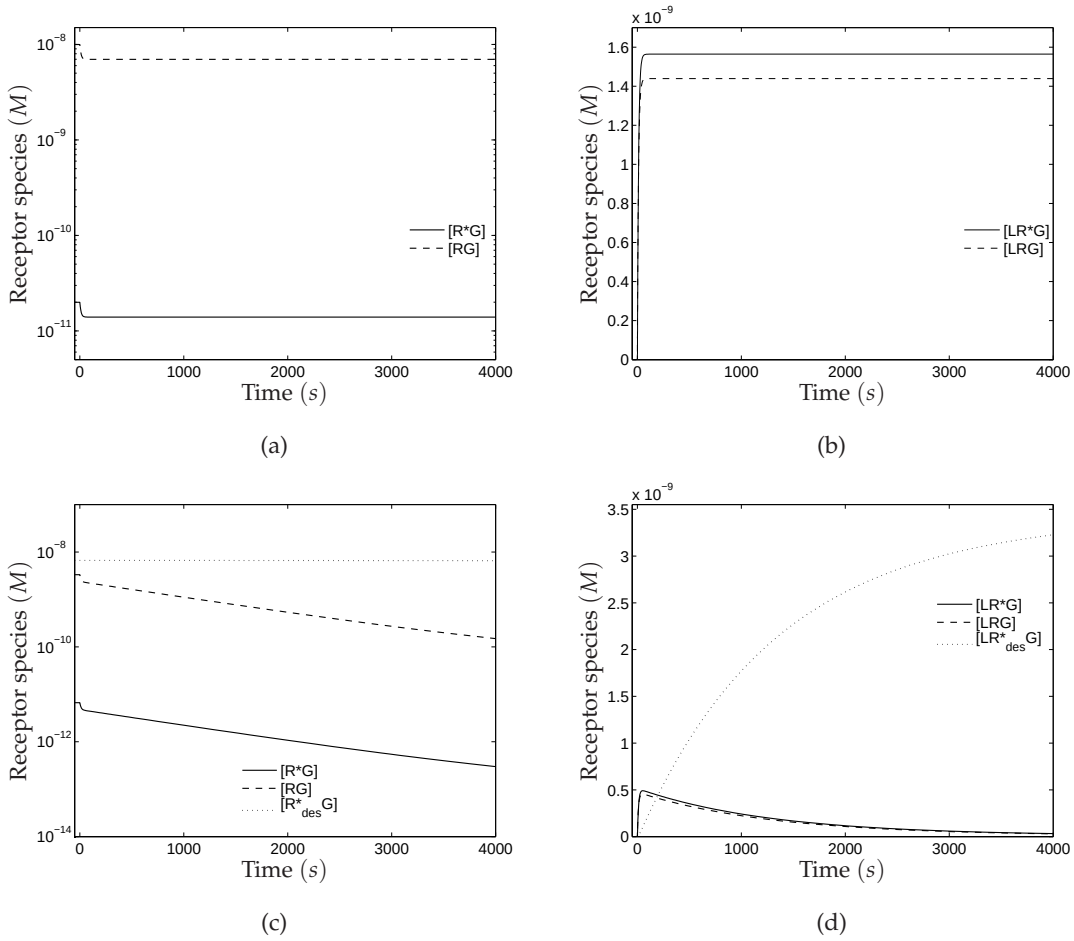
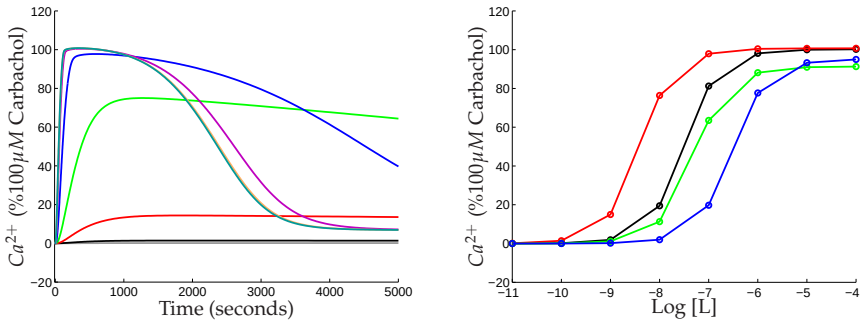


Figure 3.19: (a)-(b) Receptor states in the model without desensitisation ($k_{des+} = 0, k_{des-} = 0$) are shown; (c)-(d) With desensitisation, the states are re-distributed $k_{des+} = 5 \times 10^{-3}s^{-1}, k_{des-} = 5 \times 10^{-6}s^{-1}$, with the majority ending up desensitised.

just before stimulation. On stimulation, the agonist bound states LRG and LR^*G start to form, with the latter also being converted into the desensitised state, $LR^*_{des}G$, which accumulates, so that most receptors end up in that state. The transience of the LR^*G state then causes the Ca^{2+} response to be transient, as discussed in the next paragraph.

Figure 3.20 shows the Ca^{2+} simulations. The response is a slow transient for higher concentrations of Oxotremorine-M (Figure 3.20(a)), and is amplified for all agonists (Figure 3.20(b)). Hence, with the incorporation of a minimal receptor desensitisation process, the model predicts signal amplification *and* a transient, but not rapid, Ca^{2+} peak. This might suggest that the calcium-induced calcium release (CICR) mechanism alone is not sufficient to produce the transient Ca^{2+} peak, *in amplifying systems*. However, the derived parameters for desensitisation cannot explain the rapid time to peak, suggesting that they (and other parameters) require further analysis.



(a) (b)

Figure 3.20: The simulated time courses are now transient, with the incorporation of desensitisation into the model, (a), qualitatively resembling the data (see Figure 3.3(a)), but the peaks are not rapid. The simulated dose-response curves (b) resemble the data (see Figure 3.3(b)), with all agonists producing the maximum response.

3.9 Discussion

In this chapter a Ca^{2+} model was assembled by extending the G-protein model of Chapter 2, using relevant downstream reactions from an existing model [37]. The model predicted a rapid, transient Ca^{2+} peak, but not signal amplification, when simulated with a preliminary parameter set assembled from the literature. Using available data from experiments carried out at the University of Nottingham's Institute of cell signalling [52], and other experimental results from the literature [50, 93], adjustments were made to the preliminary set to derive a parameter set with which the model could predict signal amplification, but not the rapid or transient Ca^{2+} peak.

The model was then extended by incorporating GPCR desensitisation, which improved the predicted Ca^{2+} response; the model predicted signal amplification and the transient peak (with return to basal level, as in the data), but not its rapidity. This extension of the model is what will subsequently referred to as the Ca^{2+} mobilisation model.

GPCR pre-coupling

When the $\alpha_{GTP}\gamma S$ model (a subset of the Ca^{2+} model) was fine-tuned to qualitatively predict the $\alpha_{GTP}\gamma S$ data, the value $\nu_+ = 1$ in the preliminary parameter set had to be changed to $\nu_+ = 100$, suggesting that the agonist favours the GPCR's pre-coupled state (RG, R^*G) over the uncoupled one (R, R^*). This suggests that agonist efficacy may be dampened if pre-coupling (via reactions 1.13 and 1.14) is inhibited, a model result that can be tested experimentally.

Receptor desensitisation

To enable the simultaneous prediction of signal amplification and the transient peak, receptor desensitisation had to be incorporated into the model. This might suggest that the Ca^{2+}-induced Ca^{2+} release mechanism alone is not sufficient to cause a transient peak, when the Ca^{2+} response is amplified. This can be tested by measuring Ca^{2+} time courses in experiments where kinase/*PKC*-mediated desensitisation is blocked by using commercially available PKC/kinase inhibitors [57, 88], and checking whether the transient peak occurs.

Hints of a fine-tuned Ca^{2+} response

The fact that the model could not simultaneously predict signal amplification and the rapid Ca^{2+} time-to-peak with any of the parameter sets that were tested raises the question of how the system fine-tunes the Ca^{2+} response, amplifying a signal without delaying the response. The method of model simulation interacting with experimental data has not been sufficient to provide an explanation for this seemingly optimised mechanism. However, it has provided a base parameter set around which to define a reasonable parameter space, in which a more systematic parameter analysis can be carried out; this will be the focus of the next chapter.

CHAPTER 4

Global sensitivity analysis of parameters

4.1 Introduction to sensitivity analysis

In Chapter 3 an initial parameter analysis was carried out using a heuristic approach that involved making data-guided adjustments to relevant parameters. With these adjustments, the model's prediction of the Ca^{2+} response was improved, matching the data more closely, but not completely. Since the parameter changes were not sufficient to enable the model to reflect all the features of the Ca^{2+} data, a more systematic parameter analysis is carried out in this chapter using a method of sensitivity analysis.

Sensitivity analysis methods can quantify the impact of varying a parameter's value on specific model outputs, providing a parameter significance ranking. Such a method of analysis is ideal for the Ca^{2+} model for the following reasons. Firstly, the Ca^{2+} model, like many large models of biological systems, has many parameters whose values are uncertain. A sensitivity analysis would identify the subset of parameters which are key drivers of particular model output features like the rapid, transient peak response and signal amplification. Hence, unlike steady state analysis methods, sensitivity analysis can be used to analyse such time-dependent behaviour as exhibited in the Ca^{2+} data (see Figure 3.20). Secondly, if appropriate non-influential parameters are identified, sensitivity analysis may open up the possibility of model simplification, and perhaps, simplify or enable the use of other methods of analysis, simply by reducing the number of parameters requiring further analysis. For instance, a process like parameter estimation can potentially be significantly simplified, because the dimensionality of the parameter space can be reduced by fixing the values of relatively non-influential parameters, and only estimating the rest.

Local sensitivity analysis (LSA) methods are carried out at a point in parameter

space, and generally involve the use of partial derivatives to identify which parameters have the greatest influence on the output. Given that the output of interest is represented by Y, and the ith parameter by X_i, a commonly used measure of sensitivity is [80]:

$$C_i = \frac{\delta Y / Y}{\delta X_i / X_i}, \tag{4.1}$$

which quantifies how changes in the parameter's value influence the output; also, if $|C_i > 1|$, the relative change in the output is greater than in the parameter. LSA methods are appropriate for situations in which the model parameters are known with reasonable certainty. However, for nonlinear models with considerable uncertainty in the parameters, sensitivity analysis methods applied to parameters as they change across space are more appropriate [2, 81]. These are referred to as global sensitivity analysis (GSA) methods.

4.2 GSA methods and applications

Global sensitivity analysis methods measure sensitivity at different points in parameter space, and so give an indication of each parameter's average effect as it changes simultaneously with other parameters. There are a variety of established GSA methods in the literature, as well as instances of their application to problems in various fields of study.

Variance-based methods form a group of GSA methods that quantify the percentage that each parameter or combination of parameters (from pairs to highest order interactions) contributes to the variance in the model output [81]. The parameters which contribute little can thus be identified. Variance-based methods are called *quantitative* because the contribution of each parameter to the variance of the output has a quantitative measure, but they are computationally expensive and can become impracticable when the number of parameters is large.

The method of elementary effects (EE) calculates for each parameter, at multiple points in parameter space, the change in output relative to changes in that parameter's value, using the average of these ratios as the sensitivity measure. This measure has been shown to be as effective as variance-based measures in identifying parameters that contribute significantly to the variance of the output [15], but does not quantify their exact contribution. The EE method is able to distinguish between parameters which have effects that are (a) negligible, (b) linear and additive, or (c) nonlinear and/or due to interactions (correlation) with other parameters [63]. The method of distinguishing between these effects will subsequently be clarified (see section 4.3.2). The

EE method is computationally cheaper than variance-based methods, and suitable for models with a high number of parameters.

The method of Partial rank correlation coefficients (PRCC) is another global method; it measures the strength of the linear relationship between the output and a parameter after the effects of the other parameters have been removed, and ranks the parameters accordingly [41]. The method assumes that the relationship between parameters and the model output is monotonic [41], and so is not applicable to every type of model.

Global sensitivity analysis methods have routinely been applied in the fields of environmental engineering [15] and finance [81]. Recently they have been applied to mathematical models of biological systems, including the cubic ternary complex activation model (cTCAM) discussed in Chapter 2 [41], a gene transcription model [35], a T cell receptor signalling model [107], and a GPCR-mediated IP_3 activation model [17]. It does not appear that a (global) sensitivity analysis has previously been carried out on an agonist-induced Ca^{2+} model, especially not with regard to pharmacological features such as signal amplification. The applications to the cTCAM and the GPCR-mediated IP_3 activation model are of relevance, and are discussed in more detail below.

Kinzer-Ursem and Linderman [41] used the PRCC method to identify the parameters most correlated to the α_{GTP} response in the cTCAM, and found both ligand-specific and cell-specific parameters to be influential. The ligand-specific parameter most correlated with response generation was ζ, the extent to which the ligand favours the receptor's active state. The cell-specific parameters highly correlated with response generation were: the total concentrations of receptor and G-protein, R_{TOT}, G_{TOT}; the equilibrium of ratio of active (R) to inactive receptor (R^*), quantified by $\frac{k_{act}}{k_{deact}}$; the extent to which the active receptor binds more or less preferably to the G-protein than the inactive receptor, μ; and the rate constants in the G-protein activation loop. Guided by these results, it was found that changes in the *cell-specific* parameters, G_{TOT} and $\frac{k_{act}}{k_{deact}}$, could produce the experimentally observed phenomenon of protean agonism (which refers to the ability of a ligand to induce both positive and negative responses).

To understand why two different agonists, Endothelin-1 (ET-1) and angiotensin-II (Ang-2), could produce different IP_3 transients through essentially identical pathways in the cardiac myocyte, Cooling *et al.* [17] used the method of elementary effects to identify parameters that controlled the features of the transient. They found that the rate constant for the phosphorylation of the active receptor and the rate constant for the binding of ligand to pre-coupled receptors were among the most significant. The differences in the ET-1- and Ang-2-induced IP_3 responses could then be explained by adjusting these two parameters (and a third, chosen because of its individual effect on the IP_3 peak).

4.3 The method of elementary effects

Of the methods discussed above, the method of elementary effects is chosen to analyse the Ca^{2+} model parameters because it is computationally efficient, hence suitable for the model which has many (48) parameters. Also, no prior assumption needs to be made about the model's parameter-output relationship. The method is described in detail below.

A parameter space is defined by specifying a range of uncertainty for each parameter. Let k be the number of parameters of interest, then the vector $\mathbf{X} = (X_1, X_2, ..., X_k)$ represents a point in parameter space and the output of interest is represented by $Y(\mathbf{X})$. The parameters are not sampled directly from the actual parameter space, $\mathbf{A} \subseteq \mathbb{R}^k$. Initial values are randomly sampled from a discrete uniform distribution in the [0,1] interval, so that the *sampling space* is a k-dimensional unit hypercube, $\mathbf{\Omega}$. Let a point in $\mathbf{\Omega}$ be given by $\mathbf{x}$; its ith component, x_i, is a randomly chosen value from $\{0, \frac{1}{p-1}, \frac{2}{p-1}, ..., 1\}$ ($p \in \mathbb{N}$). Thus, the number of possible values is given by the choice of p. The $\mathbf{x}$ vector is then transformed into a point in actual parameter space, $\mathbf{X} = (X_1, X_2, ..., X_k)$, according to a specified distribution. For instance, if each parameter is to be selected from a uniform distribution, the transformation to each X_i is as follows:

$$X_i = x_i(UB_i - LB_i) + LB_i, \tag{4.2}$$

where LB_i is the lower bound of the ith parameter's range of values, and UB_i its upper bound.

A ratio pertaining to the ith parameter, referred to as an *elementary effect* (EE), is calculated using two points in $\mathbf{A}$, identical, except in their ith components which differ by Δ_i. For a given value of $\mathbf{X}$, the elementary effect of the ith input factor on an output of interest Y is defined as:

$$EE_i(\mathbf{X}) = \frac{[Y(X_1, X_2, ..., X_{i-1}, X_i + \Delta_i, ..., X_k) - Y(X_1, X_2, ..., X_{i-1}, X_i, ..., X_k)]}{\Delta}, \tag{4.3}$$

where $X_i \leq 1 - \Delta$.

If $X_i > 1 - \Delta$, the elementary effect is calculated as follows,

$$EE_i(\mathbf{X}^{\mathbf{\Delta}-}) = \frac{[Y(X_1, X_2, ..., X_{i-1}, X_i, ..., X_k) - Y(X_1, X_2, ..., X_{i-1}, X_i - \Delta, ..., X_k)]}{\Delta_i}, \tag{4.4}$$

where $\mathbf{X}^{\mathbf{\Delta}-}$ indicates that the elementary effect calculated is for $\mathbf{X}$ with Δ_i subtracted from its ith dimension.

Let $\delta \in \{\frac{1}{p-1}, \frac{2}{p-1}, ..., 1 - \frac{1}{p-1}\}$ be a fixed amount by which x_i can be increased (note that this means the parameter change is not necessarily the same as the lattice sampling,

since $\delta \neq \frac{1}{p-1}$ is possible); the corresponding increase of X_i in the actual parameter range is then given by:

$$\Delta_i = \delta(UB_i - LB_i). \tag{4.5}$$

The ith parameter has a distribution, F_i, of $p^{k-1}(p - \delta(p-1))$ possible elementary effects, if calculated at all possible sample points in $\mathbf{\Omega}$ [63]. For high values of p and k, calculating all the elements of F_i is an impractical computational effort. (For instance, if for a model with $k = 40$ parameters, $p = 10$ is chosen, $10^{39}(10 - 9\delta)$ total elementary effects can be calculated.) Instead, a random sample of r representative elementary effects is selected from F_i. The average of these elementary effects then provides a measure of each parameter's global influence on an output feature of interest.

4.3.1 Sampling strategy

Sampling the parameter space effectively involves selecting a sufficient number (r) and spread of sample points, while minimising the number of model evaluations required for the calculation of corresponding elementary effects. Each elementary effect requires the output to be calculated at two points in parameter space (see equations (4.3) and (4.4), thus requiring the model to be evaluated two times. The most straightforward sampling strategy would require $2rk$ model evaluations, a linear function of the number of parameters being tested, k. This is an advantage of the method of elementary effects; the required number of model evaluations always varies linearly with the number of parameters.

An even smaller number of model evaluations than $2rk$ can be used to calculate the same number of elementary effects using a sampling strategy that requires only $r(k+1)$ model evaluations [63]. This strategy considerably reduces the computational expense ($r(k-1)$ fewer model evaluations required), and is especially advantageous when applied to the Ca^{2+} model, which has many parameters. The strategy, which exploits the fact that one of the 2 points used to calculate an elementary effect is used in the calculation of the next one, is explained below.

A random starting point, $\mathbf{X}^*$, is selected in parameter space, from which k other points are generated, forming a trajectory. A trajectory can be represented by a matrix, $\mathbf{B}'$, with dimension $(k+1) \times k$, whose rows are the vectors $\mathbf{X_1^*}$,$\mathbf{X_1^{(1)}}$,$\mathbf{X_1^{(2)}}$,...,$\mathbf{X_1^{(k)}}$, $\mathbf{X_2^*}$,$\mathbf{X_2^{(1)}}$,... representing consecutive points in the trajectory. k elementary effects can be calculated using this matrix if every pair of consecutive rows only differs (by Δ_i) in one unique column. An example of such a matrix is

$$\mathbf{B'} = \mathbf{J}_{k+1,k}\mathbf{X}^* + \Delta\mathbf{B}, \tag{4.6}$$

where $\mathbf{J}_{k+1,k}$ is a $(k+1)$-by-k matrix of 1's and $\mathbf{B}$ is the $(k+1) \times k$ strictly lower triangular matrix of 1's:

$$\mathbf{B} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & \dots & 0 \\ 1 & 1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix}. \tag{4.7}$$

$\mathbf{B'}$ is then given by:

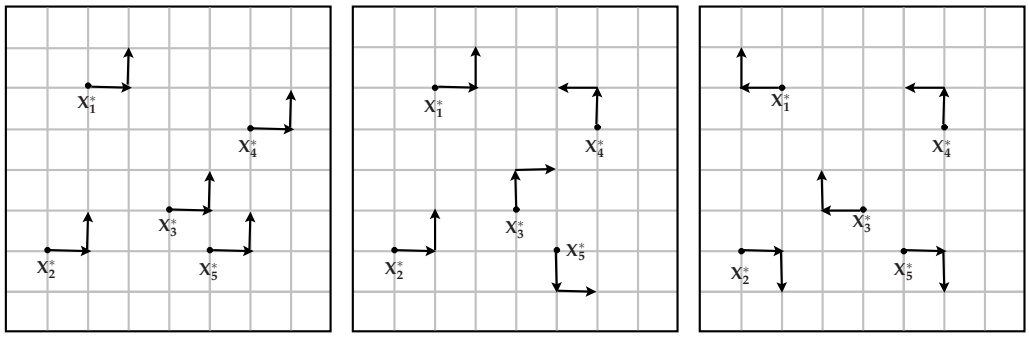
$$\mathbf{B}' = \begin{bmatrix} X_1^* & X_2^* & X_3^* & \dots & X_k^* \\ X_1^* + \Delta & X_2^* & X_3^* & \dots & X_k^* \\ X_1^* + \Delta & X_2^* + \Delta & X_3^* & \dots & X_k^* \\ X_1^* + \Delta & X_2^* + \Delta & X_3^* + \Delta & \dots & X_k^* \\ \dots & \dots & \dots & \dots & \\ X_1^* + \Delta & X_2^* + \Delta & X_3^* + \Delta & \dots & X_k^* + \Delta \end{bmatrix} \tag{4.8}$$

Note that sampling strategies can be chosen which generate points along the trajectory that decrease by Δ; that is, a parameter's value may change from X_i to $X_i - \Delta$. However, the elementary effect is always calculated as the effect of increasing a parameter by Δ; in such cases, equation (4.4) is used to calculate the elementary effect.

Given that $\mathbf{D}^*$ is a k-dimensional diagonal matrix whose elements are either $+1$ or -1 with equal probability, and $\mathbf{P}^*$, a k-by-k random permutation matrix, a randomised version of $\mathbf{B'}$ is given by:

$$\mathbf{B}^* = (\mathbf{J}_{k+1,1}\mathbf{X}^* + (\Delta/2)[(2\mathbf{B} - \mathbf{J}_{k+1,k})\mathbf{D}^* + \mathbf{J}_{k+1,k}])\mathbf{P}^*. \tag{4.9}$$

$\mathbf{D}^*$ determines, in a random manner, whether the ith parameter changing by Δ from point to point along the trajectory increases, or whether it decreases, while $\mathbf{P}^*$ randomly determines the order in which the parameters change in each trajectory. This randomisation ensures that the shape of each trajectory is randomly selected, in keeping with the statistical idea of random sampling. For illustrative purposes, the effects of the randomisation scheme are considered in a design for which the number of parameters is $k = 2$, number of grid points, $p = 9$, and number of trajectories, $r = 5$. Without the random permutation matrix, $\mathbf{P}^*$, or the random diagonal matrix, $\mathbf{D}^*$, the design is pre-determined. All trajectories from the starting points, $\mathbf{X_1^*}$-$\mathbf{X_5^*}$ would have to take



(a) (b) (c)

Figure 4.1: Possible trajectories from the starting points $\mathbf{X_1^*}$-$\mathbf{X_5^*}$ for a sampling scheme in which $k = 2$, $p = 9$, $r = 5$. (b) and (c) are possible results with the randomisation scheme, while (a) is the result without randomisation.

the same shape, as in Figure 4.1(a), but with the randomisation scheme each trajectory is able to take any of eight possible shapes, allowing the design to be randomly determined. Figures 4.1(b) and 4.1(c) show forms that the randomisation scheme can produce.

For r trajectories the matrix for the entire sample is given by:

$$\begin{bmatrix} \mathbf{B}_1^* \\ \mathbf{B}_2^* \\ \cdots \\ \mathbf{B}_r^* \end{bmatrix}. \tag{4.10}$$

Improving coverage of parameter space

Whatever the value of r chosen, trajectories can be chosen whose spread across parameter space is maximised with a strategy developed by Campolongo *et al.* [15]. A high number ($M > r$) of trajectories is generated using the above strategy, and then the subset of r trajectories with the maximum spread in the parameter space are selected, after which model evaluations can be carried out using the optimised trajectories. Thus the sampling of parameter space can be improved without increasing the number of model evaluations required.

The concept of spread is based on the Euclidean distance between a pair of trajectories, m and l:

$$d_{z,l} = \begin{cases} \sum_{j=1}^{k+1} \sqrt{\sum_{i=1}^{k} \left[X_j^z(i) - X_j^l(i) \right]^2} & \text{for } z \neq l, \\ 0 & \text{otherwise,} \end{cases} \tag{4.11}$$

where $X_j^z(i)$ is the ith coordinate of the jth point of the mth trajectory. This measure $d_{z,l}$ is used to calculate, for each possible combination of r trajectories, the quantity D, which is the square root of the sum of the squared distances $(d_{z,l})^2$ between all possible pairs of trajectories in that combination. For example, if there are $M = \{1, 2, 3, 4, 5, 6, 7, 8\}$ initial trajectories, D for the combination of trajectories $2, 3, 7$ ($r = 3$) is given by:

$$D_{2,3,7} = \sqrt{(d_{2,3})^2 + (d_{2,7})^2 + (d_{3,7})^2}. \tag{4.12}$$

The optimal sample matrix (of size $r(k+1) \times k$) for the whole design is the combination of r trajectories out of

$$\frac{M!}{(M-r)!r!} \tag{4.13}$$

possible choices with the highest value of D [81].

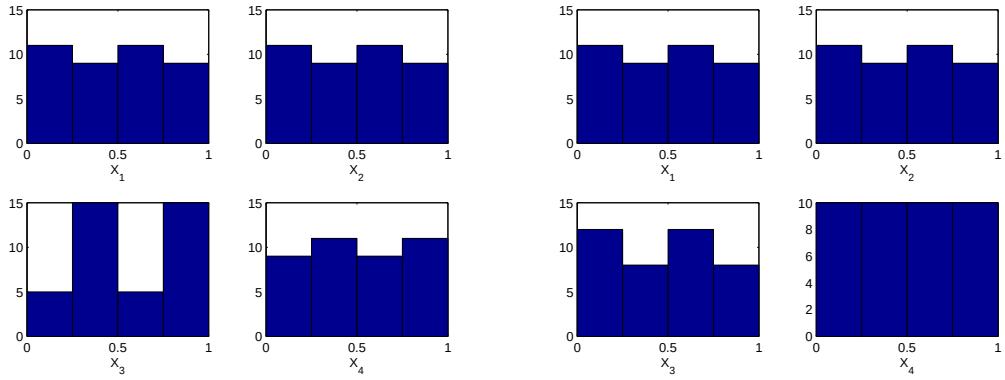
The sample of r elementary effects that are obtained per parameter will usually be few compared to the $p^{k-1}[p - \Delta(p-1)]$ number of possible effects; maximising the spread of the trajectories across parameter space is therefore important, since the trajectories should be as representative as possible. When the sensitivity analysis of the model parameters is carried out later, two different values of r will be compared in order to explore the effects of the sample size on the results obtained, and the utility of this sampling strategy.

Figure 4.2 shows the samples obtained for each X_i by the optimised sampling strategy and the original sampling strategy of Morris [63] for a scheme in which $k = 4$, $r = 20$ and $p = 4$. The optimised strategy can sometimes reduce clustering across dimensions; this can be seen in the sample for X_3, across which the spread is evidently more uniform when the optimised strategy is used.

4.3.2 Sensitivity measures

The original method of elementary effects [63] uses two measures to quantify parameter sensitivity: the average of elementary effects pertaining to the ith parameter,

$$m_i = \frac{1}{r} \sum_{j=1}^{r} EE_i^j, \tag{4.14}$$



(a) (b)

Figure 4.2: Histograms representing sample obtained for $k = 4$ parameters when $M = 1000$, $r = 20$, $p = 4$ using the (a) original Morris sampling strategy [63] (b) sampling strategy of Campolongo et al. [81]. The parameter samples have a more uniform distribution in (b).

and its standard deviation,

$$\sigma_i = \sqrt{\frac{1}{r-1}\sum_{j=1}^{r}(EE_i^j - m_i)^2}. \tag{4.15}$$

A σ_i value equal to or approximately zero, indicates that the EE_i^j values are constant or (approximately constant) over all values of the ith parameter; hence that parameter's effects are linear. A value of $\sigma_i > 0$ implies that the EE_i^j values are a nonconstant function of X_i or one or more of the other parameters ($X_j, j \neq i$). Hence, the parameter's effects are nonlinear and/or due to interaction with other parameters. The latter interpretation can be understood as follows. A high σ_i value indicates that a parameter's elementary effects are rather different from each other, which implies that its influence is dependent on its location in parameter space, and might thus be due to the particular choice of one or more other parameters.

When a parameter has inverse and non-inverse effects on the output, m_i can underestimate the parameter's significance, due to the cancellation of positive and negative values of EE_i^j. However, such a parameter would have a high σ_i value, and so, if considered together with the value of m_i, the parameter can still be identified as influential [63]. A revised version of m_i was proposed by Campolongo [15], which calculates the average of the *absolute* values of each elementary effect, serving as a single measure to quantify and rank the influence of parameters,

	High σ_i	Low σ_i
High m_i/m_i^*	Non-linear effects and/or interactions	Linear effects
Low m_i	Non-linear effects and/or interactions	Negligible effects
Low m_i^*	—	Negligible effects

Table 4.3.1: A summary of how the sensitivity measures m and σ are interpreted.

$$m_i^* = \frac{1}{r}\sum_{j=1}^{r}|EE_i^j|. \tag{4.16}$$

m_i^* identifies significant parameters even when their effect on the model output is non-monotonic, since cancellation effects cannot occur with $|EE_i^j|$. Hence, m_i^* will be the main measure used to rank the parameters in this chapter. A high value of m_i^* indicates that the parameter has considerable effects on the output, while a low value unambiguously indicates that the ith parameter has negligible effects on the output. However, m_i is still useful in some cases, where $m_i^* > m_i$ reveals that the parameter's effects occur in alternating directions (non-monotonic).

m_i is a particularly useful measure for analysing the parameters of the Ca^{2+} model, because it will be important to identify the direction in which the parameters are significant. $m_i > 0$ means that, on average, an increase in the ith parameter leads to an increase in the model output of interest, while $m_i < 0$ means that, on average, an increase in the parameter leads to a decrease in the model output. If $|m_i| = m_i^*$, then $m_i > 0$ means that, at all tested points, an increase in the ith parameter leads to an increase in the model output of interest, while $m_i < 0$ means that an increase in the parameter leads to a decrease in the model output; hence, the effects of the ith parameter are monotonic across parameter space. Parameters with opposite signed m_i values for the Ca^{2+} Time-to-Peak and Left-Shift are of special interest because changes in their values might enable the simultaneous prediction of the rapid Time-to-Peak and signal amplification, as discussed in Chapter 3.

The parameters of the Ca^{2+} model are ranked using m_i^*, while m_i and σ_i are used to give qualitative information about the effects of the parameters. Table 4.3.1 summarises how the sensitivity measures are interpreted.

4.4 Application of the EE method to the Ca^{2+} model

In the Ca^{2+}-mobilisation experiments being modelled, the agonist concentration is regulated by the experimentalist. The same experiment is repeated for n agonist concentrations $[L_1], [L_2], ..., [L_n]$, and the corresponding Ca^{2+} peak values $P_1, P_2, ..., P_n$ are used

to construct the dose-response curves (see Figure 3.3(b)). In the same way, the sensitivity analysis is carried out n times, using each agonist concentration. This makes the required number of model evaluations for the sensitivity analysis $nr(k+1)$ instead of $r(k+1)$. Hence the value of r must be even more carefully chosen, in order for the computational expense of the sensitivity analysis to stay affordable. The computational cost of implementing the method for the Ca^{2+} model will be discussed in a later section (see section 4.4.4).

The sensitivity measures (4.14) - (4.16) are then re-defined as an average over n values:

$$m_i = \frac{1}{n}\sum_{L=1}^{n}\frac{1}{r}\sum_{j=1}^{r} EE_{i,L}^{j} \tag{4.17}$$

$$m_i^* = \frac{1}{n}\sum_{L=1}^{n}\frac{1}{r}\sum_{j=1}^{r} |EE_{i,L}^{j}| \tag{4.18}$$

$$\sigma_i = \frac{1}{n}\sum_{L=1}^{n}\sqrt{\frac{1}{r-1}\sum_{j=1}^{r}(EE_{i,L}^{j} - m_{i,L})^2}. \tag{4.19}$$

The key features of the Ca^{2+} response have been identified from the data, and will be the outputs for which elementary effects are calculated. These are the Transient-Peak, the Time-to-Peak, Left-Shift of the Ca^{2+} peak dose-response curves relative to the α_{GTP} ones (indicative of signal amplification) and Peak-over-Basal. These four features can be quantified by $f = 4$ objective functions (Y_1, Y_2, Y_3, Y_4), explained below.

Ca^{2+} Peak-Over-Basal

The Ca^{2+} peaks in the data are significant because they represent a considerable rise over the basal level (see Figure 3.3(a)); hence, the Ca^{2+} peak output is considered relative to the basal response, as defined by the following:

$$Y_1 = \frac{\text{Peak} - \text{Basal}}{\text{Basal}}, \tag{4.20}$$

where 'Peak' and 'Basal' are illustrated in Figure 4.3. Considering the peak value on its own can be misleading, since a small percentage rise over a high basal value is likely to make a higher contribution to the absolute peak value than a large percentage rise over a lower basal value. Hence, the potency of an agonist is better quantified by Y_1 than the absolute peak value, as the higher the value of Y_1, the more potent the agonist. Y_1 can also be used to distinguish between types of drugs, since $Y_1 > 0$ holds for agonists, $Y_1 = 0$ for antagonists, and $Y_1 < 0$ for inverse agonists.

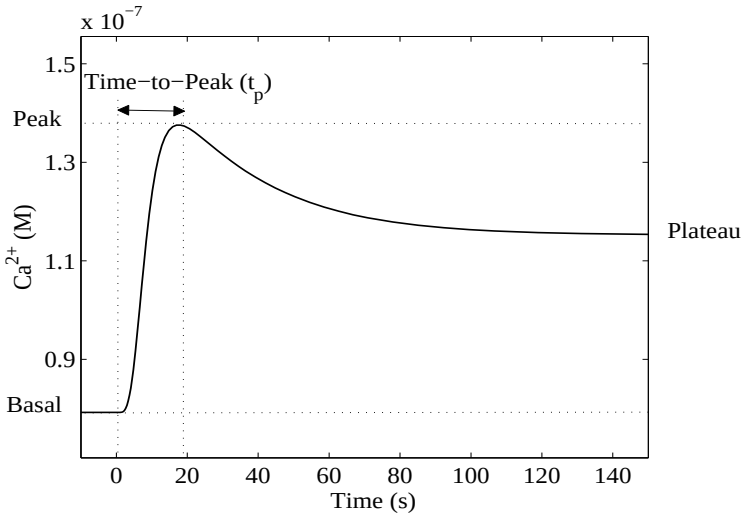


Figure 4.3: Features of the Ca^{2+} peak response which are used to define output features for the sensitivity analysis.

Peak values for which $Y_1 < 0.05$ are not taken into consideration. In such cases, 'Peak' is taken to be 0. This distinguishes between peaks that actually arise from the model, and 'peaks' that arise from small numerical errors of the ODE solver.

Ca^{2+} Transient-Peak response

The transience of the Ca^{2+} peaks in the data is also an important feature, and is quantified as:

$$Y_2 = \text{Peak} - \text{Plateau}, \tag{4.21}$$

where 'Peak' and 'Plateau' are illustrated in Figure 4.3. Y_2 indicates whether the transient peak seen in the data occurs in a model simulation. When the peak response is sustained, $Y_2 = 0$, but when it is transient, $(Peak - Plateau) > 0$.

Ca^{2+} Time-to-Peak

The time taken to reach the peak Ca^{2+} concentration after agonist stimulation is also an important feature. In the data it is rapid, $5 - 20s$, depending on $[L]$ (see Figure 3.3(a)). The Time-to-Peak,

$$Y_3 = t_p, \tag{4.22}$$

is illustrated in Figure 4.3. In those cases where $Y_1 < 0.05$, $Y_3 = 0$.

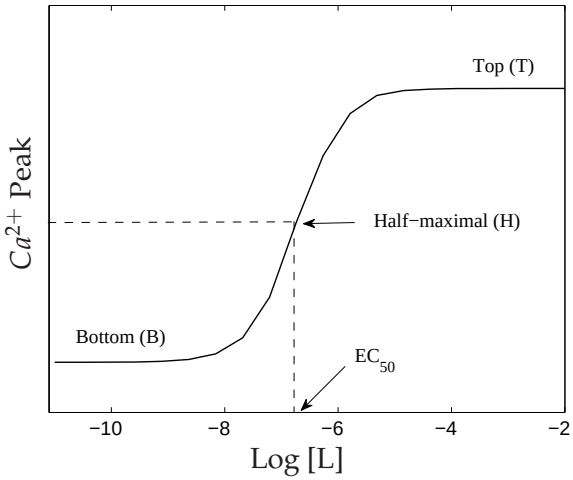


Figure 4.4: An illustration of features related to the dose-response curves, used in the quantification of the Left-Shift (see equations (4.23) and (4.25)).

Left-Shift

Signal amplification is quantified by measuring how much the dose response curve shifts from the α_{GTP} level in the pathway to the Ca^{2+} level; it is calculated by subtracting the absolute value of the EC_{50} of the α_{GTP} dose response curve from the EC_{50} of the Ca^{2+} one, as follows:

$$Y_4 = |\log(EC_{50}^{Ca^{2+}})| - |\log(EC_{50}^{\alpha_{GTP}})|, \tag{4.23}$$

where $\log(EC_{50}^{Ca^{2+}})$ is the $\log(EC_{50})$ of the Ca^{2+} peak response and $\log(EC_{50}^{\alpha_{GTP}})$ is that of the α_{GTP} peak response. Recall that the EC_{50} is the concentration of agonist that induces the peak response halfway between the baseline (B) and the maximal response (T) of the dose-response curves, as illustrated by Figure 4.4). $Y_4 > 0$, this indicates a left shift from the α_{GTP} level to the Ca^{2+} level, while a negative value implies a right shift. A left shift is an indication of signal amplification, since it means a smaller agonist concentration is required to generate the half-maximal Ca^{2+} response than the α_{GTP} one.

Y_1, Y_2 and Y_3 above are easily calculated from outputs of the model's numerical solution, but the EC_{50} values used to calculate the dose-response shift require an additional estimation process, since in many simulations, $EC_{50} \notin [L_1], [L_2], ..., [L_n]$, but will normally lie between two of the discrete values. The EC_{50} values are estimated by fitting the peak values for each parameter set (run n times, for each $[L]$ value) to the following sigmoidal function:

$$S([L_c]; B, T, EC_{50}, h) = B + \frac{T - B}{1 + (\frac{EC_{50}}{[L]})^h} \tag{4.24}$$

where $c = 1, 2, ...n$ and h is the Hill slope, which describes the steepness of the curve; this equation is standard in pharmacological practice. In pharmacology, the curve with $h = 1$ is referred to as the standard curve which describes most receptor-response systems, and is used to fit dose response curves with few points [65]; it was used to fit the M3 Ca^{2+} data, hence h is fixed at 1. The *fminsearch* function of MATLAB, which uses the Nelder-Mead simplex search method to estimate the local minimum of a function without the use of gradients [51], is used to find values that minimises the following function:

$$\sum_{c=1}^{n} (P_c - S([L_c]; B, T, EC_{50}, h))^2, \tag{4.25}$$

where P_c is the c^{th} element of the Ca^{2+} peak values, $\mathbf{P} = [P_1, P_2, ...P_c, ...P_n]$, corresponding to the c^{th} ligand value $[L_c]$. Throughout this chapter $n = 10$ ligand concentrations will be used.

The half maximal peak is given by:

$$P_{\frac{1}{2}} = \frac{T + B}{2}, \tag{4.26}$$

where B and T are illustrated in Figure 4.4. Given that $P_c < P_{\frac{1}{2}} < P_{c+1}$, either L_c or L_{c+1} is chosen as the initial guess EC_{50} for the search method, depending on whether P_c or P_{c+1} is closer in value to $P_{\frac{1}{2}}$. L_1 and L_n are chosen to ensure a wide range, so that $B = P_1$, and $T = P_n$ are expected; hence, only EC_{50} needs estimation.

4.4.1 Selection of the GSA parameters

The choices made for the EE method and sampling strategy parameters, p, δ, M and r require careful consideration in order for the sensitivity analysis to be efficiently implemented, because they influence how representative a sample of elementary effects will be.

The value of δ will depend on that selected for p. It has been shown that choosing p to be even and

$$\delta = \frac{p/2}{(p-1)}, \tag{4.27}$$

is a convenient choice [63]. This can be illustrated as follows. When $\delta = p/2(p-1)$, the [0,1] interval can be divided into two subsets of equal length $\frac{p}{2}$, as follows:

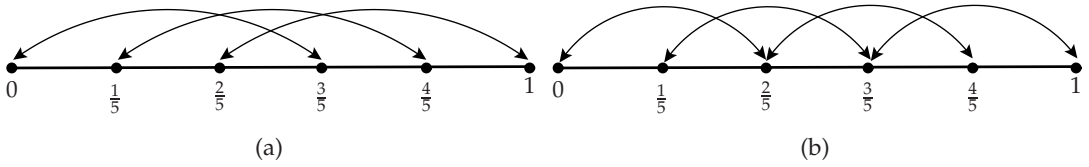


Figure 4.5: (a) A sampling scheme in which $p = 6$, and $\delta = \frac{p/2}{(p-1)} = \frac{3}{5}$ allows each point to be sampled with equal likelihood. (b) If, instead, $\delta = \frac{p/3}{(p-1)} = \frac{2}{5}$ is chosen, points $\frac{2}{5}$ and $\frac{3}{5}$ are more likely to be sampled than the others.

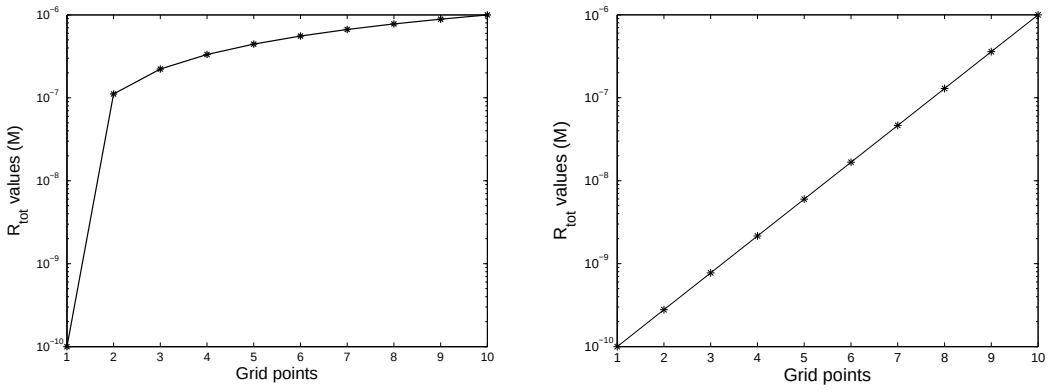
$$\begin{aligned} S_1 &= \{0, \quad 1/(p-1), \quad 2/(p-1), \quad ..., 1-\delta \quad \} \\ S_2 &= \{0+\delta, \quad 1/(p-1)+\delta, 2/(p-1)+\delta, ..., (1-\delta)+\delta \} \end{aligned} \tag{4.28}$$

It becomes clear that the pair of initial samples from $\mathbf{\Omega}$ which will be transformed into the i^{th} components of the two points in parameter space required for the calculation of any elementary effect (see equations (4.3) and (4.4)) will either be ($x_i \in S_1$, $x_i + \delta \in S_2$) or ($x_i \in S_2$ and $x_i - \delta \in S_1$). Thus all elements in S_1 and S_2 have an equal probability of being chosen.

A scheme in which $p = 6$ is used to further illustrate. Figure 4.5(a) shows the resultant spacing in the $[0, 1]$ interval. Each double-headed arrow points to a pair of possible points for the calculation of an elementary effect when $\delta = \frac{p/2}{(p-1)} = \frac{3}{5}$ is chosen; the possible pairs are $\left(0, 0+\frac{3}{5}\right)$, $\left(\frac{1}{5}, \frac{1}{5}+\frac{3}{5}\right)$, $\left(\frac{2}{5}, \frac{2}{5}+\frac{3}{5}\right)$ and there is no bias to sample any particular point more than any others. Figure 4.5(b) shows that if $\delta = \frac{p/3}{(p-1)} = \frac{2}{5}$ is chosen instead, $\frac{2}{5}$ and $\frac{3}{5}$ are more likely to be sampled than the others, illustrated by the fact that they have two arrows pointed at them.

A choice of $p = 10$ ($\delta = \frac{p/2}{(p-1)} = \frac{5}{9}$) for the Ca^{2+} model parameters affords a relatively dense grid which allows even wide ranges (e.g., four orders of magnitude) to be well explored with a log spacing (see next section). The choices made for r and M are discussed below in section 4.4.4.

The value of k is the number of parameters included in the sensitivity analysis. Some initial assumptions are made to reduce the number of parameters for consideration. The thermodynamic constants $\theta_{\nu\mu}$, $\theta_{\zeta\mu}$, $\theta_{\zeta\nu}$ are fixed and assumed to be 1. In a sensitivity analysis carried out on the cubic ternary complex model (cTCAM)'s 16 parameters, these parameters were shown to be non-influential to the G-protein activation response [41]. Also, based on discussions with experimental collaborators, $[Ca^{2+}_{AV}]$ is fixed at its value in the base parameter set, which gives an initial ER Ca^{2+} concentration in middle of the micromolar range ($4.88 \times 10^{-6} M$); the basal permeability of the ER, γ_0, chosen to be small, is also fixed, while the ER/cytosol ratio, ν_r, is also fixed at



(a) (b)

Figure 4.6: The advantage of using the log-uniform distribution (b) over the uniform distribution (a), illustrated using the values obtained with the upper and lower bounds of the parameter R_{TOT}. Note the log scale of the y-axes.

its value in the base parameter set, as published in [1]. The Ca^{2+} model has 47 parameters in total, but with the previously discussed assumptions, $k = 40$. With $p = 10$, the sampling space is then a 40-dimensional grid with 10 levels.

4.4.2 Log-uniform sampling

The ranges of uncertainty for some of the Ca^{2+} model parameters can be several orders of magnitude, therefore a log-uniform spacing is preferred over a uniform one to transform the initial $\mathbf{x}$ samples from $\mathbf{\Omega}$ into $\mathbf{X}$ values in the actual parameter space $\mathbf{A}$ [35]. The random sample $x_i \in \{0, 1/(p-1), 2/(p-1), ..., 1\}$ is transformed to a corresponding value, $X_i \in [LB_i, UB_i]$, in the ith parameter's log-uniform spaced range:

$$X_i = 10^{x_i(\log UB_i - \log LB_i) + \log LB_i}. \tag{4.29}$$

Using a range of four orders of magnitude, defined for the total receptor concentration, R_{TOT}, Figure 4.6 illustrates how the log-uniform distribution samples the parameter range more effectively than the uniform distribution, when the range is wide.

The spacing δ is also transformed from its value in the sampling space, $\mathbf{\Omega}$, into the dimensionless

$$\Delta_i = \delta(logUB_i - logLB_i), \tag{4.30}$$

whose value is thus dependent on the width of the i^{th} parameter's range.

The elementary effect of the i^{th} parameter is then given by

$$EE_i(\mathbf{X}) = \frac{[Y(X_1, X_2, ..., X_{i-1}, X_i + \Delta_i, ..., X_k) - Y(\mathbf{X})]}{\Delta_i}, \tag{4.31}$$

so that its unit is the same as that of the output Y.

4.4.3 Constraining parameter space

The particular way that the parameter space is constrained could influence the results of the sensitivity analysis. Some parameters, such as molecular concentrations, may be more likely to vary from cell to cell, requiring a wider range to be defined. Sensitivity analyses of the Ca^{2+} model parameters will be carried out in two differently constrained parameter subspaces. The effect that the different constraints have on the sensitivity analysis will then be explored by comparing results from the two subspaces, which are constrained as follows.

Parameter Subspace 1

The parameters are grouped according to their type, which determines the widths of the ranges from which they are sampled (about their values in the base parameter set, unless doing so would take them out of biologically reasonable range) as follows:

- **Molecular concentrations:** 4 orders of magnitude
- **Rate constants:** 2 orders of magnitude
- **Thermodynamic constants:** 1 order of magnitude
- **Ca^{2+} flux parameters** [1]: 4 orders of magnitude

Here, the range of values defined for molecular concentrations is wider because it can be argued that expression levels can vary from cell to cell [41], and less routinely characterised than rate constants [75]. The Ca^{2+} flux parameters, γ_1, p_1, p_2, are varied over four orders of magnitude because they have not necessarily be experimentally measured; their base values were taken from the Kang and Othmer model [37], where they were chosen to match experimental data; hence, they are considered to be more uncertain. The thermodynamic constants μ_+, μ_-, ν_+, ν_-, ζ_+ and ζ_- are only varied over one order of magnitude because there are constraints on the values their ratios $\mu = \frac{\mu_+}{\mu_-}$, $\nu = \frac{\nu_+}{\nu_-}$ and $\zeta = \frac{\zeta_+}{\zeta_-}$ can take. These ratios should be > 1; in the case of ζ, only (positive) agonists are used in the data (see section 1.3.2); in the case of ν, it is assumed that the agonist binds preferably to the pre-coupled states of the receptor

[1] These are: γ_1, p_1, p_2.

(RG, R^*G) over the uncoupled states (based on Chapter 3's results); in the case of μ, it is assumed that the receptor's inactive state is not able to bind more preferably than its active state to the G-protein. The rate constants are only varied over 2 orders of magnitude because they are less likely to vary from cell to cell. This actually allows corresponding K_D values to vary over four orders of magnitude.

Parameter Subspace 2

The ranges defined for the parameters are as follows:

- **Molecular concentrations:** 2 orders of magnitude
- **Rate constants:** 2 orders of magnitude
- **Thermodynamic constants:** 1 order of magnitude
- **Ca^{2+} flux parameters:** 2 orders of magnitude

Here, all parameters are varied over two orders of magnitude, except the thermodynamic constants (for the same reasons outlined above). Hence, all parameters except the thermodynamic constants are assumed to have the same level of uncertainty.

4.4.4 Computational expense

The size of F_i, that is, the number of possible elementary effects for the ith parameter, $p^{k-1}[p - \Delta(p-1)]$, increases with p, so presumably, the higher the value of p chosen, the higher the value of r (the sample size of F_i) should be. However, the Campolongo sampling strategy, which maximises the spread of the samples, ought to minimise the need for higher and higher values of r, by producing samples that are good representatives even at low values of r. To further explore this, two values of r, high and low, are used in two separate sensitivity analyses and the results are checked for consistency.

$M = 2000$ was chosen as a high number of initial trajectories. Sensitivity results from $r = 20$, $r = 50$, $r = 100$ and $r = 150$ trajectories were initially compared, and acceptable agreement was found between results from $r = 100$ and $r = 150$. These trials also gave an idea of the computational cost. Hence, $r_1 = 150$ and an even more thorough $r_2 = 1000$ were chosen as values with which computationally affordable analyses could be carried out, while covering parameter space well. Two sets of r_1 and r_2 trajectories were generated.

The computation time of the analysis may be divided into three: the time taken to obtain r optimised trajectories, t_{ot}, which is a function of M, r, p, k ($t_{ot}(M, r, p, k)$), the

time taken to carry out $nr(k+1)$ model evaluations, t_{me}, and the time taken to calculate the elementary effects and sensitivity measures (m^*, m, σ), t_{sm}.

The time taken to obtain $r_1 = 150$ optimised trajectories from an initial sample of $M = 2000$ on a Dell PowerEdge R610 with two Intel Xeon X5570 processors (8 processor cores total) with 48GB memory was $t_{ot}^1 \approx 5.3h$, while the time taken to generate $r_2 = 1000$ optimised trajectories was $t_{ot}^2 \approx 41.2h$ (approximately $t_{ot}^2 = \frac{t_{ot}^1 r_2}{r_1}$). The time taken to carry out the model evaluations is highly dependent on the parameter sets that are obtained with the trajectories – a few parameter sets may require the ODE solver to take far smaller time steps. For instance, for the first set of $r = 150$ trajectories, t_{me} was approximately 67.6 hours, while the second set had $t_{me} \approx 131.9$ hours. On the other hand, for the first set of $r = 1000$ trajectories, t_{me} was lower than for both $r = 150$ runs at approximately 29.2 hours, while the second set took 354.1 hours. Since the parameter sets are randomly chosen, the model evaluation times are unpredictable. The times taken to calculate the sensitivity measures, for both r_1 and r_2, were trivial.

4.5 GSA results: Influential parameters

The forty parameters for analysis are labelled $1 - 40$, as specified in Table C.1.1. In each subspace, two sets of trajectories are used to produce two sets of sensitivity results (run 1 and run 2) and the average sensitivity measures taken, so that they are given by

$$m_i = \frac{1}{2}(m_{i,1} + m_{i,2}), \tag{4.32}$$

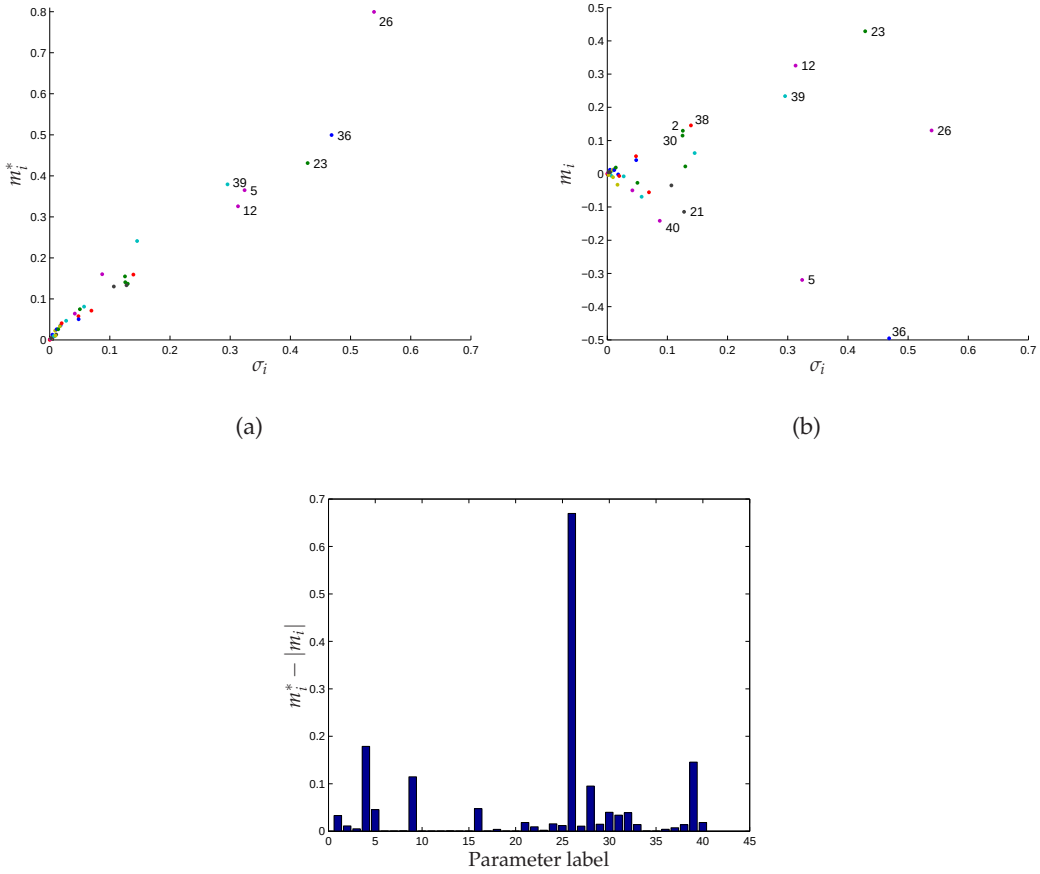
$$m_i^* = \frac{1}{2}(m_{i,1}^* + m_{i,2}^*), \tag{4.33}$$

$$\sigma_i = \frac{1}{2}(\sigma_{i,1} + \sigma_{i,2}), \tag{4.34}$$

where the subscript i, o refers to the ith parameter and the oth run.

4.5.1 Key drivers of individual features

Below, the parameters to which the individual output features are most sensitive are discussed. The results from parameter subspace 1 are considered here; they are later compared to those from parameter subspace 2. The sensitivity results obtained when $r_2 = 1000$ are analysed here, as they provide the more thorough coverage of parameter space; the results from that scheme will later be compared with results from one in which $r_1 = 150$ (see section 4.8).



(a) (b)

(c)

Figure 4.7: Sensitivity analysis results for the Ca^{2+} Peak-over-Basal. m_i^* ranks the parameters (a), the sign of m_i indicates the direction of a parameter's average effect (b), $m_i^* - |m_i| > 0$ indicates that a parameter exerts non-monotonic effects (c). σ_i is the standard deviation of the elementary effects.

Sensitivity of the Ca^{2+} Peak-over-Basal

The m_i^* values for the Ca^{2+} Peak-over-Basal are shown in Figure 4.7(a), from which it may be said that the feature is relatively insensitive to all but six parameters. These are (parameter numbers indicated before colon): the rate constant for IP_3 degradation, 26:k_{IP3deg}; the rate constant of the GPCR desensitisation reaction, 36:k_{des+}; the rate constant of PLC activation, 23:k_{PLCact}; the maximal rate of the SERCA pumps, 39:p_1; the concentration of PIP_2, 5:$[PIP_2]$; and the rate constant for GPCR deactivation, 12:k_{deact}. As mentioned previously, the Ca^{2+} Peak-over-Basal quantifies the ability of an drug to induce a response; hence changes in these parameters can alter agonist potency.

Figure 4.7(c) shows that 26:k_{IP3deg}, 36:k_{des+}, 39:p_1, 5:$[PIP_2]$ can all exert nonmonotonic effects, since $m_{26}^* - |m_{26}|$, $m_{36}^* - |m_{36}|$, $m_{39}^* - |m_{39}|$ and $m_5^* - |m_5|$ are greater than

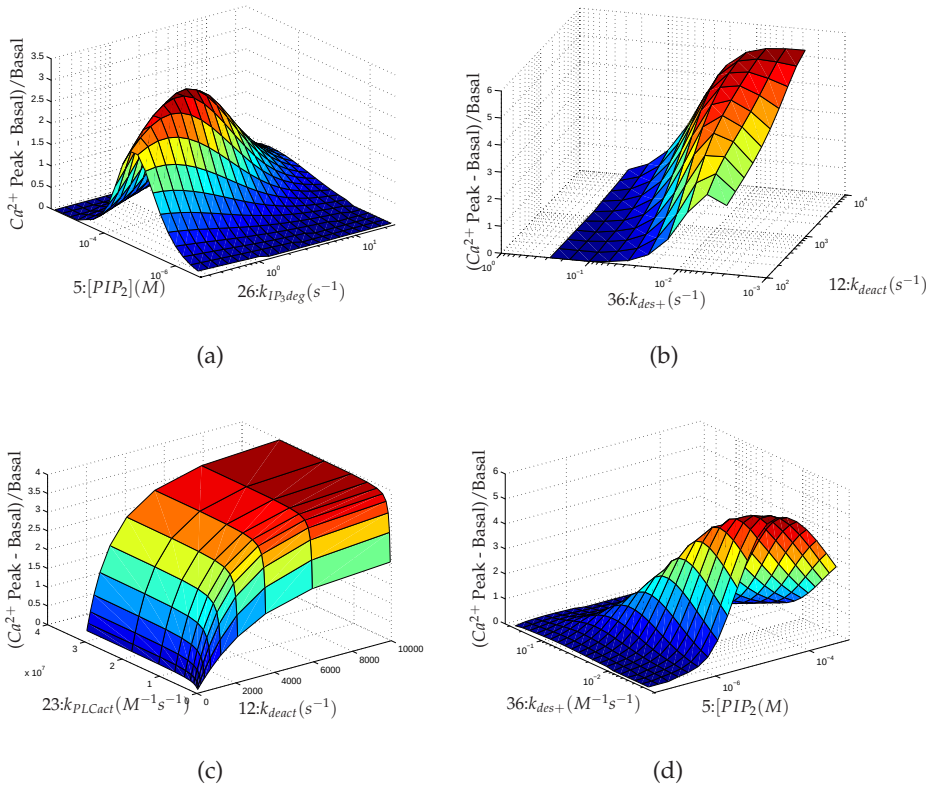
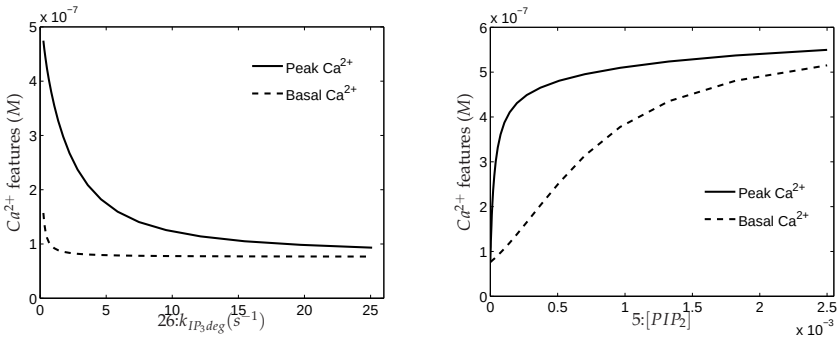


Figure 4.8: The Ca^{2+} % Peak-over-Basal changes nonlinearly in response to variations in some of the significant parameters. These parameters were selected to illustrate monotonic and biphasic nonlinearity.

0. This ability to exert either inverse or non-inverse effects may depend on the values of other parameters, and/or on the fact that they themselves act in a nonlinear fashion which is nonmonotonic. This can be seen in Figures 4.8(a), (b) and (d), where varying those four parameters from the base parameter set influences the Peak-over-Basal in a biphasic manner. This suggests that the rate of IP_3 degradation, the rate of GPCR desensitisation, the maximal rate of the SERCA pumps and the concentration of PIP_2 all have optimal values (or ranges of values), outside which the nature of agonist efficacy can alter dramatically. These parameters might be able to determine whether a drug acts as an agonist (Ca^{2+} Peak-over-Basal > 0) or antagonist (Ca^{2+} Peak-over-Basal $= 0$).

The reason that these parameters act in a biphasic fashion may be understood by considering their effects on the peak and basal separately, which are shown in Figure 4.9, where 26:k_{IP3deg} and 5:$[PIP_2]$ are used to illustrate. In Figure 4.9(a), both the peak and basal decrease monotonically as 26:k_{IP3deg} increases. Initially the relative change in peak is greater than the relative change in basal, which causes the Peak-over-Basal to increase; however, the basal level soon becomes constant, while the peak continues



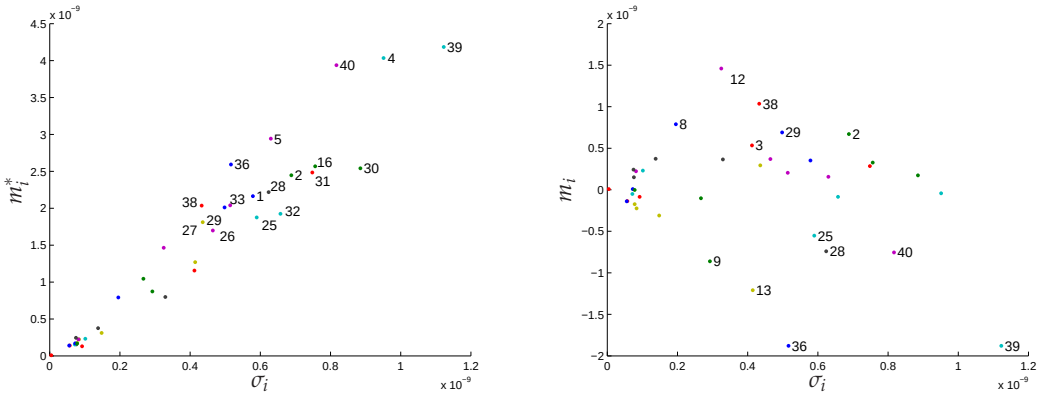
(a) (b)

Figure 4.9: The individual effects on the Ca^{2+} Peak and Basal of varying (a) 26:k_{IP3deg}, (b) 5:$[PIP_2]$.

to decrease, hence Peak-over-Basal starts to decrease. In Figure 4.9(b), both the peak and basal increase monotonically as 5:$[PIP_2]$ increases, and initially the relative change in peak is greater than the relative change in basal, causing the Peak-over-Basal to increase; however, the relative change in peak becomes small before the relative change in basal, causing the Peak-over-Basal to start to decrease.

12:k_{deact} exerts increasing monotonic effects at all points tested across parameter space, since $m_{12}^* - |m_{12}| = 0$ (see Figure 4.7(c)). When varied in the base parameter set, 12:k_{deact}, as well as, 23:k_{PLCact}, exert increasing monotonic effects on the Peak-over-Basal (Figure 4.8(c)), meaning that the reactions they are involved in are important to maintaining the Peak-to-Basal ratio. In other words, they are important determinants of an agonist's potency. 23:k_{PLCact} is the rate constant for the reaction in which Ca^{2+} activates PLC (see (3.5)), so that it can go on to hydrolyse IP_3 from PIP_2, leading to Ca^{2+} release. This Ca^{2+} release exerts a positive feedback, since it leads to the activation of more PLC, which will be more pronounced after the addition of the agonist. Hence the peak is more significantly influenced, while the basal could remain relatively constant.

12:k_{deact} is the rate constant for the reactions in which the active GPCR species are deactivated (see reactions (1.9), (1.12), (1.17), (1.18)); it determines the ratio of inactive to active GPCR's at equilibrium. The sensitivity of the Ca^{2+} Peak-over-Basal to k_{deact} suggests that an agonist's potency increases as the ratio of inactive to active GPCR's at equilibrium does. This is to be expected, as a high ratio of active to inactive GPCR's would mean that the system could generate a high response in the absence of agonist, diminishing the agonist's role. Hence the lifetime of the active GPCR species may also be a determinant of agonist potency.



(a) (b)

Figure 4.10: Sensitivity analysis results for the Ca^{2+} Transient-peak. m_i^* ranks the parameters (a), the sign of m_i indicates the direction of a parameter's average effect (b). σ_i is the standard deviation of the elementary effects. Several parameters are identified as significant.

Sensitivity of the Ca^{2+} Transient-Peak

Unlike the Ca^{2+} Peak-over-Basal, the Ca^{2+} Transient-Peak only has a few parameters to which it is relatively insensitive (Figure 4.10(a)). This means that, though there may be dominant mechanisms which control the Transient-Peak, the system has other mechanisms that can significantly contribute. The significant parameters are a combination of parameters from the G-protein activation cycle, and from further downstream in the pathway.

The most influential parameter is 39:p_1, the maximal rate of the SERCA pumps, which is unsurprising, since it is the maximal rate at which Ca^{2+} can be removed from the cytosol back into the ER, potentially making the Ca^{2+} peak transient. 40:p_2, as the Ca^{2+} concentration at which half-maximal pumping occurs, can also exert a similar influence by determining how sensitive the pumps are. Being two of the three most sensitive parameters, they imply that the mechanism of the SERCA pumps is crucial to the Transient-peak response.

Cell-specific parameters are also influential, such as: the total IP_3R concentration, 4:IP_3R_{TOT}; the total PIP_2 concentration, 5:$[PIP_2]$; the total G-protein concentration, 2:G_{TOT}; and the total receptor concentration, 1:R_{TOT}. Differences in their expression levels can occur from cell to cell [41], even in the same line, which might explain differences in Ca^{2+} time course profiles, when taken from individual cells in the same experiment [27].

Rate constants of several reactions are also significant. 36:k_{des+} affects the rate at

which desensitisation occurs; desensitisation is responsible for reducing the number of active receptor-G-protein complexes available to signal, hence it has a role in making the IP_3 response, and consequently the Ca^{2+} response, transient. Recall that PKC mediates desensitisation, so its dynamics would influence the shape of the transient. The rate constants of IP_3-IP_3R binding and unbinding, 30:k_{IP_3Rb+}, and 31:k_{IP_3Rb-} and the rate constant for the binding and unbinding of activating Ca^{2+}, 32:$k_{Ca^{2+}act+}$ and 33:$k_{Ca^{2+}act-}$, respectively are also significant. This means that the dynamics of IP_3-IP_3R binding, as well as positive and negative Ca^{2+} feedback in CICR can also contribute to the transience of the Ca^{2+} signal.

α_{GTP} dynamics are also important. The rate of α_{GTP} production is an important factor (as its rate constant, 16:k_{GTP+} is among the six most significant parameters), while the lifetime of α_{GTP} is also a relatively important factor, since the rate constant of the PLC-mediated α_{GTP} hydrolysis reaction, 25:k_{PLCdis}, has a significant m^* value.

Some of the parameters which strongly effect the Transient-peak also significantly influence the Peak-over-Basal response (5:$[PIP_2]$, 26:k_{IP3deg}, 39:p_1 and 36:k_{des+}). 23:k_{PLCact}, and 12:k_{deact} are not among the most significant parameters for the Transient peak, which means that they control the Ca^{2+} Peak-over-Basal without significantly altering the Transient-Peak.

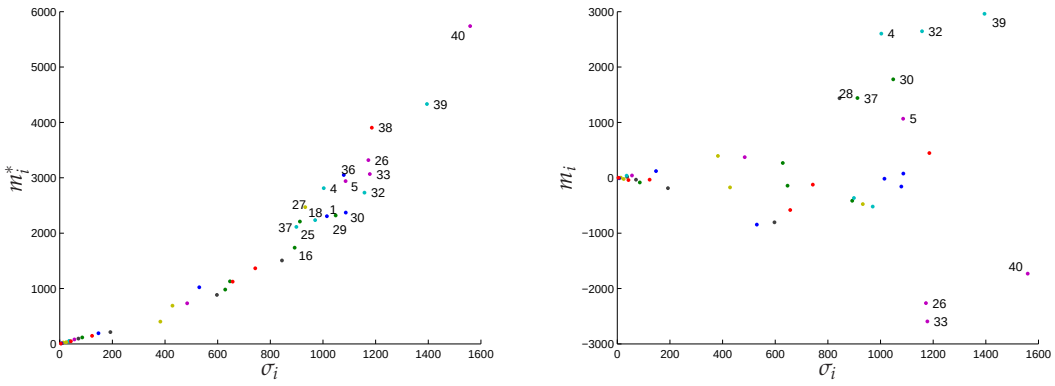
Sensitivity of the Ca^{2+} Time-to-Peak

Figure 4.11 shows the m_i^* values for the Ca^{2+} Time-to-Peak; this feature, like the Transient-Ca^{2+}-Peak, can only be considered insensitive to a few parameters (see Figure 4.11a). This means that though there are dominant mechanisms which control the timing of the peak, the system has other mechanisms that may significantly contribute. 40:p_2 and 39:p_1 are the most influential parameters, suggesting that the dominant mechanism controlling the Ca^{2+} Time-to-Peak is the action of the SERCA pumps.

The sensitivity of the IP_3 receptor to IP_3, quantified by 38:γ and the binding rate of IP_3, 30:k_{IP_3Rb+} are significant. The binding and unbinding rates of activating Ca^{2+} (32:$k_{Ca^{2+}act+}$, and 33:$k_{Ca^{2+}act-}$, respectively) are also significant, suggesting that the dynamics of the IP_3 receptor are also key to the timing of the peak.

The rates of IP_3 degradation, and of GPCR desensitisation, quantified by 26:k_{IP3deg} and 36:k_{des+} respectively, can affect the timing of the Ca^{2+} peak by preventing further IP_3-mediated Ca^{2+} release, through direct degradation of IP_3 or desensitisation of active GPCR's respectively.

Cell-specific parameters like the total IP_3R concentration, 4:IP_3RTOT, total PIP_2 concentration, 5:$[PIP_2]$ and the total receptor concentration, 1:R_{TOT} are also influential. As previously mentioned, differences in their expression levels can occur from cell to



(a) (b)

Figure 4.11: Sensitivity analysis results for the Ca^{2+} Time-to-peak. m_i^* ranks the parameters (a), the sign of m_i indicates the direction of a parameter's average effect (b). σ_i is the standard deviation of the elementary effects. Several parameters have significant effects on this feature.

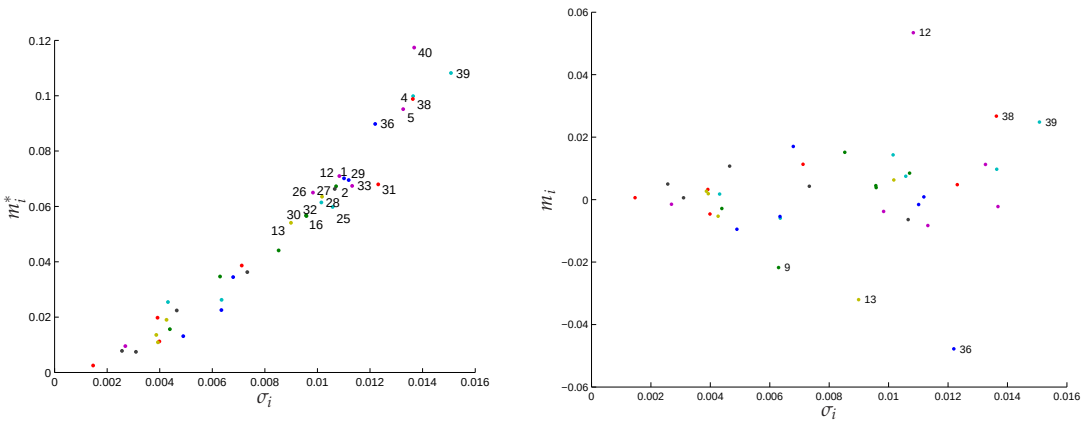
cell, and might thus explain differences in Ca^{2+} time courses, when taken from *individual* cells in the same experiment.

As with the Transient-Peak, hydrolysis of α_{GTP} can also be an important mechanism, either mediated by α_{GTP} itself (via 18:k_{gd+} in reaction 1.24) or PLC (via 25:k_{PLCdis} in reaction 3.8).

Sensitivity of the Left-Shift

Figure 4.12(a) shows the ranking of the parameters for the Left-Shift. The dominant mechanism in signal amplification is also the action of the SERCA pumps, as 40:p_2 and 39:p_1 are the highest ranked parameters. This illustrates how signal amplification is system- and not just ligand-dependent. m_{40} and m_{39} are negative, which means that on average, 40:p_2 and 39:p_1 exert inverse effects on the Left-Shift. Hence, regardless of the strength of signal applied, if the SERCA pumps are too sensitive, Ca^{2+} release can be strongly inhibited, thereby reducing the agonist's ability to generate a response irregardless of the agonist's *intrinsic* efficacy.

The dynamics of constitutive GPCR activity are also shown to be significant, since 13 : k_{act} and 12 : k_{deact} are influential. Together, these parameters determine the ratio of active to inactive GPCR's at equilibrium, implying that signal amplification is highly dependent on this ratio. This is in agreement with the sensitivity results for the Peak-over-Basal, where it was seen that the agonist potency depended on the ratio of inactive (R) to active GPCR's (R^*).



(a) (b)

Figure 4.12: Sensitivity analysis results for the Left-Shift, which is quantifies signal amplification, using equation (4.23). m_i^* ranks the parameters (a), the sign of m_i indicates the direction of a parameter's average effect (b). σ_i is the standard deviation of the elementary effects.

The total receptor concentration, 1:R_{TOT}, and the total G-protein concentration, 2:G_{TOT}, are also significant parameters. These parameters can be experimentally manipulated, indicating that these results can be readily tested.

Nonlinearity and interaction among parameters

Across all the output features, the relationship between m_i^* and σ_i is always approximately linear, which indicates that the larger a parameter's influence, the more nonlinear it is, and/or due to interaction with other parameters (see Figures 4.7(a), 4.10(a), 4.11(a), 4.12(a)). Hence, none of the influential parameters exert linear effects on the Ca^{2+} Peak-over-Basal, Transient Ca^{2+} Peak, Ca^{2+} Time-to-Peak or the Left-Shift. This highlights the importance of using a global sensitivity analysis over a local sensitivity analysis (LSA), since an LSA's ability to identify parameters of significance is highly dependent on what point in parameter space is analysed. Of course, if the parameters represented by that point are known with certainty, then an LSA should be the method of choice, as its computational expense is relatively trivial.

4.5.2 Key drivers of the overall Ca^{2+} response

A useful way to quantify the overall influence of the parameters is to define a representative sensitivity measure, across all the output features, for each parameter. Given that the vector of m^* values for the f^{th} objective function is given by $\boldsymbol{m_i^{*f}}$, the *average*

Ranking	Parameter	Description	S_{AR}
1	40:p_2	$[Ca^{2+}]$ for half-maximal rate of SERCA	0.76
2	39:p_1	Maximal rate of SERCA	0.75
3	36:k_{des+}	On-rate for GPCR desensitisation	0.72
4	4:IP_3R_{TOT}	Total IP_3R concentration	0.72
5	26:k_{IP_3deg}	Rate constant for IP_3 degradation	0.63
6	5:$[PIP_2]$	Total $[PIP_2]$ concentration	0.62
7	33:$k_{Ca^{2+}act-}$	Off rate for IP_3R-activating Ca^{2+}	0.57
8	38:γ_1	Sensitivity of IP_3R	0.57
9	28:k_{PIP2b+}	On-rate for PIP_2 and $\alpha_{GTP}PLC^*$	0.53
10	27:k_{IP_3}	Rate constant for IP_3 hydrolysis	0.52

Table 4.5.1: The ten most significant parameters of the Ca^{2+} model, as quantified by S_{AR} (see equation (4.36))

relative sensitivity of the i^{th} parameter is given by,

$$S'_{AR} = \frac{1}{q}\sum_{f=1}^{q}\frac{m_i^{*f}}{\max(\boldsymbol{m}^{*f})}, \tag{4.35}$$

which is adapted from [17]. $q = 4$, since four features are considered in the sensitivity analysis. S'_{AR} is a dimensionless measure, and is used to rank the parameters of the Ca^{2+} model according to significance. The maximum value that a parameter can score is 1, which would be the case if it was the most significant for *all* features.

S'_{AR} values can be calculated for both parameter subspace 1 (S'_{AR_1}) and subspace 2 (S'_{AR_2}). An average of the two,

$$S_{AR} = \frac{S'_{AR_1} + S'_{AR_2}}{2}, \tag{4.36}$$

is used to provide the main overall ranking for the model. The individual rankings for each subspace will be compared in section 4.5.3.

The ten parameters with the highest score (S_{AR}) are listed in Table 4.5.1, identified as the key parameters controlling the Ca^{2+} response and signal amplification. The complete list of all parameters and their average relative sensitivity scores can be found in Table C.2.1 in the appendices.

Investigating the amplified, yet rapid Ca^{2+} response

In this section, the mechanisms which might be responsible for the amplified, yet rapid Ca^{2+} response, are investigated by considering the particular effects of the ten highest ranked parameters on the Left-Shift and the Ca^{2+} Time-to-Peak. The ten parameters can be divided into those that have monotonic increasing effects on both features, shown in Figure 4.13; those that have monotonic decreasing effects on both, shown in

Figure 4.14; and those which have biphasic effects on either feature, shown in Figures 4.16 and 4.15.

The SERCA pump parameters 40:p_2 and 39:p_1 have similar sensitivity scores and are the most significant parameters. This is because the SERCA pumps play a direct role in determining the rate at which Ca^{2+} is removed from the cytosol back into the ER, and so they influence the timing and magnitude of the (basal and peak) Ca^{2+} response. Increasing the 40:p_2 means that the concentration of Ca^{2+} required for half-maximal pumping increases; this makes the pumps less sensitive and the Ca^{2+} re-entry into the ER slower, increasing both the Left-shift and the Time-to-peak (Figures 4.13(a) and 4.13(b)). On the other hand, increasing 39:p_1, the maximal rate of the SERCA pumps decreases both the Left-shift and the Time-to-peak (Figures 4.14(a) and 4.14(b)). This is because the pumping capability increases with 39:p_1, hence Ca^{2+} can be transported more quickly back into the ER, decreasing both the peak and Time-to-Peak.

Figures 4.13(c) and 4.13(d) show that increasing 4:IP_3R_{TOT} increases the Left-Shift, as well as the Ca^{2+} Time-to-Peak. The increase in the Left-Shift is explained by the fact that increasing 4:IP_3R_{TOT} increases the probability of forming IP_3IP_3R and the activated (open channel) state, $IP_3IP_3RCa^{2+}$, thus increasing Ca^{2+} release. However, it takes more time to bring down the increased Ca^{2+} peaks, probably because over time, far more receptors enter the $IP_3IP_3RCa^{2+}$ state than leave it.

The Ca^{2+} Time-to-Peak and the Left-Shift also increase with 38:γ_1 (Figures 4.13(e) and 4.13(f)). The increase in the Left-Shift is due to a higher Ca^{2+} release, because of the increased contribution of the source term involving 38:γ_1 to the Ca^{2+} ODE (see Equation (3.13)). It then takes longer for the action of the SERCA pumps to bring down the Ca^{2+} peak.

5:$[PIP_2]$ is a participant in reaction (3.6), which produces $\alpha_{GTP}PLC^*PIP_2$; therefore, as $[PIP_2]$ increases, so does $[\alpha_{GTP}PLC^*PIP_2]$; this leads to the production of more IP_3, and consequently a higher Ca^{2+} release and greater Left-Shift (Figure 4.13(g)). However, Figure 4.13(h) shows that the Ca^{2+} Time-to-peak increases in the process. Increasing another top ten parameter, 28:k_{PIP2b+}, has similar effects on the Left-Shift and the Time-to-Peak, since it is the rate constant for the binding of $[PIP_2]$ (which is assumed constant) and $\alpha_{GTP}PLC^*$ in reaction (3.6).

The identification of $[PIP_2]$ among the most significant parameters might have implications for the assumption in the model that $[PIP_2]$ is constant; in reality, $[PIP_2]$ is depleted but replenished, but there is no evidence that it is rate-limiting [104]. However, since it has been identified as one of the most important parameters, the cost of this assumption might be reconsidered, by incorporating the dynamics of $[PIP_2]$ re-synthesis into the model, and considering the effects on the Ca^{2+} response; this will be

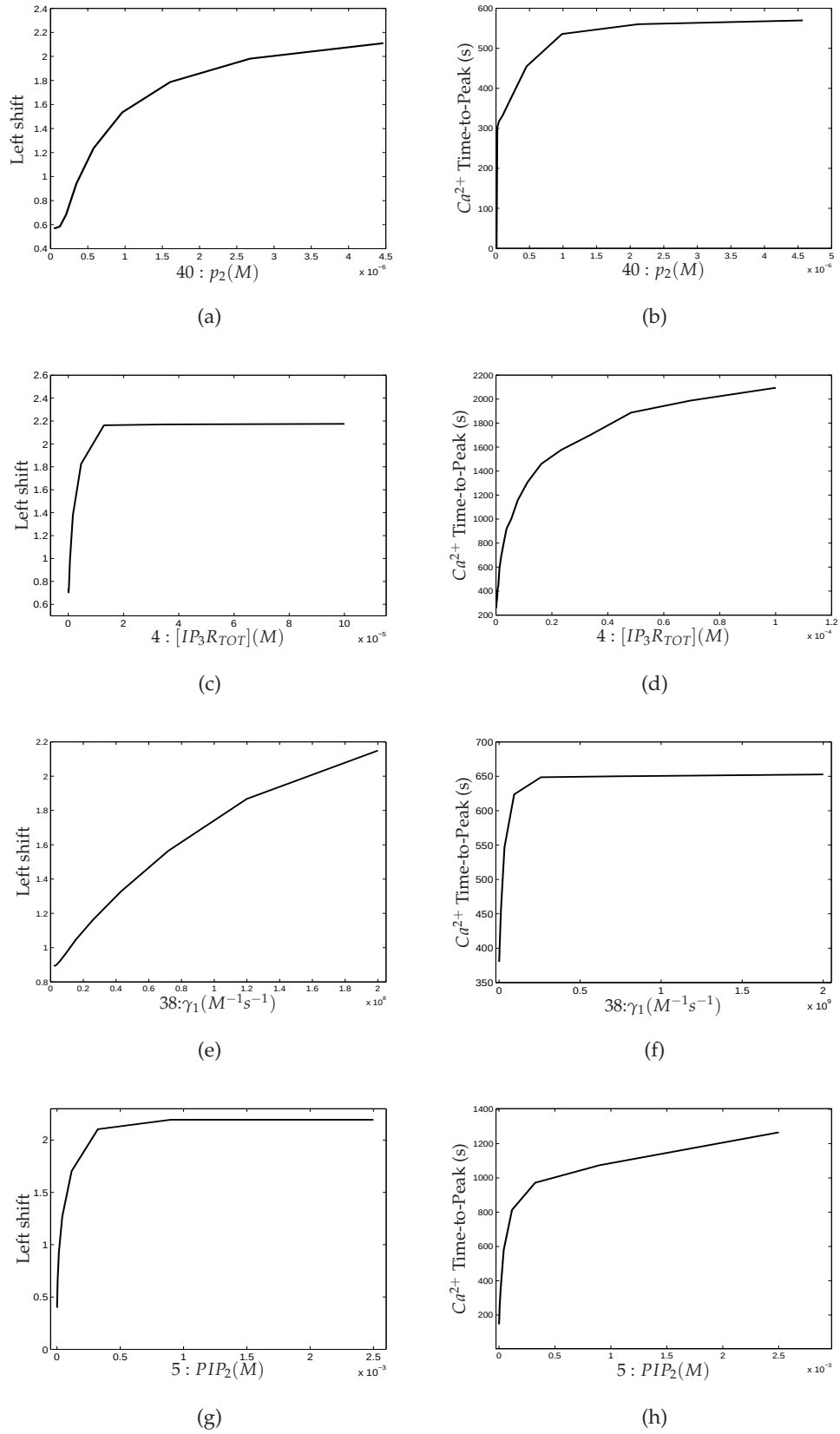


Figure 4.13: Varying significant parameters which have similar effects on the Ca^{2+} Time-to-Peak and signal amplification. Increasing 40:p_2, 5:$[PIP_2]$, 4:IP_{3Rtot} and 38:γ_1 increases the Left-Shift (1^{st} column, (a,c,e,g)), but also the Time-to-Peak (2^{nd} column, (b,d,f,h)).

done in Chapter 5.

Increasing the IP_3 degradation rate constant, 26:k_{IP_3deg}, decreases the Left-Shift and also shortens the Ca^{2+} Time-to-Peak (Figures 4.14(c) and 4.14(d)). This is as would be expected; increasing 26:k_{IP_3deg} heightens the rate of IP_3 degradation, and so Ca^{2+} release, which is mediated through IP_3, is inhibited, hence the decrease in the Left-Shift. The Time-to-Peak is also shortened, since the signal for Ca^{2+} release (IP_3), is terminated more quickly.

When 33:$k_{Ca^{2+}act-}$, the rate constant for the dissociation of activating Ca^{2+} and IP_3R (reaction (3.11)), is increased, the Left-Shift and the Time-to-Peak are increased as well. This is because increasing the rate of this reaction shortens the lifespan of the activated IP_3 receptor ($IP_3IP_3RCa^{2+}$); hence, Ca^{2+} release is inhibited, causing the decrease in the Left-Shift and Time-to-Peak.

36:k_{des+}, although a G-protein level parameter, is one of the four parameters to which the Ca^{2+} response is most sensitive (see Table 4.5.1). It is the rate constant for the reactions in which the active GPCR's are desensitised, thereby reducing the number available to signal for IP_3-mediated Ca^{2+} release. As a result, a higher 36:k_{des+} decreases the Ca^{2+} Time-to-Peak, as can be seen in Figure 4.15(b). The effect on the Left-Shift, however, is biphasic. Initially, it rapidly decreases as 36:k_{des+} increases, up to a threshold value, after which it starts to increase slowly. The Ca^{2+} Time-to-Peak decreases up to a minimal value of $\approx 200s$, so changing 36:k_{des+} cannot account for the rapid, yet amplified response.

Varying the parameter 27:k_{IP_3} influences the Ca^{2+} Time-to-Peak and Left-Shift differently than varying the other parameters. Increasing 27:k_{IP_3} has a monotonic increasing effect on the Left-Shift and a biphasic effect on the Time-to-Peak (Figure 4.16). Initially, the Ca^{2+} Time-to-Peak increases with 27:k_{IP_3}, until a threshold value, after which it begins to decrease, following the IP_3 Time-to-Peak.

Since no minimal value is reached for the Time-to-Peak within 27:k_{IP_3}'s defined range for the GSA, the Time-to-Peak can decrease further if 27:k_{IP_3} is increased beyond that range. Thus, increasing the rate of IP_3 production can potentially account for the Ca^{2+} response seen in the data, in which the signal is amplified to produce a maximal response, without delaying the Time-to-Peak. This will be explored in a later section.

Non-influential parameters

Not only is it useful for the sensitivity analysis to identify influential parameters, the non-influential parameters that come to light also provide important insight. The relatively insignificant parameters identified might not require further analysis, effectively reducing parameter space by many dimensions. In addition, they are parameters that

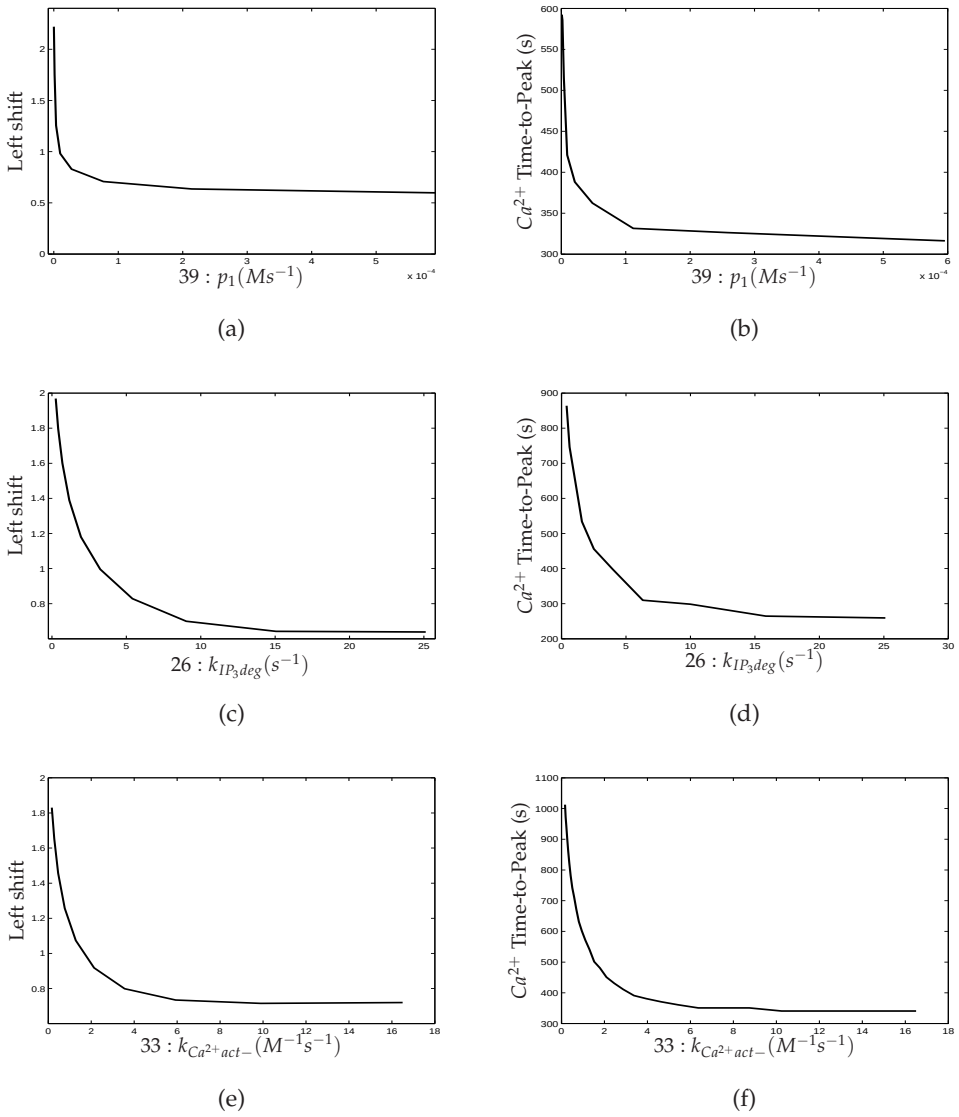
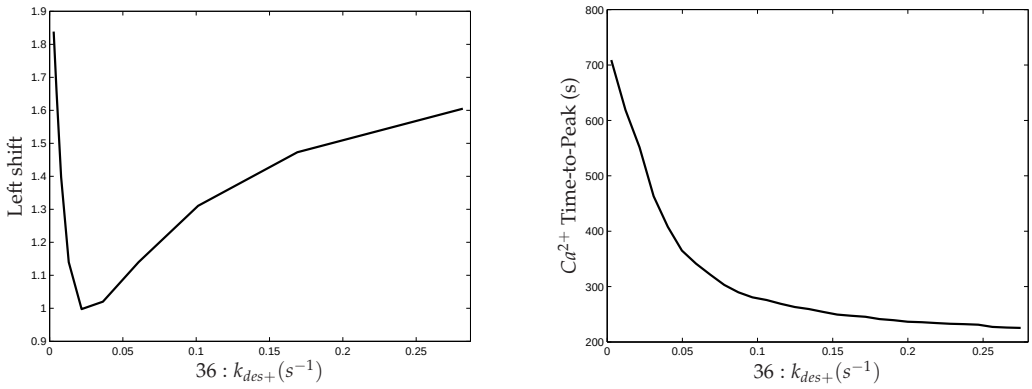


Figure 4.14: Varying significant downstream parameters which have similar effects on the Ca^{2+} Time-to-Peak and signal amplification. Increasing 39:p_1, 26:k_{IP3deg} and 33:$k_{Ca^{2+}act-}$ decreases the Time-to-Peak (1^{st} column, (a,c,e)), but also the Left-Shift (2^{nd} column, (b,d,f)).



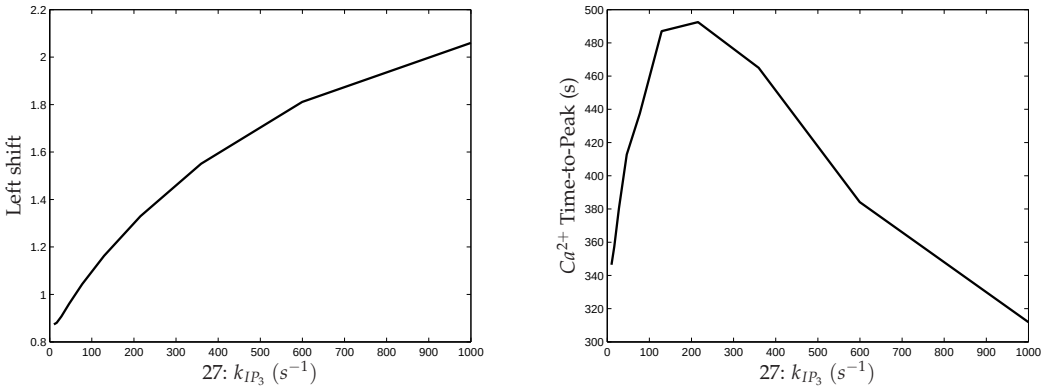
(a) (b)

Figure 4.15: The effects of varying the significant G-protein level parameter 36:k_{des+} on the Ca^{2+} Time-to-Peak and signal amplification. Increasing k_{GTP+} exerts a biphasic effect on the Left-Shift, while reducing the Time-to-Peak.

experimentalists can consider as not crucial to measure, or test as drug targets. For instance, the concentration of PLC, 3:$[PLC_{TOT}]$, is not an easily measured parameter, but since it has been identified as a relatively noninfluential parameter, it can be concluded that it is not crucial to measure. The range defined for 3:$[PLC_{TOT}]$ is $0.8 \times 10^{-6} - 0.8 \times 10^{-6} M$; as PLC works by binding to α_{GTP}, which is converted from G, it is likely that PLC is in excess if only small amounts of α_{GTP} are formed, hence variations in $[PLC_{TOT}]$ would not matter much. These results suggest that PLC might not be a worthwhile drug target, and it might be possible to test this hypothesis experimentally, by making use of available PLC inhibitors [73, 90].

The ten least significant parameters of the Ca^{2+} model are listed in Table 4.5.2. It is informative to investigate why the rate constants identified are insensitive, and what can consequently be learnt about the reactions in which they are involved. For instance, are any of these parameters insensitive because relevant reactions reach equilibrium quickly? Are their forward reactions so fast that they are essentially irreversible?

15:k_{g+} and 14:k_{g-} are the rate constants for receptor pre-coupling (reaction (1.13)), 19:k_{lb+} and 20:k_{lb-} are the rate constants for ligand binding (reaction 1.10), 34:$k_{Ca^{2+}inh+}$ and 35:$k_{Ca^{2+}inh-}$ are the rate constants for the binding and unbinding of inhibitory Ca^{2+} to IP_3R respectively (reaction (3.12)), and 24:$k_{PLCdeact}$ is the rate constant for the deactivation of PLC (reaction (3.5)). Simulations show that none of these reactions reach equilibrium quickly (see Figure C.1 in the appendices), and that setting 14:$k_{g-} = 0$ and 24:$k_{PLCdeact} = 0$ has no discernable effect on the response (see Figure 4.17), which might mean that receptor pre-coupling and the activation of PLC by Ca^{2+} are irreversible pro-



(a) (b)

Figure 4.16: The effect of varying the significant downstream parameter 27:k_{IP_3} on the Ca^{2+} Time-to-Peak and signal amplification. Increasing 27:k_{IP_3} increases the Left-Shift, while exerting a biphasic effect on the Time-to-Peak.

Ranking	Parameter	Overall score
1	k_{RA+}	0.01
2	k_{g+}	0.04
3	k_{g-}	0.05
4	k_{lb-}	0.06
5	k_{lb+}	0.06
6	$k_{PLCdeact}$	0.07
7	ζ_-	0.07
8	ν_-	0.07
9	$k_{Ca^{2+}inh+}$	0.08
10	$k_{Ca^{2+}inh-}$	0.08

Table 4.5.2: The ten least significant parameters of the Ca^{2+} model.

cesses, at least when the base parameter set is used; the model might thus be slightly simplified by excluding their backward reactions.

The insensitivity of the Ca^{2+} response to the ligand binding parameters means that the Ca^{2+} features are robust to changes in their values. Recall that the ranges of ζ_+ and ζ_- were restricted so that they only represented positive agonist values (see section 4.4.3). This result only means that positive agonists produce similar Ca^{2+} responses; if one agonist produces a particular response pattern, then changes in k_{lb+}, k_{lb-} and ζ_- do not significantly alter that pattern, due to signal amplification. This is in agreement with the data, where four different agonists induce the maximal response (see Figure 3.3(b)). However, it is important to note that the location of the Ca^{2+} dose-response curve depends on the values of k_{lb+}, k_{lb-} and ζ_- (see Figures 3.3(b) and 3.20(b)). This is probably not reflected in the sensitivity analysis results because the Left-Shift only

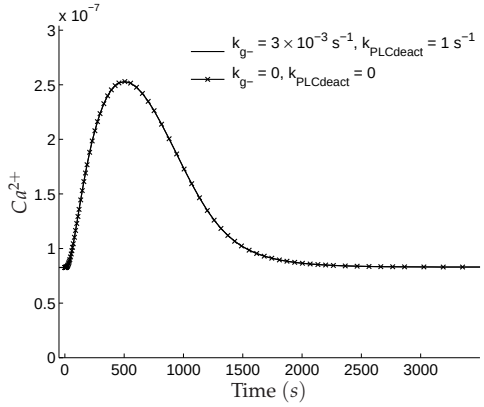


Figure 4.17: G-protein pre-coupling (reactions (1.13)-(1.16)) and Ca^2 activation of PLC (reaction (3.5)) appear to be effectively irreversible reactions, as setting the backward rates, k_{g-} and $k_{PLCdeact}$ respectively, to 0 has no distinguishable effect on the Ca^{2+} response.

quantified the location of the Ca^{2+} dose-response curve relative to the α_{GTP} one.

4.5.3 Comparing results from parameter subspaces 1 and 2

Figure 4.18 compares the parameter sensitivity rankings from the runs carried out in Subspace 1 and Subspace 2, and it can be seen that there is reasonable correlation between the rankings. The box in the bottom left contains parameters that were identified in the top 10 of both subspaces, and there are six such parameters. The rate constants 27:k_{IP_3}, 28:k_{PIP2b+}, 31:k_{IP_3Rb-} and 32:$k_{Ca^{2+}act+}$, are identified in the top ten significant parameters of Subspace 2, where all parameters, apart from the thermodynamic constants, are varied over two orders of magnitude, but not in Subspace 1, where the molecular concentrations and Ca^{2+} flux parameters are varied over four orders of magnitude (see section 4.4.3). On the other hand, the molecular concentrations, 1:R_{TOT} and 5:$[PIP_2]$, the rate constant 30:k_{IP_3Rb+}, and the Ca^{2+} flux parameter, 38:γ_1 are identified among the top ten significant parameters of Subspace 1, but not Subspace 2.

The total receptor concentration, 1:R_{TOT}, is a parameter that can vary from cell to cell, and the fact that it is higher ranked in Subspace 1 (where it has a wider range) may imply that the Ca^{2+} response is more affected when receptor expression levels vary widely. This implies that if receptor expression levels are experimentally manipulated to be approximately uniform across cells, time courses for individual cells might also be more uniform.

From the base parameter set, increasing 1:R_{TOT} increases the Left-Shift, and has the previously discussed biphasic effect on the Time-to-Peak; this is illustrated in Figure

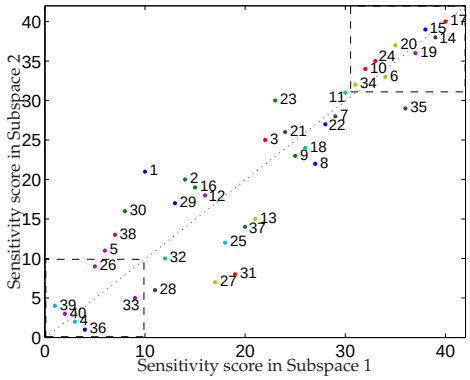
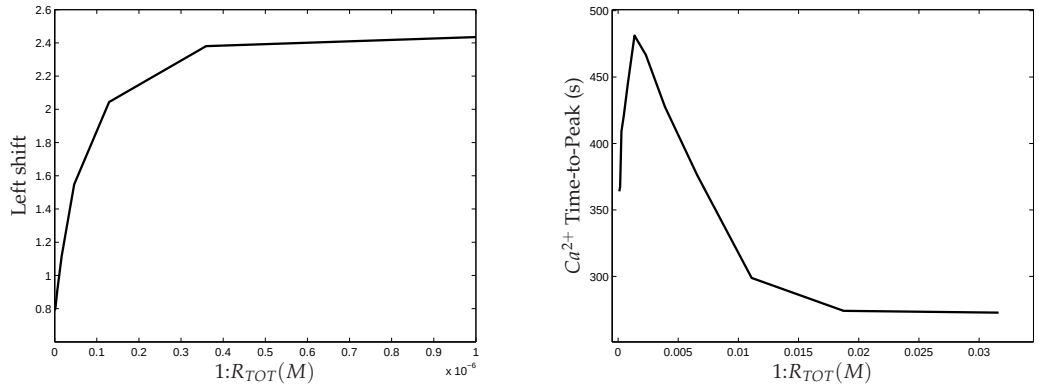


Figure 4.18: Sensitivity results (S'_{AR}) from subspaces 1 and 2 are compared. There is reasonable correlation between both sets of results.

4.19. Therefore, at least qualitatively, it can be seen how in systems where they can vary widely, the receptor expression levels can have a role in shaping the Ca^{2+} response seen in the data, in which signal is amplified without delaying the Time-to-Peak. Also, the biphasic effect that increasing 1:R_{TOT} has on the Ca^{2+} Time-to-Peak explains how low-expressing receptor systems which amplify Ca^{2+} more weakly can still have a rapid Ca^{2+} Time-to-Peak comparable to those of high-expressing systems [42]. In Chapter 2 it was seen that the receptor and G-protein expression levels influence the level of the α_{GTP} response, and it was speculated that they would similarly influence the Ca^{2+} response. The sensitivity analysis results from Subspace 1 have validated the significant sensitivity of the Ca^{2+} response to the total receptor concentration, while the medium sensitivity of 2:G_{TOT} in both subspaces confirms the influence of the G-protein concentration. However, the rate constant of desensitisation, 36:k_{des+} —another G-protein level parameter— is always higher ranked than 1:R_{TOT} and 2:G_{TOT}; kinetic models of G-protein activation alone tend not to include receptor desensitisation [13, 86, 100], but these results suggest that they should.

The box in the top right of Figure 4.18 contains parameters identified in the bottom ten ranking of both subspaces; the results are very consistent, as there are 9 parameters in the box.

Though there is good agreement between the results from parameter subspaces 1 and 2, it cannot necessarily be said that the few discrepancies are insignificant. The averaging of the overall scores from each subspace (see equation 4.36), would have evened out some of the discrepancies.



(a) (b)

Figure 4.19: The effect of varying the total receptor concentration 1:R_{TOT} on the Ca^{2+} Time-to-Peak and signal amplification. Increasing 1:R_{TOT} increases the Left-Shift, while exerting a biphasic effect on the Time-to-Peak.

4.5.4 Implications for parameter fitting

It is reasonable to assume that the most significant parameters are those which most require estimation, as variations in their values most significantly alter the Ca^{2+} response; all other parameters may then be fixed at any value in their ranges of uncertainty. A dimensionally reduced parameter space is then obtained in which a parameter estimation routine can be applied, probably at a lower computational cost, in order to locate the parameter set with which the model best fits the data. Since the computational expense of many parameter estimation routines increases with the number of parameter space dimensions, reducing the parameter space can make the parameter estimation process more practicable. This will be considered in Chapter 6.

4.6 The key driver of the amplified, yet rapid Ca^{2+} response

Of the parameters that have opposite effects on the Time-to-Peak and the Left-Shift, 27:k_{IP_3} is the only one which, when varied within its defined range of uncertainty, did not reach a lower bound for the Ca^{2+} Time-to-Peak (see Figure 4.16(b)). When 27:k_{IP_3} is increased to 50000 s^{-1} (beyond 1000 s^{-1}, the upper bound of its originally defined range), this leads to a far more rapid Time-to-Peak of 50 seconds for the higher agonist concentrations (Figure 4.20(b)), while the maximal response is predicted for all agonists (Figure 4.20(c)). This is because the rate of reaction $\alpha_{GTP}PLC^*PIP2 \xrightarrow{k_{IP_3}} \alpha_{GTP}PLC^* + IP_3 + DAG$ increases with the value of 27:k_{IP_3}. Hence, the rapid, yet amplified Ca^{2+} response cannot be explained by any of the reactions in which the other nine significant

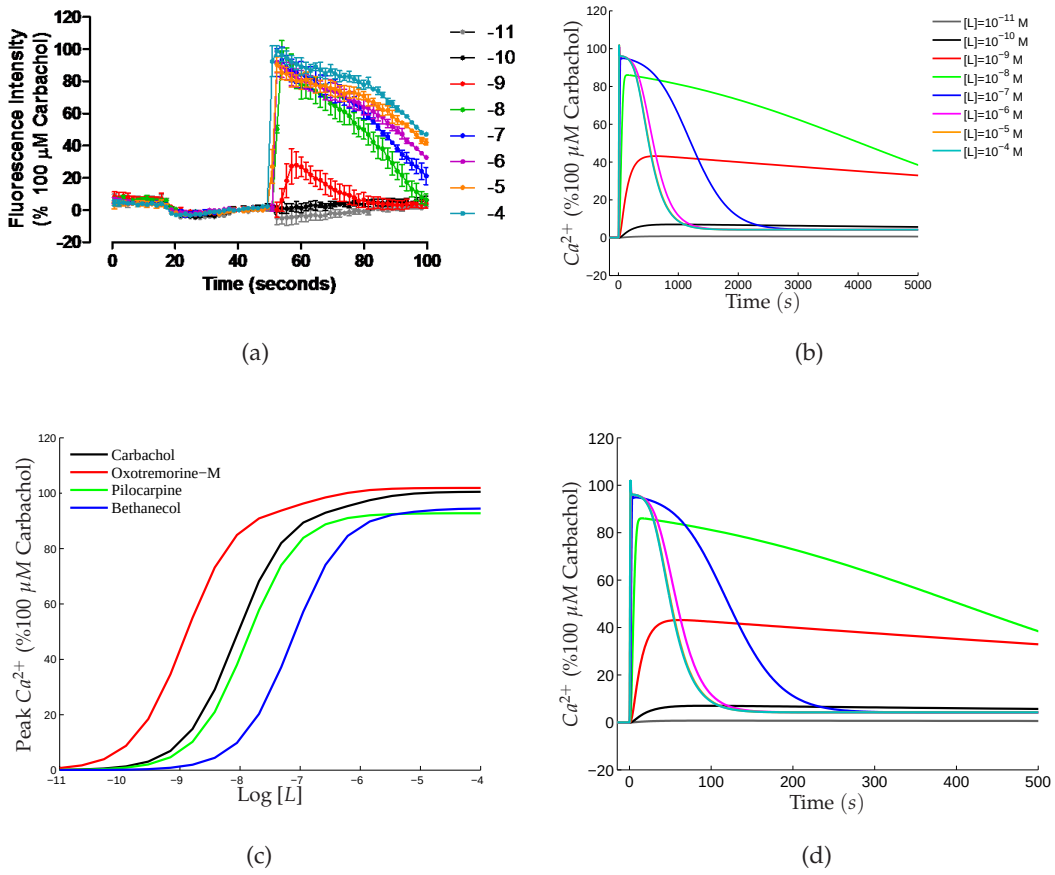
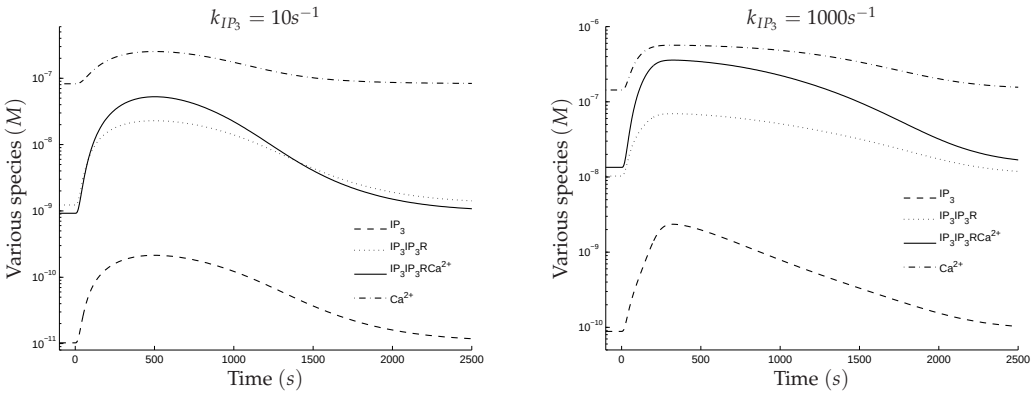


Figure 4.20: (a) Ca^{2+} time course data. When $k_{IP3} = 50000s^{-1}$, (a) the Times-to-Peak at the higher agonist concentrations are rapid, $\approx 50s$ and (c) dose-response curves show that signal amplification occurs. (d) Under a transformation of the original timescale, the time courses better resemble the timescale of the data, with Times-to-Peak $\leq 50s$.

parameters are involved, but by a rapid IP_3 production rate via a high value of 27:k_{IP_3}, through the enzymatic action of PLC^*. This means that PIP_2 is rapidly hydrolysed to produce IP_3, which binds very quickly to its receptors, to release Ca^{2+} from the ER. The fast reaction rate also increases the probability that the dissociated $\alpha_{GTP}PLC^*$ will bind (sequentially) to several free PIP_2 molecules, especially as $[PIP_2]$ is not depleted. Thus, small amounts of $\alpha_{GTP}PLC^*$ cause significant IP_3 production, which explains the significant Left-Shift.

The biphasic dependence of the Ca^{2+} Time-to-Peak on 27:k_{IP_3} is less easy to explain. The first thing to recognise is that, as Figure 4.21 shows, as 27:k_{IP_3} varies, the shape of the Ca^{2+} transient (as well as IP_3IP_3R and $IP_3IP_3RCa^{2+}$) follows the IP_3 one, so understanding the IP_3 Time-to-Peak sheds light on the Ca^{2+} Time-to-Peak. It is important to keep in mind that the IP_3 peak occurs when $\frac{d[IP_3]}{dt} = 0$, whose individual rates are given by:



(a) (b)

Figure 4.21: The shapes of the IP_3IP_3R, $Ca^{2+}IP_3IP_3R$ and Ca^{2+} transients follow that of IP_3.

$$R1 = k_{IP_3}[\alpha_{GTP}PLC^*PIP_2] \quad (4.37)$$

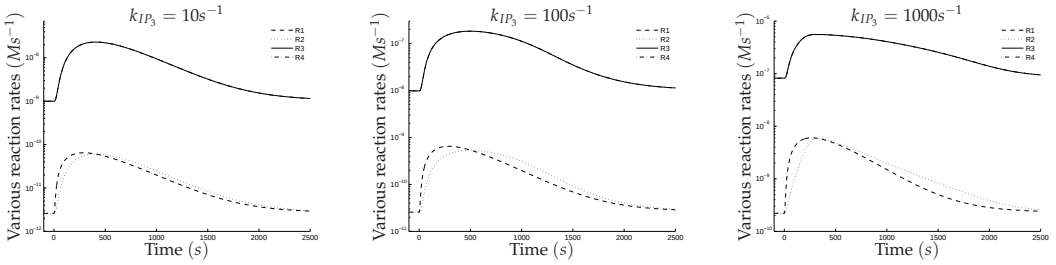
$$R2 = k_{IP_3deg}[IP_3] \quad (4.38)$$

$$R3 = k_{IP_3Rb-}[IP_3IP_3R] \quad (4.39)$$

$$R4 = k_{IP_3Rb+}[IP_3][IP_3R]. \quad (4.40)$$

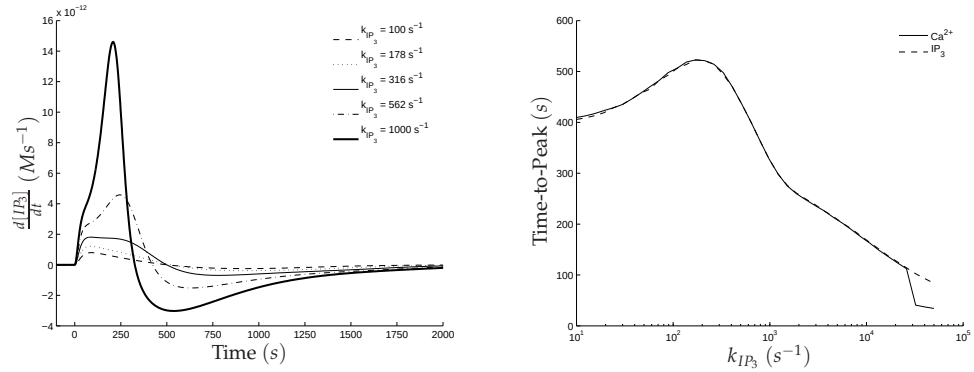
In the case of a transient peak, $\frac{d[IP_3]}{dt} = 0$ occurs the first time that $R1 = R2$ and $R3 = R4$ (see Figure 4.22). Figure 4.23 shows the time courses of $\frac{d[IP_3]}{dt}$ for values of 27:k_{IP_3} in the range $10 - 1000s^{-1}$, and it can be seen that the maximum $\frac{d[IP_3]}{dt}$ always increases with k_{IP_3} (causing increased IP_3 concentrations and signal amplification); however, the time taken for IP_3 to peak depends on the time taken to descend to zero velocity, which initially increases with k_{IP_3}, but after a threshold value (around $k_{IP_3} = 316s^{-1}$), starts to decrease, as shown in Figure 4.23. Figure 4.22(c) shows that this rapid descent occurs because at high values of k_{IP_3}, the initial velocities of IP_3 degradation ($R2$) and IP_3-IP_3R association ($R3$) and dissociation ($R4$) peak more quickly, so that $R1 = R2$ and $R3 = R4$ holds at an earlier time. Hence, increasing k_{IP_3} has a biphasic effect on the Ca^{2+} Time-to-Peak because it has the same effect on the IP_3 time-to-peak (see Figure 4.23(b)).

This explains why k_{IP_3} is the only significant parameter downstream of G-protein activation which can have this amplifying, yet quickening effect on the Ca^{2+} peak response, by speeding up $R1$ (IP_3 production) past some threshold rate which quickens the timescales of $R2 - R4$. Increasing either 5:$[PIP_2]$ or 28:k_{PIP_2b+} (two other significant parameters, see Table 4.5.1) can also promote IP_3 production by increasing the



(a) (b) (c)

Figure 4.22: The evolution of the individual terms of $\frac{d[IP_3]}{dt}$ for different values of k_{IP_3}. $R1$ and $R4$ are not visible on this scale.



(a) (b)

Figure 4.23: (a) The evolution of $\frac{d[IP_3]}{dt}$ for different values of k_{IP_3}. (b) The IP_3 and Ca^{2+} Times-to-Peak depend almost identically on k_{IP_3}.

concentration of $\alpha_{GTP}PLC^*PIP2$ but not past the threshold rate, because the concentration of $\alpha_{GTP}PLC^*PIP2$ is limited by that of $\alpha_{GTP}PLC^*$. Hence increasing 5:$[PIP_2]$ (or 28:k_{PIP_2b+}) can only mimic the first portion of the biphasic curve of the Ca^{2+} Time-to-Peak versus 27:k_{IP_3} (compare Figures 4.13(h) and 4.16(b)).

The value $k_{IP_3} = 50000s^{-1}$ suggests that the enzyme $\alpha_{GTP}PLC^*$ is in the same class as some of the fastest known enzymes, such as: carbonic anhydrase, which catalyses the hydration of carbon dioxide at rates of up to $10^6 s^{-1}$ [47]; acetylcholinesterase, which hydrolyses acetylcholine at a rate of $25,000s^{-1}$ [99]; and catalase, which hydrolyses hydrogen peroxide into water and oxygen at rates up to $1.6 \times 10^7 s^{-1}$. However, this result may be hard to test, as PLC and PIP_2 experiments are usually carried out, not in whole cells, but in buffer solutions; in one such experiment [76], the PIP_2 hydrolysis rate was $783.33s^{-1}$, which is quite different from $k_{IP_3} = 50000s^{-1}$, but this might be accounted for by the differences in experimental conditions. An informative experiment would be to observe the effects of slowing down the rate of PIP_2 hydrolysis on the features of the Ca^{2+} response.

4.6.1 Refining the Ca^{2+} Time-to-Peak prediction

The previously discussed parameter change allowed the Ca^{2+} Times-to-Peak at the three highest Oxotremorine-M concentrations to occur in $\leq 50s$. To improve the quantitative prediction, especially at the lower agonist concentrations, the timescale of the Ca^{2+} model can be transformed to quicken the dynamics without changing their pattern. If the Time-to-Peak is to be shortened by a factor of b, the time variable and the ODE system can be transformed as follows:

$$\theta = \frac{1}{b}t \tag{4.41}$$

$$\frac{d\mathbf{C}}{d\theta} = b\frac{d\mathbf{C}}{dt}, \tag{4.42}$$

where $\mathbf{C}$ is the vector of all model variables, and θ is the new time variable. When b is set to 10, the Ca^{2+} response is quickened without changing qualitatively, as shown in Figure 4.20(d)).

The Ca^{2+} time course predictions have been improved by parameter variations guided by the sensitivity analysis results. However, an aspect of the simulated response does not match the data; in the simulations, the transients for the higher agonist concentrations decay more quickly than those for the lower concentrations, whereas, in the data, the decay is slower for the higher concentrations (compare Figures 4.20(a) and (d)). Since the decay of the transient indicates extrusion of cytosolic Ca^{2+}, aspects

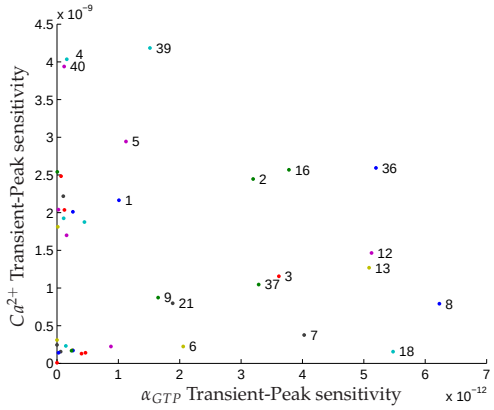


Figure 4.24: m^* values are used to quantify the sensitivities of the Ca^{2+} and α_{GTP} Transient-Peak.

of the model that regulate that process might need to be re-considered, which will be done in the next chapter.

No further manual parameter variation will be attempted at this stage; the challenge of matching the data quantitatively is left to a parameter estimation study later in Chapter 6.

4.7 Transient peak hypothesis

In Chapter 2, the peak hypothesis, which states that the transient Ca^{2+} Peak depends on the existence of a transient α_{GTP} Peak, was proposed, due to results from the analysis of the G-protein activation model. Identifying parameters to which the features are sensitive can shed light on (but not necessarily prove) the hypothesis. To do this, the sensitivity of an additional model output, *the* α_{GTP} *Transient-Peak*, which is analogous to the Ca^{2+} Transient-Peak (see equation 4.21), is considered.

Some parameters to which the Ca^{2+} Transient-Peak and the α_{GTP} Transient-Peak are both significantly sensitive are: 16:k_{GTP}, the α_{GTP} production rate constant; 36:k_{des+}, the on-rate for GPCR desensitisation; and 2:G_{TOT}, the total G-protein concentration, as shown in Figure 4.24). Their effects on both features will be analysed. Parameters to which either feature is highly sensitive will also be analysed. As Figure 4.24 shows, the Ca^{2+} Transient-Peak alone is highly sensitive to 40:p_2, the Ca^{2+} concentration for half-maximal SERCA pumping, and 4:IP_3R_{TOT}, the total concentration of IP_3 receptors, while the α_{GTP} Transient-Peak alone is highly sensitive to 18:k_{gd+}, the rate constant of α_{GTP} hydrolysis, and 7 : ζ_+, the ligand efficacy parameter.

Figure 4.25 compares $[\alpha_{GTP}]$ and $[Ca^{2+}]$ time courses for different values of the parameters to which both features are highly sensitive, while Figure 4.26 compares the

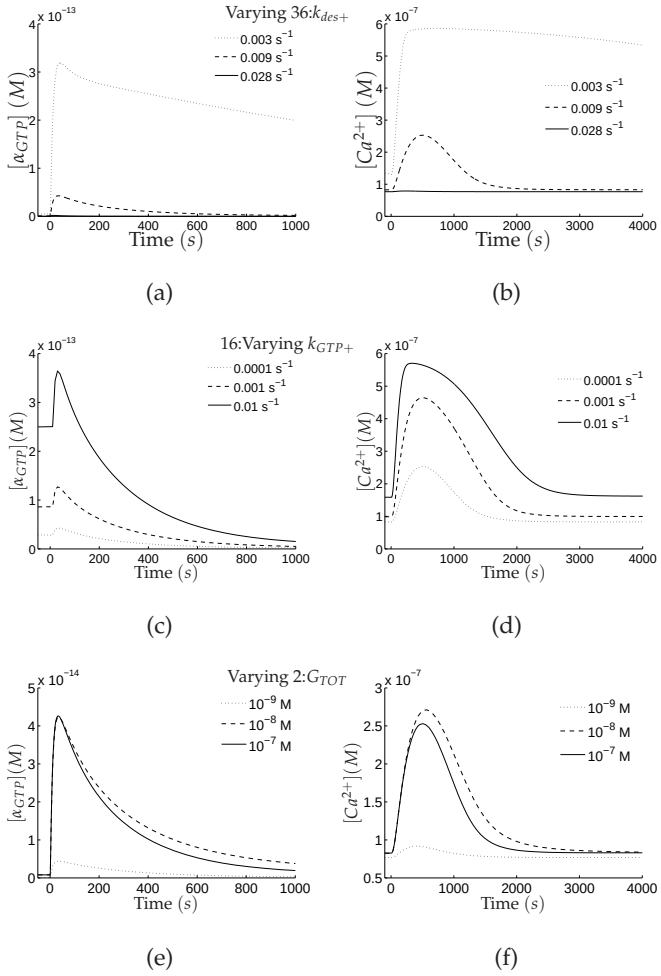


Figure 4.25: The effects of varying parameters to which the Ca^{2+} Transient-Peak and the α_{GTP} Transient-Peak are both significantly sensitive.

time courses for different values of the parameters to which only α_{GTP} is highly sensitive. In these cases, $[\alpha_{GTP}]$ and $[Ca^{2+}]$ evolve similarly; they increase together, and flatten synchronously. The uniting factor is that α_{GTP} is sensitive in both cases, which supports the peak hypothesis.

In Figure 4.27, the $[\alpha_{GTP}]$ and $[Ca^{2+}]$ time courses are compared for different values of the parameters to which only the Ca^{2+} Transient-Peak is highly sensitive, and some non-intuitive behaviour occurs. Increasing 40:p_2 and 4:IP_3R_{TOT} increases the Ca^{2+} response, but decreases the α_{GTP} one (although α_{GTP} is relatively insensitive to 4:IP_3R_{TOT}). As both parameters are rather directly linked to Ca^{2+} release, it appears that in these cases, the Ca^{2+} response modulates the α_{GTP} one. This is probably because of increased Ca^{2+} binding to the $\alpha_{GTP}PLC^*$ form, via $Ca^{2+} + \alpha_{GTP}PLC \underset{k_{PLCdeact}}{\overset{k_{PLCact}}{\rightleftharpoons}} \alpha_{GTP}PLC^*$ (see reaction (3.5)), which decreases the amount of free α_{GTP}. It is important

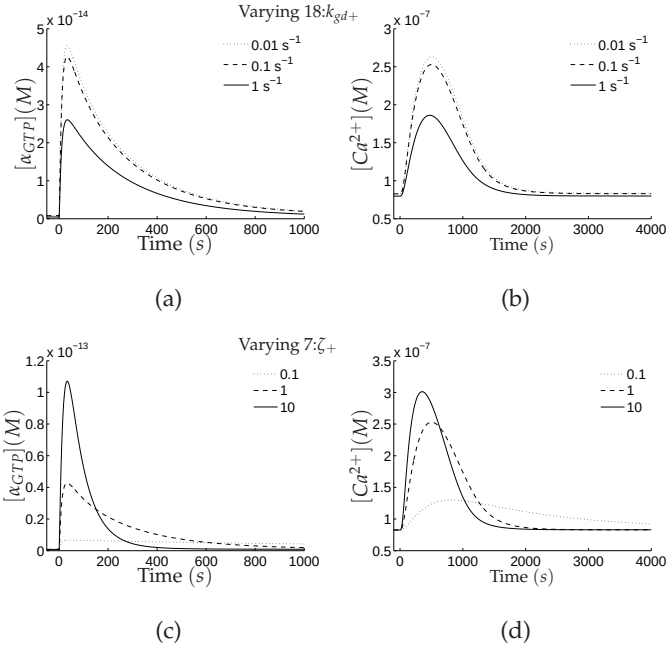


Figure 4.26: The effects of varying parameters to which only the α_{GTP} Transient-Peak is highly sensitive. Compare with Figure 4.25, where $[\alpha_{GTP}]$ and $[Ca^{2+}]$ also evolve similarly.

to note however, that the Ca^{2+} transient peak still does not occur without the α_{GTP} one; they therefore appear strongly linked.

4.8 The dependence of GSA results on the sample size, r

The sensitivity results from parameter subspace 1, which were carried out with $r_2 = 1000$ optimised trajectories, are now compared with results from the same subspace, carried out with $r_1 = 150$ trajectories, in order to consider the effect of the sample size on the results.

Figure 4.28 shows the parameter rankings obtained for both runs, using S'_{AR} (see equation 4.35), and there is reasonable correlation between both sets of results. Parameters 1, 26, 28 and 36 are significantly more highly ranked by the $r_2 = 1000$ run; on the other hand, 19, 25, 31-33, 35, and 37 are noticeably lower ranked. It could be that a smaller sample of elementary effects, such as that obtained with $r_1 = 150$ overestimates the *relative* significance of a parameter because it underestimates the significance of some of the other parameters.

In general, the $r_1 = 150$ run does well to identify the most significant parameters. Among its ten most significant parameters are seven of those found by the $r_2 = 1000$ run (see the bottom left box in Figure 4.28). The $r_1 = 150$ run identifies the least sen-

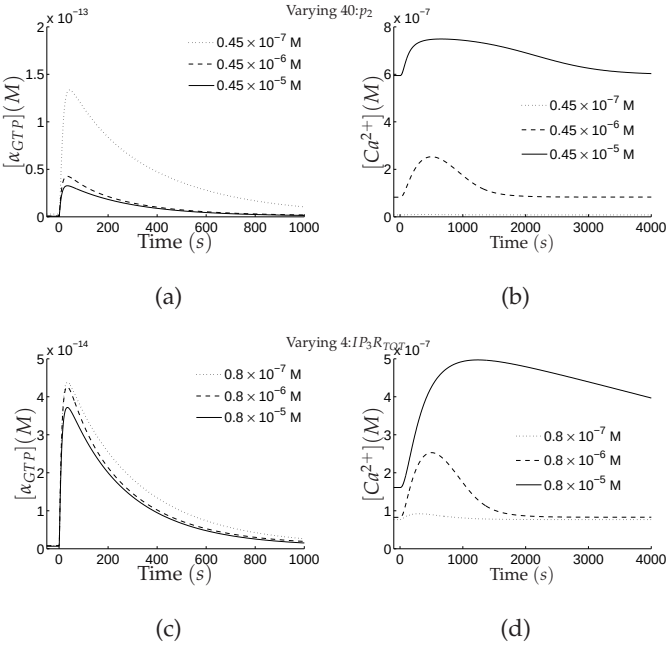


Figure 4.27: The effects of varying parameters to which only the Ca^{2+} Transient-Peak is highly sensitive.

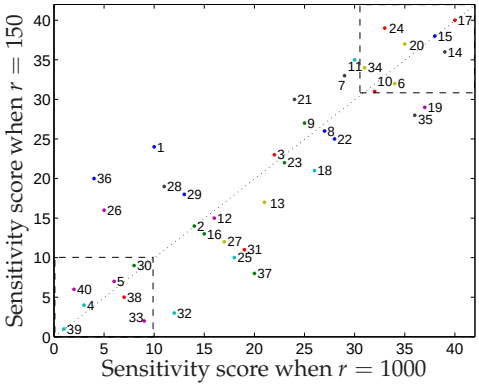


Figure 4.28: Results from parameter subspace 1, carried out with the two values $r_1 = 150$ and $r_2 = 1000$. Parameters are ranked using S'_{AR}, quantified in equation 4.35.

Optimal $r = 1000$	Optimal $r = 150$	Non-optimal $r = 150$
p_1	p_1	k_{PIP2b+}
p_2	$k_{Ca^{2+}act-}$	k_{PLCdis}
IP_3R_{TOT}	$k_{Ca^{2+}act+}$	IP_3R_{TOT}
k_{des+}	IP_3R_{TOT}	k_{IP_3Rb+}
k_{IP_3deg}	γ_1	p_2
$[PIP_2]$	p_2	γ_1
γ_1	$[PIP_2]$	p_1
k_{IP_3Rb+}	$k_{PKCdes-}$	$k_{Ca^{2+}act+}$
$k_{Ca^{2+}act-}$	k_{IP_3Rb+}	k_{GTP+}
R_{TOT}	k_{PLCdis}	$[PIP_2]$

Table 4.8.1: Sensitivity analysis results obtained using $r_2 = 1000$ optimal, $r_1 = 150$ optimal, and $\check{r}_1 = 150$ non-optimal trajectories are compared. The ten most sensitive parameters for each run (in subspace 1) are shown here.

sitive parameters almost as well as the $r_2 = 1000$ run, since nine of the ten least significant parameters of both runs are in common (see the top right box in Figure 4.28). Therefore, for this model, if the computational expense of a higher sample size run is too expensive, a lower sample size such as $r = 150$ could be used.

The importance of optimising the sampling strategy (as discussed in section 4.3.1) is also considered by comparing the results obtained from $\check{r}_1 = 150$ non-optimal trajectories with those obtained from the $r_1 = 150$ run. The results from the $r_2 = 1000$ run are taken to be the best results, and used to judge the performances of the $r_1 = 150$ and $\check{r}_1 = 150$ runs.

In terms of identifying the ten most sensitive parameters, the $r_1 = 150$ and $\check{r}_1 = 150$ runs perform similarly; the former identifies seven of the parameters found by the $r_2 = 1000$ run, while the latter identifies six, as shown in Table 4.8.1. However, the parameters identified as the two most sensitive in the $\check{r}_1 = 150$ run are not even identified among the top ten of the $r_2 = 1000$ run. Since the top two parameters identified by the $r_1 = 150$ are in the top ten of the $r_2 = 1000$ run (both runs also identify the same parameter as the most sensitive), it is concluded that the optimisation strategy results in a better *ranking* of the parameters. This becomes crucial in circumstances where only a handful of the most sensitive parameters are sought.

4.9 Discussion

In this chapter, a global sensitivity analysis has been carried out on the parameters of the Ca^{2+} model, identifying which of them are most significant, and providing insight into the pathway dynamics. The problem of finding a good balance between thorough exploration of parameter space and computational affordability when applying

a global sensitivity analysis to a model as detailed as the Ca^{2+} mobilisation model was also considered.

Global sensitivity analysis and computational efficiency

The method of elementary effects was carried out using both $r_1 = 150$ and $r_2 = 1000$ optimal trajectories, and the results from both were agreeably consistent; this means that in situations where computational resources are more limited, results generated from using a lower number of optimised trajectories such as $r_1 = 150$ can give acceptable results. Also, if the sensitivity analysis from the more expensive r_2 run leads to insight that requires a further sensitivity analysis (that is, extra, unexpected computational expense), the lower number of trajectories could be used for this.

The time taken to generate 1000 optimal trajectories from 2000 ($\approx 41.2h$) may be considered too high, but it was carried out because the available computational resources were sufficient. In addition, because each parameter is initially sampled from the [0,1] interval, the process needs to be carried out only once, and the trajectories could be used repeatedly for differently defined and constrained parameter spaces, as long as they have the same number of dimensions. For instance, the same $r_1 = 1000$ trajectories were used to generate the sensitivity results for the two subspaces earlier defined (see section 4.5.3). (The same trajectories are also used to generate new sensitivity results later, in Chapter 6.)

Insight into pathway dynamics

Parameters controlling agonist potency

Agonist potency, as quantified by the Ca^{2+} Peak-over-Basal function (equation (4.20)), was relatively insensitive to all but six parameters: 26:k_{IP3deg}, the rate constant of IP_3 degradation, 36:k_{des+}, the rate constant of the GPCR desensitisation reactions, 23:k_{PLCact}, the rate constant of PLC activation, 39:p_1, the maximal rate of the SERCA pumps, 5:$[PIP_2]$, the concentration of PIP_2 and 12:k_{deact}, the rate constant for GPCR deactivation. Guided by these results, agonist potency experiments might be designed in which any of the relevant reactions are inhibited, or one in which $[PIP_2]$ is varied. In particular, the hypothesis that such experimental manoeuvres can cause known agonists to act as antagonists should be tested. This might provide an explanation for unexpected failures when agonists are used in therapy strategies, since parameters such as $[PIP_2]$ and PKC concentration, which mediates the desensitisation process (abstracted in the Ca^{2+} mobilisation model by the reactions involving 36:k_{des+}, see reactions 3.22 and 3.23), can vary from cell, and cellular variability could be conceptualised to represent patient to

patient variability [31].

Understanding the amplified, yet rapid Ca^{2+} response

The rate constant for PIP_2 hydrolysis, 27:k_{IP_3}, was identified as the only parameter which could individually account for the amplified yet rapid response seen in the data. Individual changes in the other significant parameters could account either for the Left-Shift or the rapid Ca^{2+} Time-to-Peak, but not both. A high value of 27:k_{IP_3} ($50000s^{-1}$), was required to simultaneously predict the two features, suggesting that the enzyme $\alpha_{GTP}PLC^*$, which mediates PIP_2 hydrolysis, could be in the same class as some of the fastest known enzymes.

The most sensitive parameters

The three most sensitive parameters are the SERCA pump parameters, 39:p_1 and 40:p_2, and the GPCR desensitisation on-rate 36:k_{des+}. These are pathway components that the design of pharmacological (agonist) experiments do not typically prioritise, yet these results suggest that new experiments should be designed that explore the robustness of the agonist-induced Ca^{2+} response to changes in these particular pathway processes. The availability of inhibitors for the pump [61], and for PKC, which mediates the desensitisation process makes such experiments feasible [14].

PIP_2, is also a significant pathway component, since its concentration was ranked 6th overall; interestingly, the concentration of PIP_2's enzyme was only ranked 24th. This implies that, even though both compounds participate in the same reaction, PIP_2 might be a far more promising drug target.

The fact that there are several sensitive points in the pathway with components that are potential drug targets may also suggest co-drugging opportunities in Ca^{2+} pathways.

Model simplification

It was shown that two of the least significant parameters 14:k_{g-} and 24:$k_{PLCdeact}$ can be set to zero, without altering the Ca^{2+} response, suggesting that GPCR pre-coupling and Ca^{2+} positive feedback on PLC are irreversible reactions. Hence, in addition to providing insight into pathway dynamics, the results of the Global sensitivity analysis could be used to slightly simplify the model.

CHAPTER 5

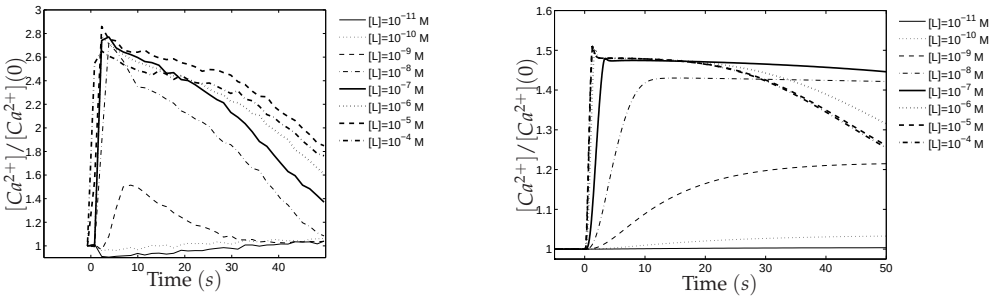
Variant Ca^{2+}-mobilisation models

5.1 Introduction

This chapter considers modelling questions that arose from the previous chapter. The results of Chapter 4's global sensitivity analysis (GSA) guided parameter changes which resulted in a parameter set (here referred to as the *GSA-guided parameter set*, found in Table B.2.2) with which the Ca^{2+}-mobilisation model predicted key features of the experimental data (the transient, rapid peak, and signal amplification, see section 4.6). However, a feature of the Ca^{2+} data came to prominence that had not previously been highlighted; while in the Ca^{2+} data the descent from peak to steady state occurs in relatively short times for $[L] = 10^{-9}$–$10^{-4}M$, the descent is slow for $[L] = 10^{-9}$–$10^{-7}M$ in the simulated time courses, as illustrated in Figure 5.1. In other words, it became clear that the steepness of the Ca^{2+} curve after the peak is also an important feature.

Since the descent from peak indicates the removal of Ca^{2+} from the cytosol, the role that extra removal mechanisms might have in determining the steepness of that descent is considered here, by incorporating plasma membrane pumps (ATPases), which facilitate Ca^{2+} efflux to the extracellular environment, into Chapter 4's Ca^{2+}-mobilisation model (hereafter referred to as the original CM). This change results in a variant of the original CM, which will be referred to as the Ca^{2+}-mobilisation model with membrane pumps (CMMP). Parameter variation will be carried out on this extended model to observe the role that membrane pumps could play in determining the steepness of the descent from peak.

The second question that arose from Chapter 4 is whether the total concentration of PIP_2 should be assumed constant, since it was identified by the GSA as the 6th most significant parameter. In reality, PIP_2 undergoes a cycle of depletion and re-synthesis (described in section 3.1) which does affect its concentration, hence this chapter also considers another extension of the original CM, which takes PIP_2 dynamics into account.



(a) (b)

Figure 5.1: A comparison of (a) experimental Ca^{2+} data and (b) corresponding simulations obtained with the GSA-guided parameter set. In the data, the time to steady state appears to increase with the agonist concentration, while the reverse is the case in the simulations.

5.2 Ca^{2+}-mobilisation model with membrane pumps

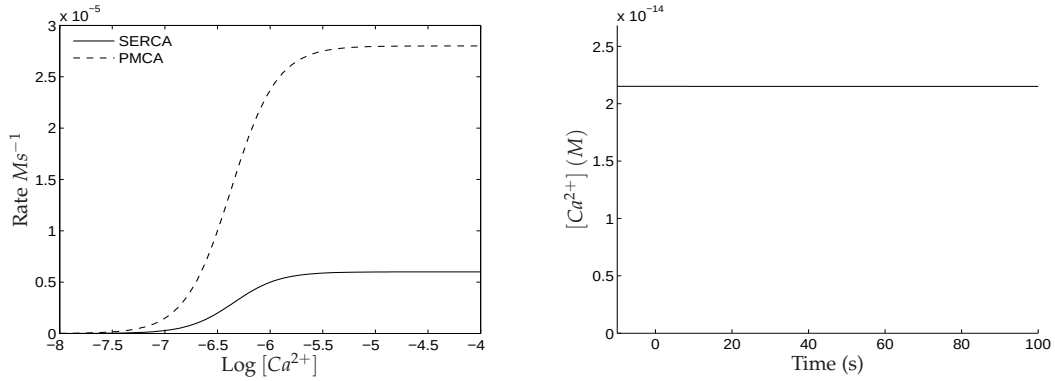
The original CM is extended into the Ca^{2+}-mobilisation model with membrane pumps (CMMP) by accounting for the fact that the cytosolic Ca^{2+} concentration can be depleted due to the action of plasma membrane Ca^{2+} ATPases (PMCA). The rate at which the pumps transport Ca^{2+} from the cytosol into the extracellular medium is modelled by a Hill function, following previous studies [87, 95, 101]. The cytosolic Ca^{2+} rate equation for this extended model is then given by (compare with equation (3.13)):

$$\begin{aligned}\frac{d[Ca^{2+}]}{dt} &= \nu_r(\gamma_0 + \gamma_1[IP_3.IP_3R.Ca^{2+}])([Ca^{2+}_{ER}] - [Ca^{2+}]) \\ &\quad - \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2} - \frac{V_P[Ca^{2+}]^2}{K_P^2 + [Ca^{2+}]^2},\end{aligned} \tag{5.1}$$

where V_P is the maximal rate of the membrane pumps, and K_P is the concentration of Ca^{2+} at which half-maximal pumping occurs. Also, note that $[Ca^{2+}_{ER}] + [Ca^{2+}]$ is no longer constant, as cytosolic Ca^{2+} is pumped out via the PMCA pumps; hence, $[Ca^{2+}_{ER}]$ is included as a variable in the model, whose ODE is given by:

$$\frac{d[Ca^{2+}_{ER}]}{dt} = \frac{1}{\nu_r}\left(\frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2}\right) - (\gamma_0 + \gamma_1[IP_3.IP_3R.Ca^{2+}])(Ca^{2+}_{ER} - [Ca^{2+}]). \tag{5.2}$$

($\frac{d[Ca^{2+}_{ER}]}{dt}$ is multiplied by the cytosol/ER volume ratio, $\frac{1}{\nu_r}$, to account for the fact that calcium flux from the cytoplasm causes a greater concentration change in the ER (because of its smaller volume) than it does in the cytoplasm.)



(a) (b)

Figure 5.2: (a) A comparison of the rates of the SERCA ($p_1 = 6 \times 10^{-6} Ms^{-1}$, $p_2 = 0.45 \times 10^{-6} M$) and membrane pumps ($V_P = 28 \times 10^{-6} Ms^{-1}$, $K_P = 0.425 \times 10^{-6} M$). (b) There is no Ca^{2+} response to $10^{-7} M$ of Oxotremorine when $V_P = 28 \times 10^{-6} Ms^{-1}$ and $K_P = 0.425 \times 10^{-6} M$.

5.2.1 Fine-tuning membrane pump parameters

The initial values given to V_P and K_P are $28 \times 10^{-6} Ms^{-1}$, and $0.425 \times 10^{-6} M$ respectively, taken from a previous study [87], in which they were estimated by fitting an IP_3-stimulated CICR model to Ca^{2+} data. The original CM and the CMMP have all other parameters in common, and their values in the GSA-derived parameter set are retained (see Table B.2.2) since they were shown to predict key features of the data in Chapter 4.

Figure 5.2(a) shows how the SERCA and plasma membrane Hill functions vary with $[Ca^{2+}]$ for the aforementioned V_P and K_P values. These values give the membrane pumps a far higher pumping capacity than the SERCA pumps over most Ca^{2+} concentrations. As a result, there is no discernible Ca^{2+} release in response to $10^{-7} M$ of Oxotremorine-M, as illustrated in Figure 5.2(b). This prediction is opposed to that of the original CM, in which (as with the data) $10^{-7} M$ of Oxotremorine-M induces a maximal Ca^{2+} release (see Figure 4.20(d)). This strongly suggests that more fitting values of V_P and/or K_P need to be found for the CMMP. Since values which allow significant responses to agonist stimulation are sought, the Ca^{2+} Peak-Over-Basal feature, as used in Chapter 4, and reproduced here:

$$P_B = \frac{\text{Peak} - \text{Basal}}{\text{Basal}}, \tag{5.3}$$

will also monitored in the process of finding suitable values for V_P and K_P. Non-zero values of P_B indicate that there is a response to agonist stimulation.

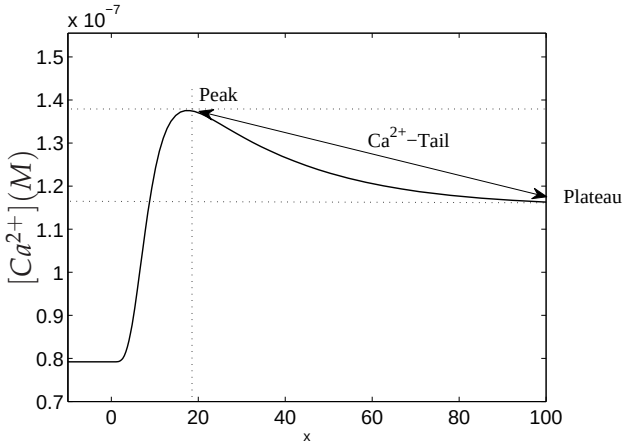


Figure 5.3: The steepness of the Ca^{2+} descent from peak is an important feature in this chapter. The descent phase of the transient is referred to as the 'Ca^{2+}-Tail'.

To also ensure that the response is transient,

$$P_{MP} = \text{Peak} - \text{Plateau}, \tag{5.4}$$

(also used in Chapter 4) will also be calculated; thus, when $P_{MP} = 0$ it can be known that the response is no longer transient.

As this chapter is concerned with fine-tuning the steepness of the descent from the Ca^{2+} peak, that feature is mathematically quantified in the next section.

Quantifying the rate of the transient decay

The Ca^{2+} descent from peak will be referred to as the steepness of the 'Ca^{2+}-Tail', which is illustrated in Figure 5.3. There are different ways to quantify the steepness of the tail; the most suitable way is more easily chosen when the exact form of the transient is known. For example, the transient may be assumed to undergo exponential decay, and then fit to an exponential function such as,

$$Ca^{2+}(t_p)e^{-zt}, \tag{5.5}$$

where time t is measured from the occurrence of the Ca^{2+} peak, t_p; z may then be taken as a measure of the Ca^{2+}-Tail steepness (which increases with z). However, there are at least two disadvantages of such an approach. Firstly, the assumption of exponential decay may not apply to some (or all) of the time courses. Secondly, a computationally inexpensive fitting method might be needed to estimate z; such a method would generally require a good initial guess for z, which might be hard to provide.

An alternative approach will be employed here; instantaneous gradients of the Ca^{2+}-Tail are used to quantify steepness, and so does not require an assumed functional form for the transient. The following equation is used to calculate a *normalised instantaneous gradient*:

$$I_G = \frac{|d[Ca^{2+}]/dt|}{[Ca^{2+}](0)}. \tag{5.6}$$

The gradients are normalised to the basal level to account for the way the Ca^{2+} data have been normalised in Figure 5.1(a); these are the same data that were used in Chapters 3 and 4, where they were normalised to the (Peak - Basal), due to stimulation by $10^{-4}M$ of Carbachol). The normalisation was changed to the present form to prevent dividing by zero when some V_P,K_P values cause Peak $-$ Basal $= 0$ (as in Figure 5.2(b)).

After the addition of the agonist, the model is run until steady state. The Ca^{2+}-Tail is measured up to when the transient falls to 95% of the (Peak $-$ Plateau) value. I_G is measured every second of the tail, and the average, T_S, is a measure of steepness, given by:

$$T_s = \frac{1}{n_t}\sum_{i=1}^{n_t} I_{G_i}, \tag{5.7}$$

where n_t is the number of time points to go from the peak value to 95% of the (Peak $-$ Plateau) value.

Parameter variation

A region of V_P–K_P space constrained by the ranges $V_P = 28 \times 10^{-14} - 28 \times 10^{-6} Ms^{-1}$ and $K_P = 0.425 \times 10^{-7} - 0.425 \times 10^{-5} M$ is searched to identify coordinates which cause significant Ca^{2+} responses. Parameter variation is carried out for two agonist concentrations; firstly, at $10^{-9}M$, which is the low agonist concentration, for which the Ca^{2+} steady state in the data is reached within the shorter timeframe of $\approx 40s$; secondly, at $10^{-5}M$, which is representative of the higher agonist concentrations ($10^{-8} - 10^{-4}M$) at which Ca^{2+} removal occurs within the longer timeframe of $\geq 50s$ (see Figure 5.1(a)).

Figures 5.4(a)–(c) show the values of the Ca^{2+}-Tail steepness (T_S), the Ca^{2+} Peak-Over-Basal (P_B) and the (Peak $-$ Plateau) (P_{MP}) in response to $10^{-5}M$ of Oxotremorine-M, for various values of K_P and V_P; for each curve K_P has a fixed value, while V_P varies. In Figure 5.4(a), all the curves have the same basic shape, with steepness initially increasing up to a maximum value, and subsequently decreasing until it reaches zero. The higher the value of K_P, the further right the curve is located. The first curve (solid lines), where $K_P = 0.425 \times 10^{-7}M$, represents the most sensitive setting of the pumps (in the defined range), where only $0.425 \times 10^{-7}M$ is required for half-maximal pumping, and the maximum Ca^{2+}-Tail steepness occurs when the maximal rate is given by

$V_P = 2.168 \times 10^{-12} Ms^{-1}$, which is on the lower end of its defined range. However, when $K_P = 0.425 \times 10^{-6} M$ (dashed lines), making the pumps less sensitive to $[Ca^{2+}]$, a higher maximal pumping rate ($V_P = 1.353 \times 10^{-11} Ms^{-1}$) is required to produce the maximum steepness. When the pumps are at their least sensitive ($K_P = 0.425 \times 10^{-5} M$, thick solid line), the curve is also shifted forward, as an even higher maximal rate ($V_P = 1.724 \times 10^{-9} Ms^{-1}$) is now required to produce the maximum steepness. A similar correlation between K_P and V_P also influence the P_B and P_{MP}, with higher maximal pumping rates compensating for lower Ca^{2+} sensitivities. For each curve, there are threshold values of V_P, after which P_B and P_{MP} start to decrease until they become zero. This means that, after the threshold value, as the maximal pumping rate increases, the transient starts to flatten, until it is completely flat.

Figures 5.4(f)–(h) show individual plots of both P_B and T_S for single K_P values, as V_P varies. This makes it easy to observe what is happening to the size of the agonist-stimulated response as the steepness of the Ca^{2+}-Tail changes (note that each curve has been normalised by its peak value to allow P_B and T_S, which are of different orders, to be seen clearly when plotted together). For all K_P values (that is, in all five plots) P_B and T_S respond similarly to changes in V_P. In general, the maximum steepness (T_S) occurs together with the maximum Peak-Over-Basal.

When $[L] = 10^{-9} M$, the changes in T_S, P_B and P_{MP}, as K_P and V_P vary, are similar to those that occur when $[L] = 10^{-5} M$, as shown in Figures 5.5(a)–(c). The aforementioned correlation between K_P and V_P is also evident here, as the curves shift forward with each increase in K_P. As would be expected, the values of P_B and P_{MP} are lower than they were for the higher agonist concentration. However, the fact that the T_S values are lower means that the Ca^{2+}-Tail tends to be less steep at the lower agonist concentration.

It is therefore clear that the incorporation of membrane pumps allows for modulation of the Ca^{2+}-Tail steepness, but also the size of the response to agonist stimulation. This is further illustrated by plotting time courses with $V_P = 0$, $K_P = 0$ (representing the original CM model), and time courses using one of the (K_P, V_P) pairs ($K_P = 0.425 \times 10^{-6} M$, $V_P = 1.6786 \times 10{-}11 Ms^{-1}$, marked '*' in Figures 5.4(a) and (b)) which simultaneously caused maximum steepness and Peak-Over-Basal at both agonist concentrations.

Figure 5.6(a) shows the response to $10^{-9} M$ of Oxotremorine-M, while Figure 5.6(b) shows the response to $10^{-5} M$. The influence of the membrane pumps is more pronounced when $[L] = 10^{-5} M$; the action of the membrane pumps causes the peak response to stimulation (relative to basal) to significantly increase, as well as the steepness of the tail, so that the time to steady state is shortened. When $[L] = 10^{-9} M$, the

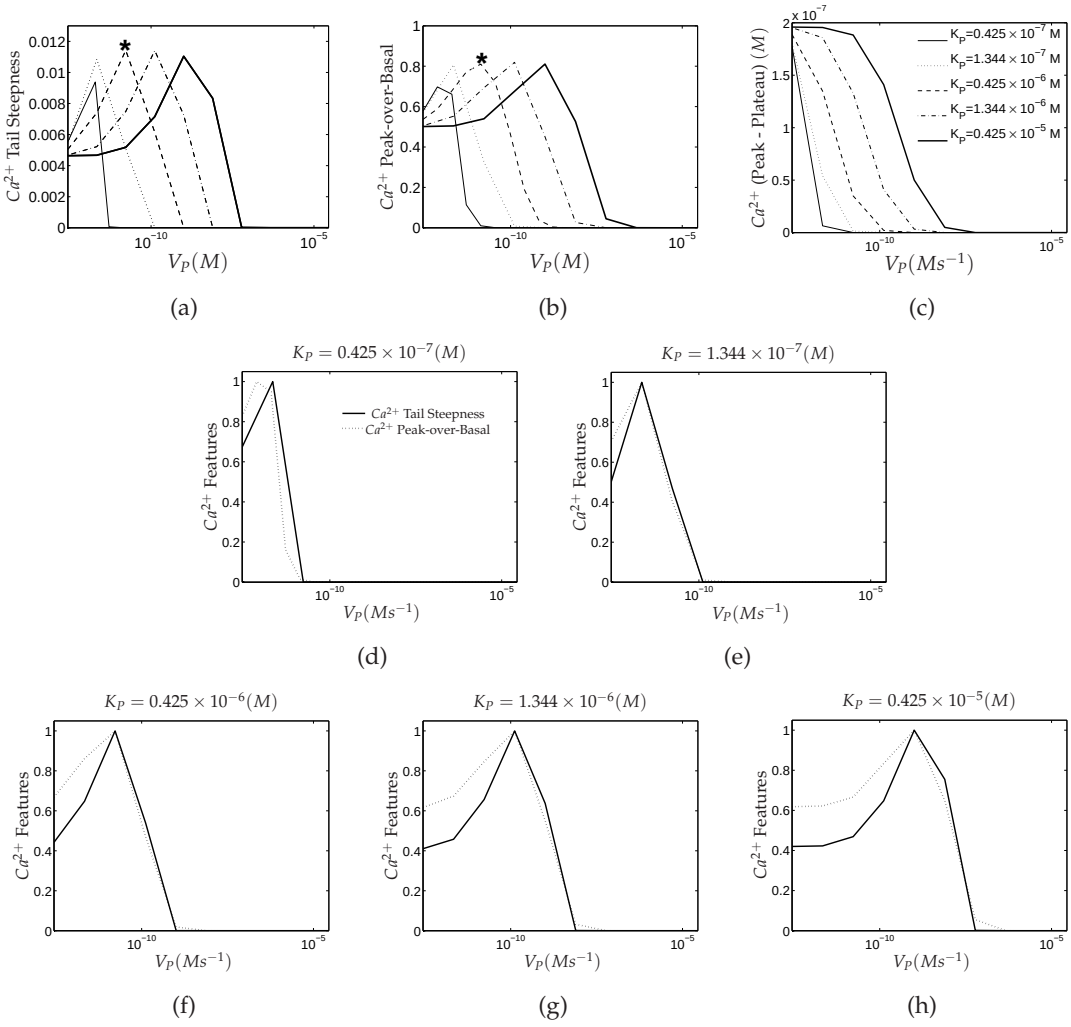


Figure 5.4: (a)-(c) Curves for the Ca^{2+}-Tail steepness (T_S), Ca^{2+} Peak-over-Basal (P_B) and Ca^{2+} (Peak - Plateau) (P_{MP}) for various K_P and V_P values, when $[L] = 10^{-5}(M)$. (d)–(h) T_S and P_B (each curve normalised to its maximum) have been plotted together for each K_P value; it can be seen that both features respond similarly to changes in V_P and K_P.

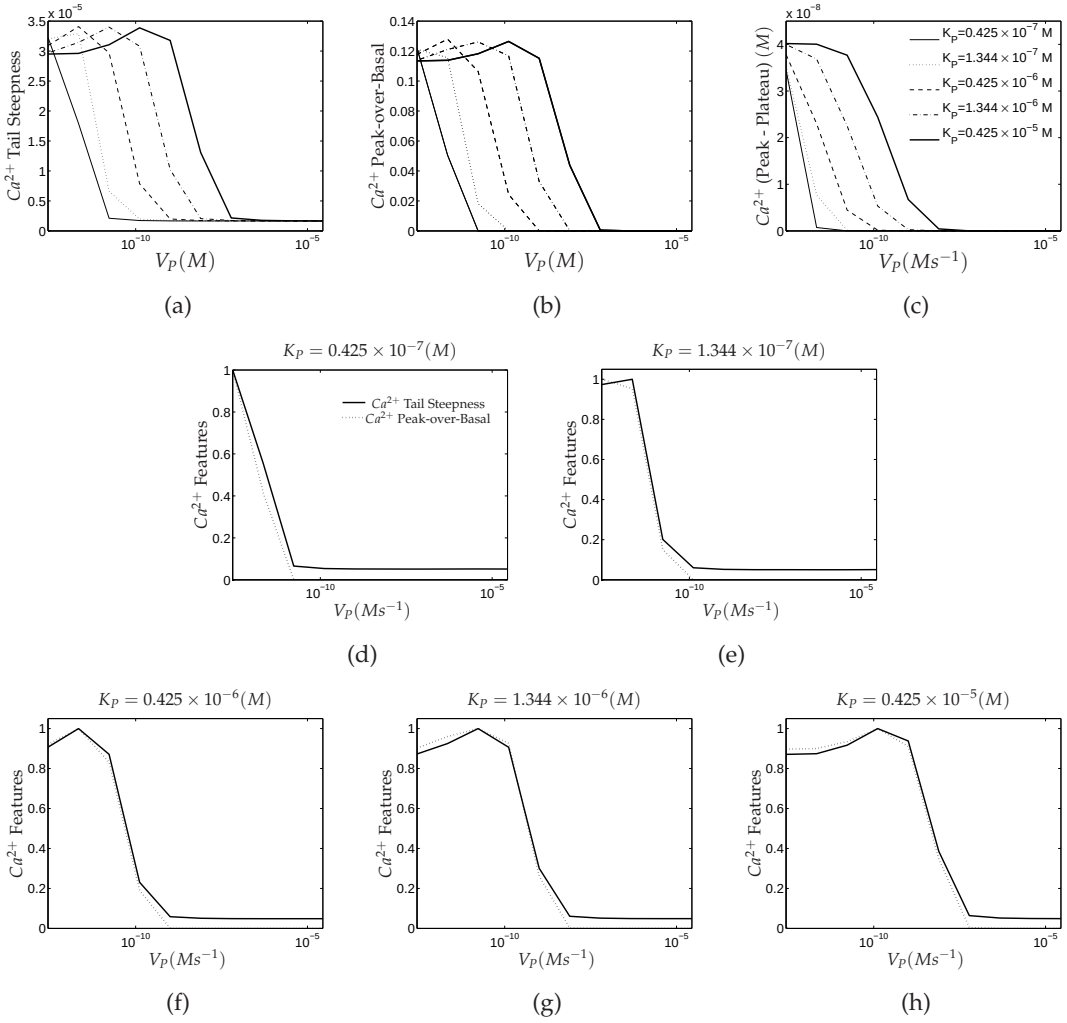


Figure 5.5: (a)-(c) Curves for the Ca^{2+}-Tail steepness (T_S), Ca^{2+} Peak-over-Basal (P_B) and Ca^{2+} (Peak - Plateau) (P_{MP}) for various K_P and V_P values, when $[L] = 10^{-9}(M)$. (d) –(h) T_S and P_B (each curve normalised to its maximum) have been plotted together for each K_P value; it can be seen that both features respond similarly to changes in V_P and K_P.

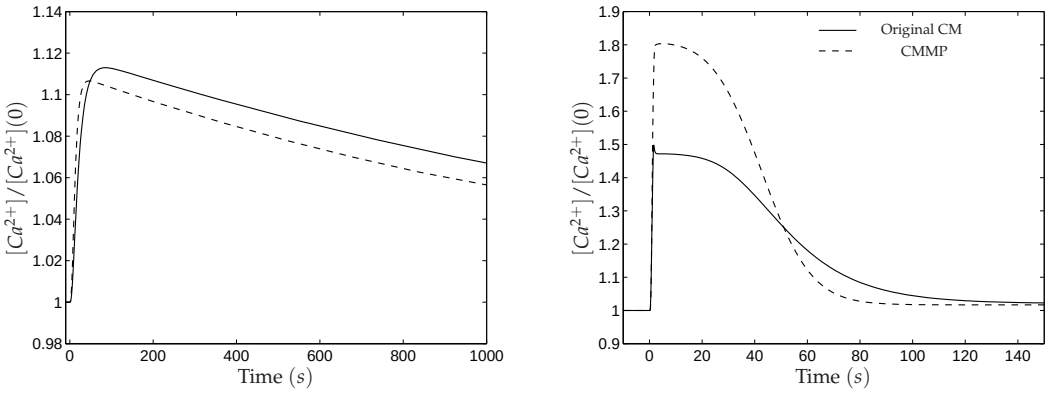


Figure 5.6: Comparing Ca^{2+} time courses from the original CM (solid line) and the CMMP (dashed line) for (a) $[L] = 10^{-9}M$ (b)$[L] = 10^{-5}M$. In both cases, the action of the membrane pumps can increase the steepness, as well as Peak-Over-Basal, but when $[L] = 10^{-9}M$, the change is not significant.

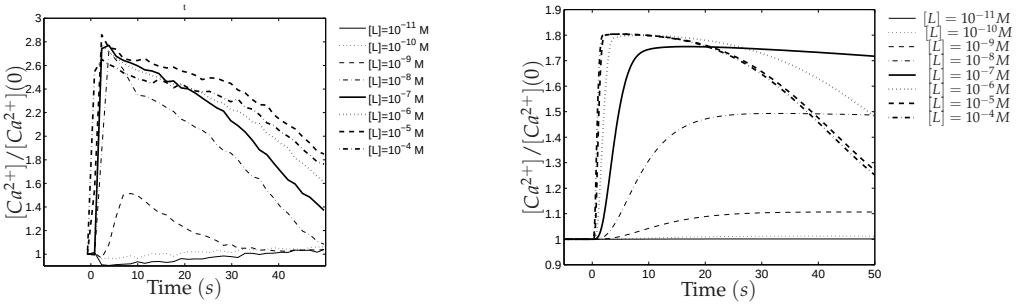
membrane pumps cause the peak response (relative to basal) to slightly decrease, but increase the steepness of the tail, although not significantly enough to resemble the data, where the steady state is reached within 40 seconds (see Figure 5.1(a)). The time courses for all agonist concentrations are plotted in Figure 5.7, where it can be seen that compared to the data (shown again in the same figure), the descent from peak is still much slower at the lower agonist concentrations.

This section has shown that although the action of membrane pumps can influence the steepness of the Ca^{2+}-Tail, the effect is not enough to allow the simulated Ca^{2+} transients for the lower agonist concentrations resemble corresponding ones in the data. It therefore seems that the extension to the original CM is not necessary.

5.3 PIP_2 depletion and replenishment

As previously mentioned, the results of Chapter 4's GSA indicated that the total concentration of PIP_2 is one of the parameters with the strongest impact on the Ca^{2+} response; hence, fluctuations in its value might alter the Ca^{2+} response. This section considers another variant Ca^{2+}-mobilisation model, which incorporates a basic PIP_2 depletion and resynthesis cycle into the original CM, allowing the effects of such a cycle to be explored.

Following a previous study [44], it is assumed that the product of PIP_2 cleavage, IP_3, is degraded into a phosphate pool, $PHOS$, from which PIP_2 can then be re-



(a) (b)

Figure 5.7: A comparison of (a) experimental Ca^{2+} data and (b) corresponding simulations of the CMMP with $V_P = 1.6786 \times 10{-}11Ms^{-1}$, $K_P = 0.425 \times 10^{-6}M$. At the lower agonist concentrations, the descent from peak is still much slower in the simulations than in the data.

synthesised, as follows:

$$IP_3 \xrightarrow{k_{IP_3deg}} PHOS \xrightarrow{k_{PIP_2rep}} PIP_2. \tag{5.8}$$

The relevant ODEs are then given by:

$$\begin{aligned} \frac{d[PIP_2]}{dt} &= k_{PIP2b-}[\alpha_{GTP}PLC^*PIP_2] - k_{PIP2b+}[\alpha_{GTP}PLC^*][PIP_2] \\ &+ k_{PIP_2rep}[PHOS] \qquad (5.9) \\ \frac{d[PHOS]}{dt} &= k_{IP_3deg}[IP_3] - k_{PIP_2rep}[PHOS] \qquad (5.10) \end{aligned}$$

This is a simplification of the cycle of PIP_2 depletion and resynthesis, but should be sufficient to address the concern of this section, which is to explore what effects fluctuations in PIP_2's concentration might have on the Ca^{2+} response. Equations (5.9) and (5.10) allow $[PIP_2]$ to drop when it is hydrolysed, and subsequently be replenished from the phosphate pool. Since PIP_2 depletion might affect the lifetime of the Ca^{2+}-activating IP_3 signal, its effect on the steepness of the Ca^{2+}-Tail is also considered.

This extended model, referred to as the Ca^{2+}-mobilisation model with $[PIP_2]$ depletion and replenishment (CMPDR), only has one additional parameter to those of the original CM: k_{PIP_2rep}, which controls the rate at which PIP_2 again becomes available after depletion (see equation 5.10). A previous study used the values $0.015, 0.1, 10s^{-1}$ [44], which are used here as a guide to vary k_{PIP_2rep} between $10^{-6} - 1s^{-1}$. (The range was adapted because the dynamics are not sensitive in the range $0.1 - 10s^{-1}$.)

Figure 5.8 shows the Ca^{2+} transients in response to both $10^{-9}M$ and $10^{-5}M$ of Oxotremorine-M, as k_{PIP_2rep} varies between $10^{-6}s^{-1}$ and $1s^{-1}$; the time courses of $[PIP_2]$

and $[PHOS]$ are also shown. When $[L] = 10^{-5}M$ and k_{PIP_2rep} is highest ($1\ s^{-1}$), the effect is virtually the same as having $[PIP_2]$ constantly available at a high concentration; as a result, the phosphate pool is hardly filled (see Figures 5.8(d) and (f)). There is significant $[PIP_2]$ depletion when $k_{PIP2rep}$ is lower, at $0.1s^{-1}$ or $0.01s^{-1}$, which is temporary as $[PIP_2]$ is subsequently replenished; however, when $k_{PIP2rep} = 10^{-3}s^{-1}$ for instance, $[PIP_2]$ is hardly recoverable by replenishment in the 120 second timeframe. Hence, one effect of $[PIP_2]$ fluctuations is on the steepness of the Ca^{2+}-Tail, which increases as $k_{PIP2rep}$ decreases from $0.1 - 10^{-4}s^{-1}$ (see Figure 5.8(b)).

In response to $10^{-9}M$ of Oxotremorine-M, the $[PIP_2]$ and $[PHOS]$ concentrations do not significantly change from their basal values, regardless of the value of k_{PIP_2rep} (Figures 5.8(c) and 5.8(e)), as there is not enough agonist to drive $[PIP_2]$ depletion. As a result, the steepness of the Ca^{2+}-Tail at $[L] = 10^{-9}M$ is not significantly impacted by changes in the value of k_{PIP_2rep}. As observed when the previous model (CMMP) was analysed, the steepness of the Ca^{2+}-Tail at $[L] = 10^{-9}M$ is not as sensitive as it should be in order to match the data.

Agonist efficacy can be diminished if the rate of PIP_2 replenishment is slow; at $k_{PIP2rep} = 10^{-5}s^{-1}$ and $k_{PIP2rep} = 10^{-6}s^{-1}$, the dose-response curves of the Ca^{2+} peak shift considerably to the right, as shown in Figure 5.9. Hence, it seems that the most reasonable values for $k_{PIP2rep}$ are between $10^{-4} - 0.1s^{-1}$. It is also important to note the effect that the rate of replenishment has on the Ca^{2+} Peak-over-Basal, which is largest when $k_{PIP2rep} = 10^{-5}s^{-4}$. However, the corresponding dose-response curve is more right-shifted than those which pertain to higher $k_{PIP2rep}$ values.

This section has shown that the original CM, which assumes that $[PIP_2]$ is constant, is roughly equivalent to the CMPDR when its rate constant of PIP_2 replenishment is very high. At intermediate values of $k_{PIP2rep}$ (when $[L] = 10^{-5}$), the Ca^{2+} Peak-over-Basal and the steepness of the Ca^{2+}-Tail can be influenced by the rate of replenishment. At the low values of $k_{PIP2rep}$ agonist efficacy is significantly diminished, suggesting that they do not represent realistic rates of PIP_2 replenishment. Hence, the dynamics of PIP_2 may be important for fine-tuning the Ca^{2+} Peak-over-Basal and the Ca^{2+}-Tail steepness at the higher agonist concentrations. However, the CMPDR does not necessarily offer improved predictions of the Ca^{2+} data, particularly at $[L] = 10^{-9}$; therefore, it might be more reasonable to find an optimal value for the constant $[PIP_2]$ in the original CM, which is the simpler model. Indeed merely by varying $[PIP_2]$ in the original model, a similar modulation of the can be achieved Ca^{2+} Peak-over-Basal and the Ca^{2+}-Tail steepness, as shown in Figure 5.10 (compare with Figures 5.8(a) and (b)).

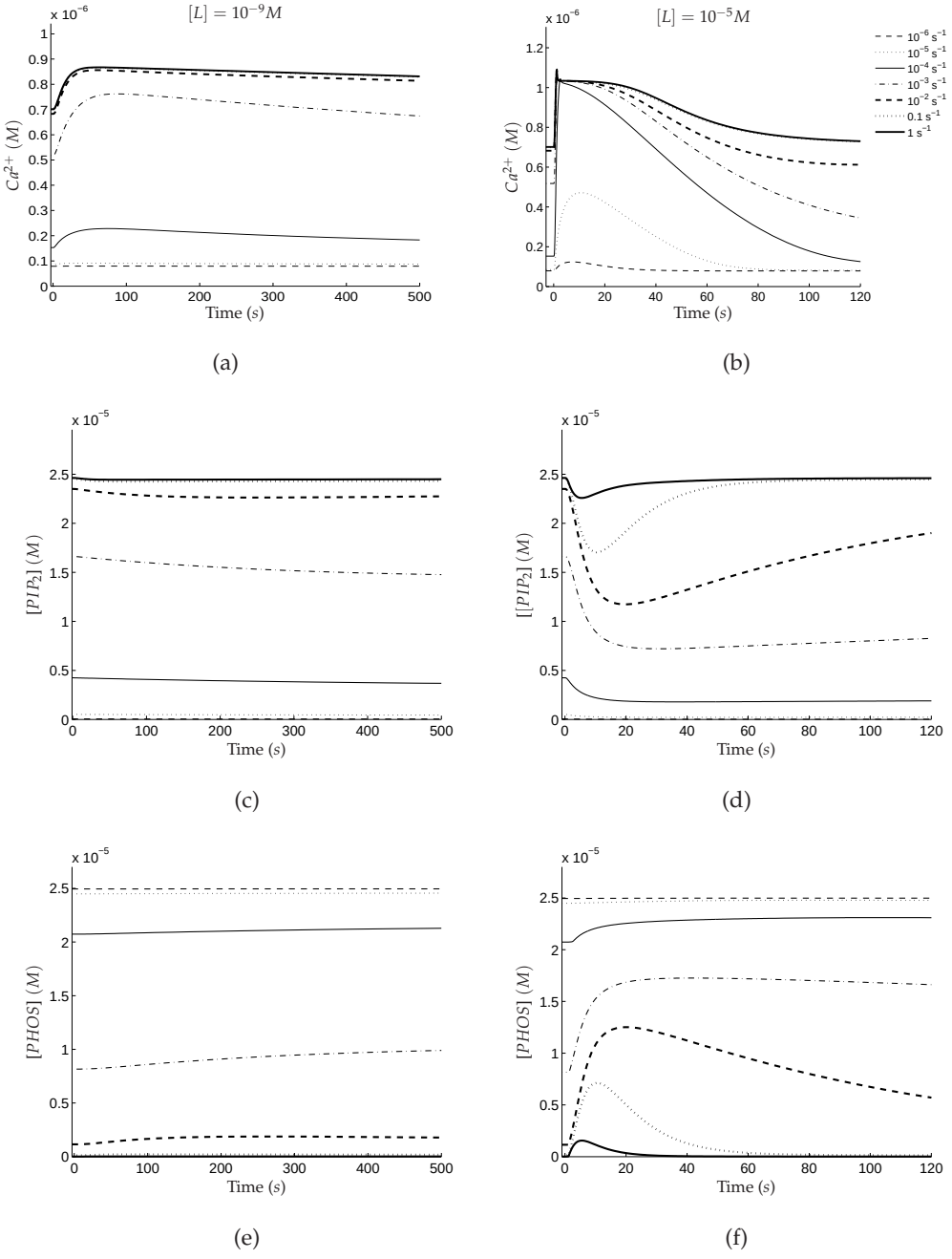


Figure 5.8: The effects of varying k_{PIP_2rep} on $[Ca^{2+}]$ (a-b), $[PIP_2]$ (c-d) and $[PHOS]$ (e-f), in response to $10^{-9}M$ and $10^{-5}M$, of Oxotremorine-M. Within a range of faster PIP_2 replenishment rates, the Ca^{2+} responses do not change much; outside this range, features like the Ca^{2+} Peak-over-Basal, and Ca^{2+}-Tail steepness can significantly change. Fast rates of PIP_2 replenishment are acceptably approximated by the assumption that $[PIP_2]$ is constant.

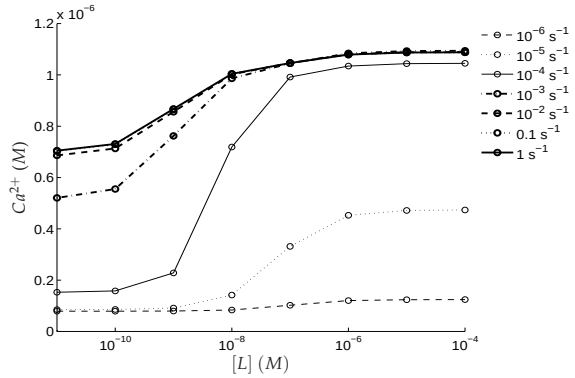
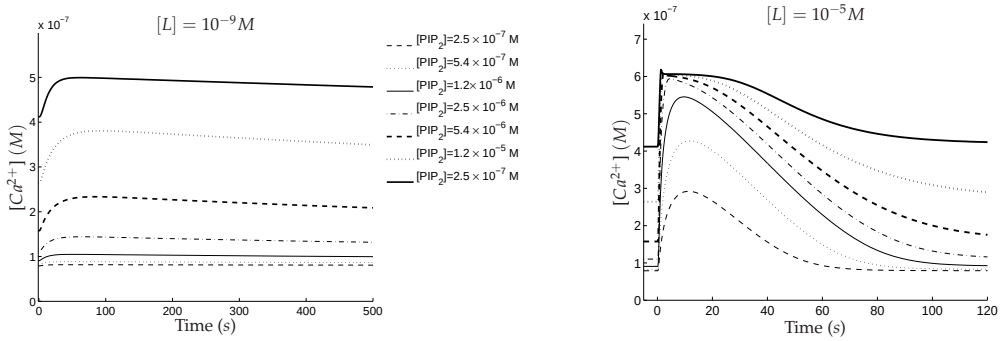


Figure 5.9: Ca^{2+} dose-response curves for Oxotremorine-M, using different values of k_{PIP_2rep} indicated in the legend. While the maximal response is robust over most values, the minimal response is quite variable.



(a) (b)

Figure 5.10: The effects of varying $[PIP_2]$ in the original CM on $[Ca^{2+}]$ in response to (a) $10^{-9}M$ and (b) $10^{-5}M$, of Oxotremorine-M. Features like the Ca^{2+} Peak-over-Basal, and Ca^{2+}-Tail steepness can be sensitive. Fast rates of PIP_2 replenishment are acceptably approximated by the assumption that $[PIP_2]$ is constant.

5.4 Discussion

In this chapter, two extended versions of the original CM were derived and analysed to address modelling questions that arose in Chapter 4.

The first variant model, the CMMP, incorporated the action of membrane pumps as an additional source of cytosolic Ca^{2+} removal, which was solely carried out by the SERCA pumps in the original CM.

Parameter variation was carried out to identify the PMCA pump parameter values (V_P and K_P) which had the strongest impact on the Ca^{2+}-Tail steepness. Even though there were V_P and K_P values which could increase the steepness of the Ca^{2+}-Tail, they were not significant enough to make the tails at the lower concentrations anything as steep as they are in the data. This suggests that it is not crucial to include the membrane pumps in the model, and that the original CM might just need parameter optimisation to better predict the Ca^{2+}-Tails.

The second variant model, the CMPDR, unlike the original CM did not assume that $[PIP_2]$ was constant; it was derived by incorporating a basic cycle of PIP_2 depletion and replenishment into the original CM. It was shown that, if not fast enough, the rate of PIP_2 replenishment can diminish the Ca^{2+} response, and thus agonist efficacy. It was also seen that at the higher agonist concentration, $[L] = 10^{-5}\ M$, the Ca^{2+} Peak-over-Basal and the Ca^{2+}-Tail steepness can be sensitive to changes in the rate constant of PIP_2 replenishment. Similar changes in those features could be achieved by varying $[PIP_2]$ in the original CM. Hence, the the models can be considered as roughly equivalent. This suggests that the better modelling strategy is to optimise the value of constant $[PIP_2]$, rather than model depletion and replenishment, which makes the model slightly more complex.

Neither of the variants considered here drastically improve the predictions of the original CM; in particular, they do not improve the prediction of the Ca^{2+}-Tail steepness when $[L] = 10^{-9}\ M$. Therefore, as it is simpler than its variants, the original CM will be retained as the model for further analysis in the next chapter, where a global parameter optimisation is performed.

CHAPTER 6

Parameter optimisation

6.1 Introduction

In Chapter 4, a global sensitivity analysis (GSA) was used to rank the parameters of the Ca^{2+}-mobilisation model according to their influence on the Ca^{2+} response. Guided by the GSA results, parameter adjustments were made to the *base parameter set* (found in Table B.2.2), which resulted in the best qualitative predictions of the Ca^{2+} response made by the model so far. The resulting parameter set (which can be found in Table B.2.2) will be referred to as the *GSA-guided parameter set*. In this chapter, a systematic parameter optimisation is carried out to improve the quantitative fit to the Ca^{2+} data, and gain more insight into pathway dynamics.

Parameter estimation is an optimisation problem, since the parameter set with which the model output least deviates from the data is sought. Specifically, it is a minimisation problem, which can be stated as follows: find the n-dimensional vector, $\boldsymbol{x}^*$ that minimises an objective function,

$$F(\boldsymbol{x}) : \mathbb{R}^n \rightarrow \mathbb{R}, \tag{6.1}$$

where, given that $\tilde{\mathbf{Y}}_\mathbf{l}$ is a vector of experimental data measured at k time points in the lth experiment,

$$F(\boldsymbol{x}) = \sum_{k=1}^{m} \sum_{l=1}^{d} (\tilde{Y}_{l,k} - Y_{l,k}(\boldsymbol{x}))^2 = \sum_{l=1}^{d} \|\tilde{\mathbf{Y}}_l - \mathbf{Y}_l(\boldsymbol{x})\|^2. \tag{6.2}$$

Equation (6.2) is also often referred to as the least squares function. For the available experimental Ca^{2+} data (discussed in previous chapters, and reproduced in Figure 6.1), the number of time points is given by $k = 34$ and the number of experiments is $d = 8$, since each time course represents an experiment with stimulation by one of eight agonist concentrations.

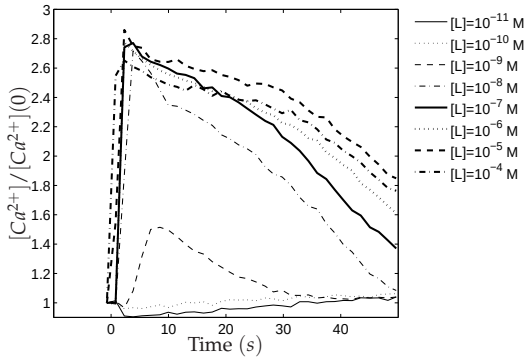


Figure 6.1: The Ca^{2+} data, to which the model will be fitted [52]; Ca^{2+} fluorescence is normalised to the basal level. The number of time points, counted from the time of agonist stimulation ($t = 0$), is $k = 34$.

6.2 Optimisation methods

For nonlinear models such as the Ca^{2+}-mobilisation model, there often exist many solutions to the minimisation problem, that satisfy the following:

$$F(\boldsymbol{x}^*) < F(\boldsymbol{x}) \qquad \forall \quad \boldsymbol{x} \in S \subseteq \mathbb{R}^n \tag{6.3}$$

$$F(\boldsymbol{x}^l) \leq F(\boldsymbol{x}) \qquad \forall \quad \boldsymbol{x} \in S_s \subset S, \tag{6.4}$$

where S (referred to as the *search space*) is the region of parameter space to which the search is confined and S_s is the set of points in the neighbourhood of $\boldsymbol{x}^l$. $\boldsymbol{x}^*$ is the global minimum, while $\boldsymbol{x}^l$ is a local minimum. It is common for some nonlinear models to have multiple local, and even global minima. In such cases, the objective function is said to be *multimodal*. The results of the global sensitivity analysis (GSA) in Chapter 4 show that there are many parameters to which the output is relatively insensitive, meaning that these parameters can vary over orders of magnitude without significantly changing the value of the cost function. The Ca^{2+} model can thus be expected to have either multiple local minima for F and/or flat sections on its surface, over which its value is approximately invariant.

6.2.1 Local optimisation methods

Local optimisation methods use local properties of the objective function to guide the search, and so when the cost function is multimodal, easily converge to the minimum closest to the starting point of the search [60]. These methods may be divided into two classes: *gradient-based*, which use function derivatives to guide the search, and *direct search*, which make use of only function evaluations.

A popular direct search method is the Nelder-Mead method, which is a typical simplex search method (a simplex is an nth-order generalisation of a triangle). A simplex of $n+1$ vertices is generated in $S \subset \mathbb{R}^n$, and the function is evaluated at each vertex. The vertex with the worst value is replaced, creating a new simplex by reflection, expansion or contraction, and the search is continued. A sequence of simplexes with improving function values at the vertices is generated, until the minimum is found, or some other specified criterion is met [66].

There are many gradient-based optimisation methods which are known to quickly converge to a local minimum, such as the Levenberg-Marquardt, and Gauss-Newton Methods, used for least squares functions. Starting from an initial guess $\boldsymbol{x}^0$, the Gauss-Newton method takes iterative steps towards the minimum, as defined by $\boldsymbol{x}^{a+1} = \boldsymbol{x}^a + \mathbf{h^a}$. Let $\tilde{\mathbf{Y}}_\mathbf{l} - \mathbf{Y_l}(\boldsymbol{x}) = \mathbf{r_l}(\boldsymbol{x})$, then equation (6.2) can be written as:

$$F(\boldsymbol{x}) = \sum_{l=1}^{d} \mathbf{r_l}(\boldsymbol{x})^T \mathbf{r_l}(\boldsymbol{x}). \tag{6.5}$$

Using a Taylor function-linearisation of $F(\boldsymbol{x}^a + \mathbf{h^a})$, it can be shown that

$$\mathbf{J}^T(\boldsymbol{x}^a)\mathbf{J}(\boldsymbol{x}^a)\mathbf{h^a} = -\mathbf{J}^T(\boldsymbol{x}^a)\mathbf{r_l}, \tag{6.6}$$

where $\mathbf{J}$ is the Jacobian matrix, which contains the first-order partial derivatives of $\mathbf{r_l}$. $\mathbf{h^a}$ is calculated from equation (6.6), and the iterations are repeated until the minimum is reached [67].

The Levenberg-Marquardt method modifies the Gauss-Newton method by adding a step size modifier as follows:

$$(\mathbf{J}(\boldsymbol{x}^a)\mathbf{J}^T(\boldsymbol{x}^a) + \boldsymbol{\lambda} \boldsymbol{I})\mathbf{h^a} = -\mathbf{J}^T(\boldsymbol{x}^a)\mathbf{r_l}, \tag{6.7}$$

where $\mathbf{I}$ is the identity matrix [67]. This is an improvement on the Gauss-Newton method because when the iterate is far from the minimum, the step size can be adjusted in the direction of the gradient. Other optimisation methods are discussed in detail elsewhere [25, 68].

Local optimisation methods, while effective in situations where there is only one minimum, or a good guess for the global minimum is known, are not appropriate otherwise, as they simply converge to the minimum closest to $\boldsymbol{x}^0$. On the other hand, global optimisation methods, which take larger parts of the search space into account, are more likely to escape local minima; hence, such a method is preferred for the Ca^{2+}-mobilisation model.

6.2.2 Global optimisation methods

Global optimisation (GO) methods can be divided into two classes: deterministic and stochastic. Deterministic methods often use information about the mathematical structure of the problem to carry out the search for the optimum. For example, trajectory methods use search strategies that generate paths such that at least one global minimum lies on one path. The paths tend to be differential equation-based, generated by an equation of motion, such as Newton's law for a particle of mass in a potential, subject to a dissipative force [21]. Trajectory methods aim to visit and enumerate all the stationary points of the objective function, thus identifying all global and local minima; this means that they are generally computationally demanding, and impractical for high order problems [72].

Another class of deterministic methods are branch and bound methods, which do not necessarily require mathematical properties of the objective function to be known. Branch and bound methods have been implemented in various ways, but are based on a general underlying principle.

An upper bound for the objective function, F^U, is set; this can be done by picking a random value, or more systematically —for example, choosing a local minimum, using a local optimisation method, such that $F^U = F(\boldsymbol{x}^l)$. A lower bound is also estimated using established methods (convex relaxation, duality, Lipschitz or other bounds [11]). The search space is then partitioned into a family of sets, using a selected method. Lower bounds and upper bounds are estimated for each subset, and those sets whose lower bounds are greater than the upper bound (hence not containing any elements that give an objective function value less than F^U) are eliminated, and permanently ignored. The upper bounds and lower bounds are then updated. Among the remaining sets, the one which contains an element giving the lowest function value is repartitioned, and the upper bound is updated with this function value. The previous steps are then repeated until no sets containing elements with lower objective function values can be found [12].

A key idea of this algorithm is that time is saved by exploring the most promising regions of the search space first, but in the worst cases, the search effort grows exponentially with the problem size.

In theory, deterministic methods are guaranteed to converge to the global optimum; in practice however, as parameter space dimensionality increases, deterministic methods are unable to converge in finite time [59]. Deterministic methods tend to carry out the search through space sequentially, while stochastic methods such as evolutionary algorithms tend to carry out parallel searches through space using a diverse set of trial solutions. On the other hand, stochastic methods use probabilistic rules to guide the

search, which involves random elements, and so are not guaranteed to find the global optimum; however, they often find very good solutions (in the vicinity of the global optimum) in finite time [59, 60].

In an informative study [59], representative GO methods were compared by testing their performance when estimating the parameters of a benchmark problem: an eighth order nonlinear ODE model of a metabolic pathway with thirty six parameters. Being a high-dimensional search space, this represents a challenging optimisation problem, with which the ability of existing GO methods to handle difficult estimation problems can be tested. The deterministic methods: Global optimisation using the DIRECT algorithm (GBLSOLVE), Multilevel Coordinate Search algorithm (MCS) were tested. Both these algorithms are related; they are similar to Branch and bound methods, as they divide the search space into smaller and smaller, non-intersecting hyperrectangles, but they do not make use of bounds [34]. MCS adds a heuristic aspect to the basic deterministic approach [59]. The following stochastic methods were also tested: Differential evolution (DE), Integrated Control Random Search algorithm (ICRS), Unconstrained evolution strategy (uES), Evolution strategy using stochastic ranking (SRES) and the Evolution strategy with covariance matrix adaptation (CMA-ES) [59].

Pseudo-experimental data were generated by simulating the benchmark problem with a fixed parameter set, so that in this case the solution to equation (6.1), $\boldsymbol{x}^*$, was known. Sixteen sets of data were produced using various initial concentrations of substrate and product. The Evolution strategy using stochastic ranking (SRES) produced the best results of the tested methods. With SRES, all 36 parameters were estimated with error less than 17%, the majority being estimated with less than 3% error, showing that convergence was towards the global minimum. The other methods, with the exception of uES, perfomed relatively poorly, as shown in Figure 6.2. Based on these results, this chapter uses SRES (which is described in the next section) to address the minimisation problem of the Ca^{2+} model.

6.3 Evolutionary algorithms

The evolution strategy using stochastic ranking (SRES) is one version of the stochastic group of GO methods called evolution strategies, which in turn belongs to the group of stochastic GO methods referred to as *Evolutionary algorithms*.

Evolutionary algorithms iteratively apply ideas from the biological theory of evolution, such as selection, reproduction, crossover (or recombination), and mutation, to improve trial solutions, in hopes of approaching the global optimum. The general scheme of an evolutionary algorithm consists of (1) below, (2)–(4) in variable order, and

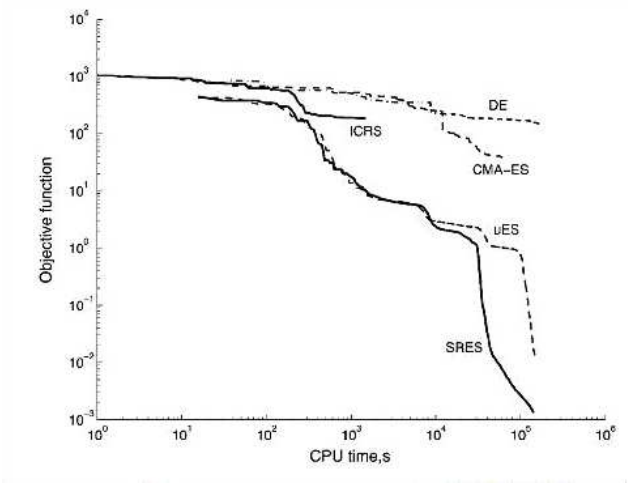


Figure 6.2: Results of a study in which several global optimisation methods were compared. The evolution strategies, uES and SRES, were by far the best performing methods. The figure was taken from [60].

(5):

1. INITIALISATION: Generate randomly or deterministically chosen trial solutions, collectively referred to as the *population*; population members are often referred to as *individuals*.

2. FITNESS: Evaluate the fitness of population members, using an objective measure such as $F(\boldsymbol{x})$.

3. REPRODUCTION: Create *offspring* from all, or selected individuals (called *parents*), using either or both of the following:

   (a) MUTATION: Produce randomly altered forms of the parents.

   (b) RECOMBINATION/CROSSOVER: Combine traits of parents to form new individuals.

4. SELECTION: Select parents and offspring with best fitness and use them to *update* the population (for evolution strategies, this is referred to as the (μ+λ) strategy); or select only highest ranked parents to reproduce for the next *generation* (for evolution strategies, referred to as the (μ,λ) strategy). The latter strategy allows for a more diverse population, as it prevents original parents from surviving into the next generation [23].

5. TERMINATION: Repeat (2)–(4) until a specified criterion, such as an acceptable error or maximum number of generations, is reached.

Many evolution strategies use all of the above steps. uES performs recombination and mutation before using selection to update the population; SRES on the other hand, performs selection before recombination and mutation, so that only the best individuals are allowed to reproduce. This might explain why SRES converged more quickly than uES in the aforementioned study [59] (see Figure 6.2).

In step 3(a), *strategy parameters* are used to determine the extent to which parents are mutated, and a unique feature of evolution strategies among evolutionary algorithms is that these mutation parameters are adaptive; they are part of the trial solutions which evolve across generations. Hence, a vector of the strategy parameters, $\boldsymbol{\sigma}$, is included in each member of the population, which is represented by $(\boldsymbol{x}, \boldsymbol{\sigma})$. Each σ_j $(j = 1, 2, ..., n)$ is subject to mutation and is used to define the mutation step size of the jth parameter. The particular way in which SRES implements steps (1)–(5) above is explained in the next section.

6.3.1 Evolution strategy using stochastic ranking (SRES)

An initial population of λ trial solutions, $\boldsymbol{x_i}$, $i = 1, ..., \lambda$, is randomly generated from an n-dimensional uniform distribution (allowing an unbiased spread across the search space), while for the strategy parameters initial guesses are given by:

$$\sigma_{i,j}^{(0)} = \frac{\overline{x}_j - \underline{x}_j}{\sqrt{n}}, \quad i \in \{1, ..., \lambda\}, \quad j \in \{1, ..., n\}, \tag{6.8}$$

where $\overline{x}_j$ is the upper bound of the jth parameter, while $\underline{x}_j$ is its lower bound; hence, the initial step size decreases as the dimension of the search space increases. The superscript in $\sigma_{i,j}^{(0)}$ in equation (6.8) is the generation counter, g; the value of zero indicates that this is the value of $\sigma_{i,j}$ before the first generation.

$F(\boldsymbol{x_i})$ is calculated for each population member, $(\boldsymbol{x_i}, \boldsymbol{\sigma_i})$, and they are ranked accordingly. The next operation is selection, by which the highest ranked μ (out of λ) individuals, $(\boldsymbol{x_i'}, \boldsymbol{\sigma_i'})$, $\forall i \in 1, 2, ..., \mu$ are chosen as parents, while all others are discarded. The population is then completely updated, with each $(\boldsymbol{x_i'}, \boldsymbol{\sigma_i'})$ replicated $P_S = \frac{\lambda}{\mu}$ times on average, as illustrated in Figure 6.3. (The population size (λ) thus remains constant.) The ratio, P_S, is normally referred to as the *selection pressure* and, because it determines the number of parents that survive to reproduce, influences the regions of space in which the search is carried out. If P_S is too high and the population size is too small, diversity will be lost, and the search will tend to be carried out in subregions of the search space [8]; the choice of these parameters can therefore influence the efficiency of the search.

After selection, *global intermediate recombination* [78] is carried out on the updated

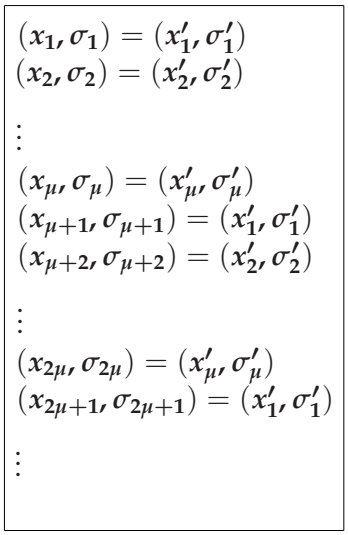


Figure 6.3: An illustration of the process by which the population is updated with the fittest individuals (selection), which then go on to reproduce for the next generation (reproduction).

strategy parameters according to the following rule:

$$\hat{\sigma}_{i,j}^{(g)} = (\sigma_{i,j}^{(g)} + \sigma_{k_j,j}^{'(g)})/2, \quad i = 1, ..., \lambda, \quad k_j \in \{1, ..., \mu\}, \tag{6.9}$$

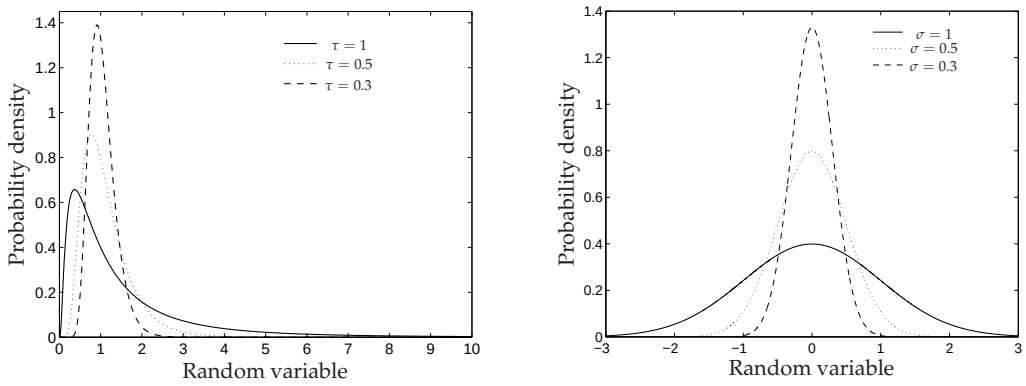
where k_j is an index generated at random and anew for each j. Hence, a strategy parameter may be replaced by its average with another randomly chosen strategy parameter from the parent population. For instance, the jth dimension (parameter) in the first individual could be updated by $\hat{\sigma}_{1,j}^{(g)} = (\sigma_{1,j}^{(g)} + \sigma_{2,j}^{'(g)})/2$ or $\hat{\sigma}_{1,j}^{(g)} = (\sigma_{1,j}^{(g)} + \sigma_{5,j}^{'(g)})/2$, etc; or it could be unchanged via the choice, $\hat{\sigma}_{1,j}^{(g)} = (\sigma_{1,j}^{(g)} + \sigma_{1,j}^{'(g)})/2$. This step allows traits of the good strategy parameters to spread across the population. It has been shown that recombination can increase the rate at which the offspring population converges to its lowest *average* fitness value [97].

The population for the next generation is obtained by mutation, which is first carried out multiplicatively on each strategy parameter, and additively to each (model) parameter, as follows:

$$\sigma_{i,j}^{(g+1)} = \hat{\sigma}_{i,j}^{(g)} \exp(N(0, \tilde{\tau}) + N_j(0, \tau)) \tag{6.10}$$

$$x_{i,j}^{(g+1)} = x_{i,j}^{(g)} + N_j(0, \sigma_{i,j}^{(g+1)}), \tag{6.11}$$

where $N(0, *)$ is a random number from a normal distribution with mean, 0, and standard deviation, $*$; $N_j(0, *)$ is similar, but generated anew for the j^{th} parameter. $\exp(N(0, \tilde{\tau}))$ and $\exp(N_j(0, \tau))$ in equation (6.10) are thus numbers from the lognormal distribution; the former applies the same mutation to each component, while the



(a) (b)

Figure 6.4: (a) Scaling factors of σ in equation (6.10) are taken from the lognormal distribution ($\exp(N(0,\tau))$), which makes smaller mutations more likely than larger ones (at the tails of the curve). Smaller values of τ increase the probability of $\exp(N(0,\tau))$ values close to 1. (b) The step sizes in equation (6.11) are taken from the normal distribution, making small values more likely. The spread of the normal distribution $N(0,\sigma)$ is controlled by its standard deviation, σ. Smaller σ values increase the probability of a step size close to 0.

latter applies dimension-wise mutation. The (multiplicative) mutation factors for $\hat{\sigma}_{i,j}$ (see equation (6.10)) are taken from the lognormal distribution because smaller scalings (gathered around the peak) are more likely than larger ones (found at the tails), as shown in Figure 6.4(a). $\acute{\tau}$ and τ are called learning rates, and as illustrated in Figure 6.4(a), the smaller they are, the more likely that $\exp(\acute{\tau}N(0,1))$ and $\exp(\tau N_j(0,1))$ will be close to 1. Hence, $\acute{\tau}$ and τ influence which the speed at which $\sigma_{i,j}^{(g+1)}$ is adapted. The values $\acute{\tau} = \frac{1}{\sqrt{2\sqrt{n}}}$ and $\tau = \frac{1}{\sqrt{2n}}$ have been recommended based on theoretical and practical considerations, and have been shown to be efficient in some multimodal problems [8, 83].

The additive step sizes (see equation (6.11)) are taken from the normal distributions because smaller values are more likely, as shown in Figure 6.4(b). If mutation takes a parameter value outside its defined bounds, it is re-tried up to ten times; if these attempts are unsuccessful, the parameter is left unmutated to limit computational expense. Since each $x_{h,j}^{(g+1)}$ is derived by taking a step size determined by a mutated value of $\hat{\sigma}_{i,j}^{g}$, the worst mutations will be rejected in the next generation when only the fittest μ offspring are selected. Hence, the step sizes evolve to optima as well as the model parameters.

The cycle is repeated until some specified user criterion is met (for example, maximum number of generations, or minimum objective function value).

SRES was designed with constrained optimisation in mind, hence it can also use a penalty function to penalise trial solutions which violate known parameter constraints. The method then adjusts the balance between the contributions of the penalty function and the objective function to the ranking of an individual's fitness [78]. The only constraints used for the Ca^{2+}-mobilisation model are the parameter ranges, thus SRES is used with the value of the penalty function as zero.

6.4 Searching with SRES for known global minima

In this section, the ability of SRES to find a known global minimum is first tested, the results of which will be used to interpret its performance when fitting the Ca^{2+} model to actual experimental data later in section 6.5.2.

6.4.1 Generation of pseudo-data

The Ca^{2+} model is used to generate what is referred to as *pseudo-data*, which is the Ca^{2+} response when the model is simulated with a chosen parameter set, whose members are then referred to as the 'nominal' values. Hence, a global minimum, $\boldsymbol{x}^*$ (chosen to be the GSA-guided parameter set in this case) for this particular estimation problem is known, and the accuracy of the estimates found by SRES can be assessed. In this case, the data in equation (6.1) is then given by $\tilde{\mathbf{Y}}_l = \mathbf{Y_1}(\mathbf{x}^*)$, and the objective function by:

$$\bar{F}(\boldsymbol{x}) = \sum_{l=1}^{d} \|\mathbf{Y_1}(\mathbf{x}^*) - \mathbf{Y}_l(\boldsymbol{x})\|^2. \tag{6.12}$$

To mimic actual experimental data, which often have measurement errors, a noisy version of the pseudo-data, $\tilde{\mathbf{Y}}_l^n$, is also generated using the following:

$$\tilde{\mathbf{Y}}_l^n = \tilde{\mathbf{Y}}_l(\boldsymbol{x}^*) + \boldsymbol{\Theta} \cdot \tilde{\mathbf{Y}}_l, \tag{6.13}$$

where Θ is a k-length vector of random numbers between -0.05 and 0.05, randomly chosen from the uniform distribution. The Ca^{2+} pseudo-data are generated at the same time points, and for the same ligand concentrations, as the experimental data, in order to set up a test case as close as possible to the actual optimisation problem; it is then more straightforward to make some direct comparisons. The pseudo-data have been normalised using the basal value, so that F is dimensionless; hence $\boldsymbol{Y}_l = [Ca^{2+}]/[Ca^{2+}(0)]$.

Ranking	Parameter
1	$k_{Ca^{2+}act-}$
2	k_{PIP2b+}
3	k_{IP_3Rb-}
4	p_1
5	k_{IP_3Rb+}
6	k_{des+}
7	$[PIP_2]$
8	k_{deact}
9	p_2
10	R_{TOT}

Table 6.4.1: The ten most sensitive parameters, identified by a GSA carried out in a parameter space confined around the GSA-guided parameter set. These 10 parameters are referred to as the GSA Top 10.

6.4.2 Reduction of search space dimension

Using the method of elementary effects (as described in Chapter 4), a global sensitivity analysis (GSA) was carried out in a parameter subspace constrained around the GSA-guided parameter set, since it will serve as the base set around which the search space S is constrained. Table 6.4.1 lists the ten parameters (out of forty) to which the Ca^{2+} response is most sensitive, and, as discussed in section 4.5.4, they are assumed to be the ones whose values most need to be determined by parameter estimation; consequently, the dimensions of the search space can be significantly reduced, so that $S \subseteq \mathbb{R}^{10}$ instead of $S \subseteq \mathbb{R}^{40}$. This assumption will be scrutinised in a later section, by comparing optimisation results from both search spaces.

The results are similar to those obtained for the previous base set in Chapter 3, with six parameters in common. The newly identified parameters are k_{IP3Rb+}, k_{IP3Rb-}, k_{deact} and R_{TOT}, replacing IP_3R_{TOT}, k_{IP_3deg}, γ_1 and k_{IP_3}.

Let the $W \in \{1, 2, ..., 10\}$ most sensitive parameters in Table 6.4.1 be referred to as the GSA Top W, then corresponding subspaces $S_G^W \subset \mathbb{R}^W$ (or hyperplanes in $\mathbb{R}^{40}$) can be defined, by fixing all other parameters at their nominal values (the subscript G refers to the fact that the GSA Top W serve as the axes in the subspace). The higher the value of W, the more dimensions there are to explore, hence lower values of W simplify the estimation problem.

The main minimisation problem in this section will be that solved in S_G^{10}. However, SRES is first used to search in $S_G^2 \subset \mathbb{R}^2$ for the nominal values of the GSA Top 2. The objective function surface in this subproblem can be visualised, and thus provide insight into the performance of the algorithm, as well as the topology of the search space. The search is then carried out in S_G^4, for the nominal values of the GSA Top 4, before

finally being carried out in S_G^{10} for all the values in Table 6.4.1. Hence, the dimensions of the solution, $\boldsymbol{x}^*$, vary in this section. By first tackling these subproblems, a sense of the computational cost attached to the main problem and of suitable SRES search parameter values, such as the population size and number of generations, can be gained.

The full search space, $S \subset \mathbb{R}^{40}$ is defined according to parameter space 2 in section 4.4.3, except with molecular concentrations varying over four orders of magnitude around their base values in the GSA-guided parameter set in Table B.2.2.

6.4.3 SRES search parameters

As with most stochastic algorithms, there are no fixed rules for choosing values for the SRES search parameters: the population size, λ, the number of generations, G, and the selection pressure P_S (or equivalently, the number of parents, μ); there is an element of trial and error involved in selecting them. It is also important, when choosing them, to remember certain factors that may influence the performance of the routine. For instance (as noted before) the value of P_S can influence how well the search space is explored. Several publications have recommended a selection pressure of 7 [43, 78, 97], which will be used as a starting value here. The size of the population also influences how well the search space is explored [8], and should be adjusted according to the dimension of the search space.

The aforementioned study [60], in which SRES was the best performing global optimisation method, used the search parameters $\lambda = 350$, $G = 8000$ and $P_S \approx 11$ for the eight order benchmark problem, which has thirty six parameters. These values are not practical for the Ca^{2+} model, which is being fit to data from $d = 8$ experiments (see Figure 6.1), and would require $\lambda \times G \times l = 22,400,000$ model evaluations. Moreover, when the sensitivity analysis was implemented in Chapter 4, it was seen that there are parameter combinations that significantly increase model evaluation time (see section 4.4.4); hence, a model evaluation could take anything from one tenth of a second to tens of seconds. Large values of λ and G would thus make the optimisation computationally unaffordable. However (at least initially) only estimates for the ≤ 10 most sensitive parameters are sought, so it is unlikely that a population size as large as $\lambda = 350$ is required. Different λ values will be tested, depending on the number of parameters being estimated. The choices for G will be discussed as the runs are implemented.

The Ca^{2+}-mobilisation model has 24 variables, and available data are for 1 variable ($[Ca^{2+}]$), while the aforementioned benchmark model had 10 variables, and data for each variable, generated from 16 different initial conditions. Hence, parameter estimation for the Ca^{2+} model is a significantly more challenging problem.

Ranking	Parameter
1	p_2
2	$k_{Ca^{2+}act+}$
3	k_{GTP+}
4	k_{IP_3Rb+}
5	R_{TOT}
6	k_{PIP2b+}
7	$[PIP_2]$
8	$k_{Ca^{2+}act-}$
9	k_{IP_3Rb-}
10	k_{IP_3deg}

Table 6.4.2: The ten most sensitive parameters, identified by a local sensitivity analysis (LSA) carried out at the point in space which corresponds to the GSA-guided parameter set. These 10 parameters are referred to as the LSA Top 10.

6.4.4 Local sensitivity of $\bar{F}$

In this section a local sensitivity analysis (LSA) is carried out at the point in parameter space which corresponds to the GSA-guided parameter set. Determining the most sensitive parameters at that point gives an indication of the coordinate-wise landscape of $\bar{F}$ in the vicinity of the global minimum, $\boldsymbol{x}^*$; thus, the steepest $\bar{F}$ changes along the axes can be estimated. It is useful to have such information because flatness in some parameter directions around the global minimum can make those parameters hard to estimate [60], as many possible values would give roughly the same low value of F; thus, numerous individuals with similarly high fitness exist and can be favoured by the (SRES) selection step, which therefore becomes of little use in focusing the progression to $\boldsymbol{x}^*$; the routine can thus come to resemble a purely random search in the flat regions.

The output feature of interest for the LSA is $\bar{F}$ itself; the sensitivity with respect to the j^{th} parameter is then given by:

$$\bar{C}_j = \frac{\bar{F}(\boldsymbol{x}^* + \mathbf{0}_{\delta_j}) - \bar{F}(\boldsymbol{x}^*)}{\delta_j / x_j^*}, \tag{6.14}$$

where $\mathbf{0}_{\delta_j}$ is an n-dimensional vector of zeros, but with its jth element equal to δ_j, which is taken to be $\pm 0.1 x_j^*$ (note that $\bar{F}(\boldsymbol{x}^*) = 0$, hence it is not used to divide the numerator). The average of $|\bar{C}_j|$ due to both $\pm 0.1 x_j^*$ is calculated and used to rank the parameters, the results of which are given in Table 6.4.2.

The $W \in \{1, 2, ..., 10\}$ most sensitive parameters in Table 6.4.2 are then referred to as the LSA Top W. Although many of the GSA Top 10 are also among the locally most sensitive, they tend to be ranked differently (compare with Table 6.4.1); this is because the GSA ranks the model parameters using an *average* of sensitivities calculated at se-

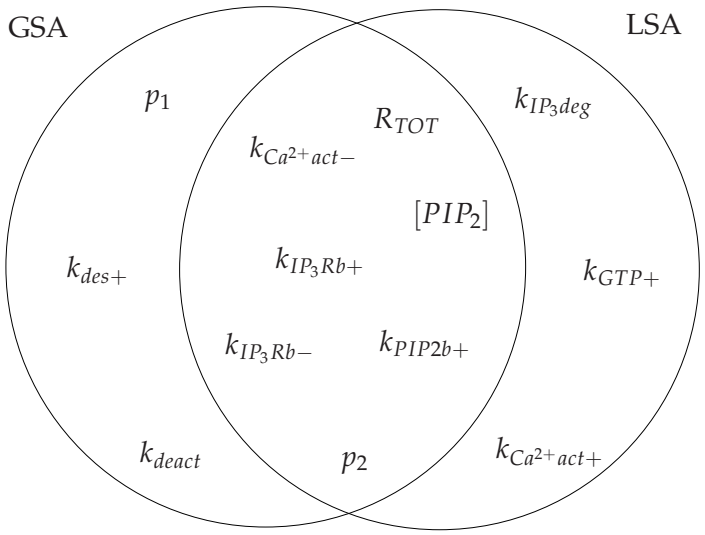


Figure 6.5: A comparison of the globally and locally most sensitive parameters. There are seven parameters which are identified by both sensitivity analyses.

lected points in parameter space, and so the GSA ranking, though representative of the whole space, may be different at individual points. Probably for that reason, there are parameters in the LSA Top 10 not represented in the GSA Top 10 (and vice versa), as illustrated in Figure 6.5, their average significance across parameter space being dissimilar to their local significance. Hence, some of the steepest directions around $\boldsymbol{x}^*$, $k_{Ca^{2+}act+}$, k_{GTP+} and k_{IP_3deg} are not represented in the GSA Top 10. When the GSA Top 10 (or any subset within) are being estimated, $k_{Ca^{2+}act+}$ and k_{GTP+} will be fixed at their nominal values, and since they are ranked high (2 and 3 respectively) in terms of local sensitivity, it might be possible to get good fits without getting close estimates for other (locally less sensitive) parameters in the GSA Top 10.

It should be noted that due to model nonlinearity (which was clearly seen in Chapter 4) the LSA results might not apply to more global regions; also, as sensitivity coefficients are calculated one parameter at a time, the LSA results do not account for effects due to parameter correlation, even within the local region.

6.4.5 Estimating two parameters

In this section the search is first carried out in the hyperplane $S_G^2 \subset \mathbb{R}^2$, to find the nominal values of the GSA Top 2, $k_{Ca^{2+}act-}$ and k_{PIP2b+}; for comparison, S_L^2 is then searched for the nominal values of the LSA Top 2. Since there are only 2 parameters to estimate,

a population of $\lambda = 20$, and number of generations, $G = 40$ are initially chosen; selection pressure is set as $P_S = 7$, so that $\mu = 3$ parents survive for reproduction in every generation. The accuracy of the estimates will be quantified by calculating a percentage error as follows:

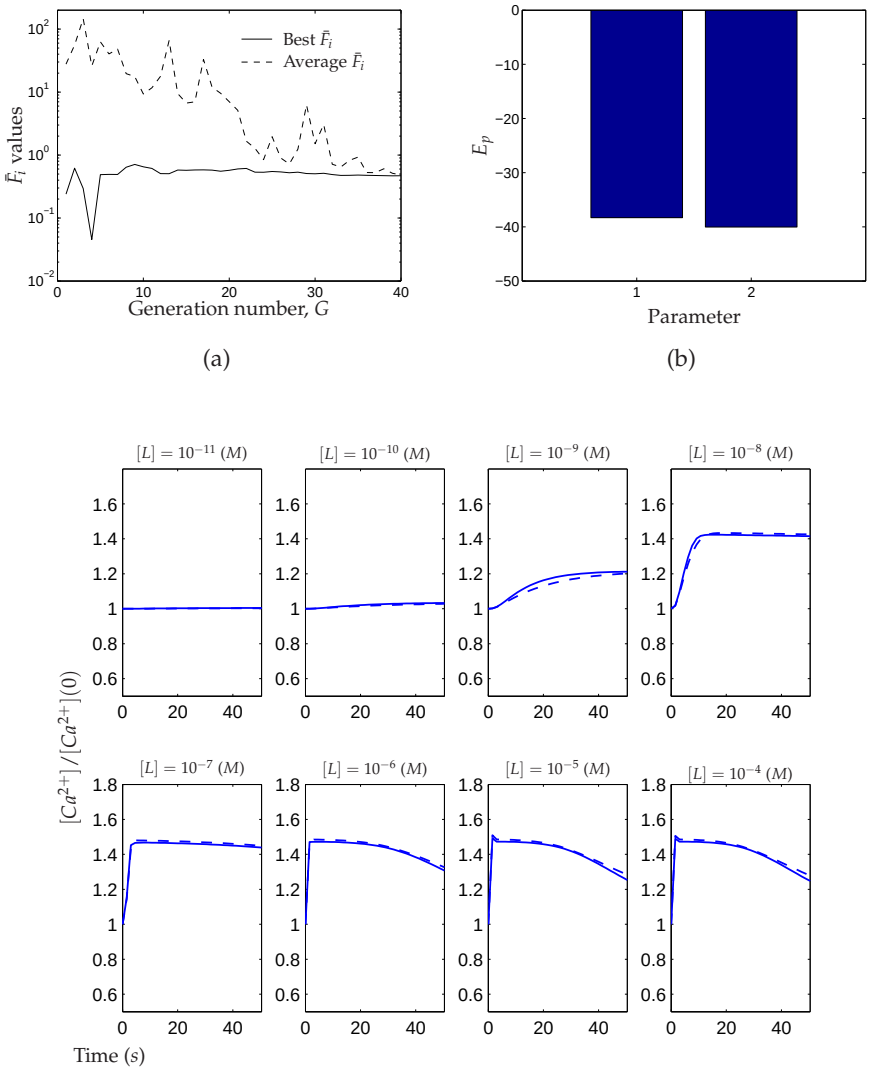
$$E_p = \frac{100(x_j^e - x_j^*)}{x_j^*}, \tag{6.15}$$

where x_j^e is the estimated value of the jth parameter, and x_j^* is its nominal value.

The results of the search are shown in Figure 6.6. Figure 6.6(a) shows that the best solution is found after 4 generations with an error of $\bar{F} \approx 0.046$; the pseudo-data and the time course predictions of the best estimates are very close (Figure 6.6(c)), because $\bar{F}$ is low, giving an average squared error of 0.000169117647 per data point, each of which is ≥ 1. However, Figure 6.6(b) shows that the estimates are $\approx -40\%$ less than their nominal values, which again raises the question of how sensitive the Ca^{2+} response is to particular parameters. It can be seen in Table 6.4.2 that $k_{Ca^{2+}act-}$ and k_{PIP2b+} are only the 8th and 6th ranked parameters in terms of local sensitivity; it follows that since the higher ranked parameters are fixed at their nominal values, good fits can be obtained even though $k_{Ca^{2+}act-}$ and k_{PIP2b+} are imperfectly estimated.

After the fourth generation, the search seems to be drawn to subregions of S which map on to similar values of $\bar{F} \approx 0.5$ (Figure 6.6(a)). The average population's fitness is also converging towards that value; hence the members of the population are becoming more similar, and little new information would be gained by increasing the number of generations, since the selection operation becomes redundant in an increasingly homogeneous population. (A large mutation step size could re-introduce diversity in the population, however, since the lognormal distribution is biased towards small mutations, such a step size has a very low probability, and would generally require many more generations to appear.) Two strategies are thus attempted to encourage further exploration of the search space; first, a lower selection pressure, $P_S = \frac{\lambda}{\mu} = \frac{20}{7} \approx 3$, is used, to allow more parents (7) to survive to reproduce; second, a higher population $\lambda = 50$ is tried, to allow a more diverse population to be considered for selection.

In Figures 6.7(a)–(b) the results from choosing $P_S \approx 3$ are shown. As intended, the rather low selection pressure keeps the average fitness of the population from converging (Figure 6.7(a)). The error is improved (from ≈ 0.046 to ≈ 0.035), but as the improvement is not a drastic one, it is hard to conclude that the lower selection pressure improves the search, because the random elements involved in the algorithm make small variations inevitable across independent searches. It can however be concluded that the lower selection pressure does not solve the problem of finding global mini-



(a) (b)

(c)

Figure 6.6: Searching in $S_G^2 \subset \mathbb{R}^2$ for the nominal values of the GSA Top 2, using the search parameters $\lambda = 20$, $G = 40$, and $P_S = 7$. (a) The best and the average $\bar{F}_i$ in each generation. (b) The errors associated with the estimates, calculated using equation (6.15). (c) The fits obtained with the best individual. The pseudo-data and solution are almost indistinguishable, even though the estimates are considerably different ($\approx \frac{2}{5}$) of the nominal values.

Subspace	λ	P_S	G	$\bar{F}$
S_G^2	20	7	40	0.045592
S_G^2	20	3	40	0.035251
S_G^2	20	3	100	0.031787
S_G^2	50	7	40	0.012738
S_G^2	50	5	40	6.813200×10^{-3}
S_L^2	50	7	40	9.444200×10^{-5}
S_G^4	50	7	100	0.016938
S_G^4	50	5	100	0.026576
S_L^4	50	7	100	1.490000×10^{-3}
S_G^{10}	50	7	100	0.125090
S_G^{10}	50	7	300	0.070371
S_G^{10}	50	5	100	0.155420
S_G^{10}*	50	5	100	0.446665
S_L^{10}	50	7	100	0.307190
S_L^{10}	50	7	200	0.198360

Table 6.4.3: A summary of the optimisation results for the trial runs with the pseudo-data. *This solution was found fitting to the noisy pseudodata.

mum's vicinity, as the errors of the estimates in this run are even larger, at 97.4% and 103% (Figure 6.7(b)). (Another run that allowed 100 generations was also carried out, but did not significantly improve the solution, giving a value of $\bar{F} \approx 0.032$.)

Figures 6.7(c)–(d) show the results of increasing the population size to $\lambda = 50$ ($P_S = 7$). The best solution is found in the first generation and is never improved thereafter (Figure 6.7(c)). In this case, SRES does not improve the best initial trial solution, whose fitness is high due to chance, helped by the large sample. The larger population size does not guide the search beyond subregions where $\bar{F} \approx 0.4$ (Figure 6.7(c)). However, the fact that an average population $\bar{F}$ value of greater than 100 is possible (Figure 6.7(c)), means that there are regions of space with high $\bar{F}$ values; the SRES algorithm does well to steer away from those regions, but in low $\bar{F}$ subregions it seems indiscriminate.

A combination of both diversity-promoting strategies is also tried by choosing $\lambda = 50$ and $P_S = 5$, the results of which are shown in Figures 6.7(e) – (f). This run produces the best solution, finding an individual which gives $\bar{F} \approx 0.0068$, but the estimates are still not very close (41.4% error for parameter 1 and 42.7% for parameter 2). It should be noted that these strategies are elaborate schemes for the estimation of only two parameters, an indication of how difficult the optimisation problem might be. However, an indication of why it is difficult can be gained; since the search space here is in $\mathbb{R}^2$, the objective function surface can be visualised (this is done at the end of this section, after all the optimisation runs have been implemented).

So far the searches have found several solutions with low $\bar{F}$ values (see Table 6.4.3 for a summary of the results from all previous (and subsequent) searches), suggesting

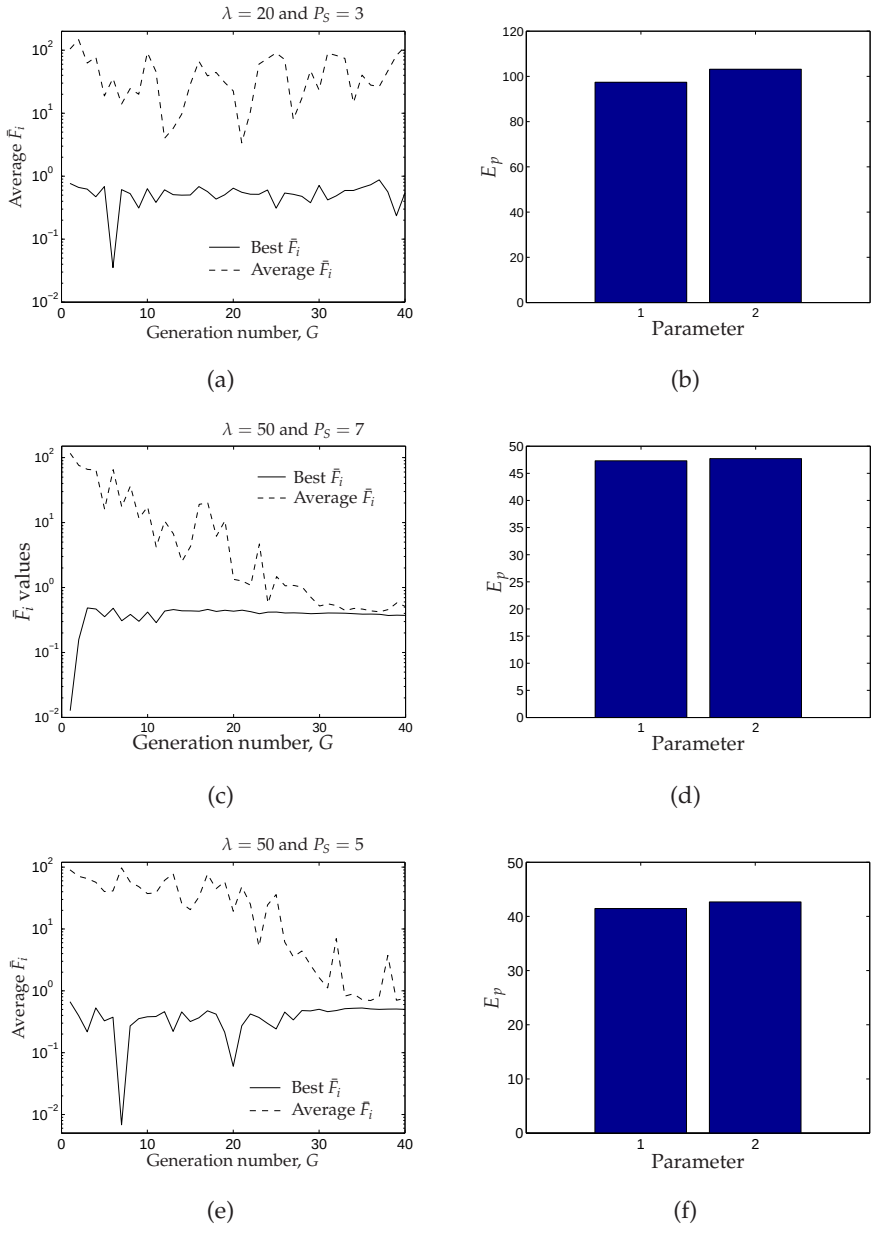


Figure 6.7: Searching in $S_G^2 \subset \mathbb{R}^2$ for the nominal values of the GSA Top 2; search parameters (specified at the top of each row) are varied to promote population diversity. (a), (c), (e) The best and the average $\bar{F}_i$ in each generation; (b), (d), (f) the errors associated with the estimates, calculated using equation (6.15). Decreasing the selection pressure seems to improve the solutions.

that there are flat regions on the objective function surface not too far from the minimum (as two of the runs found estimates with $41 - 48\%$ error (see Figures 6.7(d) and (f)). To gain perspective on how the topology of $\bar{F}$ influences the performance of the search, the LSA Top 2, p_2 and $k_{Ca^{2+}act+}$, are now estimated, while all the other parameters are fixed at their nominal values. The search is thus carried out in a subspace in which the LSA Top 2 serve as the independent axes, namely, $S_L^2 \subset \mathbb{R}^2$ (the L subscript indicates that the LSA Top 2 serve as the independent axes). Because the LSA Top 2 are the locally most sensitive parameters, around $\boldsymbol{x}^*$ the objective function surface is steepest along the axes of this hyperplane. Hence, it can be observed whether the selection step performs better on a less flat region. This might also give an indication of whether the locally sensitive parameters have to be correctly estimated to achieve a good fit.

Searching in S_L^2

This search is carried out with the higher population size $\lambda = 50$ and higher selection pressure $P_S = 7$, while $G = 40$. Figure 6.8(c) shows that for this estimation problem, SRES does perform considerably better, finding estimates for the LSA Top 2 with very small percentage error: -0.35% and $\approx 0.73\%$. The best solution gives a value of $\bar{F} = 9.4442 \times 10^{-5}$ (Figure 6.8(a)), orders of magnitude lower than was found with the best estimates of the GSA Top two; consequently, the simulations fit the pseudo data extremely well (Figure 6.8(d)). The average population fitness does not converge (it actually significantly fluctuates) towards the end of the search (see Figure 6.8(b)); this means that the selection step continues to provide information that directs the search (in fact, the majority of progress is made towards the latter half of the search, after G=23 (see Figure 6.8(b)).

Visualising $\bar{F}$

The surfaces of $\bar{F}$ in S_G^2 and S_L^2 are compared in Figure 6.9. In both subspaces there are regions of high and low $\bar{F}$ values (compare Figures 6.9(a) and 6.9(b)). Viewing the axes on log scales allows the low $\bar{F}$-value regions to be seen in more detail: the nominal values of the GSA Top 2, $(16.5s^{-1}, 10^{12}M^{-1}s^{-1})$, form the only minimum in S_G^2 (Figure 6.9(c)), while in S_L^2 there are other (high fitness) minima besides that formed by the nominal values (Figure 6.9(d)). This shows that even the LSA Top 2 do not need to be correctly estimated to achieve an acceptable fit.

The colour maps of Figures 6.9(c) and (d) are defined using a log scale, so that the low valued regions can be better observed. Both surfaces have valleys containing low value solutions (coloured dark and light blue, for $\bar{F} \leq 10^0$). All previous solutions

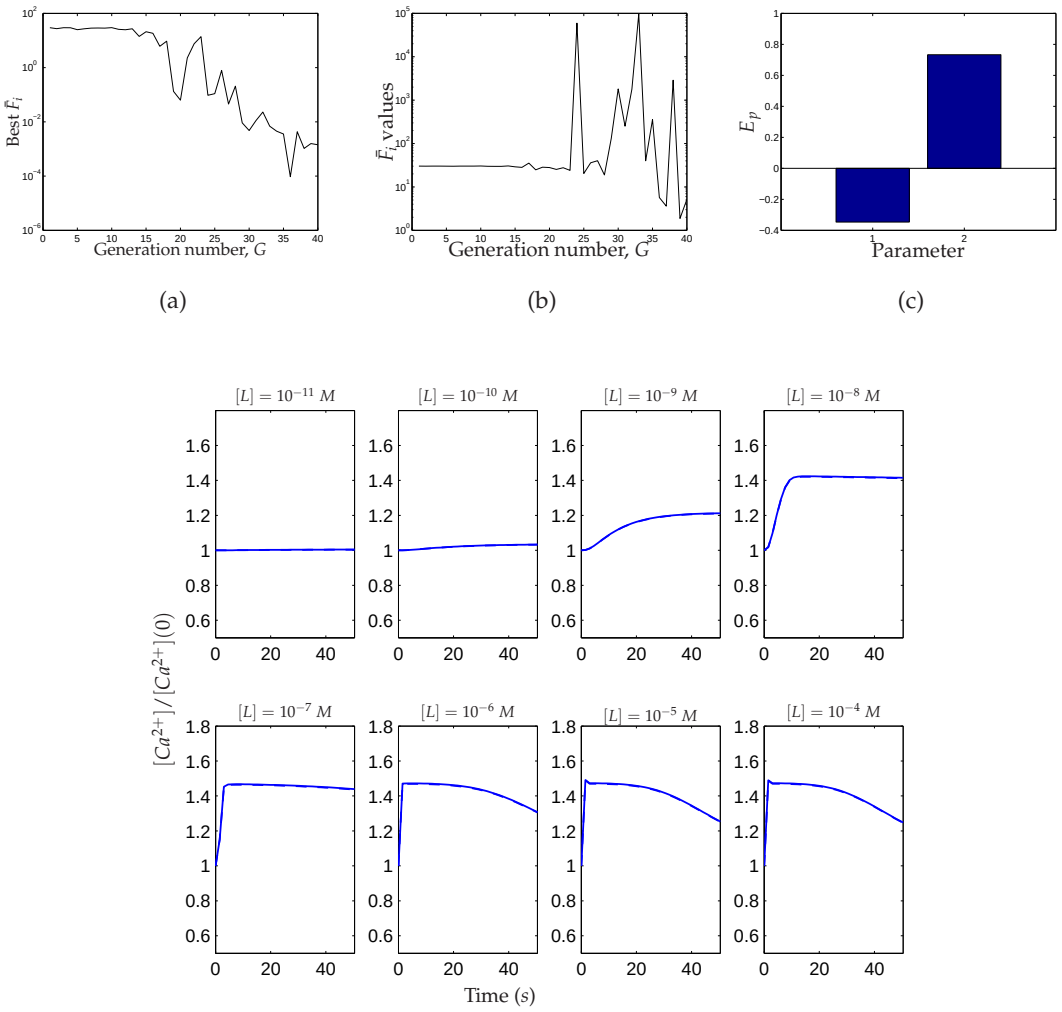


Figure 6.8: Searching in $S_L^2 \subset \mathbb{R}^2$ for the nominal values of the LSA Top 2, using the search parameters $\lambda = 50$, $G = 40$, and $P_S = 7$. (Note that the parameter labels are as in Table 6.4.2.) (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b) The average $\bar{F}_i$ of the whole population. (d) The errors associated with the estimates, calculated using equation (6.15). (d) The fits obtained with the best individual. The pseudo-data and solution are indistinguishable, and, unlike when searches were carried out S_G^2, the parameter estimates are very close to the nominal values.

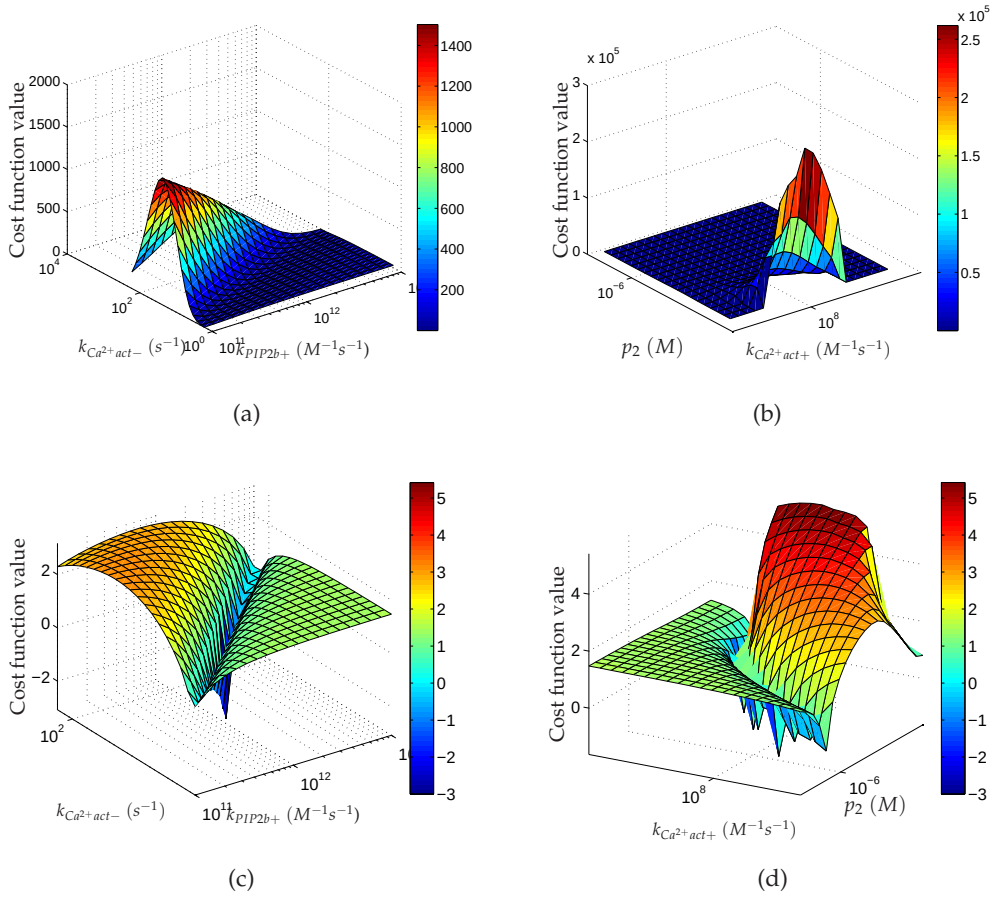


Figure 6.9: $\bar{F}$ is shown in (a) S_G^2, and (b) S_L^2. In (c) and (d), the same plots are shown using log axes; it might be inferred that optimal step sizes are hard (or require more time) to find in S_G^2 because of its flat valley; on the other hand the multiple minima in S_L^2 allow more step size information to be gained because of the more significant changes when travelling along its valley.

found in S_G^2 lie in its flat diagonal valley, where the global minimum, $\boldsymbol{x}^*$, also lies. S_L^2 also contains a curved valley which is littered with local minima. It can thus be understood that, because of its gradual descent, a variety of step sizes along S_G^2's valley might survive the selection step, while in S_L^2's valley, many would be rejected because they more significantly change the $\bar{F}$ value (by jumping in and out of minima). This seems to be the reason that SRES performs better in S_L^2.

The comparison of searches in these low-dimensional subspaces gives an indication of the performance of SRES. It is able to escape higher valued regions of $\bar{F}$, even when they occupy the majority of the search space (see Figures 6.9(c) and (d), where all solutions found were in the low-valued regions). Also, flatness of the objective function surface seems to be a harder problem for SRES than multi-modality. This also illus-

trates the importance of including more than a handful of parameters from the GSA Top ranks, when reducing the dimensions of the search space; as the GSA results pinpoint parameters that are most sensitive *on average* across space, the rankings are not necessarily representative of each individual point in space, including the global minimum, $\boldsymbol{x}^*$. Hence, the most locally sensitive parameters may not be represented if only a few are chosen for estimation, causing flatness in $\bar{F}$ around $\boldsymbol{x}^*$.

6.4.6 Estimating four parameters

In this section, the search is carried out in the hyperplane $S_G^4 \subset \mathbb{R}^4$ for the nominal values of the GSA Top 4, namely, $k_{Ca^{2+}act-}$, k_{PIP2b+}, k_{IP_3Rb-} and p_1 (with all other parameter values fixed); for comparison, S_L^4 is later searched for the nominal values of the LSA Top 4. The larger population size, $\lambda = 50$ is used and the initial selection pressure is given by $P_S = 7$, while the search is carried out for longer ($G = 100$) because more time might be needed to better explore a higher-dimensional space.

The parameter estimates found in this run have percentage errors of -70%, 25%, 75% and 250%, as shown in Figure 6.10(b), and yield a value of $\bar{F} = 0.016938$ (Figure 6.10(a)). As in the previous section, a very close fit to the pseudo-data is obtained with these estimates, even though three of them are not very close to the nominal values. It seems that the surface of $\bar{F}$ in this subspace is similar to a 4-dimensional version of that in Figure 6.9(c), with substantial flat regions, possibly in a long hypervalley. The average population fitness progressively decreases to a value of $\bar{F} \approx 0.4$ (Figure 6.10(a)), indicating that individuals are becoming more alike, since the lowest $\bar{F}$ per generation also converges to around the same value. The selection step is practically redundant in a homogeneous population, and so it does not seem that increasing the number of generations (except drastically) would improve the search.

Without increasing the number of function evaluations, population diversity (a more thorough exploration of space) is encouraged by using a lower selection pressure, $P_S = 5$, which allows 10 parents, instead of 7, to survive for reproduction. Figure 6.11 shows that this scheme does not provide a better solution, with $\bar{F} = 0.026576$ (Figure 6.11(a)), and the errors in the estimates still significant (see Figure 6.11(b)). This is perhaps unsurprising, since a similar scheme in the lower dimensional S_G^2 did not make any improvements.

Searching in S_L^4

The search is now carried out in S_L^4, for the nominal values of the LSA Top 4, namely, p_2, $k_{Ca^{2+}act+}$, k_{GTP+}, k_{IP_3Rb+} (with all other parameters fixed at their true values). The

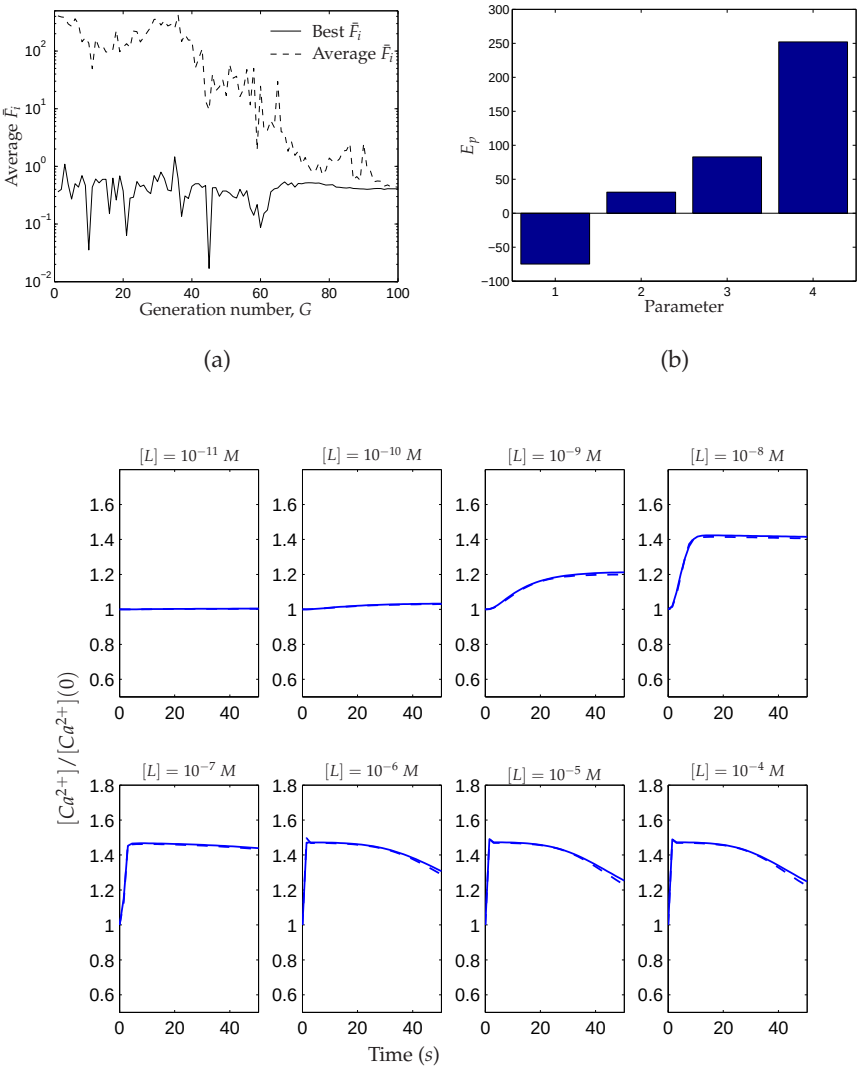
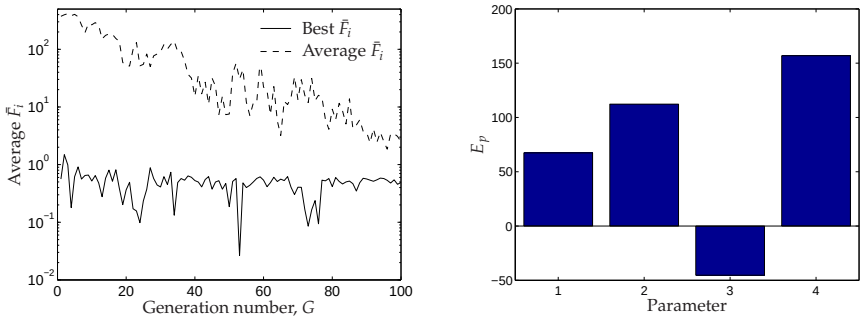


Figure 6.10: Searching in $S_G^4 \subset \mathbb{R}^4$ for the nominal values of the GSA Top 4, using the search parameters $\lambda = 50$, $G = 100$, and $P_S = 7$. (a) The best and the average $\bar{F}_i$ in each generation. (b) The errors associated with the estimates, calculated using equation (6.15). (c) The fits obtained with the best individual. As when the search was carried out in S_G^2, the pseudo-data and solution are almost indistinguishable, even though the estimates are considerably different to the nominal values.



(a) (b)

Figure 6.11: Searching in $S_G^4 \subset \mathbb{R}^4$ for the nominal values of the GSA Top 4, using the search parameters $\lambda = 50$, $G = 100$, but with a lower selection pressure $P_S = 5$ than in Figure 6.10. (a) The best and the average $\bar{F}_i$ in each generation. The lower selection pressure does not improve on the solution in Figure 6.10. (b) The errors associated with the estimates, calculated using equation (6.15).

search parameters are set to $\lambda = 50$, $G = 100$ and $P_S = 7$. The results of the search are shown in Figures 6.12(a)–(c). The best solution gives fits that are virtually indistinguishable from the pseudo-data (Figure 6.12(c)), with an error $\bar{F} = 1.49 \times 10^{-3}$ (see Figure 6.12(a)); this $\bar{F}$ value is only about one order of magnitude greater than the solution in S_L^2, which was extremely close to the global minimum ($\bar{F} = 9.4442 \times 10^{-5}$). However, in this case, the errors associated with parameters 2 and 3 are rather high (-68.1%, 369% respectively, see Figure 6.12(b)), indicating that the solution is some distance from the global minimum. This illustrates that in S_L^4 (as in S_L^2, where the function surface was actually visualised, low error estimates are not required to achieve a good fit. In other words, despite the fact that all other parameters are fixed at their nominal values, there are combinations of the LSA Top 4, apart from that of the nominal values, which give low $\bar{F}$ values. (An independent run was carried out which found a different solution that also had a low $\bar{F}$ value, given by 3.4459×10^{-3}.)

6.4.7 Estimating ten parameters

The search is now carried out in $S_G^{10} \subset \mathbb{R}^{10}$ for the nominal values of the GSA Top 10; recall that this is the main optimisation problem. For comparison, S_L^{10} will then be searched for the nominal values of the LSA Top 10. With search parameters that seemed thorough, the true solutions were not found in the lower dimension hyperplanes, S_G^2, S_G^4 and S_L^4, therefore, the true solutions are not expected to be found in (the higher dimensional) S_G^{10} and S_L^{10}. However, it is still informative to analyse the nature of solutions that will be found. Initially, the search parameters are set as $\lambda = 50$, $G = 100$

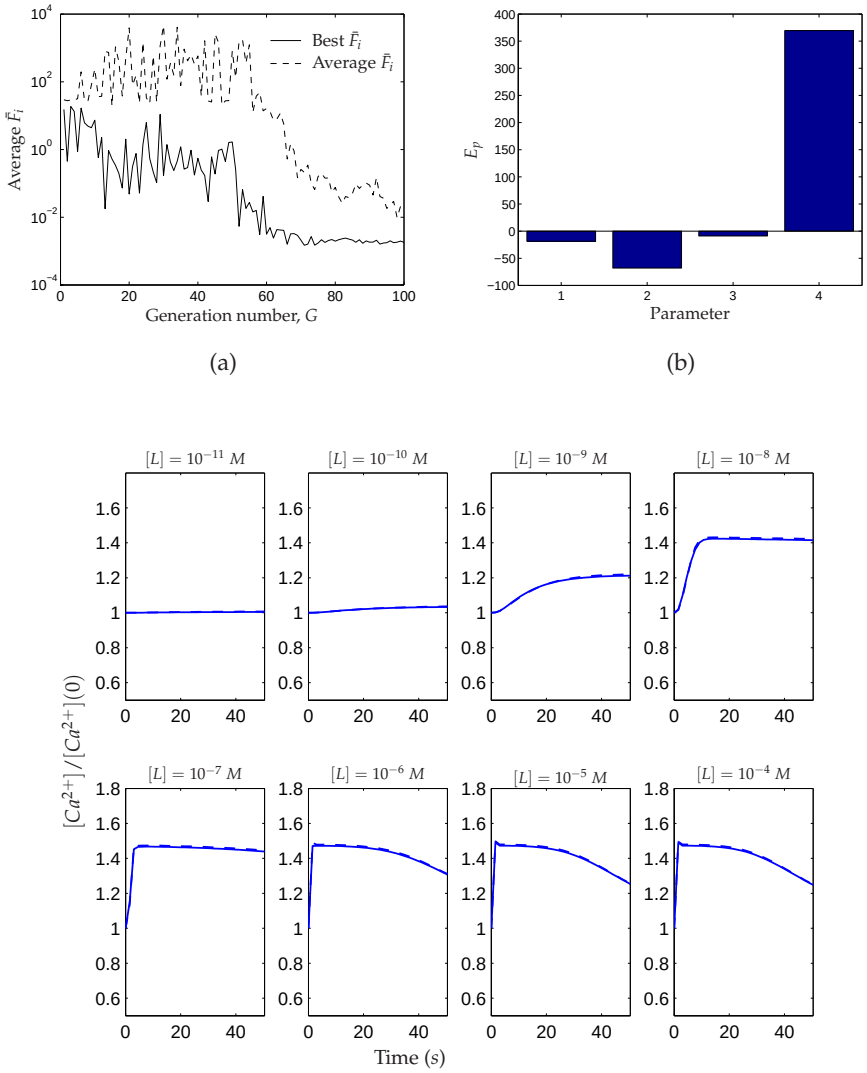


Figure 6.12: Searching in $S_L^4 \subset \mathbb{R}^4$ for the nominal values of the LSA Top 4, using the search parameters specified at the top of the rows. (a) The best and the average $\bar{F}_i$ in each generation. (b) The errors associated with the estimates, calculated using equation (6.15). (c) The fits obtained with the best individual. As when the search was carried out in S_G^2, the pseudo-data and solution are almost indistinguishable, even though the estimates are considerably different from the nominal values.

and $P_S = 7$.

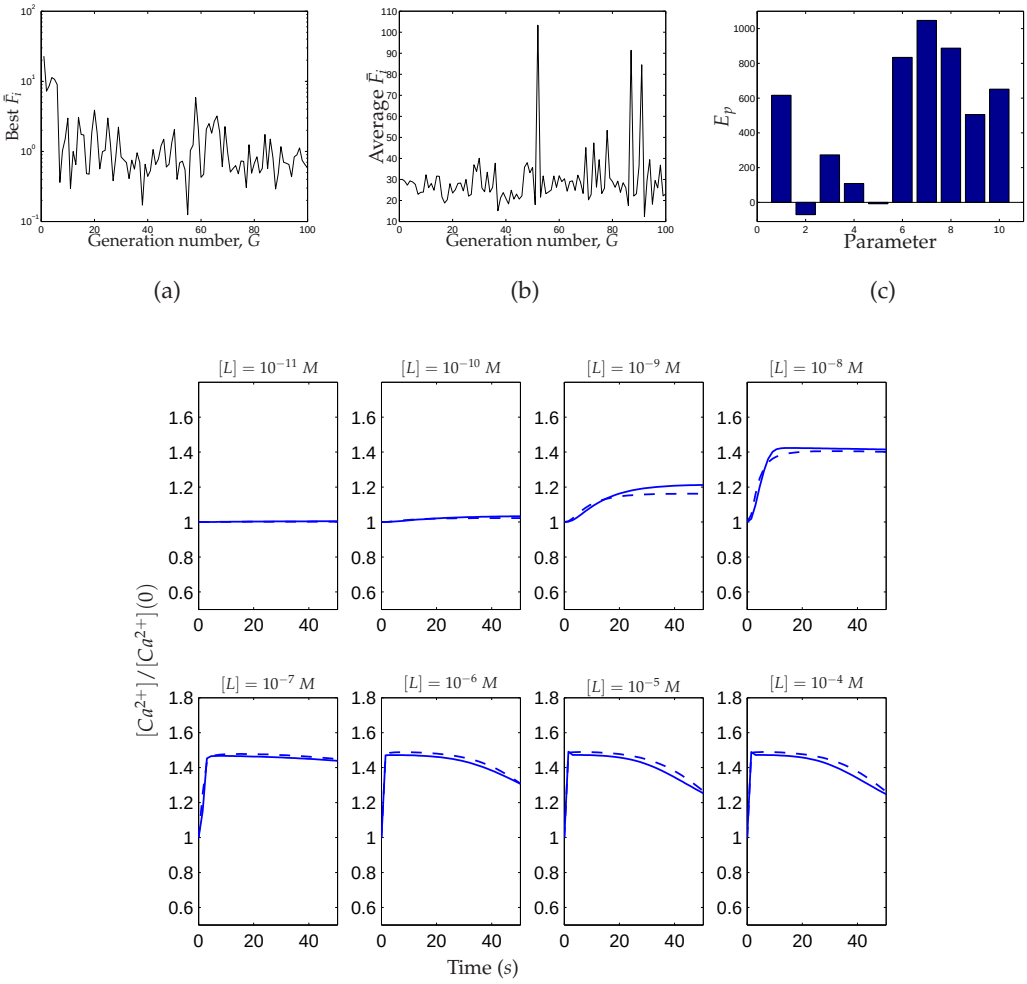
The best solution found gives $\bar{F} = 0.12509$ (see Figure 6.13(a)); searches in the 2-dimensional subspace, S_G^2, found solutions with values as low as $\bar{F} = 0.0068132$, but $\bar{F} = 0.12509$ still yields close fits (Figure 6.13(d)), since it gives an average squared error per data point of ≈ 0.00046. In this run, six of the ten parameters are estimated with $> 500\%$ error (Figure 6.13(c)), so the solution is quite far from the global minimum. For the most part, the population's average $\bar{F}$ value hovers around 30, and does not show convergence. Hence, the selection step has not necessarily become redundant, and the search might be improved by allowing more generations. This might also be necessary because this is a higher dimensional subspace that needs more time to be thoroughly explored.

Figure 6.14 shows the results of carrying out the search with $G = 300$ ($P_S = 7$). Although the solution is better, with $\bar{F} = 0.070371$, this is not due to the increase in G, as the solution was found in the 14th generation, so nothing has been gained by increasing the number of generations from $G = 100$. At $G = 300$, the selection step has not necessarily become redundant, since the average $\bar{F}$ (≈ 22) is still much higher than the lowest $\bar{F}$ (≈ 2) (compare Figures 6.14(a) and 6.14(b)); the search might thus improve with more generations, but would get computationally expensive, since this run took 181.51 hours (7.56 days), while the $G = 100$ search only took 13.72 hours (0.57 days). Like the previous solution, these estimates also have generally high percentage errors (6 estimates have $> 500\%$ error), meaning that the solution is again far from the global minimum.

To promote further exploration of space, a higher population size, $\lambda = 100$, as well as a lower selection pressure, $P_S = 5$, are tried, allowing 20 parents to survive for reproduction in each generation. The results are displayed in Figures 6.15(a)–(c), from which it can be seen that the solution is not improved, because the best individual now gives $\bar{F} = 0.15542$. Also, as when the population size was 50, the errors associated with many of the parameters are large.

Searching in S_L^{10}

The search is now carried out in S_L^{10} for the nominal values of the LSA Top 10, using the search parameters $\lambda = 50$, $G = 100$ and $P_S = 7$. The best solution has $\bar{F} = 0.30719$, as shown in Figure 6.16(a); compared to the initial search in S_G^{10} (see Figure 6.13(a), where $\bar{F} = 0.12509$), the fits are visibly worse (Figure 6.16(d)). The large percentage errors associated with some of the estimates (Figure 6.16(c)), indicate that the solution is far from the global minimum. The fact that the solution found in S_G^{10} is considerably better than that found in S_L^{10} might be due to the fact that there is more global sensitivity



(a) (b) (c)

(d)

Figure 6.13: Searching in $S_G^{10} \subset \mathbb{R}^{10}$ for the nominal values of the GSA Top 10, using the search parameters $\lambda = 50$, $G = 100$. (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b) The average $\bar{F}_i$ of the whole population. (c) The errors associated with the estimates, calculated using equation (6.15). (d) The fits obtained with the best individual. As previously seen, the fits are very close to the pseudo-data, even though the estimates are considerably different from the nominal parameter values.

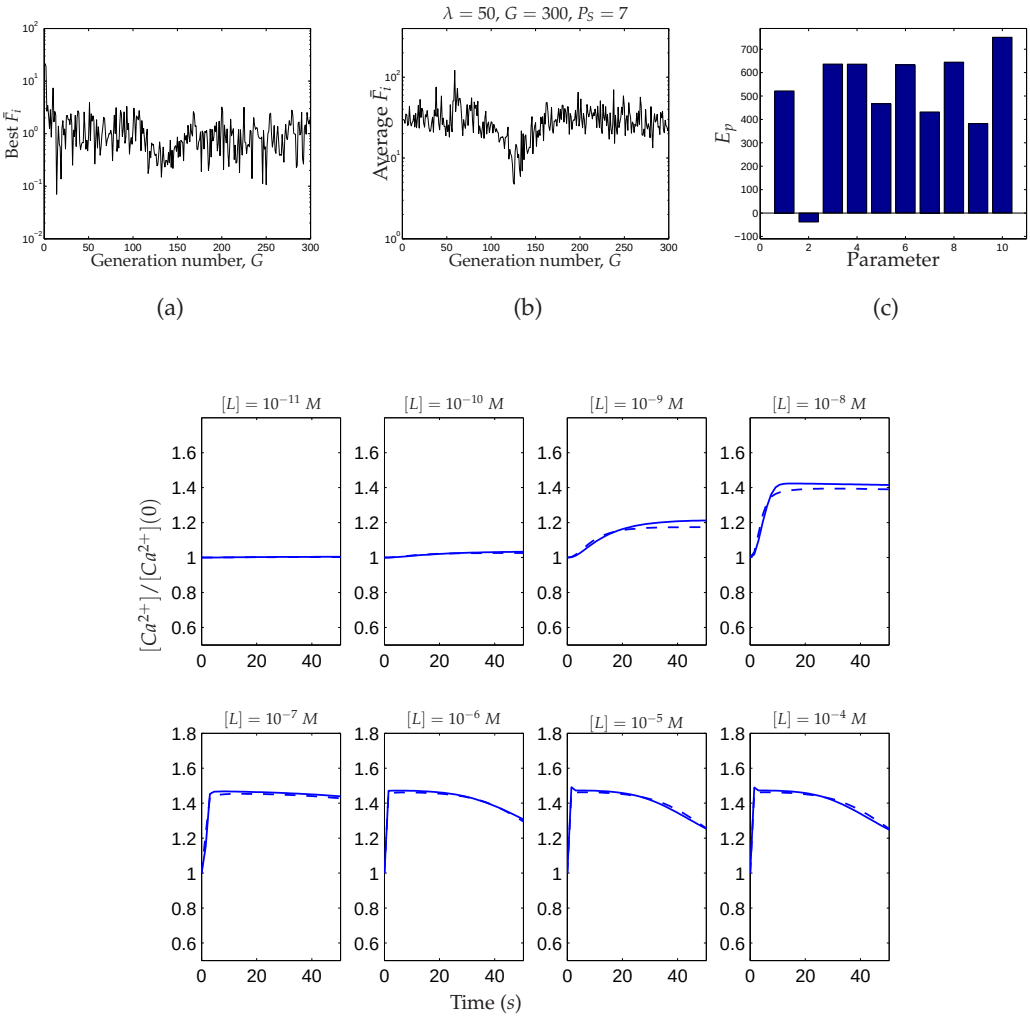
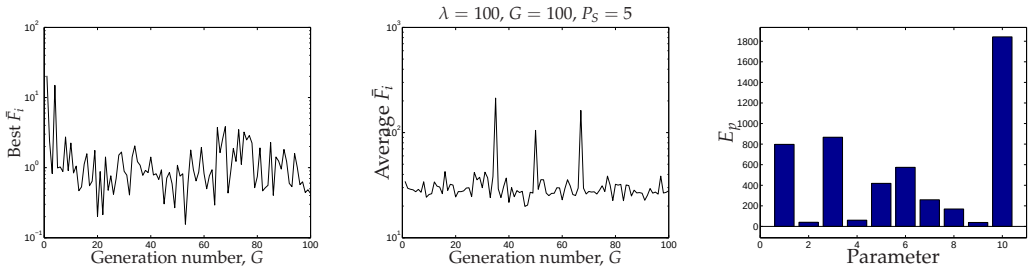


Figure 6.14: Searching in $S_G^{10} \subset \mathbb{R}^{10}$ for the nominal values of the GSA Top 10, using the search parameters specified at the top of the rows. (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b) The average $\bar{F}_i$ of the whole population. (c) The errors associated with the estimates, calculated using equation (6.15).



(a) (b) (c)

Figure 6.15: Searching in $S_G^{10} \subset \mathbb{R}^{10}$ for the nominal values of the LSA Top 10, using the search parameters specified at the top of the rows. (a), (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b), (b) The average $\bar{F}_i$ of the whole population. (c), (c) The errors associated with the estimates, calculated using equation (6.15).

information in S_G^{10} to guide the search (recall that the LSA Top 10 includes only seven parameters from the GSA Top 10 (see Figure 6.5)).

The fact that the lowest and average $\bar{F}$ values do not converge (compare Figures 6.16(a) and (b)) might mean that the selection step has not necessarily become redundant, and that a better solution might be found with more generations; however, it has previously been seen in S_G^{10} that 300 generations were not sufficient to allow this, and that beyond this the search starts to be computationally unaffordable. Hence, a smaller value of $G = 200$ was tried in S_L^{10}, which resulted in a better $\bar{F}$ value of 0.19836, but this was found just outside $G = 100$, at $G = 108$; in addition, a close value of $\bar{F} = 0.2017$ had already been found at $G = 41$ (results not shown). The computational time for this scheme was 54.94 hrs (2.29 days).

Testing the effects of noise

A search is now carried out in S_G^{10}, but using $\tilde{\boldsymbol{Y}}_{k,l}^{n}$ (equation (6.13)) in equation (6.12) to measure the sum of the squared deviations from the noisy version of the pseudo-data. The concern is what quality of fit will be obtained when the data are noisy, not how well the global minimum will be estimated, since the previous runs with the noiseless pseudo-data generally did not result in good estimates (though they produced good fits).

Figure 6.17 shows that the overall results are similar to those obtained with the noise-less pseudo-data. Figure 6.17(d) shows that a good fit is obtained; and as would be expected, based on the previous runs, the errors in the estimates are large (Figure 6.17(c)). The best solution's $\bar{F}$ value, at 0.44665, is higher than those found with the noiseless pseudo-data, which is unsurprising, as the squared deviations generally in-

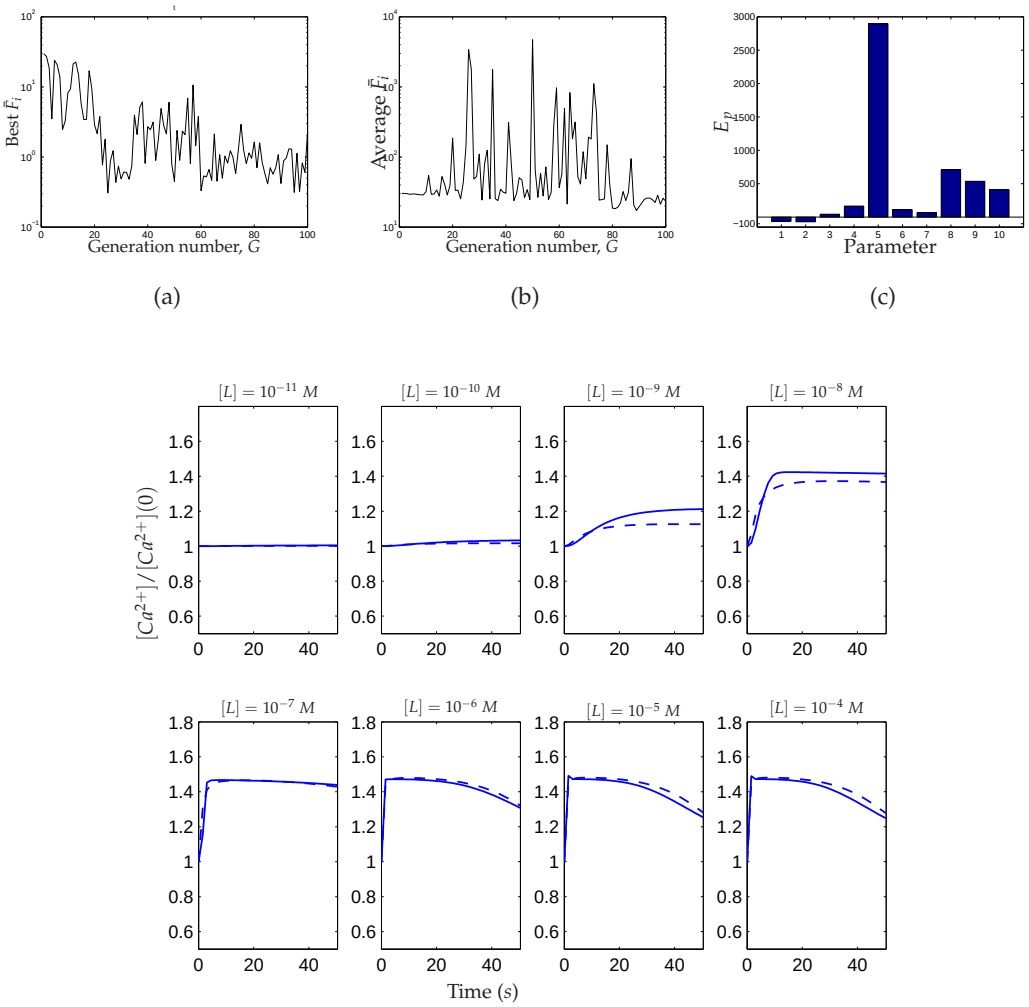


Figure 6.16: Searching in $S_L^{10} \subset \mathbb{R}^{10}$ for the nominal values of the LSA Top 10, using the search parameters $\lambda = 50$, $G = 100$, $P_S = 7$. (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b) The average $\bar{F}_i$ of the whole population. (c) The errors associated with the estimates, calculated using equation (6.15). (d) The fits obtained with the best individual. As previously seen, the fits are very close to the pseudo-data, even though the estimates are considerably different from the nominal parameter values. This occurs despite the fact that these are the locally most sensitive parameters.

crease because of the noise.

The addition of noise therefore, does not noticeably worsen the fits. This means that when fitting to the experimental data, a quality of fit can be expected, which is similar to one that would be obtained using clean data.

6.4.8 Parameter uncertainty

The errors that accompany many of the estimates have generally been large. These are examples of a phenomenon that has been referred to as 'sloppiness of parameter sensitivities'. A study on seventeen systems biology models revealed that in each, there were a few parameter combinations to which the model behaviour was sensitive (called 'stiff directions'), but that there were many more parameter combinations ('sloppy directions') over which huge changes could occur without significantly altering the response. Thus, even with a comprehensive set of data, most parameter estimates had large errors (uncertainties ranged from a factor of 50 up to around 1,000,000) [29]. This section has shown that the Ca^{2+} model is similar, as the qualities of fit have been good, even while the parameter estimates, have been poorly constrained.

A more recent study showed that, by carefully choosing the combination of experiments so that they provide complementary information, parameters can be well estimated [3]. The problem was posed as finding the subset of all possible experiments, which though individually characterised by sloppy directions/sensitivities, collectively combine to form stiff directions. Hence, comprehensive data are not necessarily key to correct estimation, and a data set that might be much smaller, but provide *complementary* information may be more important. The authors were able to use data from only five experiments (knockdown, overexpression and single and two drug stimulation) to estimate 48 parameters of a signalling model to 10% accuracy. Hence, a similar process of computational (and actual) experimental design may be needed to accurately estimate the parameters of the Ca^{2+}-mobilisation model.

Another approach might be to use, together with a comprehensive pseudo-data set, very large SRES search parameters, similar to $\lambda = 350$ and $G = 8000$, as successfully used in [60]. Even though parameter sloppiness causes flatness on the objective function surface, such large parameters might allow more time for the algorithm to adapt to flat regions on the objective function surface, and to approach the global minimum. Only pseudo-data for Ca^{2+} were generated here however, to reflect the available experimental Ca^{2+} data.

In summary, the trial runs for SRES indicate that the global minimum is not easily found with a set of data for one model variable ($[Ca^{2+}]$), but that comparably good parameter sets (in terms of goodness of fit) will generally be found.

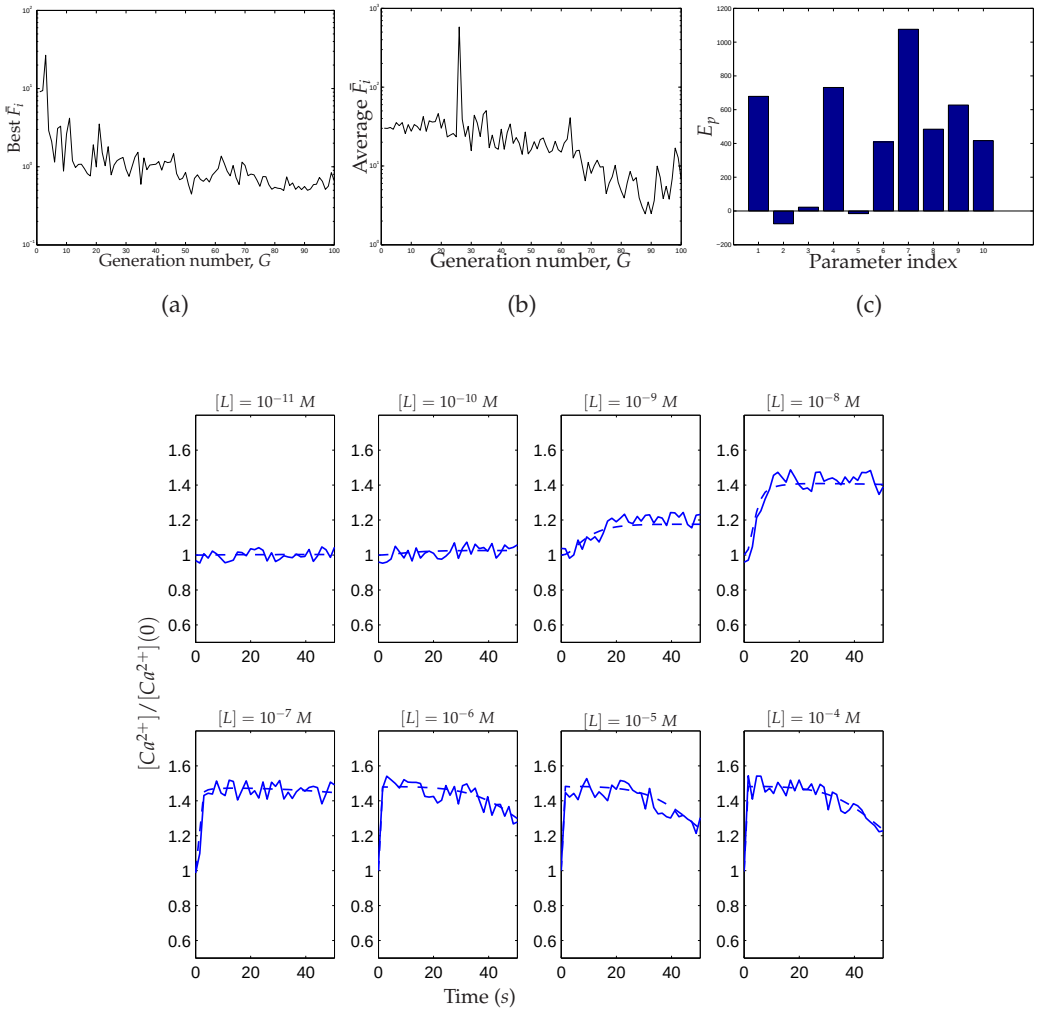


Figure 6.17: Searching in $S_G^{10} \subset \mathbb{R}^{10}$ for the nominal values of the GSA Top 10, using the search parameters $\lambda = 50$, $G = 100$, $P_S = 7$. This case is different from those in Figures 6.13 and 6.14, because the noisy pseudo-data are used here. (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b) The average $\bar{F}_i$ of the whole population. (c) The errors associated with the estimates, calculated using equation (6.15). (d) The fits obtained with the best individual. The fits are close to the pseudo-data, even though the estimates are considerably different from the nominal parameter values.

6.5 Parameter estimation for experimental data

The parameters are now estimated by fitting to equation (6.2), the objective function (F) which pertains to the experimental Ca^{2+} data. Recall that the data has been normalised (see Figure 6.1), so that F is dimensionless. Note that the global minimum for this problem is not known, hence, there will be no way of knowing how good the estimates are; however, the results of the previous section strongly suggest that the global minimum will not be found.

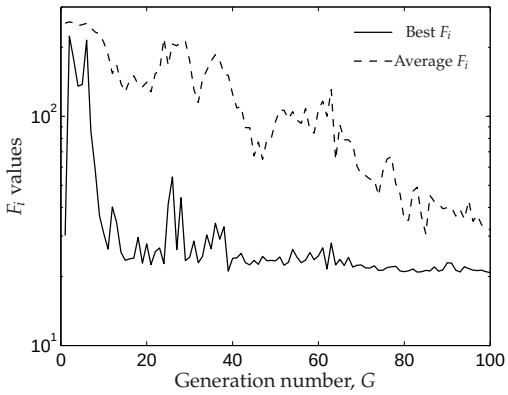
6.5.1 Searching in S_G^{10}

In this section, the search is carried out in S_G^{10}, seeking optimal values for the GSA Top 10, while the other parameters are fixed at their nominal values. Based on the previous section's results, the search parameters $G = 100$, $\lambda = 50$ and $P_S = 7$ are used.

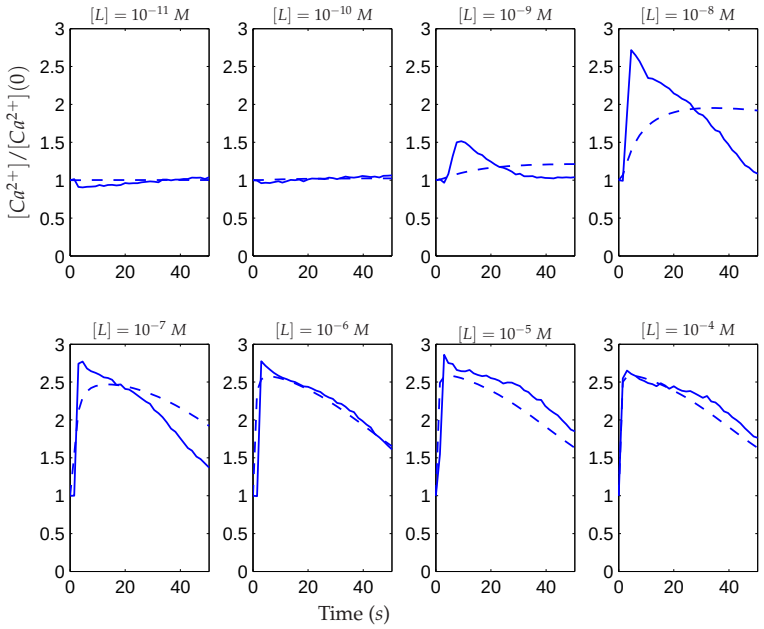
Figure 6.18 shows the results of the search. The best solution gives $F = 20.794$ (Figure 6.18(a)), and the resulting fits are shown in Figure 6.18(b). The fits obtained at the four highest agonist concentrations, $[L] = 10^{-7} - 10^{-4}M$, are good, while the fits obtained for $[L] = 10^{-11}M$ and $[L] = 10^{-10}M$ are good but trivial, because, to these concentrations, there is no essentially no response. At $[L] = 10^{-9}M$ and $[L] = 10^{-8}M$ the fits are poor; the times-to-peak are considerably slower in the simulations, and the peaks considerably lower. Towards the end of the search, the average F seems to be converging towards around the same value as the lowest F (Figure 6.18(a)). This indicates that in the last stages, the population is becoming increasingly homogeneous, so that the selection step is getting redundant. Thus, increasing the number of generations might not yield a better solution; the population size is increased instead, to promote diversity.

A population size, $\lambda = 100$, is tried, together with a selection pressure, $P_S = 7$. Figure 6.19(a) shows that the best solution found in this run has $F = 21.381$, which is slightly higher than was found in the previous run, hence no improvement has been made by increasing the population size from $\lambda = 50$; the fits from both runs are of a similar quality: good for the higher agonist concentrations, but not when $[L] = 10^{-9}M$ and $[L] = 10^{-8}M$ (compare Figures 6.19(b) and Figure 6.18(b)).

If the model is a very close reflection of the actual Ca^{2+} system, good fits to the experimental data should have been obtained for all agonist concentrations, as in the trial runs with the pseudo-data. SRES was good at converging to low-valued regions in $\bar{F}$, so there are probably no values much lower than the $F = 20.794$ and $F = 21.381$ values found here. It is therefore possible that the Ca^{2+} model, though detailed, requires some extension. There is also the possibility that more appropriate parameter ranges should

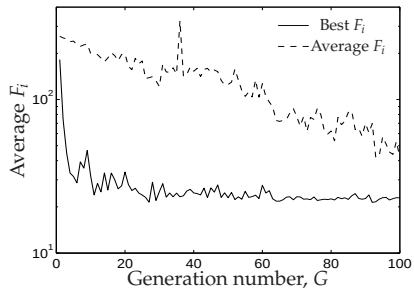


(a)

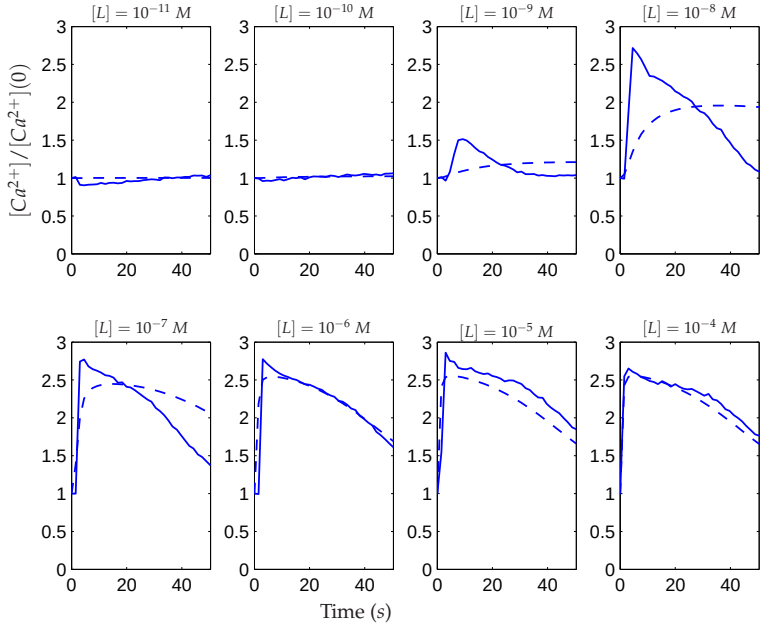


(b)

Figure 6.18: Searching in $S_G^{10} \subset \mathbb{R}^{10}$ for values of the GSA Top 10 that give good fits to the experimental data, using the search parameters $\lambda = 50$, $G = 100$, $P_S = 7$. This case is different from those in Figures 6.13, 6.14 and 6.17, because actual experimental data are used here. (a) The best and the average F_i for each generation. (b) The fits obtained with the best individual.



(a)



(b)

Figure 6.19: Searching in $S_G^{10} \subset \mathbb{R}^{10}$ optimal values of the GSA Top 10 that give good fits to the experimental data, using the search parameters $G = 100$, $P_S = 7$, and an increased population size $\lambda = 100$ (compare with Figure 6.18, where $\lambda = 50$). (a) The best and the average F_i for each generation. (b) The fits obtained with the best individual, which are similar to those found with $\lambda = 50$ (compare Figure 6.18(b)), good at the higher agonist concentrations, but not at $[L] = 10^{-9}M$ and $[L] = 10^{-8}M$.

be used in the search (which could contain significantly lower values of F). However, there are no clear indicators of what those would be. Before making such conclusions however, section 6.5.2 will check that the reduction of parameter space does not have an adverse effect on the fit.

Local sensitivity of F

A local sensitivity analysis is now carried out at the two solution points found in S_G^{10} (henceforth referred to as $\boldsymbol{x}_{\lambda_{50}}^{\mathbf{10}}$ and $\boldsymbol{x}_{\lambda_{100}}^{\mathbf{10}}$, where the superscript a indicates the dimension of the subspace in which the solution was found, and the subscript λ_b indicates that a population size of b was used). Since some of the fits could do with improvement, the LSA can identify sensitive points in the pathway which might pertain to processes whose modelling should be revisited, should the model require extension.

Sensitivity coefficients are calculated using the following:

$$C_j = \frac{(F(\boldsymbol{x}_{\lambda_a}^{\boldsymbol{b}} + \mathbf{0}_{\delta_j}) - F(\boldsymbol{x}_{\lambda_a}^{\boldsymbol{b}}))/F(\boldsymbol{x})}{\delta_j / x_{\lambda_a,j}^b}, \tag{6.16}$$

where $\mathbf{0}_{\delta_j}$ is an n-dimensional vector of zeros, but with its jth element equal to δ_j, and $x_{\lambda_a,j}^b$ is the jth estimated parameter. δ_j will be calculated as $\pm 0.1 x_{\lambda_a,j}^b$ and $\pm 0.5 x_{\lambda_a,j}^b$. The larger the value of $|C_j|$, the greater the sensitivity of F to that parameter; in particular, $|C_j| > 1$ indicates a bigger relative change in F than in the parameter.

Figures 6.20(a) and 6.20(b) show the results of the sensitivity analysis. Each 4-bar group represents the sensitivity coefficients for parameter perturbations of -50%, -10%, $+10\%$, $+50\%$ consecutively. Most of the parameters have at least one relatively significant sensitivity coefficient (defined as $|C_j| \geq 1$), except 5:k_{IP_3Rb+}, 8:k_{deact}, and 10:R_{TOT}, when perturbed from $x_{\lambda_{100}}^{10}$ (see Figure 6.20(b)). Among the significant parameters, the ones that do not just pertain to simple binding or unbinding, but to processes that could be more extensively modelled are: 2:k_{PIP_2b+}, 7:$[PIP_2]$, 6:k_{des+} and 9:p_2. The simplifying assumptions underlying the modelling of these processes, and ways in which it can be revisited, will be discussed later in section 6.6. Before that, it is important to check that there are no better solutions in the full search space that would be missed by searching in the reduced space.

6.5.2 Searching in the full search space

The search is now carried out in the full search space, $S \subset \mathbb{R}^{40}$, for optimal values of all 40 parameters originally included in Chapter 4's global sensitivity analysis (see Table C.4.1); as previously stated, this is done to understand how well the search space S_G^{10}

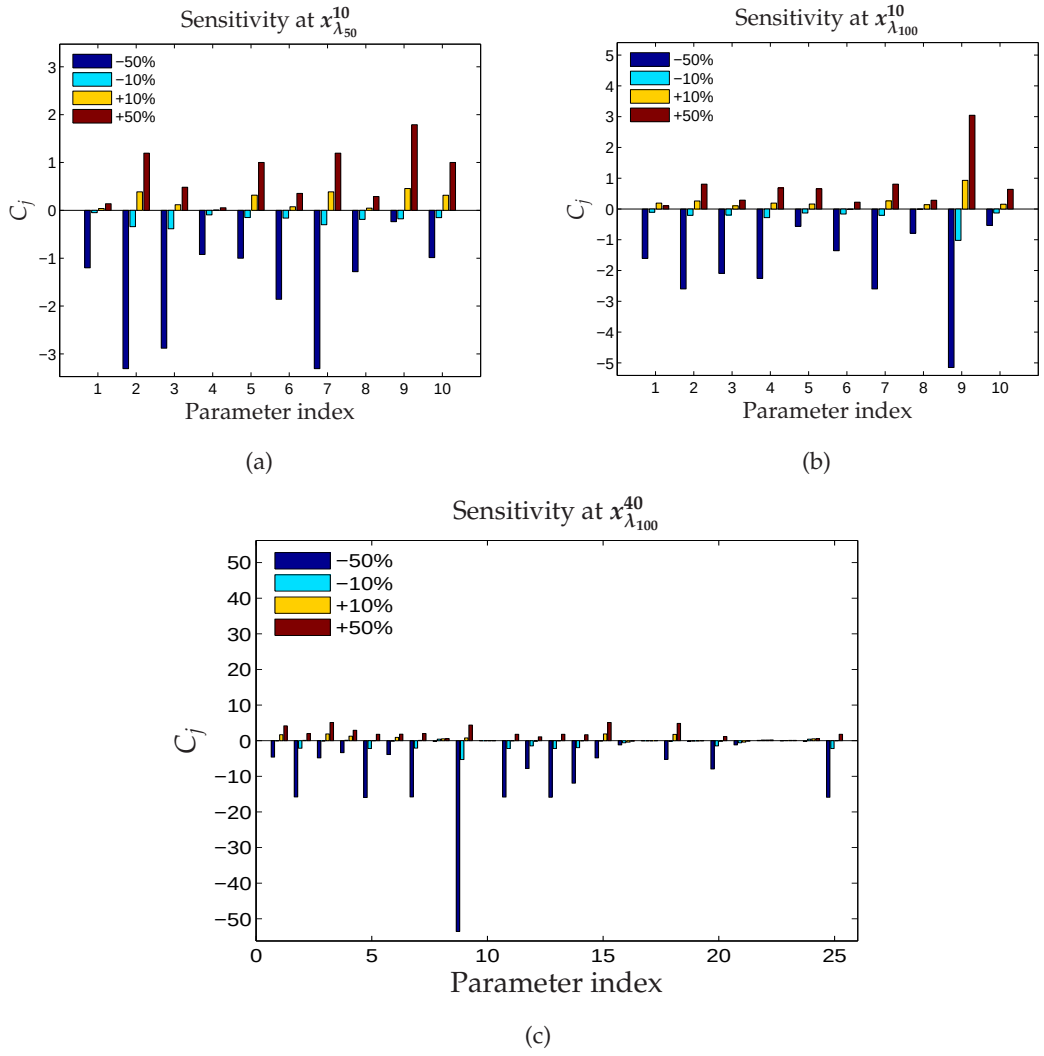


Figure 6.20: Sensitivity coefficients calculated (using equation 6.16) at the best three parameters sets found by SRES when fitting to the experimental data. (a) $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda_{50}}}$, (b) $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda_{100}}}$, (c) $\boldsymbol{x}^{\mathbf{40}}_{\boldsymbol{\lambda_{100}}}$. Several of the same parameters are sensitive at all three points, including 2:k_{PIP_2b+}, 7:[PIP_2], and 9:p_2.

substitutes the full search space. The results are then compared with those obtained in the previous section.

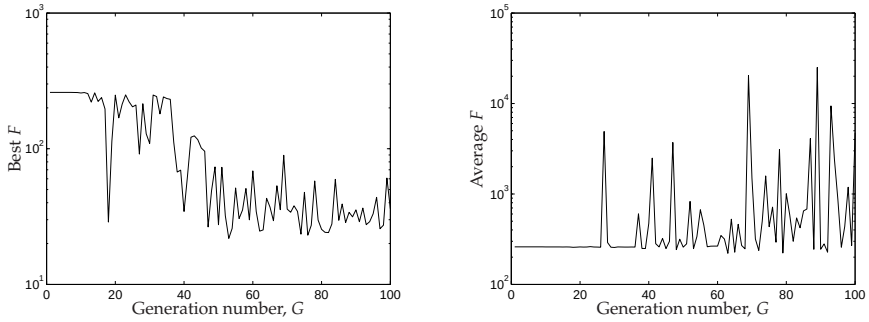
Figure 6.21 shows the results of the search. The best solution (referred to as $\boldsymbol{x}^{\mathbf{40}}_{\boldsymbol{\lambda}_{\mathbf{100}}}$) gives an F value of 21.824 (Figure 6.21(a)), slightly higher than those found by searching in S^{10}_G. Hence, the fit is still not good at the lower ligand concentrations. Unlike the previous search, however, the average population is not converging towards the lowest F value (compare Figures 6.21(a) and (b)), implying that the selection step is not necessarily redundant, and that improvements could possibly be made by increasing the number of generations. However, the computational cost of searching in S can be prohibitive; the search took 231.9 hours, an average of 10.4355s per model evaluation, compared to 18.09 hours in S^{10}_G, which is an average of only 1.6282s per model evaluation. It should also be noted that in the trial runs, increasing the number of generations never improved the results (see sections 6.4.5 and 6.4.7).

The three solutions sets, $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda}_{\mathbf{50}}}$, $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda}_{\mathbf{100}}}$, and $\boldsymbol{x}^{\mathbf{40}}_{\boldsymbol{\lambda}_{\mathbf{100}}}$ produce similar fits, but it is not necessarily because the parameters are all similar to each other. Figure 6.22 plots each $x^a_{\lambda_b,j}$, normalised to the maximum from $\{x^{10}_{\lambda_{50},j}, x^{10}_{\lambda_{100},j}, x^{40}_{\lambda_{100},j}\}$; therefore at least one $x^a_{\lambda_b,j}$ will have a value of 1. If the other two values are close to 1, then the estimates are similar; if either or both of the two other values are $<<$ 1, then there is notable dissimilarity across the estimates. Parameters 2, 4, 7 and 10 each take values as low as $\approx$ 0.1, with parameter 10 getting as low as $\approx$ 0.001, so their values in each solution can take values that differ up to 3 orders of magnitude. (The absolute values are given in Table C.4.1.) 'Parameter sloppiness' is thus exhibited across the solution sets, as would have been expected, based on the trial runs, which showed that parameter sets that were rather different from the true solution could give very similar predictions.

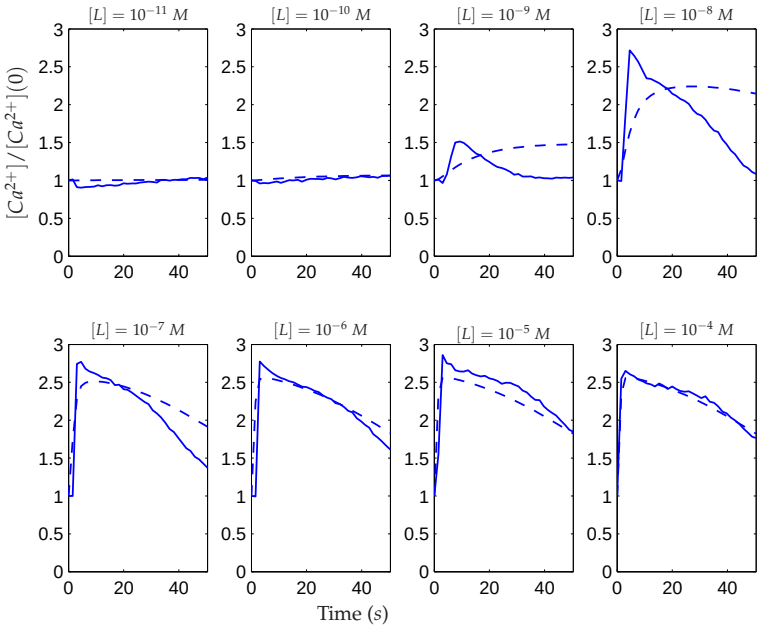
It is clear that it is reasonable to search in S^{10}_G, in place of S, since the quality of the fits found in both spaces are very similar; in other words, it is very reasonable to focus on the GSA Top 10, and this constitutes a significant simplification of the analysis. It should be noted however, that the success of a GSA-informed reduction of parameter space likely depends on how well the output features defined for the GSA represent the overall output. It seems that this chapter's parameter space reduction was successful because the Ca^{2+} Peak-over-Basal, Ca^{2+} Time-to-Peak, Ca^{2+} Transient-Peak and Left-Shift, as defined for the GSA in Chapter 4, constitute a suitable representation of the overall Ca^{2+} response (as was intended when they were selected).

Local sensitivity of F

Using equation (6.16), sensitivity coefficients are also calculated for the parameter set obtained in the full search space and the results are shown in Figure 6.20(c). Chapter



(a) (b)



(c)

Figure 6.21: Searching in the full search space, $S \subset \mathbb{R}^{40}$, for values that give good fits to the experimental data, using the search parameters $\lambda = 100$, $G = 100$, $P_S = 7$. (a) The value of $\bar{F}_i$ for the best individual found in a generation. (b) The average $\bar{F}_i$ value of the whole population. (c) The fits obtained with the best individual, are very similar to those found searching in S_G^{10} (compare Figures 6.18(b) and 6.19(b)). The fits are good at the higher agonist concentrations, but not at $[L] = 10^{-9}M$ and $[L] = 10^{-8}M$.

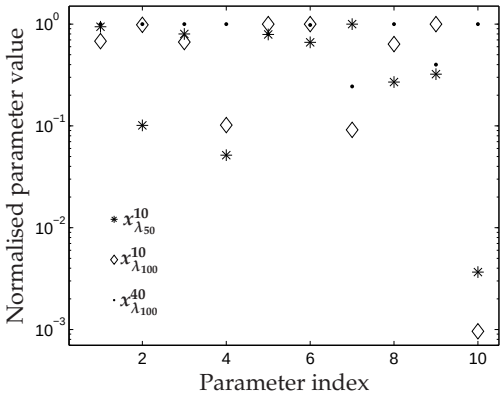


Figure 6.22: Each parameter from each solution set, $x^a_{\lambda_b,j}$, normalised to the maximum from $\{x^{10}_{\lambda_{50},j}, x^{10}_{\lambda_{100},j}, x^{40}_{\lambda_{100},j}\}$ is shown. Some parameters range between one and < 0.1, illustrating that they can take on a wide range of values.

4 showed that there are many parameters to which the Ca^{2+} response is insensitive, so such parameters will have very small coefficients; the bottom 15 parameters in Table C.4.1 (where the GSA rankings are shown) thus have insignificant coefficients, and are not shown in Figure 6.20(c). The parameters are arranged according to their sensitivity ranking (see Table C.4.1), so the first 10 parameters correspond to the GSA Top 10.

In this case, a relatively significant coefficient is defined as $|C_j| \geq 10$. Other parameters outside the GSA Top 10 show relatively significant sensitivity, including 11:k_{des-}, 13:k_{GTP+}, 14:$k_{Ca^{2+}act+}$, and 25:G_{TOT}. Of the parameters that have significant sensitivity coefficients, the ones that pertain to processes that can be more extensively modelled are 2:k_{PIP_2b+}, 7:$[PIP_2]$, 9:p_2, and 11:k_{des-}. These are consistent with those identified by the LSA carried out at $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda_{50}}}$ and $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda_{100}}}$ (see section 6.5.1). Potential ways to revisit the modelling of those processes are discussed in the next section.

6.6 Potential model extensions

It was previously mentioned that some parameters identified as locally sensitive at the three solution points ($\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda_{50}}}$, $\boldsymbol{x}^{\mathbf{10}}_{\boldsymbol{\lambda_{100}}}$, and $\boldsymbol{x}^{\mathbf{40}}_{\boldsymbol{\lambda_{100}}}$) pertain to processes that can be modelled more extensively. Those parameters are 2:k_{PIP_2b+}, 6:k_{des+}, 7:$[PIP_2]$, 9:p_2, and 11:k_{des-}.

7:$[PIP_2]$ and its rate constant of binding, 2:k_{PIP_2b+}, serve essentially the same function in reaction (3.6), since $[PIP_2]$ is assumed constant. The modelling of $[PIP_2]$ dynamics has already been considered in Chapter 4; there, a simplified cycle of PIP_2 depletion and replenishment was incorporated, but did not significantly improve the predicted Ca^{2+} response. It therefore does not seem that the cycle needs to be further described,

especially since significant PIP_2 depletion would typically occur only for strong agonist stimulation, and the model already reflects those cases well (see Figures 6.18, 6.19, 6.21).

9:p_2, which shows very high sensitivity in $\boldsymbol{x}^{\mathbf{40}}_{\boldsymbol{\lambda}_{\mathbf{100}}}$ (see Figure 6.20(c)), pertains to the Hill function that models the SERCA pump dynamics in the Ca^{2+} ODE, reproduced below:

$$\frac{d[Ca^{2+}]}{dt} = (1+\nu_r)(\gamma_0+\gamma_1[IP_3IP_3RCa^{2+}])(Ca^{2+}_{AV}-[Ca^{2+}]) - \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2+p_2^2}. \quad (6.17)$$

Equation 6.17 is a simple representation of processes that might be modelled in more detail, since the SERCA pump actually undergoes a multi-step reaction cycle (involving association and dissociation from Ca^{2+}, and phosphorylation) [22]; indeed, a realistic representation of the reaction scheme can include up to twelve states [28]. Other models have also described the pump with terms that account for modulation of the pumping rate by the concentration of Ca^{2+} in the ER [95]. As the pump plays a central role in the removal of cytosolic Ca^{2+}, it can influence the steepness of the Ca^{2+} descent from peak. Hence, modelling its dynamics in more detail might enable better reflection of the data at the lower agonist concentrations, where the descent needs to be steeper to fit the data (see Figures 6.18(b), 6.19(b) and 6.20(c)).

6:k_{des+} and 11:k_{des-} pertain to receptor desensitisation, which is an intricate set of PKC-mediated processes which, in the Ca^{2+} mobilisation model, has been minimally represented, for the sake of simplicity, as follows:

$$R^*G \underset{k_{des-}}{\overset{k_{des+}}{\rightleftharpoons}} R^*_{des}G \quad (6.18)$$

$$LR^*G \underset{\rho_- k_{des-}}{\overset{\rho_+ k_{des+}}{\rightleftharpoons}} LR^*_{des}G. \quad (6.19)$$

There is definitely more scope for re-modelling here, since the desensitisation mechanisms are not definitely understood. M3 receptors can also be desensitised as a result of being phosphorylated by G-protein receptor kinases (GRKs), after which the binding of Arrestin to the phosphorylated receptors inhibits further G-protein-receptor interaction. Alternatively, some GRKs can directly inhibit signalling without phosphorylating the receptor, possibly via binding to α_{GTP} (or $\beta\gamma$) to inhibit phospholipase C activity, and consequently, Ca^{2+} mobilisation [105]. These are all mechanisms that can be incorporated into the Ca^{2+} model and investigated. As desensitisation affects the lifetime of the active GPCR's, and consequently, the stimulus for Ca^{2+} release, these mecha-

nisms are promising extensions to the Ca^{2+} model that could improve the time course predictions at the lower agonist concentrations.

6.7 Discussion

This chapter first tested the performance of the evolutionary algorithm, SRES, in finding known global minima of a minimisation problem for the Ca^{2+} mobilisation model. It was seen that, with computationally moderate optimisation parameters (λ, G), high quality fits were obtained, but not good parameter estimates. This corresponds with findings that systems biology models can have many 'sloppy parameter sensitivities' such that several parameters values can vary widely without significantly altering the response [29]. The presence of many high quality solutions can thus make the SRES algorithm indiscriminate in choosing which individuals survive the selection step. Even in a two-dimensional search space, the global minimum could not be found when the combination of parameters formed a sloppy direction (GSA Top 2); when the search was carried out along a less sloppy direction (LSA Top 2), a solution very close to the global minimum was found. Promoting diversity in the population by increasing the population size (λ) or decreasing the selection pressure (P_S) did not generally cause significant improvements.

It is possible that even higher SRES parameter values, such as $\lambda = 350$ and $G = 8000$ (used in a previous study [60]), can furnish the algorithm with more time and information to adapt to the topology of the objective function surface, and allow convergence to the vicinity of the global minimum (with respect to the pseudodata). The computational expense of such a scheme would be considerable, and might need to be carried out using programming languages, such as Fortran and C, which can be at least an order of magnitude faster than MATLAB [60], which was used for this chapter. Also, parallel computing, which exploits the availability of multiple processors in a computing unit [7], would allow individual members of the population to be assessed concurrently (instead of sequentially) per generation, and thus significantly reduce the computation time, allowing more elaborate search parameters (high population sizes and number of generations) to be used.

Sensitivity analysis and parameter estimation

The results of a global parameter sensitivity analysis (GSA) were used to reduce the dimensions of the search space from 40 to 10. This significantly simplified the optimisation problem by reducing the computational expense, as a smaller population size can be used in the smaller subspace (S_G^{10}); in addition, the average time per model

evaluation was much lower, presumably because the probability of extreme parameter combinations is lower with only 10 parameters. When the model was fit to the experimental data, a very similar quality of fit was obtained by searching in S_G^{10}, as in the full space. The power of a global sensitivity analysis to simplify the analysis of complex models such as the Ca^{2+} mobilisation model has thus been demonstrated. However, it should be stressed, again, that this likely depends on how well the output features defined for the GSA represent the overall output.

Possible model extensions

The trial runs in which SRES was used to fit the pseudo-data showed that SRES easily finds, not the global minimum, but solution parameter sets which cause similar Ca^{2+} time course predictions. It thus seems reasonable that predictions of the global minimum of F would not be very different from those of the three solutions that produced the best fit to the experimental data. For all three solutions, good quality fits to the experimental Ca^{2+} data were only obtained for the higher agonist concentrations. If those data were the only ones available, it would have been easy to conclude that the model was sufficient, however the inability to fit the data at the lower agonist concentrations suggests otherwise. (This highlights the importance of having significant amounts of experimental data.) Thus, as comprehensive as the Ca^{2+} mobilisation model is, and successful at predicting key features like the rapid time-to-peak, the transient peak and signal amplification, it still lacks some mechanistic detail.

The local sensitivity analysis results in Figure 6.20 point to possible aspects of the model which may be adapted. The dynamics of $[PIP_2]$, SERCA and GPCR desensitisation were all highlighted as processes that can be modelled in more detail. Since $[PIP_2]$ dynamics were already minimally modelled in Chapter 5, without significantly altering the predictions of the model, the most promising model extensions should be those applied to SERCA and GPCR desensitisation mechanisms.

CHAPTER 7

Conclusion

The main aim of this thesis was to assemble a detailed kinetic model of GPCR mediated Ca^{2+} mobilisation in the M3 muscarinic receptor system, and, guided by available data, to gain further insight into the pathway dynamics. Such insight could be used to improve existing therapies for disease conditions associated with abnormal Ca^{2+} signalling, and new therapies could be developed based on the improved understanding of the pathway.

As a first step in the construction of the comprehensive model, a base G-protein activation model (the cTCAM) [86] was first analysed in Chapter 2. The total G-protein (G_{TOT}) and receptor concentrations (R_{TOT}) were varied, based on results from a sensitivity analysis [41] that identified them as key controlling parameters with respect to the G-protein response. It was seen that increasing R_{TOT} and G_{TOT} can enhance agonist stimulation, increasing G-protein activation but altering the qualitative nature of the response; in particular, removing the transient peak. It was also shown that agonist stimulated G-protein activation was more enhanced by G-protein overexpression than by receptor overexpression, due to the receptor's ability to bind sequentially to multiple G-proteins. Subsequent sensitivity analysis results on the full Ca^{2+} model ranked the desensitisation rate constant (which is not included in the cTCAM) higher than R_{TOT} and G_{TOT} (see Tables 4.5.1 and 6.4.1), which highlights a need for an extension to the cTCAM (and related G-protein activation models) that at least takes GPCR desensitisation into account.

A novel use of the cTCAM was in the simulation of specific agonists used in standard pharmacological agonist stimulation experiments on the M3 muscarinic receptor system [52, 54, 93]. This involved using recent experimental results [92] to adapt the affinity and efficacy parameters for each drug. Simulated dose-response curves then revealed partial agonism by the drugs for the G-protein response. Also in Chapter 2, a variant of the cTCAM was used to explore the hypothesis that the G-protein's subunits do not dissociate, and it was shown that the cTCAM (in which the subunits do

dissociate) is equivalent to the variant, if re-association happens quickly; otherwise, agonist efficacy was diminished. This led to the testable hypothesis that agonist efficacy depends on fast subunit reassociation. The cTCAM with fast re-association was then used throughout the thesis.

In Chapter 3, the cTCAM was extended into the Ca^{2+} mobilisation model, using downstream details from an existing model [37]. Data-guided numerical simulation and parameter variation were carried out on the model in an initial attempt to reflect the key features of the data, namely, signal amplification, the rapid time-to-peak, and the transient peak. The publication of new results from $\alpha_{GTP}\gamma S$ (a non-hydrolysable analogue of α_{GTP}) stimulation experiments allowed some cTCAM parameters to be fine-tuned, and these were then used in the full Ca^{2+} model. To better reflect the data, a parameter change that enabled agonists to bind preferably to the GPCR's pre-coupled state (over its uncoupled one) had to be made; this might suggest that agonist efficacy depends on receptor pre-coupling.

In the same chapter, receptor desensitisation had to be incorporated into the Ca^{2+} model to predict the transient peak when signal amplification occurred. This further suggests that G-protein activation models need to take receptor desensitisation into account. However, signal amplification and the rapid Ca^{2+} time-to-peak could not be predicted simultaneously, hinting that the underlying mechanisms were highly fine-tuned. This seemed to be confirmed through a global sensitivity analysis carried out on the model parameters in Chapter 4, which identified a change in k_{IP_3}, the rate constant for PIP_2 hydrolysis, as the only measure which could single-handedly account for the amplified, yet rapid Ca^{2+} response. A value of $k_{IP_3} = 50,000s^{-1}$ produced the desired change, suggesting that PIP_2 hydrolysis could be as quick as some of the fastest known catalytic reactions.

The sensitivity analysis also revealed that agonist potency, as quantified by the extent to which the Ca^{2+} peak was raised over its basal value (see equation 4.20), was mainly sensitive to six parameters, namely, the rate constant of PLC activation, the rate constant for GPCR deactivation, the rate constant of IP_3 degradation, the rate constant of (PKC-mediated) GPCR desensitisation, the maximal rate of the SERCA pumps and the PIP_2 concentration. The last four parameters could exert a biphasic effect. Agonist potency experiments were thus recommended which involve experimentally intruding on those processes, with the hypothesis that such experimental manoeuvres can cause known agonists to act as antagonists. The results might also shed light on why known agonists can sometimes fail in therapy.

Another use of the sensitivity analysis was the identification of relatively insignificant parameters. The total concentration of Phospholipase C (PLC), the enzyme that

hydrolyses PIP_2 (via $\alpha_{GTP}PLC^*PIP2 \xrightarrow{k_{IP_3}} \alpha_{GTP}PLC^* + IP_3 + DAG$) was shown to be one such parameter. Hence —even though the rate of PIP_2 hydrolysis drives the rapid, amplified response— underexpression or overexpression of the mediating enzyme does not significantly influence the response, implying that PLC may not be a worthwhile drug target. This hypothesis could possibly be tested experimentally, since there are commercially available PLC inhibitors [73, 90]. On the other hand, the concentration of PLC's substrate, PIP_2, was one of the most sensitive parameters, suggesting that PIP_2 is a more promising drug target.

After the results of the sensitivity analysis were used to make parameter changes through which the model qualitatively resembled the data, attempts were made to produce a tighter quantitative fit. In particular, the steepness of the Ca^{2+} descent from peak (which represents the rate of Ca^{2+} removal) needed to reflect the data more closely. In Chapter 5, the influence that the membrane Ca^{2+} ATP-ase might have on the rate of Ca^{2+} removal was considered, by extending the Ca^{2+} mobilisation model. In another extension, the assumption of constant $[PIP_2]$ was reconsidered —due its high sensitivity ranking from Chapter 4—by incorporating a simple cycle of $[PIP_2]$ depletion and replenishment. Neither variant model significantly improved the predictions of the Ca^{2+} response. Hence, it was concluded that neither the membrane Ca^{2+} ATPases nor PIP_2 dynamics in fact needed to be included in the model, and the job of improving the quantitative fit was left to a global parameter optimisation, using the original Ca^{2+} mobilisation model, in Chapter 6.

An evolutionary algorithm, the Evolution Strategy using Stochastic ranking (SRES), was used to tackle the parameter optimisation problem, by minimising a least squares function representing the deviation of the model from the Ca^{2+} data. Global sensitivity analysis results were used to reduce the dimensionality of the parameter space by focusing only on the ten most influential parameters and fixing the values of all others. It was shown that solutions found in the reduced subspace produced a similar fit to that found in the full search space, highlighting the role of a global sensitivity analysis in simplifying the analysis of complex, multi-parameter models like the Ca^{2+} mobilisation model.

The best parameter sets found by SRES produced good quality fits to the Ca^{2+} data at the four highest agonist concentrations; however, at $[L] = 10^{-9}M$ and $[L] = 10^{-8}M$, the fits were poor. In initial trial runs fitting 'pseudo-data' generated from model simulation, SRES had been shown to find, not the known global minimum, but solutions which fit the pseudo-data well; thus (apart from the possibility that more appropriate ranges need to be defined) it seemed most likely to conclude that no significantly better fit to the model can be found. A key result of that chapter, therefore, was that additional

mechanistic detail needs to be incorporated into the model. Local sensitivity analyses of the least squares function, at the three points in parameter space corresponding to the best solution parameter sets, identified parameters related to the dynamics of $[PIP_2]$, SERCA pumps and PKC-mediated GPCR desensitisation as influential; hence, revisiting the modelling of these aspects might improve the response. The ubiquity of these parameters in the top ranks of all sensitivity analyses carried out in this thesis (see sections 4.5.2, 6.4.2 and 6.6) pinpoints the processes they mediate as central in Ca^{2+} signalling; hence, in addition to $[PIP_2]$ (see above), the SERCA pump and PKC may also be counted as potentially good drug targets.

The importance of the SERCA pump function is supported by experimental findings that reduction (or loss) of its expression are associated with colon and prostrate cancer [61]. Such experimental findings provide extra motivation to model the pump in more detail (see below). The availability of pharmacological inhibitors of the pump [61] could also make predictions from an extended model testable.

7.1 Future work

The role of a global sensitivity analysis (GSA) in reducing the number of parameters for analysis has already been mentioned. A computationally efficient GSA method (the method of elementary effects (EE)) was used because of the large number of parameters; however, this method could only identify the significant parameters, without quantifying the contribution of each parameter to the variance of the output, something which other more computationally expensive (variance-based methods) methods are able to do [81]. As the EE method already provided a ranking for the parameters, and as it was confirmed that it was reasonable to focus on just the ten highest ranked parameters, a variance based method could be applied to quantify the percentage which each of these parameters contribute to the variance output, potentially reducing the subspace of interest further.

Following the GSA and parameter optimisation results, an important next step would be to extend the Ca^{2+}- mobilisation model by including a more mechanistic representation of Ca^{2+}'s interaction with the SERCA pump. Also, variant extensions of the model can be derived by incorporating different GPCR desensitisation hypotheses from experimental evidence [105]. The fact that these extensions would increase model complexity should not be taken lightly. Section 4.8 showed that the method of elementary effects (especially when combined with an optimised sampling strategy) still identified significant model parameters even when the number of trajectories was significantly reduced. Less expensive, but useful, global sensitivity analyses could

therefore be designed for an extension of the model.

The full potential of the Evolution Strategy with Stochastic Ranking (SRES) strategy may not have been realised with the search parameters used in Chapter 6. The largest population size ever used was $\lambda = 100$, while the biggest number of generations was $G = 300$, whereas $\lambda = 350$ and $G = 8000$ have been successfully used for a simpler optimisation problem [59]. Therefore, even larger parameters might be required for a more efficient search, which would considerably increase computational expense. However, the algorithm could be implemented using parallel computing, which can significantly reduce the computational time, scaling the number of runs by the number of available processors. It is possible then, that due to a more thorough exploration of space, a better fit to the Ca^{2+} data at the intermediate agonist concentrations might be obtained.

It has been shown that fitting to data from a few different experiments, which provide complementary information, can allow parameters to be well estimated [3]. The data that was being fit here was from agonist stimulation experiments by a single drug. Experiments in which two drugs are used to stimulate (e.g., agonist and antagonist) in Ca^{2+}-activating pathways are already routinely carried out at the University of Nottingham's Institute of Cell Signalling, so such data can be easily obtained; receptor or G-protein overexpression experiments are also possible. The latter experiments can be simulated merely by varying the total receptor or G-protein concentrations, while the former require extension of the Ca^{2+} model to include interaction with another drug. Before attempting to fit such data, tests could be carried out with similar pseudo-data, to see if the accuracy of parameter estimates from Chapter 6 is improved.

This thesis has demonstrated that, guided by experimental data, computational methods can be used to glean insight from complex signal transduction models. The collaborators from the experimental aspect of this work have confirmed that many of the model predictions are experimentally testable; the results from such experiments might lead to actual improvements of present therapy strategies.

APPENDIX A

Glossary of pharmacological terms

Ligand: The ligand is a molecule that binds to another one to form a complex. The ligand starts the signalling cascade and is also referred to as the first messenger.

Receptor: The receptor acts as a sensor for the cell. It binds selectively to ligands and transduces the signal intracellularly.

GPCR: G-protein coupled receptor.

Affinity: This refers to the willingness of one molecule to bind to another [106].

Efficacy: This refers to the ability of a ligand to cause a response by binding.

Second Messenger: This is the term used to refer to an intracellular molecule activated as a result of binding, which relays the signal into the interior of the cell. It triggeres a chain of events that eventually lead to a change in cell behaviour.

Agonist: An agonist is a ligand which has a higher affinity for active receptors. It will shift the equilibrium in favour of the active receptors. Different agonists can have different efficacies.

Antagonist: This is a ligand which binds indiscriminately to both states of the receptor. It does not affect conformational equilibrium, but competes for binding sites with other ligand types. In essence an antagonist has affinity but no efficacy.

Inverse agonist: This is a ligand which has a higher affinity for inactive receptors.It will shift the equilibrium in favour of the inactive receptors.

Dissociation constant (K_D): This is the ratio of the dissociation rate constant of a ligand–protein reaction to its association rate constant.

EC_{50} : This is the concentration of agonist it takes to produce half the maximal response.

Desensitisation: This refers to a process by which the receptor is made unable to respond to ligand application.

APPENDIX B

Model parameters

B.1 Description

Here, the parameters for the various models in Chapters 2–6 are described.

Parameter	Description	Chapter
k_{act}	Receptor activation rate constant	2,3,4,5,6
k_{deact}	Receptor deactivation rate constant	2,3,4,5,6
k_{lb+}	Drug/ligand binding rate constant	2,3,4,5,6
k_{lb-}	Drug/ligand unbinding rate constant	2,3,4,5,6
k_{g+}	G-protein binding rate constant	2,3,4,5,6
k_{g-}	G-protein unbinding rate constant	2,3,4,5,6
k_{GTP+}	G protein activation rate constant	2,3,4,5,6
k_{gd+}	GTP hydrolysis rate constant	2,3,4,5,6
k_{RA+}	G-protein reassociation rate	2,3,4,5,6
k_{PLCb+}	α_{GTP} and PLC association rate constant	3,4,5,6
k_{PLCb-}	α_{GTP} and PLC dissociation rate constant	3,4,5,6
k_{PLCact}	$\alpha_{GTP}PLC$ and Ca^{2+} association rate constant	3,4,5,6
$k_{PLCdeact}$	$\alpha_{GTP}PLC$ and Ca^{2+} dissociation rate constant	3,4,5,6
k_{PIP2b+}	$\alpha_{GTP}PLC^*$ and PIP_2 association rate constant	3,4,5,6
k_{PIP2b-}	$\alpha_{GTP}PLC^*$ and PIP_2 dissociation rate constant	3,4,5,6
k_{IP_3}	Rate constant for PIP_2 hydrolysis	3,4,5,6
k_{PLCdis}	α_{GDP} and PLC dissociation rate constant	3,4,5,6
k_{IP_3deg}	Rate constant for IP_3 degradation	3,4,5,6
k_{IP_3Rb+}	IP_3 and IP_3R association rate constant	3,4,5,6
k_{IP_3Rb-}	IP_3 and IP_3IP_3R association rate constant	3,4,5,6
$k_{Ca^{2+}act+}$	Ca^{2+} and IP_3IP_3R association rate constant (activating)	3,4,5,6
$k_{Ca^{2+}act-}$	Ca^{2+} and IP_3IP_3R dissociation rate constant	3,4,5,6

Continued on the next page...

...continued from the previous page

Parameter	Description	Chapter
$k_{Ca^{2+}inh+}$	Ca^{2+} and $Ca^{2+}IP_3IP_3R$ association rate constant	3,4,5,6
$k_{Ca^{2+}inh-}$	Ca^{2+} and $Ca^{2+}IP_3IP_3R$ dissociation rate constant	3,4,5,6
k_{des+}	Forward desensitisation rate constant	3,4,5,6
k_{des-}	Reverse desensitisation rate constant	3,4,5,6
$k_{PIP2rep}$	$[PIP_2]$ replenishment rate constant	5
ζ,ν,μ,$\theta_{\nu\mu}$,$\theta_{\zeta\mu}$, $\theta_{\zeta\nu}$	Thermodynamic constants (explained in section 1.3.2)	2,3,4,5,6
v_r	ER to cytosol volume ratio	3,4,5,6
γ_0	Basal permeability of the ER	3,4,5,6
γ_1	Sensitivity of IP_3R to IP_3	3,4,5,6
p_1	Maximal rate of SERCA	3,4,5,6
p_2	$[Ca^{2+}]$ for half-maximal SERCA pumping	3,4,5,6
V_P	Maximal rate of PMCA	5
K_P	$[Ca^{2+}]$ for half-maximal PMCA pumping	5
$[L]$	Ligand (Drug) concentration	2,3,4,5,6
R_{TOT}	Total receptor concentration	2,3,4,5,6
G_{TOT}	Total G-protein concentration	2,3,4,5,6
PLC_{TOT}	Total PLC concentration	3,4,5,6
$[PIP_2]$	Total PIP_2 concentration	3,4,5,6
$[Ca^{2+}_{AV}]$	Volume average Ca^{2+} concentration, $\frac{Ca^{2+}+v_r[Ca^{2+}_{ER}]}{1+v_r}$	3,4,6
$[IP_3R_{TOT}]$	Total concentration of IP_3R	3,4,5,6

Table B.1.1: Descriptions for all parameters.

B.2 Values

Here, the parameter values used in Chapters 2–6 are given.

Parameter	Value
k_{act}	$1s^{-1}$
k_{deact}	$10^3 s^{-1}$
k_{lb+}	$8.4 \times 10^7 M^{-1} s^{-1}$
k_{lb-}	$0.37 s^{-1}$
k_{g+}	$3.6 \times 10^7 M^{-1} s^{-1}$
k_{g-}	$3 \times 10^{-3} s^{-1}$
k_{GTP+}	$1s^{-1}$
k_{gd+}	$0.1 s^{-1}$
k_{RA+}	$1.2 \times 10^{10} M^{-1} s^{-1}$
ζ_+	1
ζ_-	10^{-3}
μ_+	1
μ_-	0.5
ν_+	1
ν_-	1
$\theta_{\nu\mu}$	1
$\theta_{\zeta\mu}$	1
$\theta_{\zeta\nu}$	1

Table B.2.1: Parameter values specific to the G-protein activation models, as used in Chapter 2.

Parameter	Preliminary	Base	GSA-guided
k_{act}	$1s^{-1}$		$10s^{-1}$
k_{deact}	10^3s^{-1}		10^4s^{-1}
k_{g+}	$3.6 \times 10^7 M^{-1}s^{-1}$		$3.6 \times 10^8 M^{-1}s^{-1}$
k_{g-}	$3 \times 10^{-3}s^{-1}$		$3 \times 10^{-2}s^{-1}$
k_{GTP+}	$1s^{-1}$	$0.001s^{-1}$	$0.01s^{-1}$
k_{gd+}	$0.1s^{-1}$		$1s^{-1}$
k_{RA+}	$1.2 \times 10^{10} M^{-1}s^{-1}$		$1.2 \times 10^{11} M^{-1}s^{-1}$
ζ_+	1		
ζ_-	10^{-3}		
μ_+	1		
μ_-	0.5		
ν_+	1	100	
ν_-	1		
$\theta_{\nu\mu}$	1		
$\theta_{\zeta\mu}$	1		
$\theta_{\zeta\nu}$	1		
k_{PLCb+}	$2.25 \times 10^6 M^{-1}s^{-1}$		$2.25 \times 10^7 M^{-1}s^{-1}$
k_{PLCb-}	$1s^{-1}$		$10s^{-1}$
k_{PLCact}	$30 \times 10^6 M^{-1}s^{-1}$		$30 \times 10^7 M^{-1}s^{-1}$
$k_{PLCdeact}$	$1s^{-1}$		$10s^{-1}$
k_{PIP2b+}	$10^9 M^{-1}s^{-1}$	$10^{11} M^{-1}s^{-1}$	$10^{12} M^{-1}s^{-1}$
k_{PIP2b-}	$5 \times 10^3 s^{-1}$		$5 \times 10^4 s^{-1}$
k_{IP_3}	$100s^{-1}$		$50000, 500000s^{-1}$
k_{PLCdis}	$1.667s^{-1}$		$16.67s^{-1}$
k_{IP_3deg}	$2.5s^{-1}$		$25s^{-1}$
k_{IP_3Rb+}	$1.2 \times 10^7 M^{-1}s^{-1}$	$1.2 \times 10^9 M^{-1}s^{-1}$	$1.2 \times 10^{10} M^{-1}s^{-1}$
k_{IP_3Rb-}	$8s^{-1}$		$80s^{-1}$
$k_{Ca^{2+}act+}$	$15 \times 10^6 M^{-1}s^{-1}$		$15 \times 10^7 M^{-1}s^{-1}$
$k_{Ca^{2+}act-}$	$1.65s^{-1}$		$16.5s^{-1}$
$k_{Ca^{2+}inh+}$	$1.8 \times 10^6 M^{-1}s^{-1}$	$1.8 \times 10^5 M^{-1}s^{-1}$	$1.8 \times 10^6 M^{-1}s^{-1}$
$k_{Ca^{2+}inh-}$	$0.21s^{-1}$		$2.1s^{-1}$
k_{des+}	0	$2.75 \times 10^{-2}s^{-1}$	$2.75 \times 10^{-1}s^{-1}$
$k_{PKCdes-}$	0	$2.75 \times 10^{-6}s^{-1}$	$2.75 \times 10^{-5}s^{-1}$
v_r	0.185		
γ_0	$0.2s^{-1}$		
γ_1	$40 \times 10^6 M^{-1}s^{-1}$		$40 \times 10^7 M^{-1}s^{-1}$
p_1	$6 \times 10^{-6} Ms^{-1}$		$6 \times 10^{-5} Ms^{-1}$
p_2	$0.45 \times 10^{-6} M$		
$[G_{TOT}]$	$4.15 \times 10^{-10} M$	$4.15 \times 10^{-6} M$	
$[R_{TOT}]$	$4.15 \times 10^{-10} M$	$10^{-8} M$	
$[PLC_{TOT}]$	$0.8 \times 10^{-6} M$		
$[PIP_2]$	$25 \times 10^{-6} M$		
$[Ca^{2+}_{AV}]$	$0.78 \times 10^{-6} M$		
$[IP_3R_{TOT}]$	$0.8 \times 10^{-6} M$		
K_P	$0.425 \times 10^{-6} M$		
V_P	$28 \times 10^{-6} Ms^{-1}$	$28 \times 10^{-13} Ms^{-1}$	
k_{PIP_2rep}	$0.015, 0.1s^{-1}$		

Table B.2.2: Parameter values for the Ca^{2+} models in Chapters 3, 4, 5 and 6.

B.3 The Ca^{2+} mobilisation model

The Ca^{2+} mobilisation model assembled in Chapter 3 is given here. Note that in the first half of that chapter, GPCR desensitisation was not included ($k_{des+} = k_{des-} = 0$).

$$\begin{aligned}
\frac{d[R]}{dt} &= -k_{act}[R] + k_{deact}[R^*] - k_{lb+}[L][R] + k_{lb-}[LR] \\
&\quad -k_{g+}[R][G] + k_{g-}[RG] && \text{(B.1)} \\
\frac{d[R^*]}{dt} &= k_{act}[R] - k_{deact}[R^*] - \zeta_+ k_{lb+}[L][R^*] + \zeta_- k_{lb-}[LR^*] \\
&\quad -\mu_+ k_{g+}[R^*][G] + \mu_- k_{g-}[R^*G] \\
&\quad +k_{GTP+}[R^*G] && \text{(B.2)} \\
\frac{d[LR]}{dt} &= k_{lb+}[L][R] - k_{lb-}[LR] - \zeta_+ k_{act}[LR] + \zeta_- k_{deact}[LR^*] \\
&\quad -\nu_+ k_{g+}[LR][G] + \nu_- k_{g-}[LRG] && \text{(B.3)} \\
\frac{d[LR^*]}{dt} &= \zeta_+ k_{lb+}[L][R^*] - \zeta_- k_{lb-}[LR^*] + \zeta_+ k_{act}[LR] - \zeta_- k_{deact}[LR^*] \\
&\quad -\theta_{\nu\mu}\mu_+\nu_+ k_{g+}[LR^*][G] + \theta_{\nu\mu}\mu_-\nu_- k_{g-}[LR^*G] \\
&\quad +\nu_- k_{GTP+}[LR^*G] && \text{(B.4)} \\
\frac{d[RG]}{dt} &= k_{g+}[R][G] - k_{g-}[RG] - \mu_+ k_{act}[RG] + \mu_- k_{deact}[R^*G] \\
&\quad -\nu_+ k_{lb+}[L][RG] + \nu_- k_{lb-}[LRG] && \text{(B.5)} \\
\frac{d[R^*G]}{dt} &= \mu_+ k_{g+}[R^*][G] - \mu_- k_{g-}[R^*G] + \mu_+ k_{act}[RG] - \mu_- k_{deact}[R^*G] \\
&\quad -\theta_{\zeta\nu}\zeta_+\nu_+ k_{lb+}[L][R^*G] + \theta_{\zeta\nu}\zeta_-\nu_- k_{lb-}[LR^*G] \\
&\quad -k_{GTP+}[R^*G] - k_{des+}[R^*G] + k_{des-}[R^*_{des}G] && \text{(B.6)} \\
\frac{d[LRG]}{dt} &= \nu_+ k_{g+}[LR][G] - \nu_- k_{g-}[LRG] - \theta_{\zeta\mu}\zeta_+\mu_+ k_{act}[LRG] \\
&\quad +\theta_{\zeta\mu}\zeta_-\mu_- k_{deact}[LR^*G] + \nu_+ k_{lb+}[L][RG] - \nu_- k_{lb-}[LRG] && \text{(B.7)} \\
\frac{d[LR^*G]}{dt} &= \theta_{\nu\mu}\mu_+\nu_+ k_{g+}[LR^*][G] - \theta_{\nu\mu}\mu_-\nu_- k_{g-}[LR^*G] \\
&\quad +\theta_{\zeta\mu}\zeta_+\mu_+ k_{act}[LRG] - \theta_{\zeta\mu}\zeta_-\mu_- k_{deact}[LR^*G] \\
&\quad +\theta_{\zeta\nu}\zeta_+\nu_+ k_{lb+}[L][R^*G] - \theta_{\zeta\nu}\zeta_-\nu_- k_{lb-}[LR^*G] \\
&\quad -\nu_- k_{GTP+}[LR^*G] - k_{des+}[LR^*G] + k_{des-}[LR^*_{des}G] && \text{(B.8)} \\
\frac{d[G]}{dt} &= -\nu_+ k_{g+}[LR][G] + \nu_- k_{g-}[LRG] \\
&\quad -\theta_{\nu\mu}\mu_+\nu_+ k_{g+}[LR^*][G] + \theta_{\nu\mu}\mu_-\nu_- k_{g-}[LR^*G] \\
&\quad -\mu_+ k_{g+}[R^*][G] + \mu_- k_{g-}[R^*G] \\
&\quad -k_{g+}[R][G] + k_{g-}[RG] + k_{RA+}[\alpha_{GDP}][\beta\gamma] - k_{RA-}[G] && \text{(B.9)} \\
\frac{d[\alpha_{GTP}]}{dt} &= k_{GTP+}[R^*G] - k_{GTP-}[R^*][\alpha_{GTP}][\beta\gamma] - k_{gd+}[\alpha_{GTP}] + k_{gd-}[\alpha_{GDP}] \\
&\quad -\nu_+ k_{GTP-}[LR^*][\alpha_{GTP}][\beta\gamma] + \nu_- k_{GTP+}[LR^*G]
\end{aligned}$$

$$-k_{PLCb+}[\alpha_{GTP}][PLC] + k_{PLCb-}[\alpha_{GTP}PLC] \tag{B.10}$$

$$\frac{d[\beta\gamma]}{dt} = = k_{GTP+}[R^*G] + \nu_- k_{GTP+}[LR^*G] - k_{RA+}[\alpha_{GDP}][\beta\gamma] \tag{B.11}$$

$$\frac{d[\alpha_{GDP}]}{dt} = k_{gd+}[\alpha_{GTP}] - k_{RA+}[\alpha_{GDP}][\beta\gamma] + k_{PLCdis}[\alpha_{GTP}PLC^*] \tag{B.12}$$

$$\frac{d[PLC]}{dt} = -k_{PLCb+}[\alpha_{GTP}][PLC] + k_{PLCb-}[\alpha_{GTP}PLC] + k_{PLCdis}[\alpha_{GTP}PLC^*] \tag{B.13}$$

$$\frac{d[\alpha_{GTP}PLC]}{dt} = k_{PLCb+}[\alpha_{GTP}][PLC] - k_{PLCb-}[\alpha_{GTP}PLC] - k_{PLCact}[\alpha_{GTP}PLC][Ca^{2+}] + k_{PLCdeact}[\alpha_{GTP}PLC^*] \tag{B.14}$$

$$\frac{d[\alpha_{GTP}PLC^*]}{dt} = k_{PLCact}[\alpha_{GTP}PLC][Ca^{2+}] - k_{PLCdeact}[\alpha_{GTP}PLC^*] - k_{PIP_2b+}[\alpha_{GTP}PLC^*][PIP_2] + k_{PIP_2b-}[\alpha_{GTP}PLC^*PIP_2] + k_{IP_3}[\alpha_{GTP}PLC^*PIP_2] - k_{PLCdis}[\alpha_{GTP}PLC^*] \tag{B.15}$$

$$\frac{d[\alpha_{GTP}PLC^*PIP_2]}{dt} = k_{PIP_2b+}[\alpha_{GTP}PLC^*][PIP_2] - k_{PIP_2b-}[\alpha_{GTP}PLC^*PIP_2] - k_{IP_3}[\alpha_{GTP}PLC^*PIP_2] \tag{B.16}$$

$$\frac{d[IP_3]}{dt} = k_{IP_3}[\alpha_{GTP}PLC^*PIP_2] - k_{IP_3deg}[IP_3] - k_{IP_3Rb+}[IP_3][IP_3R] + k_{IP_3Rb-}[IP_3IP_3R] \tag{B.17}$$

$$\frac{d[IP_3R]}{dt} = -k_{IP_3Rb+}[IP_3][IP_3R] + k_{IP_3Rb-}[IP_3IP_3R] \tag{B.18}$$

$$\frac{d[IP_3IP_3R]}{dt} = k_{IP_3Rb+}[IP_3][IP_3R] - k_{IP_3Rb-}[IP_3IP_3R] - k_{IP_3Ract+}[IP_3IP_3R][Ca^{2+}] + k_{IP_3Ract-}[IP_3IP_3RCa^{2+}] \tag{B.19}$$

$$\frac{d[Ca^{2+}]}{dt} = (1+v_r)(\gamma_0 + \gamma_1[IP_3IP_3RCa^{2+}])([C_{AV}] - [Ca^{2+}]) - \frac{p_1[Ca^{2+}]^2}{[Ca^{2+}]^2 + p_2^2} \tag{B.20}$$

$$\frac{d[IP_3IP_3RCa^{2+}]}{dt} = k_{IP_3Ract+}[IP_3IP_3R][Ca^{2+}] - k_{IP_3Ract-}[IP_3IP_3RCa^{2+}] - k_{IP_3Rinh+}[IP_3IP_3RCa^{2+}][Ca^{2+}] + k_{IP_3Rinh-}[IP_3IP_3RCa^{2+}Ca^{2+}] \tag{B.21}$$

$$\frac{d[IP_3IP_3RCa^{2+}Ca^{2+}]}{dt} = k_{IP_3Rinh+}[IP_3IP_3RCa^{2+}][Ca^{2+}] - k_{IP_3Rinh-}[IP_3IP_3RCa^{2+}Ca^{2+}] \tag{B.22}$$

$$\frac{d[R^*_{des}G]}{dt} = k_{des+}[R^*G] - k_{des-}[R^*_{des}G] \tag{B.23}$$

$$\frac{d[LR^*_{des}G]}{dt} = k_{des+}[LR^*G] - k_{des-}[LR^*_{des}G] \tag{B.24}$$

APPENDIX C

Sensitivity analysis-related information

C.1 Parameter bounds and labels

Parameter	Label	Range	Unit
R_{TOT}	1	$10^{-10} - 10^{-6}$	M
G_{TOT}	2	$10^{-10} - 10^{-6}$	M
PLC_{TOT}	3	$8 \times 10^{-9} - 8 \times 10^{-5}$	M
IP_3R_{TOT}	4	$8 \times 10^{-9} - 8 \times 10^{-5}$	M
$[PIP_2]$	5	$2.5 \times 10^{-7} - 2.5 \times 10^{-3}$	M
ζ_-	6	$10^{-3} - 10^{-2}$	Unitless
ζ_+	7	$10^{-1} - 1$	Unitless
μ_-	8	$10^{-3} - 10^{-2}$	Unitless
μ_+	9	$10^{-1} - 1$	Unitless
ν_-	10	$3.162 \times 10^{-1} - 3.162$	Unitless
ν_+	11	$31.62 - 316.2$	Unitless
k_{deact}	12	$100 - 10000$	s^{-1}
k_{act}	13	$10^{-1} - 1$	s^{-1}
k_{g-}	14	$3 \times 10^{-4} - 3 \times 10^{-2}$	s^{-1}
k_{g+}	15	$3.6 \times 10^{8} - 3.6 \times 10^{10}$	$M^{-1}s^{-1}$
k_{GTP+}	16	$10^{-4} - 10^{-2}$	s^{-1}
k_{RA+}	17	$1.2 \times 10^{9} - 3.6 \times 10^{11}$	$M^{-1}s^{-1}$
k_{gd+}	18	$10^{-2} - 1$	s^{-1}
k_{lb+}	19	$288.56 - 28856$	$M^{-1}s^{-1}$
k_{lb-}	20	$0.14 \times 10^{1} - 0.14 \times 10^{3}$	s^{-1}
k_{PLCb+}	21	$2.25 \times 10^{5} - 2.25 \times 10^{7}$	$M^{-1}s^{-1}$

Continued on the next page...

...continued from the previous page

Parameter	Label	Range	Unit
k_{PLCb-}	22	$10^{-1} - 10$	s^{-1}
k_{PLCact}	23	$30 \times 10^5 - 30 \times 10^7$	$M^{-1}s^{-1}$
$k_{PLCdeact}$	24	$10^{-1} - 10$	s^{-1}
k_{PLCdis}	25	$1.667 \times 10^{-1} - 16.67$	s^{-1}
k_{IP_3deg}	26	$2.5 \times 10^{-1} - 25$	s^{-1}
k_{IP_3}	27	$10 - 1000$	s^{-1}
k_{PIP2b+}	28	$10^{10} - 10^{12}$	$M^{-1}s^{-1}$
k_{PIP2b-}	29	$5 \times 10^2 - 5 \times 10^4$	s^{-1}
k_{IP_3Rb+}	30	$1.2 \times 10^8 - 1.2 \times 10^{10}$	$M^{-1}s^{-1}$
k_{IP_3Rb-}	31	$8 \times 10^{-1} - 8 \times 10^1$	s^{-1}
$k_{Ca^{2+}act+}$	32	$15 \times 10^5 - 15 \times 10^7$	$M^{-1}s^{-1}$
$k_{Ca^{2+}act-}$	33	$1.65 \times 10^{-1} - 16.5$	s^{-1}
$k_{Ca^{2+}inh+}$	34	$1.8 \times 10^4 - 1.8 \times 10^6$	$M^{-1}s^{-1}$
$k_{Ca^{2+}inh-}$	35	$0.21 \times 10^{-1} - 2.1$	s^{-1}
k_{des+}	36	$2.75 \times 10^{-3} - 2.75 \times 10^{-1}$	s^{-1}
$k_{PKCdes-}$	37	$2.75 \times 10^{-7} - 2.75 \times 10^{-5}$	s^{-1}
γ_1	38	$20 \times 10^4 - 20 \times 10^8$	$M^{-1}s^{-1}$
p_1	39	$6 \times 10^{-8} - 6 \times 10^{-4}$	Ms^{-1}
p_2	40	$0.45 \times 10^{-8} - 0.45 \times 10^{-4}$	M

Table C.1.1: Model parameters and their corresponding labels, as used in the sensitivity analysis in Chapter 4.

C.2 Sensitivity results for Chapter 4

Ranking	Parameter	Overall sensitivity (S'_{AR})
1	p_2	0.76
2	p_1	0.75
3	k_{des+}	0.72
4	IP_3R_{TOT}	0.72
5	k_{IP_3deg}	0.63
6	$[PIP_2]$	0.62
7	$k_{Ca^{2+}act-}$	0.57
8	γ_1	0.57
9	k_{PIP2b+}	0.53
10	k_{IP_3}	0.52
11	$k_{Ca^{2+}act+}$	0.51
12	k_{PLCdis}	0.50
13	k_{IP_3Rb+}	0.49
14	k_{IP_3Rb-}	0.49
15	k_{PIP2b-}	0.46
16	k_{GTP+}	0.46
17	k_{deact}	0.45
18	$k_{PKCdes-}$	0.41
19	k_{act}	0.41
20	G_{TOT}	0.41
21	R_{TOT}	0.38
22	μ_-	0.25
23	μ_+	0.23
24	PLC_{TOT}	0.23
25	k_{gd+}	0.21
26	k_{PLCb+}	0.21
27	k_{PLCact}	0.17
28	k_{PLCb-}	0.13
29	ζ_+	0.11
30	ν_+	0.09
31	$k_{Ca^{2+}inh-}$	0.08
32	$k_{Ca^{2+}inh+}$	0.08
33	ν_-	0.07
34	ζ_-	0.07
35	$k_{PLCdeact}$	0.07
36	k_{lb+}	0.06
37	k_{lb-}	0.06
38	k_{g-}	0.05
39	k_{g+}	0.04
40	k_{RA+}	0.01

Table C.2.1: Model parameters ranked according to their overall influence on the five objective functions, as quantified by (4.36), produced with $r = 1000$ trajectories. Averaged over parameter subspaces 1 and 2.

C.3 Extra simulations for Chapter 4

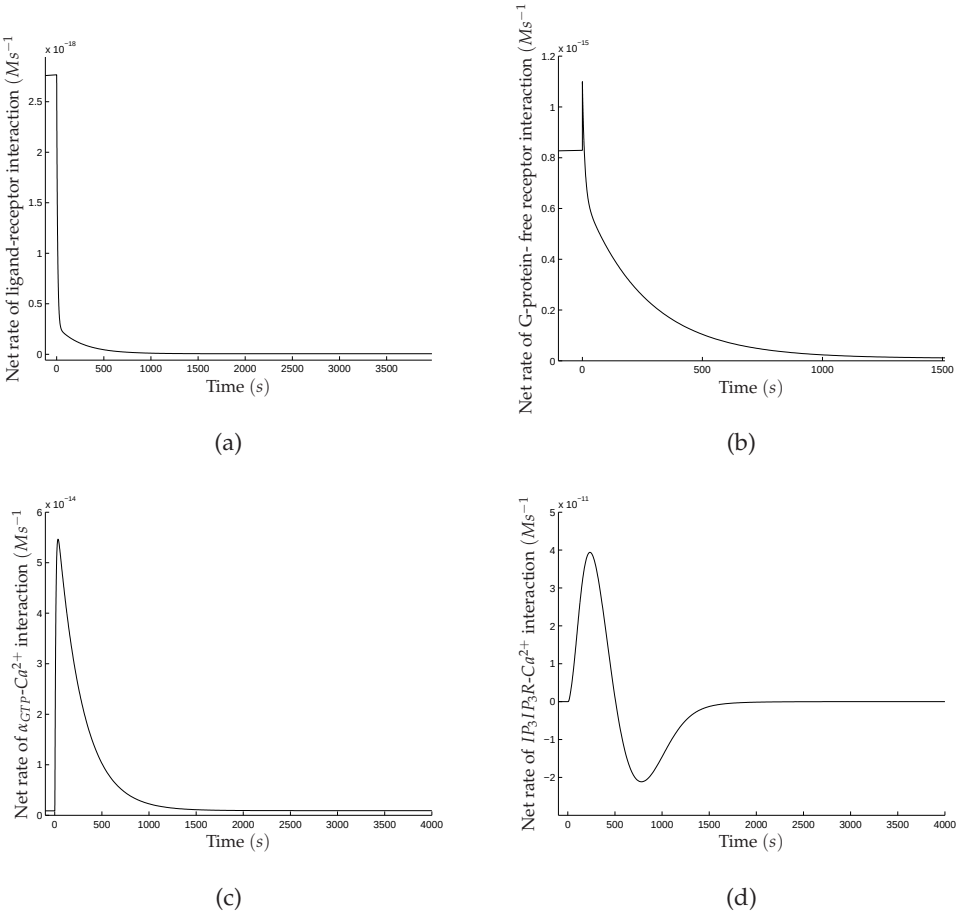


Figure C.1: The rates of the reversible reactions involving the least significant rate constants, 15:k_{g+} and 14:k_{g-}, 19:k_{lb+} and 20:k_{lb-}, 34:$k_{Ca^{2+}inh+}$ and 35:$k_{Ca^{2+}inh-}$, and 24:$k_{PLCdeact}$, do not reach steady state quickly.

C.4 GSA (and optimisation) results for Chapter 6

Ranking	Parameter	Overall score	$\boldsymbol{x}^{\mathbf{10}}_{\lambda_{50}}$	$\boldsymbol{x}^{\mathbf{10}}_{\lambda_{100}}$	$\boldsymbol{x}^{\mathbf{40}}_{\lambda_{100}}$	Unit
1	$k_{Ca^{2+}act-}$	0.74	149.95	108.30	159.18	s^{-1}
2	k_{PIP2b+}	0.63	0.395×10^{12}	3.84×10^{12}	3.8997×10^{12}	$M^{-1}s^{-1}$
3	k_{IP_3Rb-}	0.62	536.5	447.01	670.49	s^{-1}
4	p_1	0.61	3.02×10^{-4}	5.96×10^{-4}	5.8526×10^{-3}	$M^{-1}s^{-1}$
5	k_{IP_3Rb+}	0.59	0.82×10^{10}	1.04×10^{10}	0.83067×10^{10}	$M^{-1}s^{-1}$
6	k_{des+}	0.54	1.43	2.16	2.1102	s^{-1}
7	$[PIP_2]$	0.51	3.997×10^{-4}	0.366×10^{-4}	0.97612×10^{-4}	M
8	k_{deact}	0.48	15850	37358	58759	s^{-1}
9	p_2	0.47	0.9247×10^{-6}	2.8684×10^{-6}	1.1462×10^{-6}	M
10	R_{TOT}	0.46	1.682×10^{-7}	0.44141×10^{-7}	459.37×10^{-7}	M
11	$k_{PKCdes-}$	0.44				
12	γ_1	0.43				
13	k_{GTP+}	0.42				
14	$k_{Ca^{2+}act+}$	0.42				
15	k_{IP_3deg}	0.41				
16	k_{act}	0.38				
17	k_{PLCact}	0.35				
18	k_{PLCdis}	0.34				
19	$k_{Ca^{2+}inh+}$	0.33				
20	IP_3R_{TOT}	0.33				
21	μ_+	0.33				
22	ν_+	0.30				
23	$k_{Ca^{2+}inh-}$	0.30				
24	μ_-	0.28				
25	G_{TOT}	0.28				
26	k_{PLCb-}	0.28				
27	k_{g-}	0.27				
28	PLC_{TOT}	0.18				
29	ζ_+	0.16				
30	k_{PLCb+}	0.15				
31	k_{lb-}	0.15				
32	k_{gd+}	0.12				
33	ζ_-	0.12				
34	k_{IP_3}	0.11				
35	ν_-	0.11				
36	k_{lb+}	0.10				
37	k_{PIP2b-}	0.08				
38	k_{g+}	0.06				
39	$k_{PLCdeact}$	0.06				
40	k_{RA+}	0.01				

Table C.4.1: The ranking of parameters, according to a GSA carried out in a parameter space constrained around the GSA-guided parameter set (see Table B.2.2). This is the ranking that was used to reduce the search space in Chapter 6. The best parameter estimates found from fitting to the experimental data in Chapter 6 are also shown.

References

[1] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P Walter. *Molecular Biology of the Cell*. Garland Science, 2002.

[2] B.B. Aldridge, G. Haller, P.K. Sorger, and D.A. Lauffenburger. Direct Lyapunov exponent analysis enables parametric study of transient signalling governing cell behaviour. *Systems Biology, IEE Proceedings*, 153(6):425–432, 2006.

[3] J.F. Apgar, D.K. Witmer, F.M. White, and B. Tidor. Sloppy models, parameter uncertainty, and the role of experimental design. *Molecular BioSystems*, 6(10): 1890–1900, 2010.

[4] J. Ariens, E. Affinity and intrinsic activity in the theory of competitive inhibition. I. Problems and theory. *Archives internationales de pharmacodynamie et de thérapie*, 99(1):32, 1954.

[5] S. Armstrong, T. Steinfeld, M.T. Pulido-Rios, and S. S. Hegde. *In Vivo Pharmacodynamic Assays for M2 and M3 Muscarinic Receptors*, chapter 12, pages 13.1–13.10. John Wiley & Sons, Inc., 2001.

[6] S.P. Armstrong, C.J. Caunt, R.C. Fowkes, K. Tsaneva-Atanasova, and C.A. McArdle. Pulsatile and Sustained Gonadotropin-releasing Hormone (GnRH) Receptor Signaling. *Journal of Biological Chemistry*, 284(51):35746–35757, 2009.

[7] T. Back, U. Hammel, and H. Schwefel. Evolutionary Computation: Comments on the History and Current State. *IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION*, 1(1):3–17, 1997.

[8] H. Beyer and Schwefel H. Evolutionary strategies: A comprehensive introduction. *Natural Computing*, 1:3–52, 2002.

[9] J.W. Black and P. Leff. Operational models of pharmacological agonism. *Proceedings of the Royal society of London. Series B. Biological sciences*, 220(1219):141, 1983.

[10] K.T. Blackwell. Modeling Calcium Concentration and Biochemical Reactions. *Brains Minds and Media, pg*, pages 1–27, 2005.

[11] S. Boyd. Branch and bound methods. Retrieved September 23, 2011, from `http://www.stanford.edu/class/ee364b/bb_slides.pdf`, 2010. Lecture notes.

[12] S. Boyd and J. Mattingley. Branch and bound methods. Retrieved September 23, 2011, from `http://www.stanford.edu/class/ee364b/bb_notes.pdf`, 2010.

[13] L. Bridge, S. Hill, J. King, M. Owen, and P. Woodroffe. Mathematical modelling of signalling in a two-drug G-protein coupled receptor system. *Mathematical Biosciences*, 223:115–132, 2010.

[14] A.J. Bridges. Chemical inhibitors of protein kinases. *Chemical Reviews*, 101(8): 2541–2572, 2001.

[15] F. Campolongo, J. Cariboni, and A. Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, 22(10): 1509–1518, 2007.

[16] A.J. Clark. *The mode of action of drugs on cells*. The Williams & Wilkins company, 1933.

[17] M. Cooling, P. Hunter, and E. J. Crampin. Modeling Hypertrophic IP3 Transients in the Cardiac Myocyte. *Biophysical Journal*, 93:3421–3433, 2007.

[18] M. Davies, A. Secker, M. Halling-Brown, D. Moss, A. Freitas, J. Timmis, E. Clark, and D. Flower. GPCRTree: online hierarchical classification of GPCR function. *BMC research notes*, 1(1):67, 2008.

[19] A. De Lean, J.M. Stadel, and R.J. Lefkowitz. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *The Journal of biological chemistry*, 255(15):7108–7117, 1980.

[20] G.W. De Young and J. Keizer. A single-pool inositol 1, 4, 5-trisphosphate-receptor-based model for agonist-stimulated oscillations in Ca2+ concentration. *Proceedings of the National Academy of Sciences of the United States of America*, 89(20): 9895, 1992.

[21] I. Diener. Trajectory Methods In Global Optimization. In *Handbook of Global Optimization*, pages 649–668. Kluwer Academic, 1995.

[22] L. Dode, B. Vilsen, K. Van Baelen, F. Wuytack, J.D. Clausen, and J.P. Andersen. Dissection of the functional differences between sarco (endo) plasmic reticulum Ca2+-ATPase (SERCA) 1 and 3 isoforms by steady-state and transient kinetic analyses. *Journal of Biological Chemistry*, 277(47):45579, 2002.

[23] A. P. Engelbrecht. *Computational Intelligence: An Introduction*. Wiley, 2007.

[24] N. Frey and EN Olson. Cardiac hypertrophy: the good, the bad, and the ugly. *Annual review of physiology*, 65:45–79, 2003.

[25] P.E. Gill, W. Murray, and M.H. Wright. *Practical optimization*. Academic Press, 1981.

[26] A. Goldbeter, G. Dupont, and M.J. Berridge. Minimal model for signal-induced ca2+ oscillations and for their frequency encoding through protein phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America*, 87(4):1461, 1990.

[27] B.D. Gomperts, I.M. Kramer, and P.E. Tatham. *Signal Transduction*. Elsevier Academic Press, 2002.

[28] A.B. Goryachev. Modelling of intracellular Ca2+ dynamics. Retrieved October 11, 2011, from `http://www.biology.ed.ac.uk/research/groups/goryachev/Lectures/L24.pdf`, 2011. Lecture notes.

[29] R.N. Gutenkunst, J.J. Waterfall, F.P. Casey, K.S. Brown, C.R. Myers, and J.P. Sethna. Universally sloppy parameter sensitivities in systems biology models. *PLoS computational biology*, 3(10):1871–1878, 2007.

[30] E. Hazum and D. Keinan. Gonadotropin releasing hormone activation is mediated by dimerization of occupied receptors. *Biochemical and biophysical research communications*, 133(2):449–456, 1985.

[31] B.S. Hendriks, F. Hua, and J.R. Chabot. Analysis of mechanistic pathway models in drug discovery: p38 pathway. *Biotechnology Progress*, 24(1):96–109, 2008.

[32] S.J. Hill, G.J. Baker, and S. Rees. Reporter-gene systems for the study of G-protein-coupled receptors. *Current opinion in Pharmacology*, 1:526–532, 2001.

[33] Stephen J. Hill. G-protein-coupled receptors: past, present and future. *British Journal of Pharmacology*, 147:S27–S37, 2006.

[34] W. Huyer and A. Neumaier. Global optimization by multilevel coordinate search. *Journal of Global Optimization*, 14(4):331–355, 1999.

[35] Y. Jin, H. Yue, M. Brown, Y. Liang, and D. Kell. Improving data fitting of a signal transduction model by global sensitivity analysis. *Proceedings of the 2007 American Control Conference*, 11-13:2708–2713, 2007.

[36] M. Kang and H.G. Othmer. Spatiotemporal characteristics of calcium dynamics in astrocytes. *Chaos*, 19(3), 2009.

[37] M. Kang and H.G. Othmer. The variety of cytosolic calcium responses and possible roles of PLC and PKC. *Physical Biology*, 4(4):325–343, 2007.

[38] T. Kenakin. *Molecular pharmacology: A short course*. Blackwell science, 1997.

[39] T. Kenakin. The physiological significance of constitutive receptor activity. *Trends in Pharmacological Sciences*, 26(12):349–379, 2005.

[40] T.P. Kenakin. *A pharmacology primer: Theory, applications and methods*. Elsevier Academic Press, 2006.

[41] T.L. Kinzer-Ursem and J.J. Linderman. Both ligand- and cell-specific parameters control ligand agonism in a kinetic model of G protein coupled receptor signaling. *PLOS Comput. Biol.*, 3(1):e6, 2007.

[42] J.P. Kukkonen, J. Näsman, P. Ojala, C. Oker-Blom, and K.E. Akerman. Functional properties of muscarinic receptor subtypes Hm1, Hm3 and Hm5 expressed in Sf9 cells using the baculovirus expression system. *Journal of Pharmacology and Experimental Therapeutics*, 279(2):593–601, 1996.

[43] F. Kursawe. Evolution strategies — Simple "Models" of Natural Processes? *Revue Internationale de Systemique*, 7(5):627–642, 1993.

[44] G. Lemon, W.G. Gibsona, and M.R. Bennet. Metabotropic receptor activation, desensitization and sequestration–I: modelling calcium and inositol 1,4,5-trisphosphate dynamics following receptor activation. *Journal of Theoretical Biology*, 223(1):93–111, 2003.

[45] Y. Li and J. Rinzel. Equations for InsP3 Receptor-mediated [Ca2+]i Oscillations Derived from a Detailed Kinetic Model: A Hodgkin-Huxley Like Formalism, journal = Journal of Theoretical Biology. 166(4):461–473, 1994.

[46] J.J. Linderman. Modeling of G -protein coupled receptor signaling pathways. *J. Biol. Chem.*, 284:5427–5431, 2009.

[47] S. Lindskog. Structure and mechanism of carbonic anhydrase. *Pharmacology & therapeutics*, 74(1):1–20, 1997.

[48] T.J. Lukas. A Signal Transduction Pathway Model Prototype I: From Agonist to Cellular Endpoint. *Biophys. J.*, 87(3):1406–1416, 2004.

[49] F. Macian et al. NFAT proteins: key regulators of T-cell development and function. *Nature Reviews Immunology*, 5(6):472–484, 2005.

[50] J.S. Marchant, Y. Chang, S. Chung, R.F. Irvine, and C.W. Taylor. Rapid kinetic measurements of Ca2+ mobilization reveal that Ins(2,4,5)P3 is a partial agonist at hepatic InsP3 receptors. *Biochem. J.*, 321:573–576, 1997.

[51] *MATLAB R2009a Documentation*. Mathworks.

[52] L.T. May. M3-CHO mediated Ca2+ mobilisation in the absence of extracellular Ca2+. (Personal Communication), Institute of Cell Signalling, Queens's medical centre, University of Nottingham, 2010.

[53] L.T. May. Functional properties of fluorescent agonists at the Adenosine-A1 and -A3 receptor . (Personal Communication), Institute of Cell Signalling, Queens's medical centre, University of Nottingham, 2007.

[54] L.T. May. M3-CHO mediated Ca2+ mobilisation in the presence of extracellular Ca2+. (Personal Communication), Institute of Cell Signalling, Queens's medical centre, University of Nottingham, 2007.

[55] A.C. Megson, E.M. Walker, and S.J. Hill. Role of Protein Kinase C alpha in Signaling from the Histamine H1 Receptor to the Nucleus. *Mol Pharmacol*, 59(5): 1012–1021, 2001.

[56] C. Mendelson, M. Dufau, and K. Catt. Gonadotropin binding and stimulation of cyclic adenosine 3': 5'-monophosphate and testosterone production in isolated leydig cells. *Journal of Biological Chemistry*, 250(22):8818, 1975.

[57] S. Mette, K. Morten, D. Morten, B. Jean, N. Olivier, and H. Kim. Screening of protein kinase inhibitors identifies pkc inhibitors as inhibitors of osteoclastic acid secretion and bone resorption. *BMC Musculoskeletal Disorders*, 11(1):250, 2010.

[58] G. Milligan. Principles: Extending the utility of [35S]GTP gamma S binding assays. *Trends in Pharmacological Sciences*, 24(2):87 – 90, 2003.

[59] C.G. Moles, P. Mendes, and J.R. Banga. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome research*, 13 (11):2467, 2003.

[60] C.G. Moles, M Pedro, and J.R. Banga. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Genome Res.*, 13:2467–2474, 2003.

[61] G.R. Monteith, D. McAndrew, H.M. Faddy, and S.J. Roberts-Thomson. Calcium and cancer: targeting Ca2+ transport. *Nature Reviews Cancer*, 7(7):519–530, 2007.

[62] T.M. Moore, P.M. Chetham, J.J. Kelly, and T. Stevens. Signal transduction and regulation of lung endothelial cell permeability. Interaction between calcium and cAMP. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 275 (2):L203, 1998.

[63] M.D. Morris. Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33 (2):161 – 174, 1991.

[64] S.A. Morris, V. Correa, T.J.A. Cardy, G. O'Beirne, and C.W. Taylor. Interactions between inisitol trisphosphate receptors and fluorescent Ca^{2+} indicators. *Cell Calcium*, 25 (2):137–142, 1995.

[65] H. Motulsky and A. Christopoulos. *Fitting models to biological data using linear and nonlinear regression: a practical guide to curve fitting*. Oxford University Press, USA, 2004.

[66] J.A. Nelder and R. Mead. A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313, January 1965.

[67] O. Nelles. *Nonlinear system identification: from classical approaches to neural networks and fuzzy models*. Springer Verlag, 2001.

[68] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, 1999.

[69] G.R. Oliver, H. Isabell, and T. Gerald. Transcriptional response to muscarinic acetylcholine receptor stimulation: Regulation of Egr-1 biosynthesis by ERK, Elk-1, MKP-1, and calcineurin in carbachol-stimulated human neuroblastoma cells. *Archives of Biochemistry and Biophysics*, 470(1):93 – 102, 2008.

[70] H. Othmer and Y. Tang. *Oscillations and waves in a model of of InsP3-controlled calcium dynamics*. Experimental and theoretical advances in biological pattern formations. New York: Plenum, 1993.

[71] Leff P., Scaramellini C., C. Law, and McKechnie K. A three-state receptor model of agonist action. *Trends in Pharmacological Sciences*, 18(4):355–362, 1 October 1997.

[72] P.M. Pardalos, Romeijn H.E., and H. Tuy. Recent developments and trends in global optimization. *Journal of Computational and Applied Mathematics*, 124:209–228, 2000.

[73] G. Powis, M.J. Seewald, C. Gratas, D. Melder, J. Riebow, and E.J. Modest. Selective inhibition of phosphatidylinositol phospholipase c by cytotoxic ether lipid analogues. *Cancer research*, 52(10):2835, 1992.

[74] J.E. Purvis. *A systems approach to cellular signal transduction*. PhD thesis, University of Pennsylvenia, 2009.

[75] J.E. Purvis, R. Radhakrishnan, and S.L. Diamond. Steady-state kinetic modeling constrains cellular resting states and dynamic behavior. *PLoS Computational Biology*, 5(3):e1000298, 2009.

[76] M.J. Rebecchi and O.M. Rosen. Purification of a phosphoinositide-specific phospholipase C from bovine brain. *Journal of Biological Chemistry*, 262(26):12526, 1987.

[77] A. Richardson and C.W. Taylor. Effects of Ca2+ chelators on purified inositol 1,4,5-trisphosphate (InsP3) receptors and InsP3-stimulated Ca2+ mobilization. *J. Biol. Chem.*, 268:11528–11533, 1993.

[78] T.P. Runarsson and X. Yao. Stochastic Ranking for Constrained Evolutionary Optimization. *IEEE Transactions on Evolutionary Computation*, 4:284–294, 2000.

[79] Rebois R.V., Warner D.R., and Basi N.S. Does subunit dissociation necessarily accompany the activation of all heterotrimeric G proteins? *Cellular Signalling*, 9: 141–151(11), February 1997.

[80] A. Saltelli, M. Ratto, S. Tarantola, and F. Campolongo. Sensitivity analysis for chemical models. *Chemical Reviews*, 105(7):2811–2828, 2005.

[81] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola. *Global sensitivity analysis: the primer*. Wiley-Interscience, 2008.

[82] S. Schuster, M. Marhl, and T. Hofer. Modelling of simple and complex calcium oscillations. from single-cell responses to intercellular signalling. *Eur. J. Biochem.*, 269(5):1333–55., 2002.

[83] H.P.P. Schwefel. *Evolution and Optimum Seeking: The Sixth Generation*. John Wiley & Sons, Inc., 1993.

[84] R ed. Seifert. *G Protein-Coupled Receptors as Drug Targets*. C.H.I.P.S, 2006.

[85] S.E. Senogles, A.M. Spiegel, E. Padrell, R. Iyengar, and M.G. Caron. Specificity of receptor-G protein interactions. Discrimination of Gi subtypes by the D2 dopamine receptor in a reconstituted system. *J. Biol. Chem.*, 265(8):4507–4514, 1990.

[86] L.D. Shea, R.R. Neubig, and J.J. Linderman. Timing is everything: The role of kinetics in G protein activation. *Life Sciences*, 178:647–658, 2000.

[87] J. Sneyd, editor. *Tutorials in Mathematical Biosciences II: Mathematical Modelling of Calcium Dynamics and Signal Transduction*. Springer, 2003.

[88] R.F. Spurney, P.J. Flannery, S.C. Garner, K. Athirakul, S. Liu, F. Guilak, and L.D. Quarles. Anabolic effects of a G protein-coupled receptor kinase inhibitor expressed in osteoblasts. *Journal of Clinical Investigation*, 109(10):1361–1372, 2002.

[89] M.P. Strathmann and M.I. Simon. G alpha 12 and G alpha 13 subunits define a fourth class of G protein alpha subunits. *Proceedings of the National Academy of Sciences*, 88(13):5582–5586, 1991.

[90] K. Suzuki, T. Siddique, S. Ohya, M. Kaneko, S. Maruyama, S. Shoji, and M. Uyeda. A Novel Phospholipase C Inhibitor, S-PLI Produced by Streptomyces Sp. Strain No. A-6288. *Journal of Enzyme Inhibition and Medicinal Chemistry*, 10(3): 177–186, 1996.

[91] J.E. Swatton and C.W. Taylor. Fast biphasic regulation of type 3 inositol trisphosphate receptors by cytosolic calcium. *J. Biol. Chem.*, 277:17571–17579, 2002.

[92] D.A. Sykes, Dowling M.R., and Charlton S.J. Exploring the mechanism of agonist efficacy: A relationship between efficacy and agonist dissociation rate at the muscarinic M3 receptor. Poster presented at the 4th British Pharmacological Society Focused Meeting on Cell Signalling, University of Leicester, April 2007.

[93] D.A. Sykes, Dowling M.R., and Charlton S.J. Exploring the mechanism of agonist efficacy: A relationship between efficacy and agonist dissociation rate at the muscarinic M3 receptor. *Molecular Pharmacology*, 76:543 – 551, 2009.

[94] R. Thul, T.C. Bellamy, H.L. Roderick, M.D. Bootman, and S. Coombes. Calcium oscillations. *Cellular Oscillatory Mechanisms*, 641:1–27, 2009.

[95] K.T. Tsaneva-Atanasova. *A Mathematical Study of Calcium Oscillations and Waves*. PhD thesis, University of Auckland, 2004.

[96] T. Uchiyama and R. Chess-Williams. Muscarinic receptor subtypes of the bladder and gastrointestinal tract. *J. Smooth Muscle Res.*, 40 (6):237–247, 2003.

[97] H. Ulrich and B. Thomas. Evolution Strategies on Noisy Functions How to Improve Convergence Properties. In *Parallel Problem Solving from Nature*, pages 159–168. Springer, 1994.

[98] D.K. Vassilatis, J.G. Hohmann, H. Zeng, F. Li, J.E. Ranchalis, M.T. Mortrud, A. Brown, S.S. Rodriguez, J.R. Weller, A.C. Wright, et al. The G protein-coupled receptor repertoires of human and mouse. *Proceedings of the National Academy of Sciences*, 100(8):4903–4908, 2003.

[99] F. Villatte, H. Schulze, R.D. Schmid, and T.T. Bachmann. A long insertion reverts the functional effect of a substitution in acetylcholinesterase. *Protein engineering*, 16(7):463, 2003.

[100] M. Waelbroeck, L. Boufrahi, and S. Swillens. Seven helix receptors are enzymes catalysing g protein activation. what is the agonist kact? *Journal of theoretical biology*, 187(1):15–38, 1997.

[101] N.J. Warren, Tawhai M.H., and Crampin E.J. Mathematical modelling of calcium wave propagation in mammalian airway epithelium: evidence for regenerative ATP release. *Exp Phys*, 95.1:232–249, 2009.

[102] T.M. Washington, J.M. Blum, M.C. Reed, and P.M. Conn. A mathematical model for LH release in response to continuous and exposure of gonadotrophs to GnRH. *Theoretical Biology and Medical Modelling*, 1:9, 2004.

[103] J.M. Weiss, P.H. Morgan, M.W. Lutz, and T.P. Kenakin. The cubic ternary complex receptor occupancy model I. Model description. *J. theor. Biol.*, 178:151–167, 1996.

[104] G.B. Willars, S.R. Nahorski, and J.R.A. Challiss. Differential Regulation of Muscarinic Acetylcholine Receptor-sensitive Polyphosphoinositide Pools and Consequences for Signaling in Human Neuroblastoma Cells. *J. Biol. Chem.*, 273(9): 5037–5046, 1998.

[105] J.M.R. Willets, J.A. Challiss, E. Kelly, and S.R. Nahorski. G Protein-Coupled Receptor Kinases 3 and 6 Use Different Pathways to Desensitize the Endogenous M3 Muscarinic Acetylcholine Receptor in Human SH-SY5Y Cells. *Molecular Pharmacology, Vol. 60, No. 2. (1 August 2001), pp. 321-330.*, 60(2):321–330, 2001.

[106] P. Woodroffe. *Mathematical Modelling of Cell Signalling*. PhD thesis, University of Nottingham, 2004.

[107] Y. Zheng and A. Rundell. Comparative study of parameter sensitivity analyses of the TCR-activated Erk-AMPK signaling pathway. *IEE Proc Syst Biol*, 153:201–211, 2006.