Swansea University
Prifysgol Abertawe

Contributions to mathematical pharmacology: new receptor theory with dimeric receptor models

Carla White

Swansea University

Submitted to Swansea University in fulfilment of the requirements for the Degree of Doctor of Philosophy

2021

Abstract

Classical receptor theory is largely built on assumptions of monomeric receptors. In this thesis, we contribute to receptor theory by considering the now widely accepted cases of dimeric receptors. The implications of dimerisation for drug discovery and therapeutics remain unclear. Therefore, a theoretical consideration of ligand binding and signalling via receptor dimers is warranted. Here, we develop mathematical models for ligand binding at dimerised and dimerising receptors. A key factor in developing these theoretical models is *cooperativity* across the dimer, whereby binding of a ligand to one protomer affects the binding of a ligand to the other protomer. The effects of cooperativity on binding dynamics are a primary point of interest.

The first models we present focus on G protein-coupled receptors, where we assume that all receptors are pre-dimerised. Ligand binding models give linear systems of differential equations which we use to analyse time course behaviours. At equilibrium, these models may exhibit multi-phasic log dose response curves, critically depending on cooperativity factors. When considering receptor activation, we see dose response curves that are indicative of non-standard ligand-receptor interactions, giving a quantitative and qualitative platform for analysing and interpreting data when dimers are suspected. A ligand induced model for vascular endothelial growth factor receptors is developed, whereby receptors exist constitutively as monomers and dimerise in response to ligand binding. The resulting nonlinear system of differential equations is investigated using numerical computations and perturbation methods. We see an excellent fit to published data, validating the model.

The utility of our models in parameter estimation is explored theoretically using structural identifiability analysis. This determines which parameters can be theoretically estimated from fitting. This analysis is valuable but often overlooked when fitting to ligand-receptor interaction models. We explore the identifiability of some canonical ligand binding models, and our dimer binding models, providing a tutorial and results to contribute to the receptor theory toolbox.

Acknowledgements

First and foremost I am extremely grateful to my supervisors, Dr Gibin Powathil, Dr Lloyd Bridge and Dr Vivi Rottschäfer for their invaluable advice, continuous support, and patience during my PhD study. Without them I would not have been able to complete this research, and without whom I would not have made it through my PhD. Their immense knowledge and plentiful experience have encouraged me in all the time of my academic research and daily life. I would like to thank Dr Martin Crossley for being a much needed mentor, and Dr Sara Hamis for taking each step of the journey alongside me. And my biggest thanks to my family for all the support you have shown me through this research. For my kids, sorry for being even grumpier than normal whilst I wrote this thesis! For my mum, for picking me up when things got tough. And finally, for my husband Simon, thank you tremendously for being there through all of the ups and downs. Without their tremendous understanding and encouragement in the past few years, it would be impossible for me to complete my study.

Suggested Layout of Declaration/Statements page

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ..[redacted]..... (candidate)

Date .....24/3/2021.....

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Where correction services have been used, the extent and nature of the correction is clearly marked in a footnote(s).

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ....[redacted]..... (candidate)

Date .....24/3/2021.....

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....[redacted]..... (candidate)

Date .....24/3/2021.....

NB: *Candidates on whose behalf a bar on access has been approved by the University (see Note 7), should use the following version of Statement 2:*

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access approved by the Swansea University.**

Signed ....[redacted]..... (candidate)

Date .....24/3/2021.....

Contents

List of Figures

1.5.4 Schematic: The extended ternary complex model [68]. Receptors, in this model, can bind a ligand (AR), activate (R^* and AR^*), though G proteins can only bind activate receptors (R^*G and AR^*G). This model contans both cooperativity factors α (from the two-state model) and γ (from the first ternary complex model).

1.5.5 Schematic: The cubic ternary complex model [133]. All combinations of bound receptors, active receptors and receptors with G proteins attached are possible. Along with the two existing cooperativity factors α and γ, two new cooperativity factors, β and δ, are introduced.

1.6.1 Schematic for a competition binding scenario. Labelled ligand, A, competes with unlabelled ligand, B, for receptor R.

1.6.2 Concentration of u and v (equations (1.43)) are plotted on a log-log scale to highlight the time course dynamics on the different timescales. Parameters used for the plot are $\lambda = 3.4$, $\kappa = 6.7$, $\epsilon = 1e-2$.

1.6.3 The numerical solution (to equations (1.42))is plotted together with the inner and outer approximated solutions. The approximations in each region show a good approximation to the curves. Parameters used for the plot are $\lambda = 3.4$, $\kappa = 6.7$, $\epsilon = 1e-2$.

1.6.4 Plotting the concentration curves for all species individually, for a competition binding with monomeric receptor model, using three different parameter sets. Although all curves have the same measured output (concentration of $[AR]$) curve, non-identifiability can be seen from the differences in the different species curves.

2.1.1 Schematic: A ligand A binds to one side of a dimer, R, creating AR. A second ligand subsequently binds this complex, creating ARA.

2.1.2 LogDR curves (plotted using equation (2.9)) for α ranging from extreme negative to extreme positive cooperativity. We see that extra inflections appear in A_{bound} when we have extreme low cooperativity.

2.1.3 The binding of ligand A to a pre-dimerised receptor results in slight peaks in $[AR]$, while positive cooperativity leads to all dimers becoming dual bound. The plot shows a numerical solution to the system in equations (2.1). We use a ligand concentration of $[A] = 10^{-8}M$ while the cooperativity factors were set at $\alpha_+ = 2$ and $\alpha_- = 0.01$ giving positive cooperativity.

2.1.4 Plot shows a numerical solution to equations (2.1) for a varying ligand concentration. With high levels of $[A]$ in the system we no longer see peaks in $[AR]$. Cooperativity values are fixed at $\alpha_+ = 2$ and $\alpha_- = 0.01$.

2.1.5 Plot shows a numerical solution to equations (2.1) for a varying forward cooperativity factor. As we move from positive to negative cooperativity we see a more pronounced peak in $[AR]$ with A_{bound} tending to lower concentrations. Ligand concentration for the plot is $[A] = 10^{-8}M$.The value $\alpha_- = 1$ was fixed so that cooperativity depends solely on α_+.

2.2.1 Schematic: Two ligands, A and B compete for binding sites on dimer R.

2.2.2 Schematic of the two-ligand dimer model presented by May *et al.* in [83].

2.2.3 Plotting logDR curves, using equation (2.39), for varying α shows extra inflections when we have low A-A cooperativity regardless of $[B]$. Other cooperativity factors are fixed at $\beta = \gamma = 1$.

2.2.4 Individual species plots (equations (2.35)) for a varying α help to understand the effects of cooperativity on the dose response curves. Competition ligand concentration is fixed at $B = 10^{-8}M$, while $\beta = \gamma = 1$.

2.2.5 Plotting logDR curves, using equation (2.39), for varying β shows extra inflections appear when we have both low B-B cooperativity and low $[B]$. Other cooperativity factors are fixed at $\alpha = \gamma = 1$.

2.2.6 Individual species plots for a varying β help to understand the effects of cooperativity on the dose response curves. Competition ligand concentration is fixed at $B = 10^{-5}M$, while $\alpha = \gamma = 1$.

2.2.7 Plotting logDR curves, using equation (2.39), for varying γ shows extra inflections when we have high A-B cooperativity as well as low $[B]$. Other cooperativity factors are fixed at $\beta = \gamma = 1$.

2.2.8 Individual species plots for a varying γ help to understand the effects of cooperativity on the dose response curves. Competition ligand concentration is fixed at $B = 10^{-8}M$, while $\beta = \gamma = 1$.

2.2.9 In the time course plot, we have a numerical solution to the system in equations (2.28), where we see peaks in both $[AR]$ and $[BR]$. Cooperativity values are $\alpha_+ = \beta_+ = \gamma_+ = 2$, $\alpha_- = \beta_- = \gamma_- = 0.01$, while ligand concentrations used are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

2.2.10 Plotting the numerical solution to equations (2.28) with a varying α_+ while all other cooperativity factors equal 1. Increasing α_+ results in increases in $[ARA]$. Ligand concentrations are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

2.2.11 Plotting the numerical solution to equations (2.28) with a varying β_+ while all other cooperativity factors equal 1. Increasing β_+ results in increases in $[BRB]$. Ligand concentrations are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

2.2.12 Plotting the numerical solution to equations (2.28) with a varying γ_+ while all other cooperativity factors equal 1. Increasing γ_+ results in increases in $[ARB]$. Ligand concentrations are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

2.3.1 Schematic: the binding of a single ligand to a pre-formed heterodimer.

2.3.2 Plotting logDR curves, using equation (2.54), for varying α show extra inflections appear for heterodimers, similar to those in the homodimer binding curves.

2.3.3 Plotting the numerical solution to equations (2.44) shows peaks appearing in the time course curves for $[AR_1\text{-}R_2]$ and $[R_1\text{-}R_2A]$. Cooperativity values used for the plot are $\alpha_+ = 2$ and $\alpha_- = 0.01$, while ligand concentration is $[A] = 10^{-8}M$.

2.3.4 Plotting the numerical solution to equations (2.44) for a range of both cooperativity factors (α_+ and α_-) result in peaks appearing in some of the $[AR_1\text{-}R_2]$ and $[R_1\text{-}R_2A]$ curves. Ligand concentration used is $[A] = 10^{-8}M$.

3.1.1 The binding of a ligand A, and activation of pre-dimerised receptors show a spherical schematic describing interactions with ten dimer complexes. Binding is determined by K_A, while activation K_{act}. Cooperativity factors α, λ, ν and ξ describe crosstalk and the more complex interactions.

3.3.1 The logDR A_{bound} curve for this activation model (equation (3.8)) is plotted alongside the logDR A_{bound} curve for the binding model (Section 2.1, equation (3.9)). With equal binding parameters we see that the curves appear to be the equal.

3.3.2 Plotting a logDR curve (equation (3.12)) for a range of values for binding cooperativity factor α. While all signal curves have a bell shape, increasing α decreases the peak size.

3.3.3 Plotting a logDR curve (equation (3.18)) for a range of values for activation cooperativity factor λ. Increasing λ increases the signal for all concentration of $[A]$.

3.3.4 Plotting a logDR curve (equation (3.26)) for a range of values for efficacy parameter ν. This parameter determines whether the ligand acts as an agonist (when $\nu > 1$), or inverse agonist ($\nu < 1$), while $\nu = 1$ shows a curve with a small peak.

3.3.5 We plot the signal curve (equation (3.26)) for values of ν close to one, showing the smoothing out of the peak.

3.3.6 Plotting two sets of logDR curve (equation (3.33)), with $\nu = 100, \xi = 0.1$ and also $\nu = 0.1, \xi = 100$. The symmetry in the model sphere means that, although the curves of $[R_A^* R]$ and $[R_A R^*]$ change with varying ν and ξ, their sum, as well as all other concentrations, is equal.

3.3.7 Plotting a logDR curve (equation (3.6)) for a range of values for weighting parameter ω, with all cooperativity factors fixed at one. Increasing ω increases the contribution symmetric dimers have to the signal, which then increases the signal.

3.4.1 We fix $\xi = 1$ and vary λ on each row, ν in each column and α within each subplot (using equation (3.6)). We see effects such as inflections and overshoots in some of the curves.

3.4.2 We plot dose response curves (equation (3.6)) for $\alpha = 10^{-4}, \nu = 10$ and $\lambda = 0.1$, as well as for these each being set individually.

3.4.3 We fix $\xi = \nu$ and vary λ on each row, ν in each column and α within each subplot. We see some extra inflections but now no overshoots in the curve.

4.1.1 Ligand binding and dimerisation is a two step process whereby a two poled ligand (in red) first binds a monomeric receptor (blue) before then binding a second and instantaneously dimerising the monomers.

4.1.2 Schematic representing the reactions resulting from the binding of a two poled ligand to two monomeric receptors.

4.2.1 LogDR curves for varying cooperativity factor, ψ. The plots show the overall signal as well as concentrations of R (equation (4.7a)), AR (equation (4.7b)) and RAR (equation (4.7c)).

4.3.1 A numerical investigation into the effects of cooperativity factors ψ_+ and ψ_-. The columns show the overall signal as well as individual species (equations 4.1) while in each row we fix the ψ_+ value. Each plot then shows a varied ψ_-. Plots are created with ligand concentration $[A] = 10^{-10}M$.

4.5.1 Both the solution to the full system in equations (4.32) is plotted along side the solution of the reduced, approximated system (equations (4.34)) in red. Conservation of receptors is used to show the approximation for q. Parameters used to create plots are $\alpha = \beta = 1$ and $\gamma = \epsilon = 0.001$.

4.5.2 Numerical solutions of the full system (equation (4.35)), on a log-log scale. Plot created with $\alpha = 6.8$, $\beta = 1/\epsilon$, $\gamma = 1.4$ with $\epsilon = 10^{-4}$.

4.5.3 Plotting the numerical solution of equation (4.35) with inner (equation (4.40))), intermediate (equation (4.65)) and outer (equations (4.52) and (4.55)) solutions. The intermediate solution matches both the inner and outer solution, creating a full approximation to the numerical solution. Conservation of receptors is used to show the approximation for q. Plot created with $\alpha = 6.8$, $\gamma = 1.4$ and $\beta = \epsilon = 10^{-4}$.

4.5.4 Numerical solutions of the full system in equation (4.70) on a log-log scale. Plot created with α=6.8, β=1, $\gamma = \epsilon = 10^{-4}$.

4.5.5 Plotting the numerical solution of equation (4.70) with asymptotic solutions (4.88) and (4.108) show good agreement in both regions. Parameters used to create plot are α=6.8, β=1, ϵ=1e-4.

4.6.1 Data published in [101] is used to estimate the model parameters. Experiments were performed using five concentrations of three different VEGF isoforms (VEGF $_{165}$a-TMR, VEGF $_{165}$b-TMR and VEGF $_{121}$a-TMR) and are scaled with respect to R_{tot}. An excellent fit to the data is seen from fitting to all data sets simultaneously. Parameter values returned can be seen in Table 4.2.

4.6.2 Individual species curves using the estimated parameters show peaks some of the $[RAR]$ curves.

5.2.1 An algorithm for using the transfer function method to determine identifiability.

5.2.2 An algorithm for using the Taylor series method to determine identifiability.

5.2.3 An algorithm for using the similarity transformation method to determine identifiability.

5.3.1 Three sets of parameters are used to plot the system given in equations (5.23). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.1 together with $[A] = 10^{-8}M$.

5.3.2 Three sets of parameters are used to plot the system given in equations (5.45). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.2 together with $[A] = 10^{-8}M$.

5.3.3 Three sets of parameters are used to plot the system given in equations (5.90). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.3 together with $[A] = 10^{-8}M$.

5.3.4 Three sets of parameters are used to plot the system given in equations (5.125). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.4 together with $[A] = 10^{-8}M$.

5.3.5 The STRIKE-GOLDD toolbox [129] confirms the identifiability of the parameters k_+, k_- and ψ_-.

*

Chapter 1

Introduction

1.1 Motivations

In the pursuit of drug discovery and therapeutics, pharmacological analysis is required to understand the interactions of drug/ligand molecules and cell surface receptors. Classical pharmacological receptor theory falls short of analysing the dynamics of the range of these interactions, despite an increased recognition of their importance. In this thesis, we will develop new mathematical models to investigate ligand-cell interactions when receptor dimerisation is suspected, and use mathematical techniques to analyse these models in order to give important insights into the pharmacological impacts. In this chapter, we review the underlying pharmacology as well as some of the mathematical techniques used throughout the thesis.

1.2 Basic pharmacology principles

Pharmacology is a branch of medical science that studies the actions of ligands. Not only is pharmacology essential in discovering new ligands and improving the effectiveness of existing ligands, whilst reducing unwanted side effects, but is also used to help in understanding why some ligands are effective for some people but not others. Although pharmacology has been studied for over a century the underlying concepts remain the same, and these begin with the ligand. A ligand is a substance that is either man-made, natural or endogenous molecule that binds to a protein and causes a physiology or

psychology effect on the body [66]. A detailed history of the development of pharmacology can be found in [69].

1.2.1 Pharmacokinetics/pharmacodynamics

Pharmacology can be divided into two broad categories: pharmacokinetics and pharmacodynamics. Pharmacokinetics describes the effect the body has on the ligand as it moves through the body. This includes processes such as absorption, distribution, metabolism and elimination [64]. Pharmacodynamics focuses on the effect the ligand has on the body once it reaches the site of action, and so dose-dependent issues such as toxicity and efficacy are considered. Pharmacodynamics is also used to determine what the appropriate use for the ligand is and what disease or symptom it will target. In this thesis, we focus on receptor theory, which falls under pharmacodynamics.

1.2.2 Cell surface receptors

A cell surface receptor is a protein that lies on the surface of a cell, and is the component that interacts with the ligand and initiates the reactions that lead to an observed effect [66]. There are many types of receptors in human cells, but the most well characterised are G-protein coupled receptors (GPCRs), enzyme-linked receptors and ligand-gated ion channels.

A ligand is a molecule that binds with a specific receptor and can be a drug or endogenous molecules such as hormones [118]. When a ligand molecule is in close proximity to a relevant receptor they bind together, by means of a chemical bond, setting off a chain of reactions that alter the activities of the cell on which the receptor resides. For binding to occur the ligand molecule must be the same shape as the receptor's binding site, so that these fit together in a lock and key style manner [64, 107]. The 'strength' with which a ligand binds to a receptor is an important concept, known as affinity.

1.2.3 Cell signalling

To elicit a response a signal needs to be transmitted into the interior of the cell. In Figure 1.2.1 we see an example of how ligand-receptor binding leads to a physiological response. Ligands can act as agonists, antagonists or inverse agonists (see Section 1.2.6) based on their effect on a cell's signal. Upon an agonist binding, a receptor becomes activated, which initiates a conformational change in the receptor, allowing for the binding of other molecules or proteins. This subsequently initiates a chain of biochemical reactions within the cell, which ultimately leads to a cellular response, that is, a change in the cell's behaviour or characteristic. This is known as a signal transduction pathway. Examples of signalling pathways for GPCRs can be seen in Section 1.3.2.

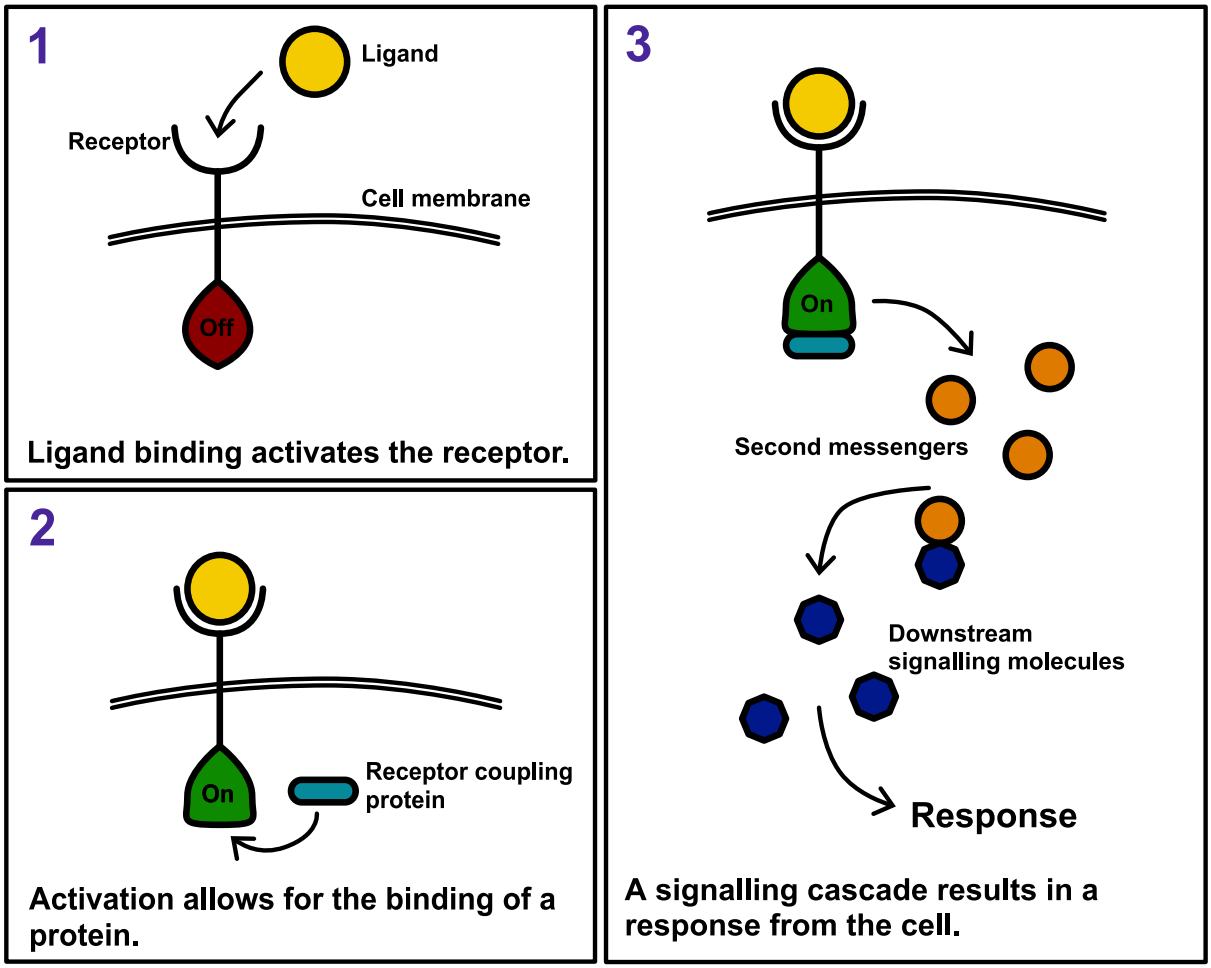

Figure 1.2.1: Cell signalling: Ligand binding initiates a signalling cascade that results in a cellular response.

1.2.4 Ligand binding assays

A ligand-binding assay is an experiment that assesses or measures ligand binding, and is used to characterise ligand-receptor interactions. There is a range of types of ligand binding assays, but most common methods are based on radioligand binding or fluorescence detection [107, 69]. Radioligand binding requires the ligand molecules to be radiolabelled and then radioactive decay is observed. Fluorescent detection techniques utilise molecules that increase in fluorescence when in close proximity to a receptor. However, all methods are based on being able to discern bound molecules from unbound ones. Ligand binding experiments can be categorised as either saturation, competition or kinetic binding experiments [66, 57].

1.2.5 Saturation binding assays and dose-response curves

Saturation binding experiments measure the amount of ligand bound once equilibrium (or steady-state) is reached, for a range of ligand concentrations. The primary aim of these experiments is to estimate the total number of receptors as well as the equilibrium binding rate of the ligand, that is, to determine the affinity of the ligand for the receptor [66]. A common way to analyse this data is a Scatchard plot, where 'bound ligand' is plotted against 'bound ligand divided by the number of free receptors'. However, this method has limitations. As the concentration of bound ligand appears on both the x and y axis, the effect of any experimental errors is compounded. Furthermore, nonlinear plots are difficult to interpret. For these reasons we do not use Scatchard plots in our analysis, however more information on these can be found in [69].

An alternative way to analyse the data is a dose-response curve, (as can be seen in Figure 1.2.2) that is, a plot relating the ligand concentration to the amount of ligand bound, or output response of a cell. We note that, we use term dose-response to also encompass concentration-response. While these curves can be used to study the affinity of a ligand, they are also used in experiments where a cell's response is measured, as opposed to ligand bound. In this case, it is a ligand's efficacy, that is, a ligand's ability to activate receptors and produce a response, and potency that is analysed, as opposed to

affinity [107]. A common measurement that is considered in dose-response curves is the EC_{50}, that is defined as the concentration of a drug that gives half-maximal response.

1.2.6 Agonists and antagonists

Ligands are classified as agonists, antagonists (also referred to as neutral agonist) or inverse agonists, depending on how they affect the activation of a receptor and therefore the resulting response [64]. We can see how each of these affect the dose-response relationship in Figure 1.2.2. Once bound an agonist increases the chances of the receptor activating, and therefore, as the ligand concentration increases so does the response, until the maximal response is reached. Agonists can be either full or partial agonists, with a full agonist able to increase the response to 100% of the maximum possible response, whereas a partial agonist can only achieve a partial response, regardless of the amount of ligand present. Antagonists bind a receptor but have no effect on the response [66, 64]. These are most commonly used in competition binding to block or reduce the activity of an agonist. Thus if we consider the dose-response relationship of an agonist with an antagonist present, it will require a higher concentration of agonist to reach maximal effect than if the antagonist wasn't present. Many receptors exist in a state of spontaneous equilibrium; a number of receptors within a cell will activate without any outside factors affecting them [78]. GPCRs, in particular, are known to behave in this way. An inverse agonist binds a receptor and forces it to become inactive, thus reducing the constitutive activity, and as such the basal receptor-induced response from the cell.

1.2.7 Competition binding assays

Competition binding experiments (sometimes referred to as displacement binding experiments) measure the binding of a single concentration of a labelled ligand whilst in the presence of a second, unlabelled ligand [95]. Ligands bind to, and therefore compete for, the same domain on the receptors. Usually, the experiment is performed for a range of concentrations of the unlabelled ligand, and measurements are taken once equilibrium is reached. For each of the unlabelled ligand concentrations, the observation of interest is, how much the binding of the labelled ligand is displaced or otherwise affected [66].

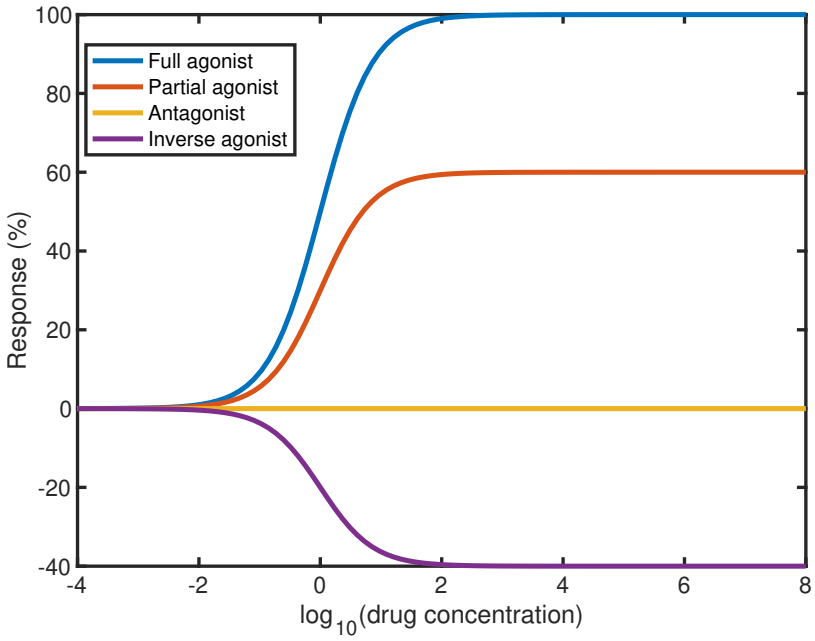

Figure 1.2.2: Drugs can be classified depending on how they affect the cell's response. Agonists increase the response and can be full or partial, antagonists produce no response and inverse agonists reduce the response. Response is considered as a percentage of the maximal response.

For each of these experiments, the IC_{50} is determined. This is the amount of unlabelled ligand that inhibits the response by 50%. This is the equivalent to the EC_{50}, where the IC_{50} relates to an antagonist ligand as opposed to an agonist.

Competition binding assays have a number of significant uses. Firstly, they are a fast and efficient way to determine whether a ligand binds to a receptor. Also, results from direct binding assays can be validated via competition binding. When competing with a ligand where the potencies are known, this can validate that the correct receptor has been identified. Finally, they are used to investigate the interactions of a ligand that has a low affinity for the receptor, as these can be difficult to observe in direct binding assays [86].

1.2.8 Kinetic binding assays

New technological developments, specifically fluorescence-based methods, have allowed for real-time observations of ligand binding in living cells [70]. This has lead to a rise

in kinetic binding studies, as more time-dependent data is available. These studies are advantageous when compared to saturation and competition assays as they provide information about the rates at which a ligand binds to and dissociates from a receptor individually, as opposed to only equilibrium values. This gives a more detailed view of the ligand-receptor interactions and can aid in determining ligand safety levels. However, kinetic binding methods can be more costly than saturation methods [51].

There are two general types of kinetic binding assays: association experiments and dissociation (or washout) experiments. In an association binding experiment, all receptors on the cell begin in an unbound form and binding is observed once the ligand is added. This is useful, not only for determining the rate of association but also how long it takes for equilibrium to be reached. Dissociation binding experiments are often performed following association experiments, that is once the system is in equilibrium. At this point the free ligand is removed by repeated washing, ensuring that no further ligand associates with the receptors and the rate of decrease of effect is then observed. An example time course of an association followed by dissociation curve can be seen in Figure 1.2.3. The total number of receptors can be ascertained from dissociation experiments, as well as the rate of dissociation [120].

1.3 G protein-coupled receptors

GPCRs are one of the largest families of cell surface receptors and one of the most diverse. They mediate most of our cellular responses to hormones, neurotransmitters and environmental stimulants, and are targets for up to approximately 50% of current ligands [14]. GPCRs can be classified into four main families, rhodopsin (family A), secretin (family B), glutamate (family C), adhesion and Frizzled/Taste2, although all have a similar structure [126, 122]. This structure consists of an extracellular N-terminus, an intracellular C-terminus and a middle section containing seven membrane-spanning helices that pass through the cell membrane seven times in a serpentine pattern, with three loops extending into the extracellular region [72]. This has led to them also being referred to as seven-transmembrane domain receptors (7TM) receptors. This structure

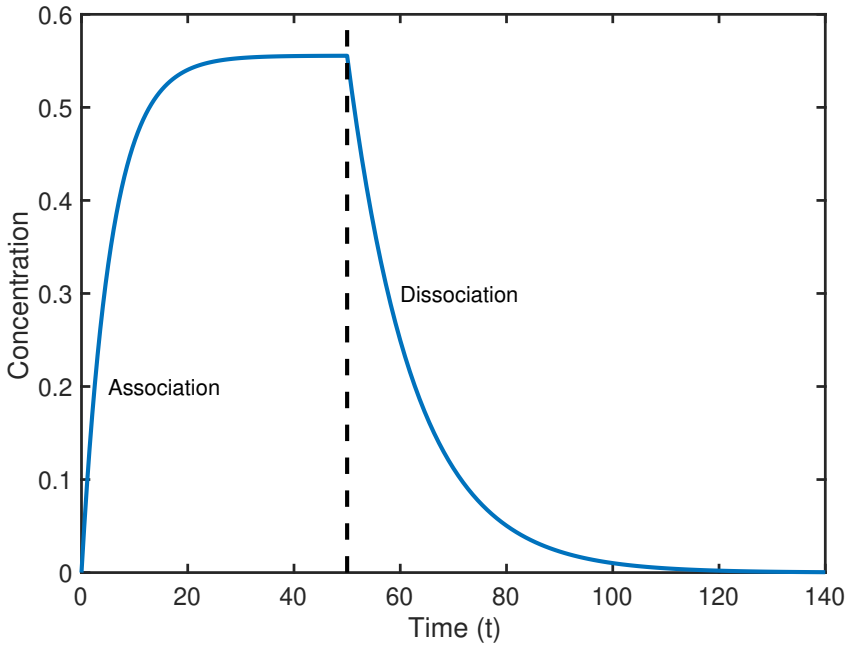

Figure 1.2.3: Ligand dissociation is observed by removing the ligand as it dissociates from receptors.

can be seen in Figure 1.3.1. Ligand binding can occur on either the N-terminus or on any of the extracellular loops [126]. Upon binding of an agonist, a conformational change in the 7TM domain is initiated, activating the C-terminus which allows for the binding of a G-protein.

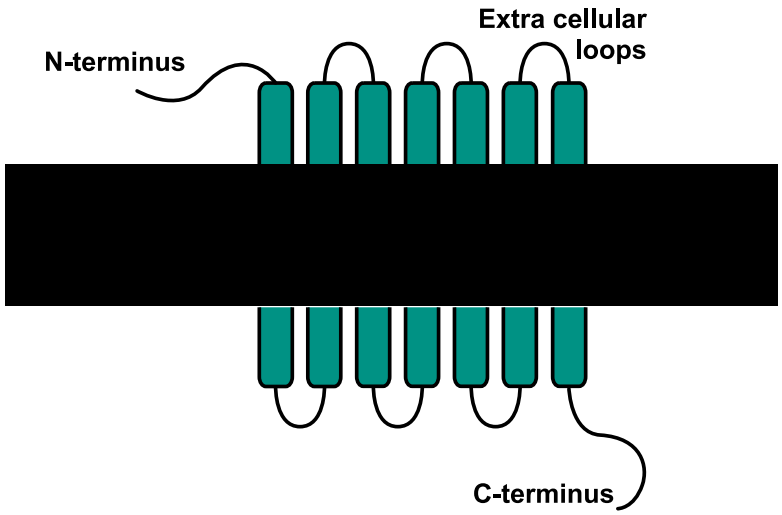

Figure 1.3.1: The GPCR is made up of a N-terminus, C-terminus and 7 membrane spanning domains [72].

1.3.1 G-proteins

G-proteins are heterotrimeric proteins, in that they consist of three subunits, namely α, β and γ. These are each encoded with a unique gene. There are at least 20 types of α subunits, and can be divided into four classes: G_s, G_q, G_i and G_t with each stimulating or inhibiting different enzymes [138]. Whilst in an inactive state a molecule of guanine diphosphate (GDP) is bound to the α subunit. When the GPCR becomes ligand bound and, therefore, activated, the G-protein is able to bind to the receptor creating a temporary complex. This causes the GDP molecule to dissociate from the G-protein, leaving the G-protein in an 'empty' state. In this form either a GDP molecule can rebind, or instead a molecule of guanosine triphosphate (GTP) is able to bind [122].

The binding of GTP forces the G-protein to dissociate from the receptor and a further conformational change is initiated. Upon dissociation, the G-protein splits into two parts, with the α subunit breaking away from the G-protein. The β and γ subunits remain bound together. The G-protein is now said to be active and the two parts are now referred to as G_α and $G_{\beta\gamma}$ [52]. Once in this form, the G_α and $G_{\beta\gamma}$ units act as effector molecules causing the release of second messengers. Deactivation occurs once the G_α unit hydrolyses the GTP molecule, that is, one phosphate molecule is removed, back into a GDP molecule. This allows for the reassociation of the G_α and $G_{\beta\gamma}$. This G-protein cycle can be seen in Figure 1.3.2 and is also explained in [138] and [137].

1.3.2 Signalling pathways

There are two primary pathways that are activated by G-proteins, namely the cyclic-adenosine monophosphate (cAMP) pathway and the phosphatidylinositol signalling pathway [52, 122]. Both G_α and $G_{\beta\gamma}$ subunits activate second messengers and subsequently, these pathways. However, while the G_α pathways have been well researched, the signal elicited from the $G_{\beta\gamma}$ unit has been considered less important, hence, there is still much unknown about their role in GPCR signalling [52]. G_α subclasses $G_{\alpha(s)}$ and $G_{\alpha(i)}$ regulate production the second messenger of cyclic-adenosine monophosphate (cAMP), with $G_{\alpha(s)}$ stimulating and $G_{\alpha(i/o)}$ inhibiting production of cAMP. Whereas subtypes $G_{\alpha(o)}$ and $G_{\alpha(q)}$

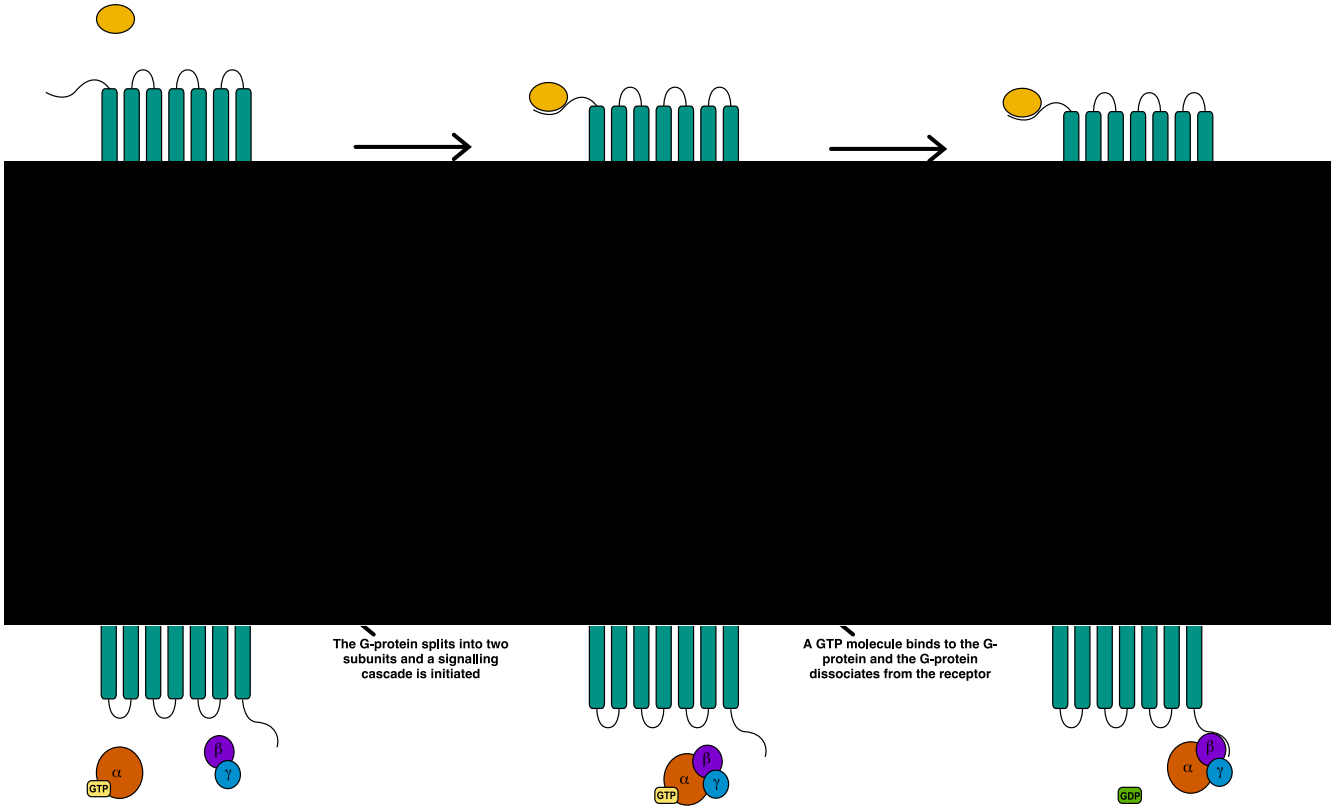

Figure 1.3.2: The G protein cycle [138].

activate phosphoinositol phospholipase C enzyme (PLC) [52].

1.3.3 Dimerised receptors

It has recently become accepted that GPCRs may exist as dimers or higher oligomers [91, 92]. An oligomer is formed by the joining of multiple receptors by molecular bonds, creating a single unit. A dimer is a special case of an oligomer consisting of two receptors. Single receptors are often referred to as monomers. A dimer can be made up of two identical monomers, then known as a homodimer, or similar but not identical monomers, creating a heterodimer. Theoretical models for GPCR binding and signalling are based on assumptions that receptors are all monomeric [138, 14]. When one site of the dimer becomes bound, this has an effect on the binding characteristics of the other sites. This is termed *cooperativity*, and the effect of cooperativity and cross-talk within the dimer is something that is yet to be fully explored. Understanding cooperativity and the resulting effects on ligand binding and signalling, as well as how these can be exploited for ligand development purposes, is one of the the main aims for this thesis.

1.4 Vascular endothelial growth factors

The vascular endothelial growth factor (VEGF) family consists of five members, VEGF-A, -B, -C, -D and placenta growth factor (PIGF), and is a sub-family of growth factors (GFs). Similar to most GFs, VEGF is a multivalent ligand, in that each molecule possesses more than one binding site, therefore, it can bind with as many receptors as there are sites [3]. Many aspects of cellular function, including survival, proliferation, migration and differentiation are regulated by VEGF [47, 99, 119]. VEGF binds to three VEGF receptors (VEGFRs), namely VEGFR-1, -2 and -3. These are receptor tyrosine kinases that are expressed predominately on endothelial cells and regulated by the VEGF ligand [99]. Like most other receptor tyrosine kinases (RTK), activation occurs upon dimerisation [80, 3, 119]. It is thought that VEGFRs diffuse across the cell membrane as monomeric receptors, although there is recent evidence suggesting that a percentage may exist as dimers [82] (also see references within). Binding of a ligand to the extracellular domain of the receptor triggers dimerisation with adjacent receptors, which leads to the receptors becoming activated, then leading to trans-autophosphorylation of the receptors. This provides a docking site for downstream signalling proteins, which results in the activation of signalling pathways, and ultimately a response [99, 82, 80].

1.5 Mathematical modelling

Quantitative systems pharmacology (QSP) is a novel discipline within biomedical research that uses mathematical and computational modelling techniques to address current problems in the discovery and the development of therapies. QSP studies aim to combine mechanistic modelling frameworks with clinical data to predict ligand efficacy and toxicity and give insights into ligand action [116]. In turn this can help guide further experiments to yield more meaningful results. The field of QSP has made many advances in the past decade, with even the United States Food and Drug Administration (FDA) accepting modelling and simulation results in evaluating ligand proposals [103, 104]. The term mathematical pharmacology was officially coined by Van der Graaf [125] to describe the recent emergence of the use of mathematical approaches to further understand phar-

macological processes, and can be seen as a sub-discipline of QSP. In this section we will review some key models, but first we present some of the underlying mathematical techniques used in ligand modelling.

1.5.1 The law of mass action

The law of mass action was was first proposed by Guldberg and Waage over the period 1864-79 [79] and is the basis of all ligand- receptor modelling [69], although it was initially derived to model only chemical reactions. Given two chemicals, the law of mass action describes the rate at which these collide to form a new chemical complex. For example, given chemicals A and B, with

$$A + B \xrightarrow{k} AB, \tag{1.1}$$

then mass action, in its simplest form, states that the rate of the reaction is proportional to the product of the concentrations of A and B. The parameter k is the rate constant of the reaction. If we consider the reaction with respect to time we can say that the rate of change of each concentration is

$$\frac{d\left[AB\right]}{dt} = k\left[A\right]\left[B\right], \tag{1.2}$$

$$\frac{d\left[A\right]}{dt} = -k\left[A\right]\left[B\right], \tag{1.3}$$

$$\frac{d\left[B\right]}{dt} = -k\left[A\right]\left[B\right], \tag{1.4}$$

where the square brackets represent chemical concentrations. A concentration of ligand is the number of moles of solute per litre of solution has units molar (M). It is important to mention however, that the law of mass action is not a law in the sense that it holds in any situation, for example, when dealing with particularly high or low concentrations [65]. Reactions that do follow the law are described as elementary reactions.

The reaction in (1.1) was used to demonstrate the law of mass action, however, many reactions are reversible, thus we can extend this to

$$\mathrm{A} + \mathrm{B} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{AB}, \tag{1.5}$$

This leads to the system of differential equations becoming

$$\frac{d\,[AB]}{dt} = k_+\,[A]\,[B] - k_-[AB], \tag{1.6}$$
$$\frac{d\,[A]}{dt} = -k_+\,[A]\,[B] + k_-[AB], \tag{1.7}$$
$$\frac{d\,[B]}{dt} = -k_+\,[A]\,[B] + k_-[AB], \tag{1.8}$$

we can see that $[AB]$ increases by consuming molecules of A and B but now also decreases proportionally to the existing concentration of AB. Once the system reaches equilibrium, thermodynamic principles (see Section 1.5.4) state that the rate of the forward reaction is equal to the rate of the backward reaction. Hence at equilibrium

$$k_+\,[A]\,[B] = k_-\,[AB]\,, \tag{1.9}$$

which can be rearranged to

$$K_D = \frac{k_-}{k_+} = \frac{[A]_{eq}\,[B]_{eq}}{[AB]_{eq}}, \tag{1.10}$$

where K_D is generally called the equilibrium dissociation constant of the reaction. So far the reactions have only accounted for single molecules of each reactant. Suppose instead

the reaction is

$$2\mathrm{A} + \mathrm{B} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{AB}, \tag{1.11}$$

In the forward reaction, two molecules of A are used, as opposed to one, to create the complex AB. As the rate of reaction is proportional to the product of each concentration the equation for A becomes

$$\frac{d\left[A\right]}{dt} = -2k_+ \left[A\right]^2 \left[B\right] + k_- \left[AB\right]. \tag{1.12}$$

As every reaction uses two molecules of A but only one molecule of B to create a single AB then the change in both B and AB is half that of the change in A, that is

$$\frac{d\left[A\right]}{dt} = \frac{1}{2}\frac{d\left[B\right]}{dt} = \frac{1}{2}\frac{d\left[AB\right]}{dt}. \tag{1.13}$$

1.5.2 The Langmuir isotherm, occupancy theory and the Hill equation

The first model of ligand binding was based on 'The Langmuir adsorption isotherm' [73], published by Irving Langmuir, although originally this was designed to describe chemicals binding to metal surfaces [66]. These ideas were extended by Clark [29] to describe agonist-receptor binding, and are encapsulated by what is currently known as occupancy theory. Assuming a reaction, with a ligand molecule, A , and receptor, R, we have

$$\mathrm{A} + \mathrm{R} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{AR}, \tag{1.14}$$

giving an equilibrium dissociation constant as

$$K_D = \frac{[A][R]}{[AR]}, \tag{1.15}$$

where $K_D = k_-/k_+$. However, Langmuir proposed that we instead consider the amount of ligand bound as a fraction of the total number of receptors. Defining ρ_A as the fraction of bound receptors, we then have $1-\rho_A$ as the fraction of unbound, or available, receptors. Langmuir's isotherm model asserts that the amount of ligand binding is proportional to the fraction of available receptors, the forward reaction rate and the ligand concentration, that is

$$k_+[A](1 - \rho_A). \tag{1.16}$$

Similarly, the amount of ligand unbinding is proportional to the amount of ligand bound to receptors and the reverse reaction rate, that is

$$k_-\rho_A. \tag{1.17}$$

At equilibrium it is assumed that these are in balance, so set

$$k_+[A](1 - \rho_A) = k_-\rho_A. \tag{1.18}$$

Solving for ρ_A, we have

$$\rho_A = \frac{k_+[A]}{k_- + k_+[A]} = \frac{[A]}{K_D + [A]}, \tag{1.19}$$

as the fraction of bound receptors. This gives a basic measurement and quantification of a ligand's affinity [66]. The Hill-Langmuir equation is an extension of this. If we instead assume n ligand molecules simultaneously binding a single receptor, and so have the reaction

$$n\mathrm{A} + \mathrm{R} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{AR}_n, \tag{1.20}$$

then the fraction of occupied receptors is given as

$$\rho_A = \frac{[A]^n}{K_D^n + [A]^n}. \tag{1.21}$$

This n is known as the Hill coefficient [106].

In 1968, Wagner [130] proposed that there is a relationship between the concentration of ligand and the cellular response, and gave a 'new' Hill equation describing this as

$$\frac{E}{E_{max}} = \frac{[A]^n}{(EC_{50})^n + [A]^n} \tag{1.22}$$

where E is the effect (or response), E_{max} is the maximal response and EC_{50} is the ligand concentration that gives half maximal response. This Hill equation is most commonly used in pharmacology for fitting to equilibrium data. The n in this equation is still referred to as the Hill coefficient, however, where in equation (1.21) it had a mechanistic, biological representation, in this case the n is more empirical. With the equation in this form, the Hill coefficient is used to measure cooperativity, with $n > 1$ giving positive cooperativity, that is, once a receptor is bound by a single molecule the affinity for ligand binding on that receptor is increased. Conversely, with $n \leq 1$, the affinity for binding is decreased once a receptor has bound one ligand molecule. These are concepts we use throughout this thesis, though in a slightly different context. Further background on the different forms and origins of the Hill equations can be found in [43].

Following the work of Clark, the occupancy model was built on by a number of authors. Ariëns [5] noticed a limitation of Clark's model was that it described only binding and did not consider the response elicited from binding. To overcome this Ariëns introduced a new parameter representing *intrinsic activity* and proposed that a theoretical response should be composed of this new parameter as well as the affinity parameter [110].

Both the models given by Clark and Ariëns are based on the assumption that the percentage of the maximal response is equal to the percentage of bound receptors, that is, 50 percent of receptors being bound leads to 50 percent of the maximal response being elicited [110]. This was later shown by Stephenson [117], Furchgott [42] and Nickerson [97] to not always be true. Stephenson [117] proposed that ligands have varying capabilities to elicit a response. Hence, some ligands are able to elicit the maximal response with only a small proportion of receptors bound. This became known as *efficacy*. Furthermore, he introduced the idea of a *partial agonist*, whereby a ligand is only able to elicit a fraction of the maximal response, even at maximal binding.

There were many modifications of these models, with one of the most notable being by Furchgott [42] who proposed that efficacy is the product of the intrinsic efficacy parameter (as defined by Ariëns [5]) and the concentration of active receptors [110]. This was a particularly important development as it highlighted the dependence of the response on the properties of both the ligand and the tissue. Mathematically, it can be stated that the response from an agonist ligand, A, as

$$Response = f\left[\frac{[A]\epsilon R_{tot}}{K_D + [A]}\right]. \tag{1.23}$$

where f is a (usually hyperbolic) function relating receptor occupancy and response, ϵ represents intrinsic efficacy and R_{tot} is the total concentration of receptors [67, 69].

1.5.3 The operational model of agonism

The operational model of agonism was developed by Black and Leff [11] and is an extension to the occupancy models. It was designed to overcome the limitations of existing occupancy models, such as the reliance on the parameter ϵ which does not correlate to any chemical identity. To derive their model they assumed a simple reversible reaction between a agonist, A and receptor R, creating the complex AR, that follows the law of mass action. At equilibrium this gives

$$[AR] = \frac{[A]R_{tot}}{K_D + [A]}, \tag{1.24}$$

where K_D is the dissociation constant and R_{tot} is the total receptor concentration. This function is a hyperbolic function (in fact, is a rectangular hyperbolic function). Black and Leff assumed that the relationship between $[AR]$ and the response must also be hyperbolic, and so defined the observed effect, E, as

$$E = \frac{E_{max}[AR]}{K_E + [AR]}, \tag{1.25}$$

where E_{max} is the maximal effect and K_E is the value of $[AR]$ that elicits half maximal effect. Combining equations (1.24) and (1.25), we have

$$E = \frac{[A]R_{tot}E_{max}}{K_D K_E + (R_{tot} + K_E)[A]}, \tag{1.26}$$

that is, a hyperbolic curve describing the agonist-effect relationship. Defining the *transducer ratio*

$$\tau = \frac{R_{tot}}{K_E}, \tag{1.27}$$

as a constant that characterises the *operational efficacy* [12], that is, the propensity of the agonist to elicit a response [69], and substituting this in to equation (1.26) gives

$$E = \frac{[A]\tau E_{max}}{K_D + (\tau + 1)[A]}, \tag{1.28}$$

This is extended to include non-hyperbolic curves so as to encompass possible receptor cooperativity by the reintroduction of the Hill coefficient, as

$$E = \frac{[A]^n \tau^n E_{max}}{(K_D + [A])^n + \tau^n [A]^n}. \tag{1.29}$$

This model is canonical in receptor theory and is still widely accepted and used in much of the pharmacology community [67]. However, there are still areas requiring further exploration. While the operational model allows for an agonist that can elicit either a full or partial response, it cannot account for ligands that have no effect on the given signal, or ligands that decrease the signal below the basal level. Furthermore, it is useful only for equilibrium studies, although Hoare *et al* [54] make some steps towards extending it to include kinetics.

1.5.4 The two state model

In 1957, del Castillo and Katz [21], and later Katz and Thesleff [63], proposed the concept of receptor activation, although originally it was applied to only ligand-gated ion channels. They proposed that receptors undergo a conformational change and can exist in two states, a resting state and an active one [76], and the response is elicited only from receptors that are active.

In their first model, del Castillo and Katz proposed that receptors remain in an inactive form, in the absence of ligand, and activation only occurs upon ligand binding. A schematic of the model can be seen in Figure 1.5.1. The model assumes a ligand A, binds a free receptor R which creates the complex AR. Activation occurs subsequently

to binding, and so AR^* represents the receptor in a bound-activated state. Parameters K_A and K_{act} represent equilibrium binding and activation constants respectively.

$$A + R \underset{}{\overset{K_A}{\rightleftharpoons}} AR \overset{K_{act}}{\rightleftharpoons} AR^*$$

Figure 1.5.1: The first two-state model. A ligand A binds a receptor R creating the complex AR. Binding causes the receptor to activate, creating AR^* and the signal is altered [76].

This simple model was later modified by Monod, Wyman and Changeux [94] who suggested that receptors also undergo a conformational change and become activated in the absence of ligand. This modified model allows for a basal response, coming from constitutive activity of the receptor, in the absence of ligand, and assumes that ligand binding changes the ratio of active to inactive receptors and, therefore, the functional response. The schematic of this model is presented in Figure 1.5.2, where α is introduced to represent the intrinsic efficacy of the ligand.

$$\begin{array}{ccccc} A & + & \mathrm{R} & \overset{K_A}{\rightleftharpoons} & AR \\ & & \Big\uparrow\Big\downarrow K_{act} & & \Big\uparrow\Big\downarrow \alpha K_A \\ A & + & R^* & \overset{\alpha K_{act}}{\rightleftharpoons} & AR^* \end{array}$$

Figure 1.5.2: The modified two-state model [94]. A modification of the original two-state model, a receptor can now also activate in the absence of ligand, represented by R^*. Cooperativity factor α represents intrinsic efficacy of the ligand.

The advantage of this model is that it allows for a ligand to act as an antagonist or inverse agonist, as well as a full or partial agonist. While an agonist will more likely bind and stabilise an active receptor, an inverse agonist would act preferentially to an inactive receptor. An antagonist would show no preference to either conformation [56].

Detailed balance

The modified two-state model also includes the introduction of the new parameter α, which multiplies the binding/activation rates of the second reactions. This parameter must appear in both secondary reaction due to a concept called detailed balance, that was introduced by Lewis [77] in 1925. Based upon the assumption that each elementary reaction is reversible and should respect thermodynamics[48], detailed balance states that:

"Corresponding to every individual process there is a reverse process, and in a state of equilibrium the average rate of every process is equal to the average rate of its reverse process" [77]

In the model given in Figure 1.5.2, the parameter α represents the intrinsic efficacy of the ligand and accounts for the change in propensity for receptor activation when ligand is bound as well as change in affinity of the ligand for active receptors over inactive receptors.

1.5.5 The ternary complex model

The first attempt to incorporate the G protein into receptor theory models was the ternary complex model. developed by De Lean, Stadel and Lefkowitz[31] in 1980. They proposed two models to describe the interactions between ligand, receptor, and G-protein. The first assumes that the binding of a membrane component, such as a G protein, occurs sequentially to ligand binding (as we saw in Figure 1.3.2). The second, and more generally accepted, is outlined schematically in Figure 1.5.3 and allows for the spontaneous binding of a G protein in the absence of ligand. In this model, a ligand A or G protein, can bind a free receptor R, creating the complexes AR and RG (with equilibrium binding rates K_A and K_G) respectively. Furthermore, ligands and G proteins can bind a receptor that is already bound by the alternate molecule, for example, a ligand can bind an already formed RG complex, thus creating ARG. A signal is then assumed to be given from any receptor with a coupled G protein.

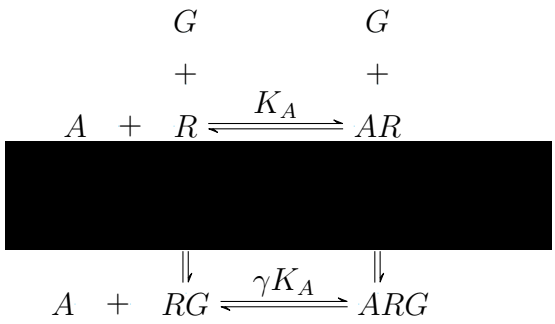

Figure 1.5.3: Schematic: The ternary complex model [31]. Ligands or G proteins can bind a free receptor, creating AR and RG, or both can bind, creating ARG. Cooperativity factor in this model is γ.

The model appears similar to the modified two-state model, with a signal being elicited from an active receptor in the two-state model, or a bound G protein complex in the ternary complex model. However, there is one key difference: in the two-state model the efficacy of an agonist depends receptor density, while in the ternary model, the amount of available G proteins also has a role in how sensitive the system is to an agonist [68].

1.5.6 The extended ternary complex model

The extended ternary complex model was developed in response to findings that showed that GPCRs were able to spontaneously activate, and thus, activate G proteins, without any ligand present [68]. The schematic of this model can be seen in Figure 1.5.4. This model essentially combines the original ternary complex model and the two-state model. In this model, a receptor can become active spontaneously, represented by R^*, or can be induced upon ligand binding, AR^*. Similarly, a G protein can bind any active receptor. It is assumed that a response is produced from active, G protein bound species, that is R^*G and AR^*G [69]. The equilibrium constant for G protein binding is K_G.

The cooperativity (or efficacy) parameters α and γ determine if a ligand acts as an agonist/inverse agonist. If $\alpha > 1$ and $\gamma > 1$, the ligand stabilises activation and G protein binding, thus the ligand has positive efficacy and is considered an agonist. Conversely, if $\alpha < 1$, the inactive receptor state is stabilised, or $\gamma < 1$, the affinity of the G protein for the receptor is reduced for the active receptor state, then the signal will be reduced from the basal level. The ligand is then showing negative efficacy and acts as an inverse

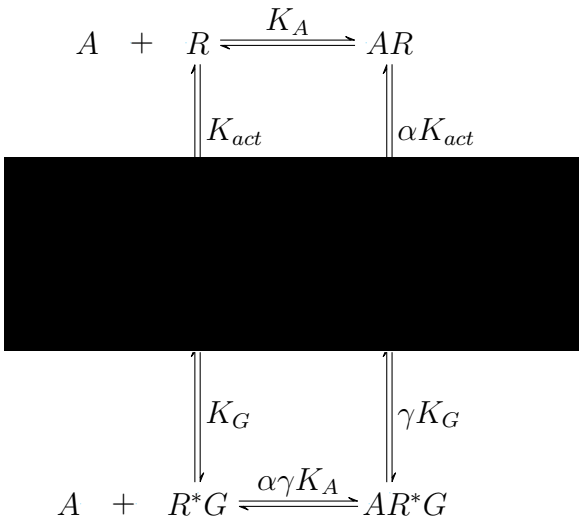

Figure 1.5.4: Schematic: The extended ternary complex model [68]. Receptors, in this model, can bind a ligand (AR), activate (R^* and AR^*), though G proteins can only bind activate receptors (R^*G and AR^*G). This model contans both cooperativity factors α (from the two-state model) and γ (from the first ternary complex model).

agonist. Although it is important to note that this only holds for systems that have a basal level of activity. While this model is an improvement in terms of representing the GPCR signalling mechanisms, the model is incomplete as interactions between inactive receptors and G proteins are not represented.

1.5.7 The cubic ternary complex model

The cubic ternary complex model is the thermodynamically complete version of the extended ternary complex model which allows for interaction between the G protein and inactive form of the receptor [133]. The model schematic can be seen in Figure 1.5.5. Equilibrium rates K_A, K_{act} and K_G determine the equilibrium ratios of ligand binding, receptor activation and G protein binding respectively. There are now four cooperativity factors. While α and γ are defined as in the previous section when detailing the extended ternary complex model, we also have the addition of two new cooperativity factors, β and δ. In summary the effect of these parameters is as follows:

- The parameter α is the effect of ligand binding on activation and vice versa.
- The parameter β represents the effect of receptor activation on the coupling of a G protein and vice versa.

– The parameter γ is the effect ligand binding has on the coupling of a G protein and vice versa.

– The parameter δ represents the effect of any two of ligand binding, receptor activation or G protein coupling, on the occurrence of the third.

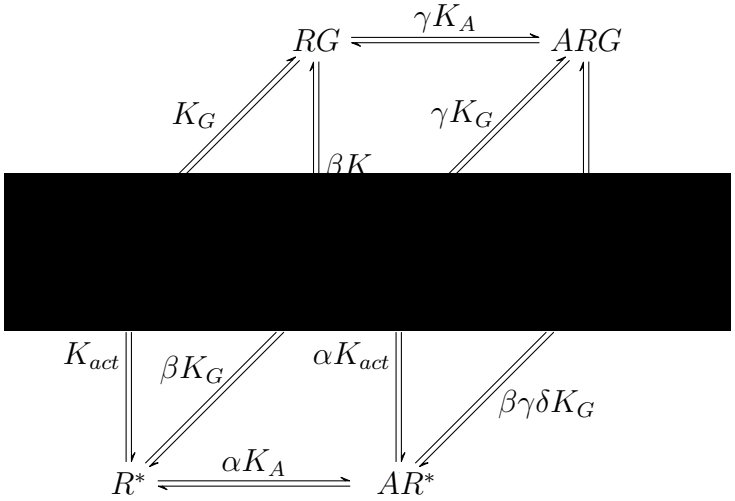

Figure 1.5.5: Schematic: The cubic ternary complex model [133]. All combinations of bound receptors, active receptors and receptors with G proteins attached are possible. Along with the two existing cooperativity factors α and γ, two new cooperativity factors, β and δ, are introduced.

These ternary complex models have formed the basis of much GPCR modelling work done in recent decades. In particular, they underpin the work by Woodroffe *et al* [138, 137] who have extended these models to include the full G protein cycle that was seen in Figure 1.3.2.

1.6 Mathematical methods

Throughout this thesis, we use a number of mathematical techniques, such as solving systems of ordinary differential equations (ODEs) and asymptotic analysis. In this section we will use relevant ligand binding models to explain the steps taken in these methods.

1.6.1 Solving a linear system of ODEs

While there are multiple methods to solve systems of ODEs, we use the method of undetermined coefficients. In order to detail the steps in this method we use the competition binding model presented by Motulsky and Mahan in [95], as this model also underlies some of our later work, in Chapter 2. In [95] the authors develop a model that represents a competition binding scenario with GPCRs. A labelled ligand, A, competes with an unlabelled ligand, B, for the same binding site on a receptor, denoted R. Association and dissociation rates for ligand A are k_{a+} and k_{a-} respectively, and k_{b+} and k_{b-} for ligand B. The schematic for the model is given in Figure 1.6.1.

$$A + R \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} AR, \qquad B + R \underset{k_{b-}}{\overset{k_{b+}}{\rightleftharpoons}} BR,$$

Figure 1.6.1: Schematic for a competition binding scenario. Labelled ligand, A, competes with unlabelled ligand, B, for receptor R.

The law of mass action gives the system of ODEs that govern the dynamics as

$$\frac{d[R]}{dt} = -(k_{a+}[A] + k_{b+}[B])[R] + k_{a-}[AR] + k_{b-}[BR], \tag{1.30a}$$

$$\frac{d[AR]}{dt} = k_{a+}[A][R] - k_{a-}[AR], \tag{1.30b}$$

$$\frac{d[BR]}{dt} = k_{b+}[B][R] - k_{b-}[BR], \tag{1.30c}$$

which, together with the initial conditions

$$[R](0) = R_{tot}, \qquad [AR](0) = 0, \qquad [BR](0) = 0, \tag{1.30d}$$

where R_{tot} is the total receptor concentration, form the initial value problem describing the kinetics of the system. We notice that

$$[R]' + [AR]' + [BR]' = 0, \tag{1.31}$$

hence, the total number of receptors, that is R_{tot}, must be a conserved quantity, and so we have

$$[R] + [AR] + [BR] = R_{tot}. \tag{1.32}$$

This can be rearranged to give an expression for $[R]$, which we substitute in to the system of ODEs, leaving the reduced system of ODEs, in the form $\mathbf{X}' = M\mathbf{X} + \mathbf{f}$, as

$$M = \begin{bmatrix} -(k_{a+}[A] + k_{a-}) & -k_{a+}[A] \\ -k_{b+}[B] & -(k_{a+}[B] + k_{b-}) \end{bmatrix}, \qquad \mathbf{f} = \begin{bmatrix} k_{a+}[A] \\ k_{b+}[B] \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} [AR] \\ [BR] \end{bmatrix}. \tag{1.33}$$

To construct a solution to this system, we first calculate the eigenvalues of M, which are

$$\lambda_{1,2} = \frac{1}{2}\left(\mathrm{Tr}(M) \pm \sqrt{(\mathrm{Tr}(M))^2 - 4\det(M)}\right), \tag{1.34}$$

where $\mathrm{Tr}(M)$ is the trace of M and $\det(M)$ is the determinant of M. We note that $\det(M) > 0$ and $\mathrm{Tr}(M) < 0$, hence we must have two real, negative eigenvalues. As such, a solution to the system will be of the form

$$\mathbf{X}(t) = c_1\mathbf{v}_1 e^{\lambda_1 t} + c_2\mathbf{v}_2 e^{\lambda_2 t} + \mathbf{X}_p(t), \tag{1.35}$$

where $\mathbf{v}_1, \mathbf{v}_2$ are the eigenvectors corresponding to eigenvalues $\lambda_{1,2}$. The coefficients c_1, c_2 are arbitrary constants which account for the initial conditions, and $\mathbf{X}_p(t)$ is a particular solution that shows the equilibrium concentrations. In this case we have

$$\mathbf{v}_1 = \begin{bmatrix} k_{b+}[B] + k_{b-} + \lambda_1 \\ k_{b+}[B] \end{bmatrix}, \qquad \mathbf{v}_2 = \begin{bmatrix} k_{b+}[B] + k_{b-} + \lambda_2 \\ k_{b+}[B] \end{bmatrix}. \tag{1.36}$$

We calculate $\mathbf{X}_p(t)$ by setting $M\mathbf{X} + \mathbf{f} = \mathbf{0}$ and solving for $\mathbf{X}$. This gives

$$\mathbf{X}_p = \frac{R_{tot}}{\det(M)} \begin{bmatrix} k_{a+}k_{b-}[A] \\ k_{a-}k_{b+}[B] \end{bmatrix}. \tag{1.37}$$

Using the initial conditions, we calculate the constants c_1 and c_2 as

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{R_{tot}}{\det(M)(\lambda_1 - \lambda_2)} \begin{bmatrix} k_{a-}k_{b+}[B] - k_{a+}k_{b-}[A] + k_{a-}k_{b-} + k_{a-}\lambda_2 \\ k_{a+}k_{b-}[A] - k_{a-}k_{b+}[B] - k_{a-}k_{b-} - k_{a-}\lambda_1 \end{bmatrix}. \tag{1.38}$$

The expressions for c_1 and c_2 can then be substituted back into (1.35), to give analytical solutions for the concentration of labelled and unlabelled ligands. Solutions can then be used to plot species curves as well as total ligand bound, under various parameter regimes. This can give insights into possible system behaviours.

1.6.2 Perturbation analysis

In Chapter 4, we use perturbation analysis (or asymptotic analysis) to analyse a system of ODEs with different parameter values. Perturbation analysis is useful when analysing the complexities in the dynamics of biological systems. It allows for asymptotic solutions of nonlinear problems dependent on a small parameter and is useful for problems requiring a separation of timescales. Perturbation analysis is particularly advantageous over other methods as it provides insights into the biological system as opposed to relying on observations. It is also used to analyse models for GPCR signalling in [138] and [137].

To illustrate the method we use the canonical model, first published by Michaelis and Menten [87, 60] in 1913, of the interactions of enzymes and substrates. The reactions in this model are known as enzymatic catalytic, where it is assumed that the concentration of substrate is greater than the concentration of enzyme. A more detailed analysis is given in [96]. The schematic of the model is given as

In these reactions a molecule of the enzyme E reacts with a substrate S to create

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E$$

the complex ES. This complex then dissociates into a product P and enzyme E. It is generally assumed that there are very low concentrations of E and so the rate at which the product can be created is limited, and the product P is removed by the body continually by some other process. Applying the law of mass action and defining $s = [S]$, $e = [E]$, $c = [ES]$ and $p = [P]$, gives the system of differential equations governing the dynamics as

$$\frac{ds}{dt} = -k_1 es + k_{-1} c, \tag{1.39a}$$
$$\frac{de}{dt} = -k_1 es + (k_{-1} + k_2) c, \tag{1.39b}$$
$$\frac{dc}{dt} = k_1 es - (k_{-1} + k_2) c, \tag{1.39c}$$
$$\frac{dp}{dt} = k_2 c, \tag{1.39d}$$

with initial conditions initial conditions

$$s(0) = s_0,\ e(0) = e_0,\ c(0) = 0,\ p(0) = 0. \tag{1.39e}$$

From these we can see that

$$\frac{de}{dt} + \frac{dc}{dt} = 0, \tag{1.40}$$

and so we can conclude that the concentration of enzymes is conserved, hence we have

$$e + c = e_0, \tag{1.41}$$

which can be used to reduce the system by one equation. Also, as p does not appear

explicitly in any of the equations, we can solve this individually once we have a solution for c. This leaves the two equations to solve, as

$$\frac{ds}{dt} = -k_1 e_0 s + (k_1 s + k_{-1})c, \tag{1.42a}$$

$$\frac{dc}{dt} = k_1 e_0 s - (k_1 s + k_{-1} + k_2)c, \tag{1.42b}$$

$$s(0) = s_0,\ c(0) = 0. \tag{1.42c}$$

In order to analyse the system, we first nondimensionalise the system by setting

$$\tau = k_1 e_0 t, \quad \lambda = \frac{k_2}{k_1 s_0}.\ \ \epsilon = \frac{e_0}{s_0}, \quad \kappa = \frac{k_{-1} + k_2}{k_1 s_0},$$

$$u(\tau) = \frac{s(t)}{s_0}, \quad v(\tau) = \frac{c(t)}{e_0}.$$

This gives the dimensionless system as

$$\frac{du}{d\tau} = -u + (u + \kappa - \lambda)v, \tag{1.43a}$$

$$\epsilon\frac{dv}{d\tau} = u - (u + \kappa)v, \tag{1.43b}$$

$$u(0) = 1,\ v(0) = 0, \tag{1.43c}$$

noting that $\kappa - \lambda > 0$ and $0 < \epsilon \ll 1$, hence ϵ is the required small parameter.

We first determine the type of perturbation problem we have. If having small, nonzero values of ϵ gives qualitatively the same system as having a zero ϵ, we have a *regular perturbation* problem. However, if this is not the case, and the problem cannot be approximated by setting $\epsilon = 0$, then we say we have a *singular perturbation* problem. Typically, this

is when the small parameter is multiplying the highest derivative. Singular perturbation problems usually indicate that a problem is evolving on multiple timescales.

The general method of perturbation analysis involves representing each variable by a power series expansion about the small parameter, so for a variable $x = (x_1...x_n)$ and small parameter ϵ, the approximation would be

$$x \approx \sum_{n=0}^{\infty} \epsilon^n x_n. \tag{1.44}$$

In equations (1.43), we have ϵ multiplying the highest derivative in the second equation and so we have a singular perturbation problem. This can also be seen in Figure 1.6.2 where we plot each concentration on a log-log scale to highlight the dynamics in the different timescales. In this, we see how we have growth in the concentration of v on a much lower timescale than the concentration of u curves. For singular perturbation prob-

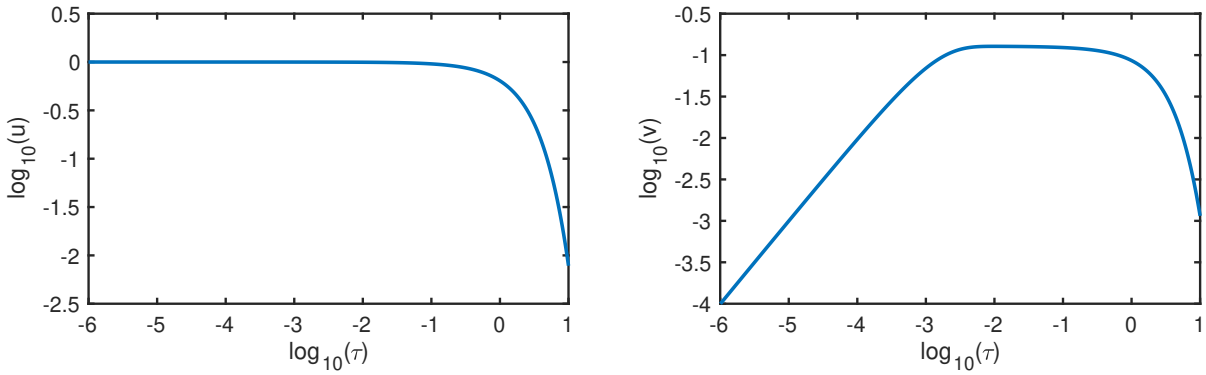

Figure 1.6.2: Concentration of u and v (equations (1.43)) are plotted on a log-log scale to highlight the time course dynamics on the different timescales. Parameters used for the plot are $\lambda = 3.4$, $\kappa = 6.7$, $\epsilon = 1e-2$.

lems such as this we seek solutions to the two timescales individually, these are known as the inner (or boundary) solution, covering the time scale where changes are rapid, and outer solution for the slower part of the dynamics.

Outer Solution

We first consider what happens in the outer region. We make use of big O notation to describe the timescale that the curve is evolving on. Big O notation describes the growth rate of a variable or function. For this outer solution, we have time increasing at $O(1)$. We first assume asymptotic approximations

$$u \approx u_0 + \epsilon u_1 + \epsilon^2 u_2 + ..., \qquad v \approx v_0 + \epsilon v_1 + \epsilon^2 v_2 + ..., \tag{1.45}$$

where $u_0 \gg \epsilon u_1 \gg \epsilon^2 u_2 \gg ...$ and $v_0 \gg \epsilon v_1 \gg \epsilon^2 v_2 \gg ....$ Substituting these into equations (1.43) and collecting the leading order terms (that is the largest terms), gives

$$\frac{du_0}{d\tau} = -u_0 + (u_0 + \kappa - \lambda)v_0, \tag{1.46a}$$

$$0 = u_0 - (u_0 + \kappa)v_0. \tag{1.46b}$$

The initial conditions for the outer system are unknown and will be determined later by matching to the inner solution. The second of these equations gives the relation

$$v_0 = \frac{u_0}{u_0 + \kappa}, \tag{1.47}$$

which in turn gives

$$\frac{du_0}{d\tau} = -\frac{\lambda u_0}{u_0 + \kappa}, \tag{1.48}$$

which solves to give the implicit solution for $u_0(\tau)$ as

$$u_0 + \kappa \log(u_0) = c_1 + \lambda\tau, \tag{1.49}$$

for some constant c_1. The relation in equation (1.47) then gives the impicit solution for

$v_0(\tau)$ as

$$\frac{\kappa v_0}{1-v_0} + \kappa \log\left(\frac{\kappa v_0}{1-v_0}\right) = c_1 + \lambda\tau. \tag{1.50}$$

Inner Solution

As neglecting the term $\epsilon dv/d\tau$ in the previous solution means that the initial condition $v(0) = 0$ cannot be satisfied by the outer solution so we rescale variables

$$\tau = \epsilon T, \qquad U(T) = u(\tau), \qquad V(T) = v(\tau), \tag{1.51}$$

to magnify the neighbourhood around $\tau = 0$. Substituting these into equation (1.43), we have the system of equations governing the dynamics in this inner region as

$$\frac{dU}{dT} = \epsilon(-U + (U + \kappa - \lambda)V), \tag{1.52a}$$

$$\frac{dV}{dT} = U - (U + \kappa)V, \tag{1.52b}$$

$$U(0) = 1,\ V(0) = 0. \tag{1.52c}$$

Assuming asymptotic approximations

$$U \approx U_0 + \epsilon U_1 + \epsilon^2 U_2 + ..., \qquad V \approx V_0 + \epsilon V_1 + \epsilon^2 V_2 + ..., \tag{1.53}$$

where $U_0 \gg \epsilon U_1 \gg \epsilon^2 U_2 \gg ...$ and $V_0 \gg \epsilon V_1 \gg \epsilon^2 V_2 \gg ...$, gives a leading order problem

of

$$\frac{dU_0}{dT} = 0, \tag{1.54a}$$

$$\frac{dV_0}{dT} = U_0 - (U_0 + \kappa)V_0, \tag{1.54b}$$

$$U_0(0) = 1,\ V_0(0) = 0. \tag{1.54c}$$

which solves to give

$$U_0(T) = 1, \qquad V_0(T) = \frac{1 - e^{-(\kappa+1)t}}{\kappa + 1}. \tag{1.55}$$

This clearly demonstrates how only the concentration of v is changing on this timescale.

Matching

In order to determine the constant, c_1 that appears in the outer region solutions (equations (1.49) and (1.50)) we use matching, where we require the inner and outer solutions to match in some intermediate timescale. As the inner solution is only valid whilst time is $O(\epsilon)$, and the outer solution valid for time $O(1)$, there must exist some intermediate region, when $\epsilon \ll \tau \ll 1$, that both of these hold. That is, we require

$$\lim_{\tau \to 0} u(\tau) = \lim_{T \to \infty} U(T) \tag{1.56}$$

$$\lim_{\tau \to 0} v(\tau) = \lim_{T \to \infty} V(T). \tag{1.57}$$

From the first condition, we take the limits of equations (1.49) and (1.55), giving

$$c_1 = 1, \tag{1.58}$$

whilst calculation of each of the limits in the second condition (equations (1.50) and (1.55)) confirms that this also matches. In Figure 1.6.3 we plot the approximations in both regions against the numerical solution, where the agreement between them is clear, hence we have good approximations for each region.

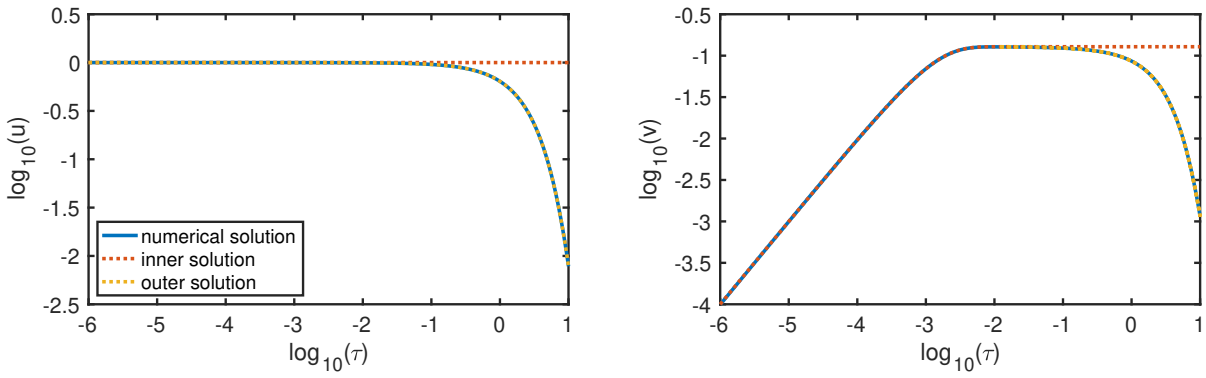

Figure 1.6.3: The numerical solution (to equations (1.42))is plotted together with the inner and outer approximated solutions. The approximations in each region show a good approximation to the curves. Parameters used for the plot are $\lambda = 3.4$, $\kappa = 6.7$, $\epsilon = 1e{-}2$.

1.6.3 Structural identifiability analysis

We previously discussed how ligand binding assays are used to observe ligand-receptor interactions. Mathematical models of pharmacodynamics systems have become key in understanding these interactions, between ligands and living cells, and as such play a significant role in the development of new therapeutic medicines. These models are often comprised of ODEs that have many mechanistic parameters that represent biological processes and are largely unknown [28]. An essential step in using these models requires establishing the values of these parameters [129] by fitting to experimental data from ligand binding assays. This helps to quantify ligand-receptor interactions. Parameter estimation for biological systems often involves global search or optimisation algorithms [36]. However, these fitting routines can result in inaccurate and unreliable estimates [89, 36].

Identifiability analysis is the process of assessing whether it is theoretically possible to estimate a set of parameters from experimental observations and the dynamic equations [46, 129]. This is therefore needed to determine the reliability of these estimates.

In particular, structural identifiability analysis (SIA) uses the model structure, together with observed outputs, to determine whether parameters can be returned successfully, given perfect, noiseless and bias-free observations [128, 8].

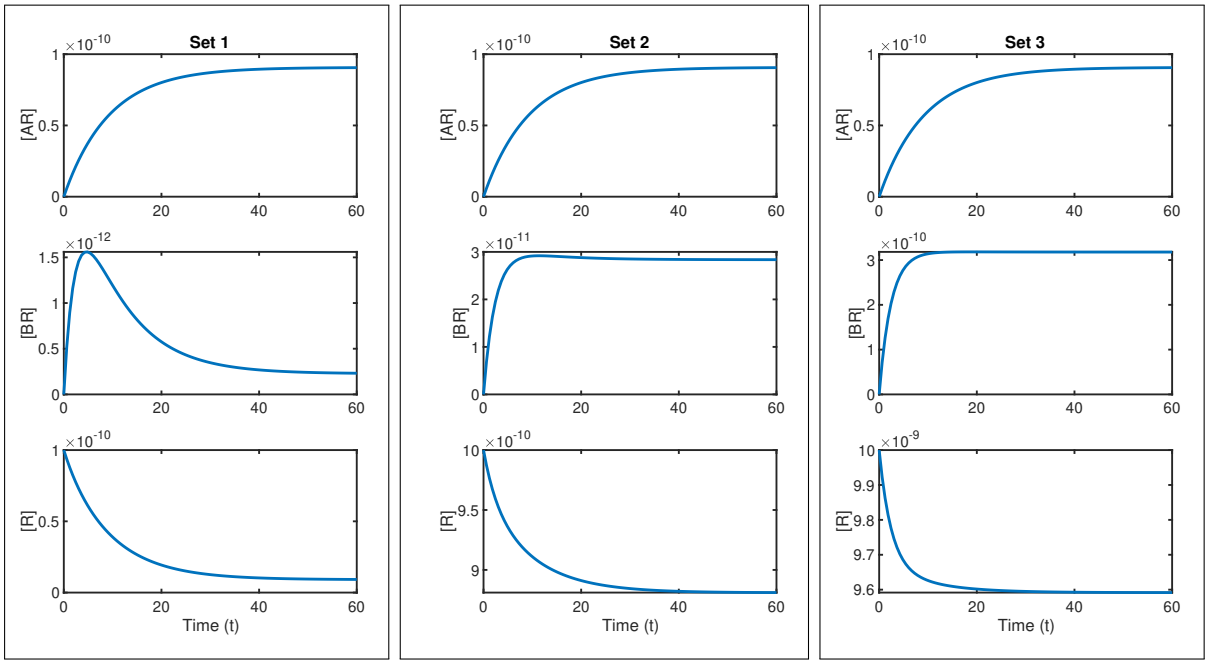

Figure 1.6.4: Plotting the concentration curves for all species individually, for a competition binding with monomeric receptor model, using three different parameter sets. Although all curves have the same measured output (concentration of $[AR]$) curve, non-identifiability can be seen from the differences in the different species curves.

An example of a system with non-identifiable parameters can be seen in Figure 1.6.4, where we plot time courses for the competition binding model that we described in Section 1.6.1. We assume that the only measured quantity is the amount of labelled ligand $[A]$ that is bound, that is the concentration $[AR]$. However, when plotting the solutions for all concentrations, using three different parameter sets, we see how these all result in the same measured output curve, even though the concentrations of $[R]$ and $[BR]$ evolve very differently. In a case such as this, it is then not possible to determine the model parameters from the measured output. We note that the parameter values for this plot along will the full analysis can be found in Chapter 5.

1.7 Thesis overview

In this thesis, we develop mathematical models, first for ligand binding and later include receptor activation, where receptors have the ability to form dimers. We follow a framework of analysis for each model, where we first look at equilibrium solutions, followed by then exploring time-dependent behaviours. The main goal for this thesis is to contribute to existing receptor theory by giving qualitative and quantitative information for ligand-receptor interactions, and provide models that can be of practical use to pharmacologists. The remainder of the thesis is laid out as follows.

In Chapter 2 we present three models for ligand binding to pre-dimerised GPCRs. In the first model we assume all dimers are homodimers and consider binding with a single ligand, focussing on the effect and quantification of cooperativity across the dimer. At equilibrium we give analytical solutions and find conditions on the model parameters that give biphasic dose-response curves (curves with multiple points of inflections). As the system is small and linear, we compute time-dependent analytical solutions and use these to understand possible time course dynamics. We then build upon this model, where we first introduce a second ligand, then also consider binding to a pre-dimerised heterodimer. In these, we again give analytical expressions for equilibrium solutions, but instead use numerical methods to examine time course behaviours under various parameter regimes.

In Chapter 3 we build on the single ligand, homodimer equilibrium model, by including receptor activation. As each receptor within the dimer has the ability to bind a ligand molecule and activate individually we treat each receptor as an individual entity, as opposed to a single dimer unit. This results in a much larger system, with more complex behaviours to explore and understand. At equilibrium, varying the different cooperativity factors leads to the emergence of some unusual signal curves, such as extra inflections and overshoots in the curve.

In Chapter 4 we look at the binding of VEGF to VEGFRs. This receptor system

is within the growth factor family of receptors and is known to dimerise in response to ligand binding. Including the dimerisation process leads to a nonlinear system, whereas our previous models have been linear. Numerical results for time course dynamics show peaks and curves evolving on multiple scales and we use asymptotic analysis to understand these. The model is validated by fitting to published data.

In Chapter 5 we study the identifiability of model parameters when fitting to data, using existing structural identifiability analysis techniques. We apply these to a range of classical ligand binding models as well as some of our existing dimer models. For each model, we apply relevant techniques to determine identifiability from a single set of time course data, in order to give a comparison of the methods. For non-identifiable models we give identifiable parameter combinations and explore ways to make all parameters, and therefore the whole system, identifiable. This includes using equilibrium data, data from washout experiments and multiple time courses to identify the parameters.

In Chapter 6, we summarise the work done throughout the thesis, discuss ongoing work, and suggest possible directions for future work.

Chapter 2

Binding Models for GPCR homodimers

We note that much of the work in this chapter is published by White and Bridge in [135]. Furthermore, some paragraphs are quoted verbatim from this paper.

Mathematical modelling of binding to dimerised GPCRs has largely focused on equilibrium binding models, such as those presented by Franco *et al* [40, 41, 39] and Durroux [33]. Typically, these equilibrium models yield analytical solutions for total bound labelled ligand, derived algebraically using mass action and receptor conservation considerations. Corresponding log dose-response (logDR) curves for bound ligand may be multiphasic, exhibiting multiple inflections [37, 19, 108]. For models of dimerised receptor binding, this departure from monophasic logDR curves (typical of monomeric receptor binding) may be quantified by a dimer cooperativity index [44, 20, 38], which relates to the apparent binding cooperativity in the Hill function sense. Within a Hill function analysis, however, there is the possibility that information is missed due to the inability of Hill coefficients to distinguish interaction mechanisms [106]. These works appear to be the state-of-the art in practical GPCR models, while more mathematically abstract approaches are taken elsewhere: an algebra of dimerisation is presented in [139], and generalised multi-site binding models are analysed in [61].

Dynamic models of binding and signalling for dimerised receptors are less common

than equilibrium models. Spatial models of the dimerisation process and subsequent signalling are developed in [85], while dynamics of receptor and transducer protein dimerisation are studied using an ordinary differential equation (ODE) model in [127], wherein it is suggested that homodimerisation may serve to regulate signalling over multiple time scales. Ligand-induced dimerisation of VEGF receptors is modelled using ODEs in [80]. In order to lay the foundations for dynamic modelling of signalling via dimerised GPCRs, we refer to a more recent study of GPCR ligand binding dynamics (May *et al.* [83]). Therein, linear ODE models of single-ligand and competition dynamics at pre-dimerised homodimeric receptors are presented, with a brief model analysis and numerical data fitting, without presenting analytical solutions to the ODE systems. In the current work, we describe a simplified formulation of the May *et al.* models, and derive their analytical solutions. For the single-ligand model, the solution structure (bi-exponential in time, reflecting two distinct eigenvalues) is reminiscent of Motulsky-Mahan competition dynamics at a monomer [95].

In this Chapter, we develop mathematical models for the dynamics of ligand binding at pre-dimerised receptors. In Section 2.1, we formulate and solve (analytically) a linear ODE model for single-ligand (A) binding kinetics at constitutively dimerised receptors. We first relate our model to existing models, and find an analytical equilibrium solution for total bound ligand. From this solution, we derive a condition under which multiple inflections in the logDR curve appear, in terms of the mechanistic binding cooperativity coefficient (α). Further, we show that the time-dependent problem has an analytical solution which may be easily constructed and computed without the need for numerical ODE solvers, and use this solution to simulate the binding dynamics. In Section 2.2, we extend our model to account for the presence of a second competing ligand (B). Again, we relate to previous models, and begin by finding the equilibrium solution. This enables us to derive a condition for multiple inflections in the logDR curve for total bound ligand A, which this time depends on multiple mechanistic binding cooperativity coefficients and the concentration of B. The ODE model for competition dynamics is linear as before, but its analytical solution is more laborious to compute. We again simulate the binding dynamics to explore the effects of the dynamic cooperativity factors. In Section 2.3 we

again consider a single ligand binding a pre-dimerised receptor, however, assume this dimer is a heterodimer, in that it consists of two similar but non-identical receptors. We follow the framework of analysis as in the previous sections to analyse results both at equilibrium and for time-course dynamics.

2.1 A single drug binding model

In this initial model, we consider a single ligand, A, binding to a pre-dimerised receptor complex. In this model, we assume all dimers are homodimers, so consist of two identical receptors bound together. We denote these dimers as R for simplicity in the model but keep the assumption that ligand molecules can bind to either side of the dimer. Association and dissociation rates are k_{a+} and k_{a-} respectively and we assume there is no bias to which protomer the ligand molecule binds to first. As the second ligand binding may be affected by one side of the dimer being already bound we introduce $\alpha = \alpha_+/\alpha_-$ which represents the equilibrium binding cooperativity, that is the factor change in affinity for the dimer when it is already ligand-bound. The value $\alpha = 1$ represents neutral cooperativity, and $\alpha > 1$ and $\alpha < 1$ represent positive and negative cooperativity respectively. Figure 2.1.1 shows the system of biochemical reactions. Since R represents a dimer, AR is the complex created by a single ligand molecule binding to a protomer and ARA is a dual bound receptor.

$$A + R \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} AR$$
$$A + AR \underset{\alpha_- k_{a-}}{\overset{\alpha_+ k_{a+}}{\rightleftharpoons}} ARA$$

Figure 2.1.1: Schematic: A ligand A binds to one side of a dimer, R, creating AR. A second ligand subsequently binds this complex, creating ARA.

Applying the law of mass action, the binding kinetics are governed by the following system

of ordinary differential equations (ODEs):

$$\frac{d[R]}{dt} = k_{a-}[AR] - k_{a+}[A][R], \tag{2.1a}$$
$$\frac{d[AR]}{dt} = k_{a+}[A][R] - k_{a-}[AR] - \alpha_+ k_{a+}[A][AR] + \alpha_- k_{a-}[ARA], \tag{2.1b}$$
$$\frac{d[ARA]}{dt} = \alpha_+ k_{a+}[A][AR] - \alpha_- k_{a-}[ARA], \tag{2.1c}$$

The number of receptors is a conserved quantity. Similarly, as we are replicating in vitro (lab experiment) conditions, the cell (or cell membrane) is also a constant. As such, the concentration of receptors is also conserved, hence, this allows us to give the total concentration of dimerised receptors as

$$R_{tot} = [R] + [AR] + [ARA], \tag{2.2}$$

which we use to reduce the system to two ODEs. We also assume that the concentration of ligand is much larger than the concentration of receptors (this is usually the case experimentally), hence the depletion of ligand is negligible and the ligand concentration can be assumed to be constant. Thus, taking $[A]$ as a constant, we state the system in the form $\frac{d}{dt}\mathbf{X}(t) = M\mathbf{X}(t) + \mathbf{f}$ where $\mathbf{X} = ([AR], [ARA])^T$,

$$M = \begin{bmatrix} -(k_{a-} + k_{a+}[A] + \alpha_+ k_{a+}[A]) & \alpha_- k_{a-} - k_{a+}[A] \\ \alpha_+ k_{a+}[A] & -\alpha_- k_{a-} \end{bmatrix} \text{ and } \mathbf{f} = \begin{bmatrix} k_{a+}[A]R_{tot} \\ 0 \end{bmatrix} \tag{2.3}$$

with initial conditions

$$\begin{bmatrix} [AR](0) \\ [ARA](0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{2.4}$$

2.1.1 Equilibrium analysis

We later look at a time course analysis and find analytical solutions to the the system, but in the spirit of classical receptor theory, we first investigate the equilibrium behaviour of the system, in particular the effect of α. The law of mass action allows us to state the equilibrium relationships as

$$[AR] = K_A[A][R], \tag{2.5a}$$

$$[ARA] = \alpha K_A[A][AR], \tag{2.5b}$$

where $K_A = k_{a+}/k_{a-}$ is the equilibrium association constant and $\alpha = \alpha_+/\alpha_-$ is the equilibrium binding cooperativity. The total concentration of dimers is

$$\begin{aligned} R_{tot} &= [R] + [AR] + [ARA] \\ &= [R](1 + K_A[A] + \alpha K_A^2[A]^2), \end{aligned} \tag{2.6}$$

which we can combine with equations (2.5) to express the equilibrium concentrations in terms of parameters only, giving us

$$[R] = \frac{1}{1 + K_A[A] + \alpha K_A^2[A]^2} R_{tot}, \tag{2.7a}$$

$$[AR] = \frac{K_A[A]}{1 + K_A[A] + \alpha K_A^2[A]^2} R_{tot}, \tag{2.7b}$$

$$[ARA] = \frac{\alpha K_A^2[A]^2}{1 + K_A[A] + \alpha K_A^2[A]^2} R_{tot}. \tag{2.7c}$$

We can clearly see that as $[A] \to \infty$ the concentrations of $[R]$ and $[AR]$ fall to 0, whereas $[ARA]$ tends to R_{tot}. The total amount of ligand bound at equilibrium is:

$$A_{bound} = [AR] + 2[ARA], \tag{2.8}$$

due to a single ligand molecule being bound in $[AR]$ and two ligand molecules being bound in $[ARA]$. Using equations (2.7) we can therefore state this quantity as

$$A_{bound} = \frac{(K_A[A] + 2\alpha K_A^2[A]^2)}{1 + K_A[A] + \alpha K_A^2[A]^2} R_{tot}. \tag{2.9}$$

This result is equivalent to the result Giraldo [44] states when discussing the mechanistic model for a single ligand binding to a dimerised receptor. We see the equivalence to his model if we take $K_A = 1/K_{D1}$ and $\alpha = K_{D1}/K_{D2}$. Letting $[A]$ tend to infinity in this result gives us the maximal ligand bound as $2R_{tot}$, thus we can state that half of the maximal A_{bound} is R_{tot}. Setting $A_{bound} = R_{tot}$ and solving for $[A]$ gives us the amount of ligand needed to get this response, that is the EC_{50} value, as

$$A_{50} = \frac{1}{K_A\sqrt{\alpha}}. \tag{2.10}$$

Figure 2.1.2 shows how the equilibrium cooperativity factor α affects the log dose-response (logDR) binding relationship. In the majority of cases the concentration curves behave as expected, with increasing cooperativity leading to dimers becoming dual-bound for lower concentrations of $[A]$ along with smaller peaks in $[AR]$. However, if we look to the plot of total ligand bound we see a curve that is not usual for monomer binding. Clearly as α increases we see a leftward shift in the curve. However, we can also see that for small values of α we get three inflections in the curve instead of just one.

Since the existence of extra inflections depends on α we seek a condition on α for when these appear. Following the calculation in Appendix B, with $a = 1, b = \alpha, c = 1$ and $X = K_A[A]$, we find that there is always an inflection point at

$$[A] = \frac{1}{K_A\sqrt{\alpha}}, \tag{2.11}$$

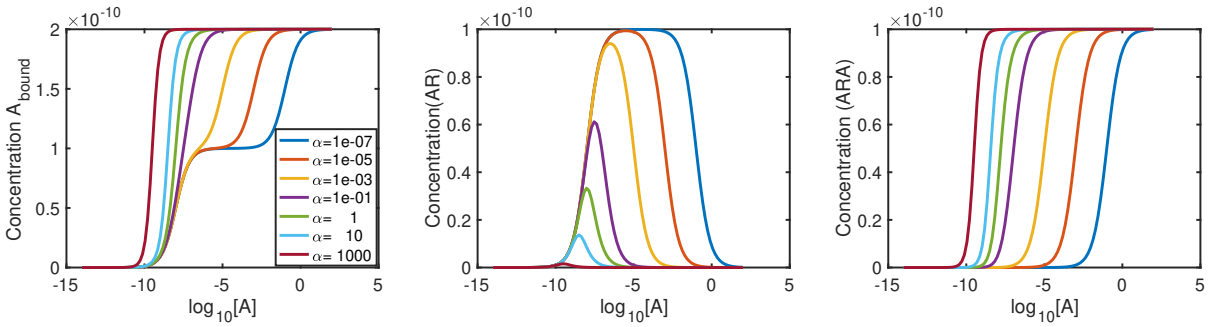

Figure 2.1.2: LogDR curves (plotted using equation (2.9)) for α ranging from extreme negative to extreme positive cooperativity. We see that extra inflections appear in A_{bound} when we have extreme low cooperativity.

that corresponds to the A_{50}, and two further inflection points at

$$[A]_{\pm} = \frac{-(8\alpha - 1) \pm \sqrt{(8\alpha - 1)^2 - 4\alpha}}{2}, \tag{2.12}$$

under the conditions of

$$0 < \alpha < 1/16. \tag{2.13}$$

When α is within this range the inflection at $\alpha = A_{50}$ changes from a rising inflection point to a falling one and we get two extra inflection points, one for $[A] < A_{50}$ and one for $[A] > A_{50}$. We see in Figure 2.1.2 that this results in a biphasic curve. We note here that a biphasic logDR curve with three inflections as shown would not be seen for monomeric receptors, and this is suggestive of the possibility of pre-dimerised receptors and very negative cooperativity. If the extra inflections are evident in a data set, and the receptors are homodimers, then we can conclude that $\alpha < 1/16$.

2.1.2 Binding dynamics - analytical solutions

Analytical solutions aid pharmacologists by allowing them to easily run simulations. Recall the differential equations representing the binding kinetics of the system $\frac{d}{dt}\mathbf{X}(t) =$

$M\mathbf{X}(t) + \mathbf{f}$ where

$$M = \begin{bmatrix} -(k_{a-} + k_{a+}[A] + \alpha_+ k_{a+}[A]) & \alpha_- k_{a-} - k_{a+}[A] \\ \alpha_+ k_{a+}[A] & -\alpha_- k_{a-} \end{bmatrix} \text{ and } \mathbf{f} = \begin{bmatrix} k_{a+}[A]R_{tot} \\ 0 \end{bmatrix}, \tag{2.14}$$

with initial conditions

$$\begin{bmatrix} [AR](0) \\ [ARA](0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{2.15}$$

As this is a linear system we are able to solve it to find exact solutions as functions of time, using the method outlined in Section 1.6.1. We note that $\det(M) > 0$ and $\mathrm{Tr}(M) < 0$, as in §1.6.1. Using this method, we find the solutions to the system to be

$$[AR](t) = \frac{k_{a+}[A]R_{tot}}{\det(M)} \left(\frac{\lambda_2(\alpha_- k_{a-} + \lambda_1)e^{\lambda_1 t} - \lambda_1(\alpha_- k_{a-} + \lambda_2)e^{\lambda_2 t}}{\lambda_1 - \lambda_2} + \alpha_- k_{a-} \right), \tag{2.16}$$

$$[ARA](t) = \frac{\alpha_+ k_{a+}^2 [A]^2 R_{tot}}{det(M)(\lambda_1 - \lambda_2)} \left(\lambda_2 e^{\lambda_1 t} - \lambda_1 e^{\lambda_2 t} + \lambda_1 - \lambda_2 \right). \tag{2.17}$$

where

$$\lambda_{1,2} = \frac{\mathrm{Tr}(M) \pm \sqrt{\mathrm{Tr}(M)^2 - 4\det(M)}}{2}. \tag{2.18}$$

are the eigenvalues of the system. From these solutions we can say that the total ligand bound is

$$A_{bound}(t) = \frac{k_{a+}[A]R_{tot}}{\det(M)} \left(\frac{\begin{array}{c} \lambda_2(\alpha_- k_{a-} + \lambda_1 + 2\alpha_+ k_{a+}[A])e^{\lambda_1 t} - \lambda_1(\alpha_- k_{a-} + \lambda_2 \\ + 2\alpha_+ k_{a+}[A])e^{\lambda_2 t} + \alpha_- k_{a-} + 2\alpha_+ k_{a+}[A] \end{array}}{\lambda_1 - \lambda_2} \right). \tag{2.19}$$

Looking more closely at the components of these solutions we see that we have two exponential components meaning we would expect to see biphasic curves. As λ_1 and λ_2 are real, distinct and always negative, providing all parameters in the model are positive we can say with certainty that as $t \longrightarrow \infty$ the exponential components will reduce to zero and we can state the equilibrium concentrations as

$$[AR] = \frac{\alpha_- k_{a+} k_{a-} [A] R_{tot}}{\det(M)}, \tag{2.20}$$

$$[ARA] = \frac{\alpha_+ k_{a+}^2 [A]^2 R_{tot}}{\det(M)}, \tag{2.21}$$

giving the total ligand bound at equilibrium as

$$A_{bound} = \frac{k_{a+}[A] R_{tot}}{\det(M)} \Big(\alpha_- k_{a-} + 2\alpha_+ k_{a+} [A] \Big). \tag{2.22}$$

Expanding $\det(M)$ and simplifying in MATLAB [2] confirms the equivalence of these to equations (2.7).

Having these analytical solutions (equations (2.16) and (2.17)) to the ODE system allows time courses to be plotted without the need for numerical ODE solvers. This is particularly useful for pharmacologists without numerics expertise, allowing them to construct exact solutions for any parameter values in software packages such as Excel [88] and Graph Pad [1].

2.1.3 Single ligand time course results

Here we present numerical results which demonstrate the effects of cooperative dynamics across dimerised receptors, particularly contrasting with the dynamics observed for monomeric receptors. Figure 2.1.3 shows the binding of ligand $[A]$ to the dimerised re-

ceptor $[R]$ with the ligand being added as a constant $10^{-8}M$. Initially, ligand molecules bind singularly to dimers, creating an increase in $[AR]$. As time increases, a positive cooperativity factor means that the chance of a second ligand molecule binding to the singularly bound dimer is increased; thus, $[ARA]$ increases. This increase in $[ARA]$ also has the effect of reducing $[AR]$ which then falls towards zero. Hence, we see a peak in $[AR]$.

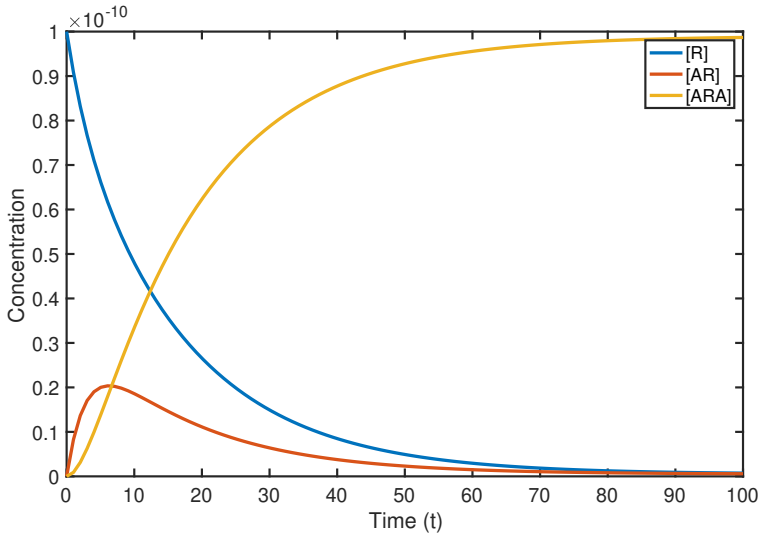

Figure 2.1.3: The binding of ligand A to a pre-dimerised receptor results in slight peaks in $[AR]$, while positive cooperativity leads to all dimers becoming dual bound. The plot shows a numerical solution to the system in equations (2.1). We use a ligand concentration of $[A] = 10^{-8}M$ while the cooperativity factors were set at $\alpha_+ = 2$ and $\alpha_- = 0.01$ giving positive cooperativity.

In fact, this peak becomes a point of interest as we look at Figure 2.1.4. In this, we consider a range of values of $[A]$ and study the effect this has on the binding dynamics. It is clear that not only does A_{bound} increase to a higher equilibrium concentration as [A] increases, but also, equilibration is reached in a shorter timescale. While $[A]$ is low, $[AR]$ approaches equilibrium monotonically, whereas a peak in $[AR]$ occurs for high concentrations of A.

We notice in Figure 2.1.4 that as $[A]$ increases there is a leftward shift in the peak as the time at which it occurs decreases. Using equation (2.16) we can find when this peak occurs exactly. To find a solution we differentiate this equation with respect to t then set

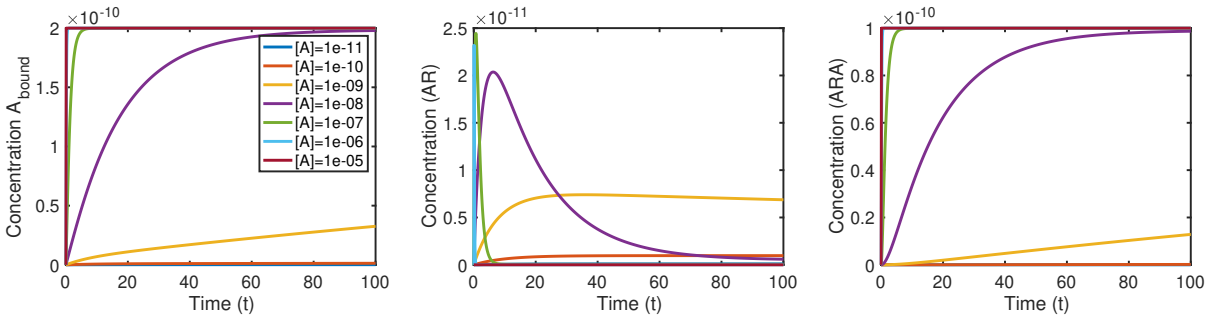

Figure 2.1.4: Plot shows a numerical solution to equations (2.1) for a varying ligand concentration. With high levels of $[A]$ in the system we no longer see peaks in $[AR]$. Cooperativity values are fixed at $\alpha_+ = 2$ and $\alpha_- = 0.01$.

this equal to zero before solving to find that the peak occurs at time

$$t = -\frac{\ln\left(\dfrac{\alpha_- k_{a-} + \lambda_1}{\alpha_- k_{a-} + \lambda_2}\right)}{\lambda_1 - \lambda_2}. \tag{2.23}$$

For peak existence, we therefore require

$$\frac{\alpha_- k_{a-} + \lambda_1}{\alpha_- k_{a-} + \lambda_2} > 0, \tag{2.24}$$

where $\lambda_2 < \lambda_1 < 0$. Hence we require, for a positive peak time,

$$0 < \frac{\alpha_- k_{a-} + \lambda_1}{\alpha_- k_{a-} + \lambda_2} < 1,$$

which requires that

$$\alpha_- k_{a-} + \lambda_1 < 0.$$

Substituting in the full expression (equation (2.18)) for λ_1, gives

$$\alpha_- < K_A[A]. \tag{2.25}$$

To evaluate the peak concentration $[AR]$, we substitute the peak time, equation (2.23), into equation (2.16) (the analytical solution for $[AR]$), giving, after some simplification

$$[AR] = \frac{k_{a+}[A]R_{tot}}{\det(M)} \left(-(\alpha_- k_{a-} + \lambda_1) \left(\frac{\alpha_- k_{a-} + \lambda_1}{\alpha_- k_{a-} + \lambda_2} \right)^{\frac{\lambda_1}{\lambda_2 - \lambda_1}} + \alpha_- k_{a-} \right) \tag{2.26}$$

as the corresponding concentration of $[AR]$.

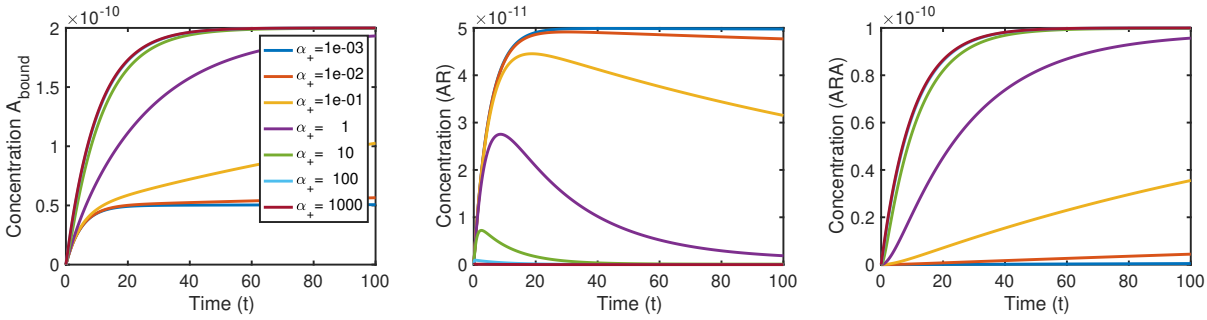

Figure 2.1.5: Plot shows a numerical solution to equations (2.1) for a varying forward cooperativity factor. As we move from positive to negative cooperativity we see a more pronounced peak in $[AR]$ with A_{bound} tending to lower concentrations. Ligand concentration for the plot is $[A] = 10^{-8}M$.The value $\alpha_- = 1$ was fixed so that cooperativity depends solely on α_+.

In Figure 2.1.5 we look at how a range of cooperativity factors affect the binding of ligand $[A]$. Clearly, the binding rate for a second ligand to the dimer increases and decreases for positive and negative cooperativity, respectively. We see a peak in $[AR]$ whose timing and concentration are dependent on α_+. As α_+ increases, the peak in $[AR]$ decreases and occurs earlier. Both $[ARA]$ and A_{bound} increase as α_+ increases.

2.2 A two drug binding model

We now introduce a second ligand, B, to the system. The rationale for this is that quantification of effects of unlabeled ligands can be achieved by competition experiments with labeled and unlabeled ligands, as seen in previous studies involving monomeric and dimeric receptors [95, 83]. The kinetics of this system are key in highlighting and quantifying allosteric interactions across dimerised receptors, as indicated by May *et al* [83], who discuss the influence of an unlabeled ligand on the dissociation (washout) kinetics of a labeled ligand, when dimers are present.

We assume the association rates remain as they were in the previous model, along with the cooperativity factor α to describe the cross talk across a dimer dual bound by $[A]$. We express the association and dissociation of ligand B by the rates k_{b+} and k_{b-} respectively and now require two extra cooperativity factors. We denote $\beta = \beta_+/\beta_-$ as the influence a protomer bound by ligand B has on a second B binding, and vice-versa, and $\gamma = \gamma_+/\gamma_-$ as the cooperativity factor describing the interaction between A and B bound receptors. This extended system gives a set of six reactions, as can be seen in Figure 2.2.1.

$$A + R \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} AR \qquad B + R \underset{k_{b-}}{\overset{k_{b+}}{\rightleftharpoons}} BR$$

$$A + AR \underset{\alpha_- k_{a-}}{\overset{\alpha_+ k_{a+}}{\rightleftharpoons}} ARA \qquad B + AR \underset{\gamma_- k_{b-}}{\overset{\gamma_+ k_{b+}}{\rightleftharpoons}} ARB$$

$$A + BR \underset{\gamma_- k_{a-}}{\overset{\gamma_+ k_{a+}}{\rightleftharpoons}} ARB \qquad B + BR \underset{\beta_- k_{b-}}{\overset{\beta_+ k_{b+}}{\rightleftharpoons}} BRB$$

Figure 2.2.1: Schematic: Two ligands, A and B compete for binding sites on dimer R.

2.2.1 Differential equations

We use the law of mass action to derive from these reactions a system of differential equations that govern the binding kinetics of the system:

As the total concentration of receptors is conserved we can state that

$$R_{tot} = [R] + [AR] + [BR] + [ARA] + [ARB] + [BRB]. \tag{2.27}$$

Using this conservation law allows us to reduce the system to 5 equations by eliminating $[R]$, leaving

$$\begin{aligned}\frac{d[AR]}{dt} &= -(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A] - \gamma_+ k_{b+}[B])[AR] - k_{a+}[A][BR] \\ &\quad + (\alpha_- k_{a-} - k_{a+}[A])[ARA] + (\gamma_- k_{b-} - k_{a+}[A])[ARB] - k_{a+}[A][BRB] + k_{a+}[A]R_{tot},\end{aligned} \tag{2.28a}$$

$$\begin{aligned}\frac{d[BR]}{dt} &= -k_{b+}[B][AR] - (k_{b+}[B] + k_{b-} + \gamma_+ k_{b+}[A] + \beta_+ k_{b+}[B])[BR] \\ &\quad - k_{b+}[B][ARA] + (\gamma_- k_{b-} - k_{b+}[B])[ARB] + (\beta_- k_{b-} - k_{b+}[B])[BRB] + k_{b+}[B]R_{tot},\end{aligned} \tag{2.28b}$$

$$\frac{d[ARA]}{dt} = \alpha_+ k_{a+}[A][AR] - \alpha_- k_{a-}[ARA], \tag{2.28c}$$

$$\frac{d[ARB]}{dt} = \gamma_+ k_{b+}[B][AR] + \gamma_+ k_{a+}[A][BR] - \gamma_- k_{b-}[ARB] - \gamma_- k_{a-}[ARB], \tag{2.28d}$$

$$\frac{d[BRB]}{dt} = \beta_+ k_{b+}[B][BR] - \beta_- k_{b-}[BRB]. \tag{2.28e}$$

with initial conditions

$$\begin{pmatrix} [AR](0) \\ [BR](0) \\ [ARA](0) \\ [ARB](0) \\ [BRB](0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2.29}$$

We note that this system is a simplification of the one developed by May *et al* [83]. We recall the model presented in [83] in Figure 2.2.2. When translating the schematic to a system of ODEs, May *et al.* assume assuming symmetry across the dimer, thus assuming

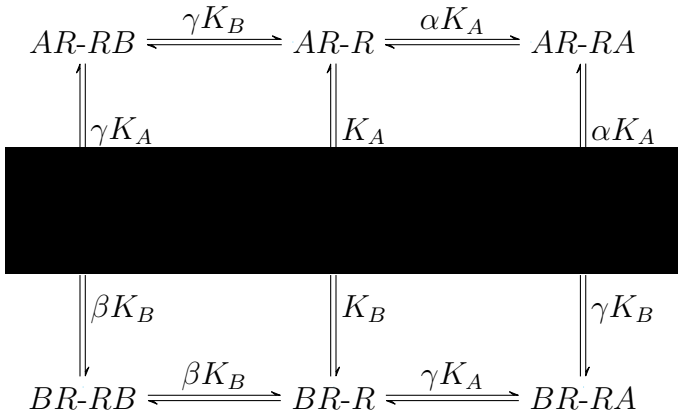

Figure 2.2.2: Schematic of the two-ligand dimer model presented by May *et al.* in [83].

AR-R is the same as *R-RA*, for example, and denoting this $[AR]$. We take this one step further by assuming symmetry in the model initially, thus reducing the both the number of reactions, and number of species. The two models are equivalent by taking

$$k_{a+} = \frac{\widetilde{k}_{a+}}{2}, \quad k_{a-} = \widetilde{k}_{a-} \quad k_{b+} = \frac{\widetilde{k}_{b+}}{2}, \quad k_{b-} = \widetilde{k}_{b-}, \tag{2.30}$$

$$\alpha_+ = 2\widetilde{\alpha}_+, \quad \alpha_- = \frac{\widetilde{\alpha}_-}{2}, \quad \beta_+ = 2\widetilde{\beta}_+, \quad \beta_- = \frac{\widetilde{\beta}_-}{2}, \quad \gamma_+ = 2\widetilde{\gamma}_+, \quad \gamma_- = \frac{\widetilde{\gamma}_-}{2}, \tag{2.31}$$

noting that

$$[AR] = \frac{\widetilde{[AR]}}{2}, \quad [BR] = \frac{\widetilde{[BR]}}{2}, \quad [ARB] = \frac{\widetilde{[ARB]}}{2}, \tag{2.32}$$

where a tilde denotes a parameter or variable in the May *et al* [83] model.

2.2.2 Equilibrium analysis

Once the system reaches equilibrium we have the set of relations

$$[AR] = K_A[A][R], \qquad [ARA] = \alpha K_A[A][AR] = \alpha K_A^2[A]^2[R], \tag{2.33a}$$

$$[BR] = K_B[B][R], \qquad [BRB] = \beta K_B[B][BR] = \beta K_B^2[B]^2[R]. \tag{2.33b}$$

$$[ARB] = \gamma K_A[A][BR] = \gamma K_A K_B[A][B][R], \tag{2.33c}$$

Since we can write the total concentration of dimerised receptors as

$$\begin{aligned} R_{tot} &= [R] + [AR] + [BR] + [ARA] + [ARB] + [BRB] \\ &= [R](1 + K_A[A] + K_B[B] + \alpha K_A^2[A]^2 + \gamma K_A K_B[A][B] + \beta K_B^2[B]^2), \end{aligned} \tag{2.34}$$

we find the concentrations of each of the molecular complexes in terms of parameters as follows:

$$\begin{aligned} [R] &= \frac{1}{D} R_{tot}, & [AR] &= \frac{K_A[A]}{D} R_{tot}, \\ [BR] &= \frac{K_B[B]}{D} R_{tot}, & [ARA] &= \frac{\alpha K_A^2[A]^2}{D} R_{tot}, \\ [ARB] &= \frac{\gamma K_A K_B[A][B]}{D} R_{tot}, & [BRB] &= \frac{\beta K_B^2[B]^2}{D} R_{tot}. \end{aligned} \tag{2.35}$$

where

$$D = 1 + K_A[A] + K_B[B] + \alpha K_A^2[A]^2 + \gamma K_A K_B[A][B] + \beta K_B^2[B]^2 \tag{2.36}$$

Here we consider, as in May *et al* [83], competition experiments where ligand A is

labelled and ligand B is unlabelled. We can clearly see that, providing $[B]$ is fixed

$$[R], [AR], [BR], [ARB], [BRB] \to 0, \text{ and } [ARA] \to R_{tot} \text{ as } [A] \to \infty. \tag{2.37}$$

The total concentration of bound ligand A (assumed an experimentally measurable quantity) is

$$\begin{aligned} A_{bound} &= [AR] + 2[ARA] + [ARB] \\ &= [R](K_A[A] + 2\alpha K_A^2[A]^2 + \gamma K_A K_B[A][B]). \end{aligned} \tag{2.38}$$

which, using equation (2.33), gives us

$$A_{bound} = \frac{K_A[A] + 2\alpha K_A^2[A]^2 + \gamma K_A K_B[A][B]}{1 + K_A[A] + K_B[B] + \alpha K_A^2[A]^2 + \gamma K_A K_B[A][B] + \beta K_B^2[B]^2} R_{tot}. \tag{2.39}$$

As the maximal A_{bound}, for varying $[A]$, remains as $2R_{tot}$ we can calculate the EC_{50} of A_{bound} as

$$A_{50} = \frac{1}{K_A}\sqrt{\frac{1 + K_B[B] + \beta K_B^2[B]^2}{\alpha}}, \tag{2.40}$$

which we note is independent of the A-B cooperativity factor γ (we return to this point when varying γ). In the following sections we show the effects of each of the cooperativity factors α, β, γ in turn. We vary each of the equilibrium cooperativity factors individually while keeping the others set to 1 to give neutral cooperativity to be able to study the effects the individual cooperativity factors has on the system. In each case we plot logDR curves for A_{bound} for a range of values for $[B]$. To begin we consider a range of values of α, hence fix β and γ.

The effects of α cooperativity

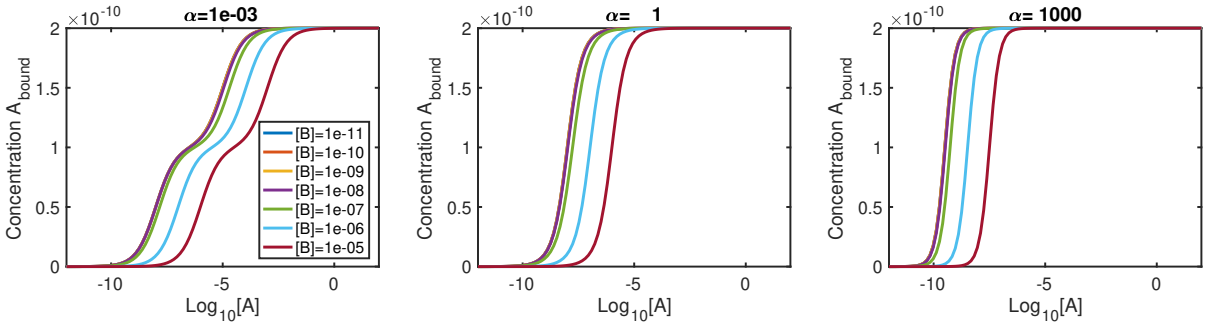

Figure 2.2.3: Plotting logDR curves, using equation (2.39), for varying α shows extra inflections when we have low A-A cooperativity regardless of $[B]$. Other cooperativity factors are fixed at $\beta = \gamma = 1$.

Again we see extra inflections for low values of α, similarly to the single ligand system, as shown in Figure 2.2.3. While $[A] \ll [B]$ dimers are primarily bound by molecules of B. As $[A]$ increases we see increases in concentrations of AR. However, low cooperativity means that these become bound by molecules of $[B]$, forming ARB complexes, but few become bound by a second molecule of A, and as such, $[ARA]$ remains low. It is not until $[A] \gg [B]$ that the low cooperativity is countered, where we then see increases in $[ARA]$, corresponding to a fall in dimers bound by a single molecule of $[A]$, thus creating peaks in these. It is this that changes the nature of the original inflection point for A_{bound}, as well as creating two extra inflections. This is clear if we look to the individual species plots in Figure 2.2.4.

The effects of β cooperativity

Figure 2.2.5 shows the effect of varying the B-B cooperativity factor, β. We see inflections still appear when we have both low β and high $[B]$. In Figure 2.2.6 we look at the individual species curves, with $[B]$ fixed at a high concentration, showing inflections for all A_{bound} curves where $\beta < 1$. With $\beta < 1$, these become then bound by a molecule of A, giving an increase in $[ARB]$. However, as $[A] \to \infty$ this is countered and all dimers become dual bound by A, so $[ARA] \to R_{tot}$ while all other complexes fall to zero. This increased peak in $[ARB]$ results in an extra inflection in the A_{bound} curve.

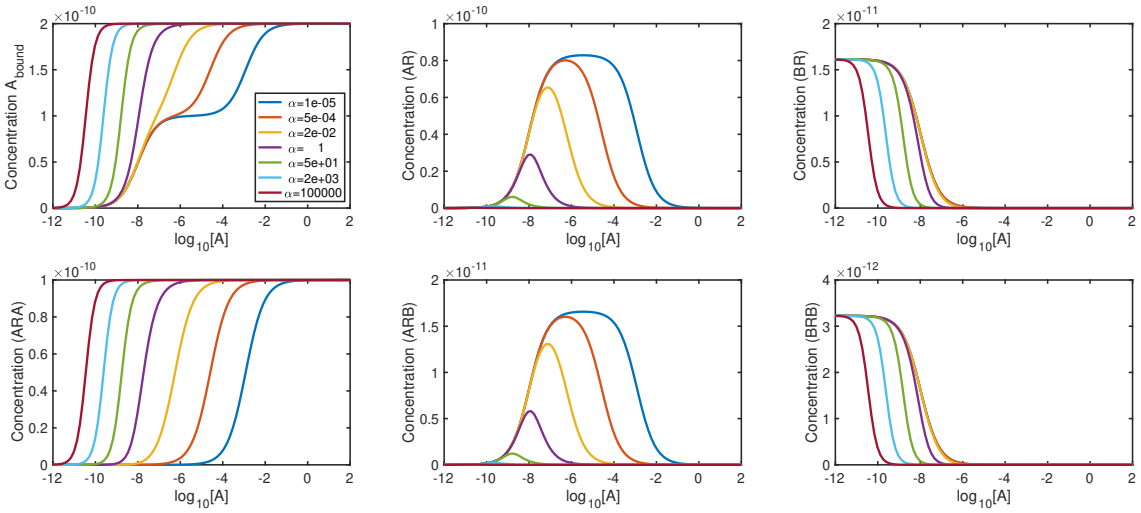

Figure 2.2.4: Individual species plots (equations (2.35)) for a varying α help to understand the effects of cooperativity on the dose response curves. Competition ligand concentration is fixed at $B = 10^{-8}M$, while $\beta = \gamma = 1$.

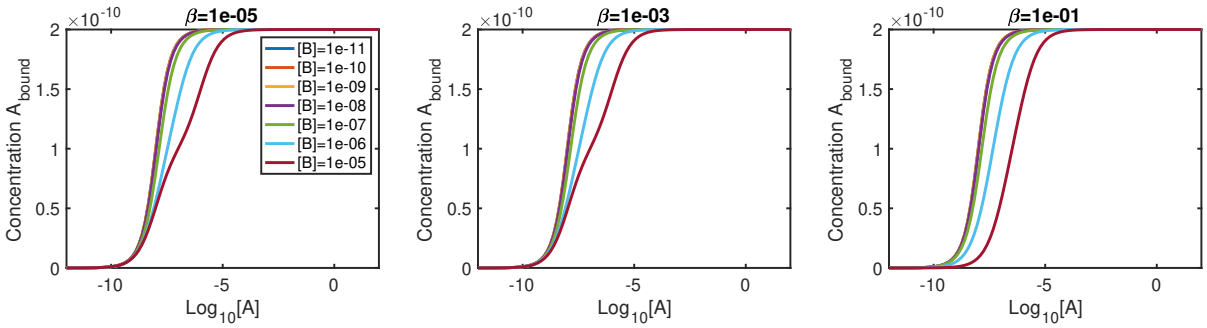

Figure 2.2.5: Plotting logDR curves, using equation (2.39), for varying β shows extra inflections appear when we have both low B-B cooperativity and low $[B]$. Other cooperativity factors are fixed at $\alpha = \gamma = 1$.

The effects of γ cooperativity

In Figure 2.2.7 we look at the effects of A-B cooperativity factor γ and see slightly different results. To get extra inflections, γ is required to be high as opposed to low. We again require there to be a high concentration of B in the system for inflections to appear. We again fix B at a high concentration to study the individual species curves, in Figure 2.2.8. With a large γ (and sufficiently large $[B]$), we see large increases in $[ARB]$, creating a peak once $[A] \approx [B]$. A much higher concentration of A is required for concentrations of ARB to fall. This large peak creates the extra inflections in the A_{bound} curve. We also note that, γ does not shift the curve, as α and β do. Varying γ changes the slope of the curve, but the point of inflection and thus the A_{50} value, remain the same for all curves.

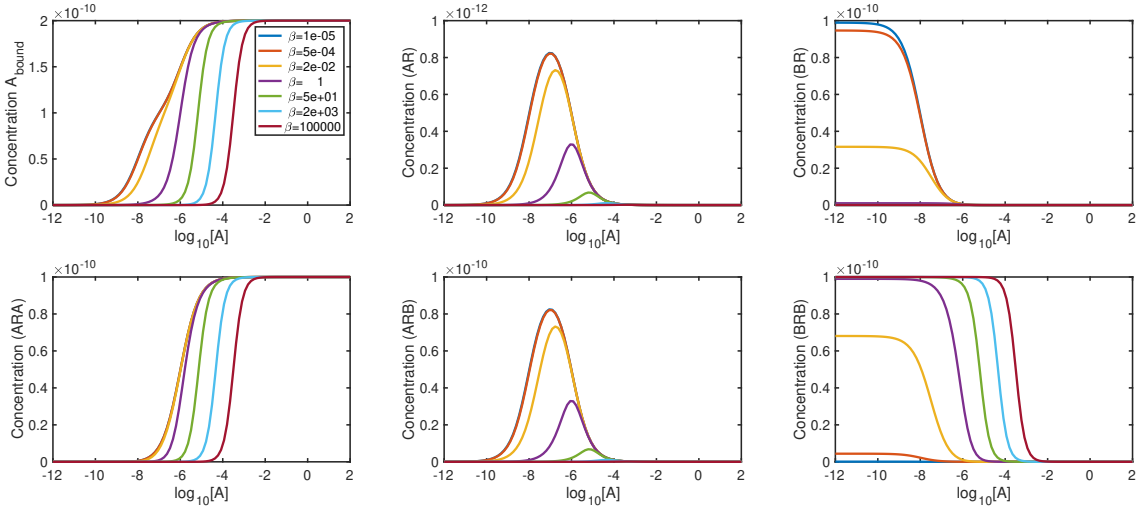

Figure 2.2.6: Individual species plots for a varying β help to understand the effects of cooperativity on the dose response curves. Competition ligand concentration is fixed at $B = 10^{-5}M$, while $\alpha = \gamma = 1$.

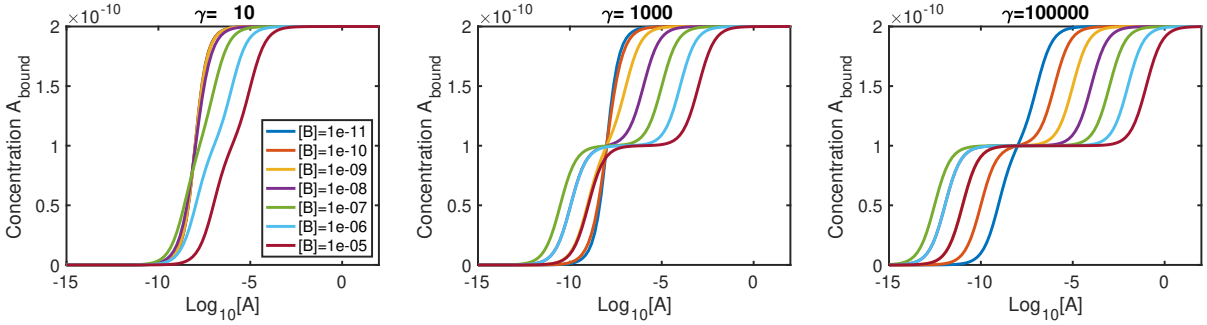

Figure 2.2.7: Plotting logDR curves, using equation (2.39), for varying γ shows extra inflections when we have high A-B cooperativity as well as low $[B]$. Other cooperativity factors are fixed at $\beta = \gamma = 1$.

This can be seen in the A_{bound} curve in Figure 2.2.8.

Investigating the inflections

Similarly to the single ligand system we also get extra inflections in the A_{bound} curve when we have two ligands in the system. Investigating these extra inflections in Abound (Appendix B with $a = 1 + \gamma K_A[B], b = \alpha, c = 1 + K_B[B] + \beta K_B^2[B]^2$ and $X = K_A[A]$),

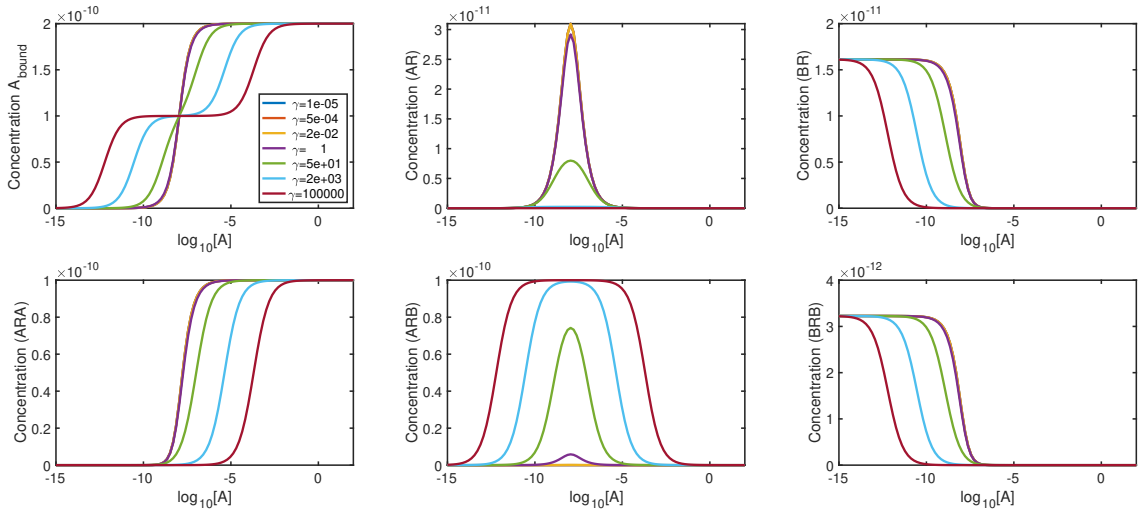

Figure 2.2.8: Individual species plots for a varying γ help to understand the effects of cooperativity on the dose response curves. Competition ligand concentration is fixed at $B = 10^{-8}M$, while $\beta = \gamma = 1$.

we find that there is always an inflection point at

$$[A] = \frac{1}{K_A}\sqrt{\frac{1 + K_B[B] + \beta K_B^2[B]^2}{\alpha}}. \tag{2.41}$$

We further find that extra inflections appear under the condition

$$16\alpha(1 + K_B[B] + \beta K_B^2[B]^2) < (\gamma K_B[B] + 1)^2, \tag{2.42}$$

with an extra inflection at either side of the original one, at the points

$$[A]_\pm = \frac{-(8bc - a^2) \pm \sqrt{(a^2 - 16bc)(a^2 - 4bc)}}{2abK_A}. \tag{2.43}$$

2.2.3 Time course results

Analytical solutions for the system (2.28) are theoretically possible, given that the system is linear. However, the task of computing eigenvalues exactly becomes laborious and impractical; we instead construct the solutions by numerical evaluation of the eigenvalues.

A numerical ODE solver could also be used.

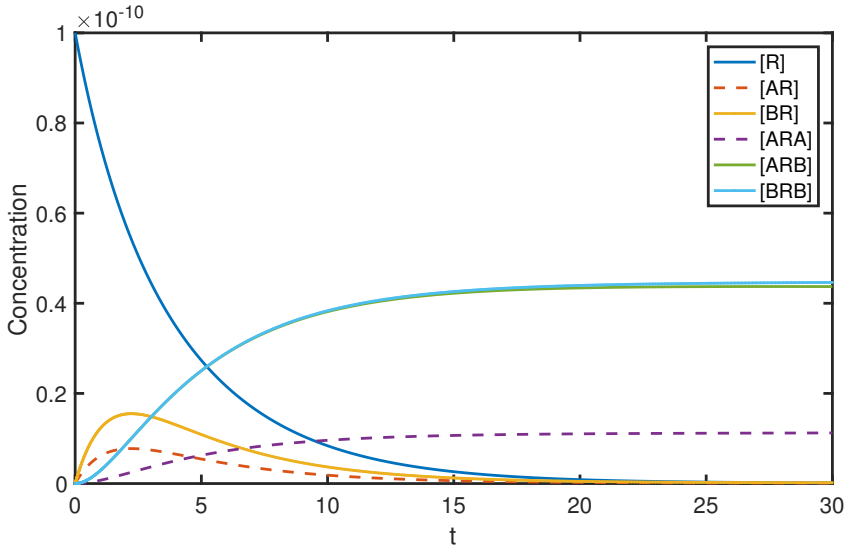

Figure 2.2.9: In the time course plot, we have a numerical solution to the system in equations (2.28), where we see peaks in both $[AR]$ and $[BR]$. Cooperativity values are $\alpha_+ = \beta_+ = \gamma_+ = 2$, $\alpha_- = \beta_- = \gamma_- = 0.01$, while ligand concentrations used are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

We now consider two-ligand competition in time course simulations. The concentration of $[A]$ remains the same as in the single ligand binding section, and a second ligand $[B]$ is introduced, with $[B] > [A]$. Association and dissociation rates k_{b+} and k_{b-} are set such that $k_{b+} < k_{a+}$ and $k_{b-} > k_{a-}$. Figure 2.2.9, with all forward and reverse cooperativity factors fixed at 2 and 0.01 respectively, shows $[AR]$ and $[ARA]$ dynamics similar to that of the single ligand problem.

While the curves of $[BR]$ and $[BRB]$ are similar to those of $[AR]$ and $[ARA]$, as $[B] > [A]$, this results in an increased peak in $[BR]$ and an increased equilibrium concentration of BRB. As $[ARB]$ has a contribution from both ligands A and B, it increases in the same way but tends to an equilibrium between $[ARA]$ and $[BRB]$.

In Figure 2.2.10 we see the effect α_+ has on the time course dynamics of the system. Varying α_+, while keeping all other cooperativity factors fixed equal to one isolates the effects of forward $A - A$ binding cooperativity. Clearly, peaks appear in $[AR]$ and $[BR]$

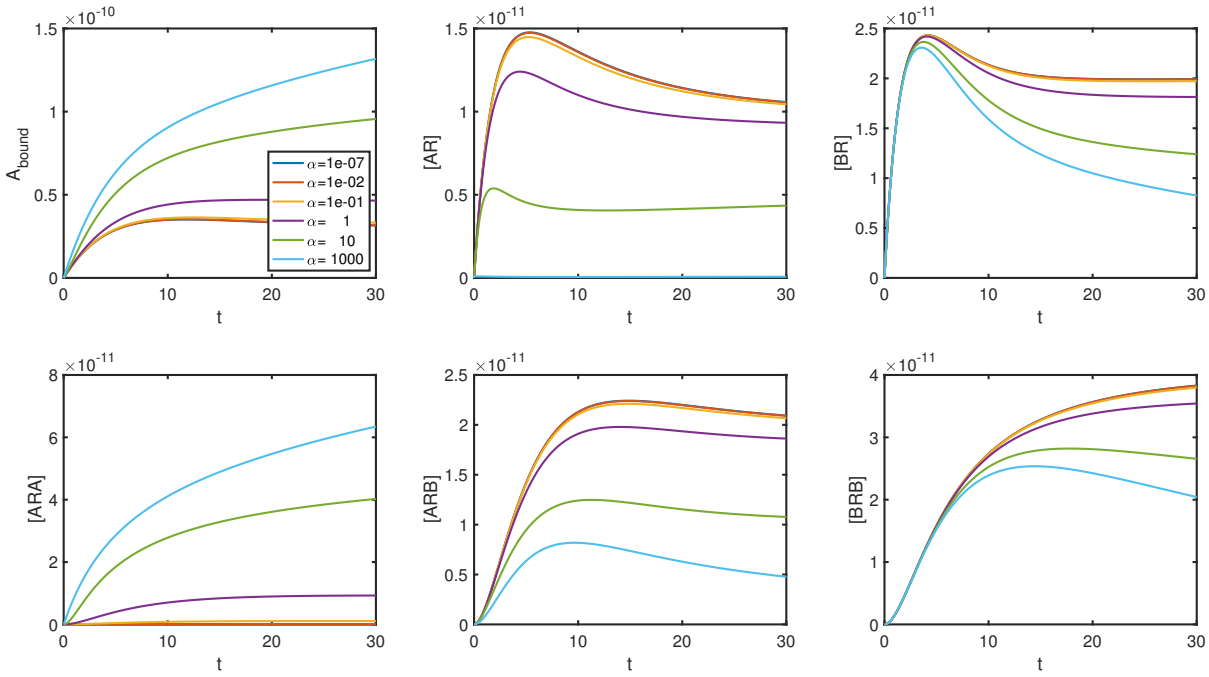

Figure 2.2.10: Plotting the numerical solution to equations (2.28) with a varying α_+ while all other cooperativity factors equal 1. Increasing α_+ results in increases in $[ARA]$. Ligand concentrations are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

for all values of α_+. However, increasing α_+ results in a smaller peak in $[AR]$, as these quickly become bound by two molecules of A, thus we see $[ARA]$ increasing more quickly and to a higher concentration. This in turn means that we see decreases in $[BR]$, $[ARB]$ and $[BRB]$ due to less available receptors.

In Figure 2.2.11 we vary β_+ while all other cooperativity factors are equal to one. In this, we see the results mirrored in $[A]$ and $[B]$ from when we varied α_+. Increasing β_+ shows a decreased peak in $[BR]$ and consequently, an increase in $[BRB]$. Concentrations of $[AR]$, $[ARA]$ and $[ARB]$ are reduced as there are less available free receptors for the binding of $[A]$.

In Figure 2.2.12 we demonstrate the effect of varying γ_+. As molecules of A and B become bound to free receptors we again see $[AR]$ and $[BR]$ increase. Once a dimer becomes bound by a ligand molecule, a small γ_+ means that it is less likely to become bound by an alternate molecule, so dimers become bound by two of the same ligand, hence $[ARA]$ and $[BRB]$ increase while $[ARB]$ remains low. As γ_+ increases, the converse happens and instead we see increases in $[ARB]$.

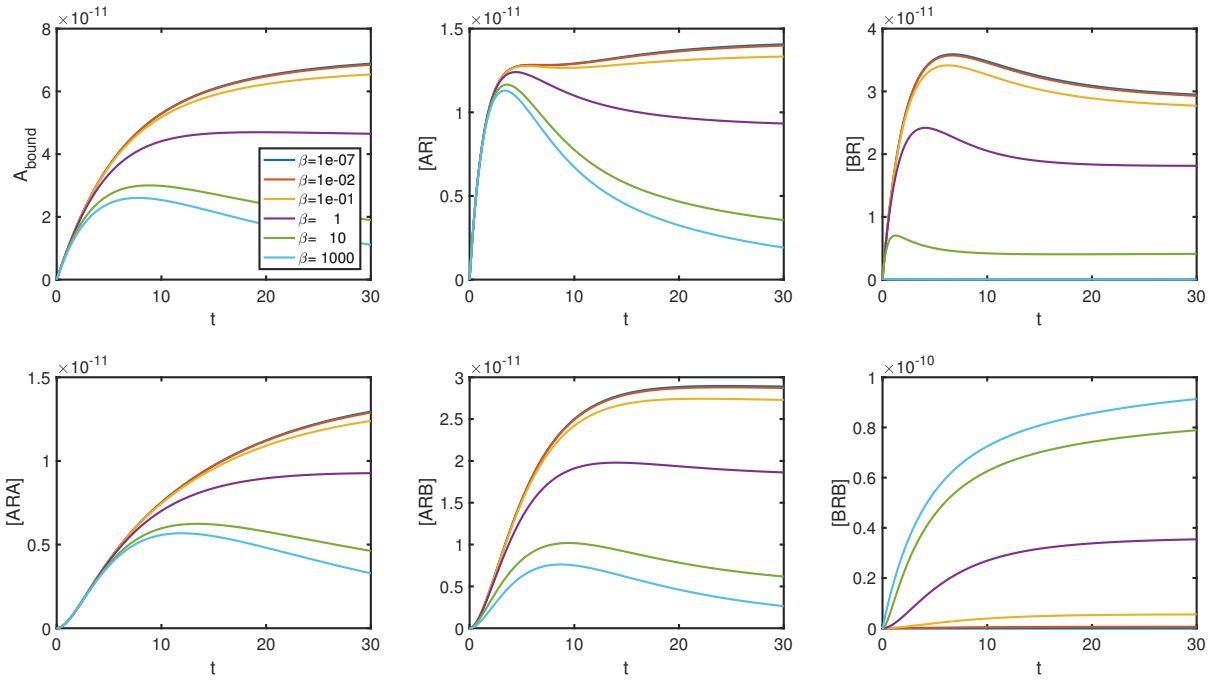

Figure 2.2.11: Plotting the numerical solution to equations (2.28) with a varying β_+ while all other cooperativity factors equal 1. Increasing β_+ results in increases in $[BRB]$. Ligand concentrations are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

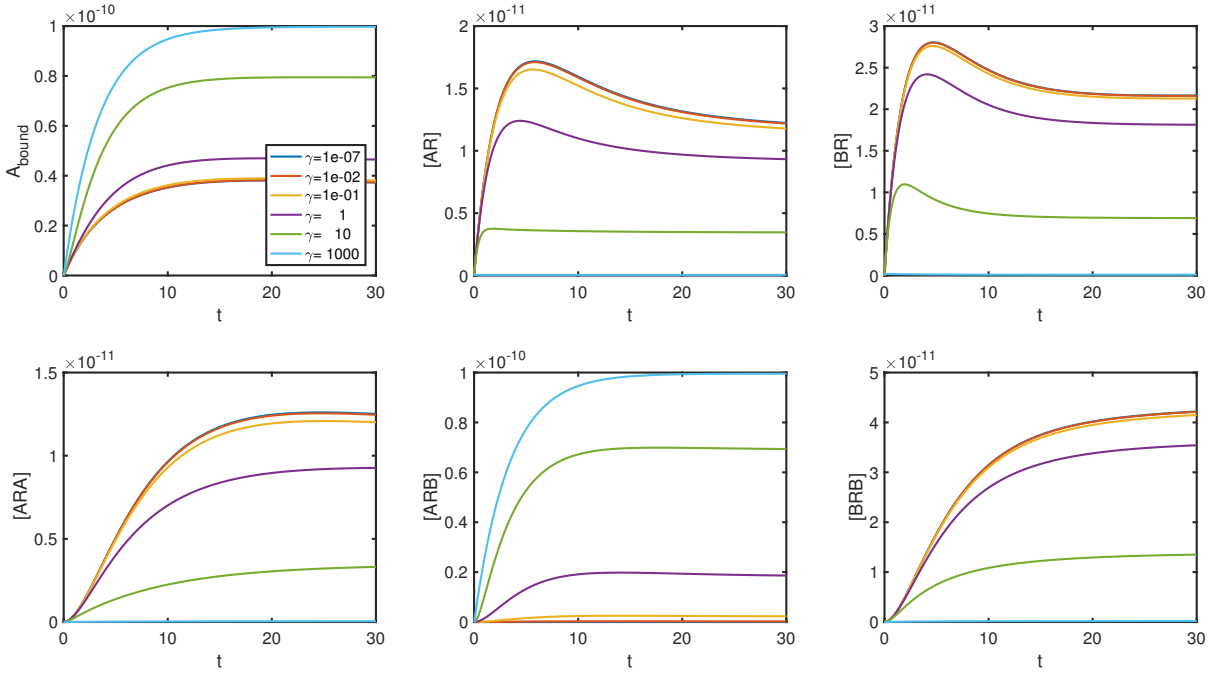

Figure 2.2.12: Plotting the numerical solution to equations (2.28) with a varying γ_+ while all other cooperativity factors equal 1. Increasing γ_+ results in increases in $[ARB]$. Ligand concentrations are $[A] = 10^{-8}M$, $[B] = 2 \times 10^{-8}M$.

2.3 A heterodimer binding model

It is widely acknowledged that GPCRs can also exist as heterodimers, that is a dimer consisting of two non-identical receptors [91, 92, 93, 115]. In the previous dimer model, we used the symmetry inherent in the homodimer model to reduce the model, resulting in a clean formulation amenable to algebraic manipulation. In this section we extend this to a more general model which will encompass heterodimers. This model is similar in form to the May *et al* schematic [83], with an implicit asymmetry which will allow for an "extra level of complexity" [6].

The proposed model consists of a single ligand, A, binding to a pre-bound heterodimer, R_1-R_2 and we assume that the ligand is able to bind to either of the receptors. The first reaction consists of a molecule of ligand A binding to either of these receptors creating the complexes AR_1-R_2 or R_1-R_2A. Association and dissociation rates of ligand binding to the R_1 receptor are k_{a_1+} and k_{a_1-} respectively, while binding rates of the R_2 receptor are k_{a_2+} and k_{a_2-}. We also introduce a cooperativity factor $\alpha = \alpha_+/\alpha_-$ which describes the crosstalk between the receptors. This affects the binding rates of the second ligand molecule to the dimer, upon when the complex AR_1-R_2A is created. We have positive cooperativity when $\alpha > 1$, so the equilibrium binding affinity of the second ligand binding is increased and negative cooperativity for $\alpha < 1$. We have a total of four possible reactions as can be seen in Figure 2.3.1. These reactions can be transcribed to a system

$$A + R_1\text{-}R_2 \underset{k_{a_1-}}{\overset{k_{a_1+}}{\rightleftharpoons}} AR_1\text{-}R_2 \qquad A + R_1\text{-}R_2 \underset{k_{a_2-}}{\overset{k_{a_2+}}{\rightleftharpoons}} R_1\text{-}R_2A$$

$$A + AR_1\text{-}R_2 \underset{\alpha_- k_{a_2-}}{\overset{\alpha_+ k_{a_2+}}{\rightleftharpoons}} AR_1\text{-}R_2A \qquad A + R_1\text{-}R_2A \underset{\alpha_- k_{a_1-}}{\overset{\alpha_+ k_{a_1+}}{\rightleftharpoons}} AR_1\text{-}R_2A$$

Figure 2.3.1: Schematic: the binding of a single ligand to a pre-formed heterodimer.

of ordinary differential equations (ODEs) that govern the binding kinetics of the system

by applying the law of mass action. Fixing $[A]$ as a constant we express these as

$$\frac{d[R_1\text{-}R_2]}{dt} = -(k_{a_1+} + k_{a_2+})[A][R_1\text{-}R_2] + k_{a_1-}[AR_1\text{-}R_2] + k_{a_2-}[R_1\text{-}R_2A], \tag{2.44a}$$

$$\frac{d[AR_1\text{-}R_2]}{dt} = k_{a_1+}[A][R_1\text{-}R_2] - (k_{a_1-} + \alpha_+ k_{a_2+}[A])[R_1\text{-}R_2A] + \alpha_- k_{a_2-}[AR_1\text{-}R_2A], \tag{2.44b}$$

$$\frac{d[R_1\text{-}R_2A]}{dt} = k_{a_2+}[A][R_1\text{-}R_2] - (k_{a_2-} + \alpha_+ k_{a_1+}[A])[AR_1\text{-}R_2] + \alpha_- k_{a_1-}[AR_1\text{-}R_2A], \tag{2.44c}$$

$$\frac{d[AR_1\text{-}R_2A]}{dt} = \alpha_+ k_{a_2+}[A][AR_1\text{-}R_2] + \alpha_+ k_{a_1+}[A][R_1\text{-}R_2A] - (\alpha_- k_{a_1-} + \alpha_- k_{a_2-})[AR_1\text{-}R_2A]. \tag{2.44d}$$

The total concentration of receptors is

$$R_{tot} = [R_1\text{-}R_2] + [AR_1\text{-}R_2] + [R_1\text{-}R_2A] + [AR_1\text{-}R_2A], \tag{2.45}$$

which is a conserved quantity, hence we can use this to reduce the system of ODEs by one, which we can write in the form $\frac{d}{dt}\mathbf{X}(t) = M\mathbf{X}(t) + \mathbf{f}$, where

$$\mathbf{X}(t) = \begin{pmatrix} [AR_1\text{-}R_2] \\ [R_1\text{-}R_2A] \\ [AR_1\text{-}R_2A] \end{pmatrix}, \qquad \mathbf{f} = \begin{pmatrix} k_{a_1+}[A]R_{tot} \\ k_{a_2+}[A]R_{tot} \\ 0 \end{pmatrix}, \tag{2.46}$$

and

$$M = \begin{pmatrix} -(k_{a_1-}[A] + k_{a_1-} + \alpha_+ k_{a_2+}[A]) & -k_{a_1+}[A] & \alpha_- k_{a_2-} - k_{a_1+}[A] \\ -k_{a_2+}[A] & -(k_{a_2+}[A] + k_{a_2-} + \alpha_+ k_{a_1+}[A]) & \alpha_- k_{a_1-} - k_{a_2+}[A] \\ \alpha_+ k_{a_2+}[A] & \alpha_+ k_{a_1+}[A] & -(\alpha_- k_{a_1-} + \alpha_- k_{a_2-}) \end{pmatrix}, \tag{2.47}$$

The system has the initial conditions

$$\begin{pmatrix} [AR_1\text{-}R_2] \\ [R_1\text{-}R_2A] \\ [AR_1\text{-}R_2A] \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{2.48}$$

2.3.1 Equilibrium analysis

Investigating the behaviour of the system at equilibrium allows us to see the effect of α. At equilibrium we have the relationships

$$[AR_1\text{-}R_2] = K_{A1}[A][R_1\text{-}R_2], \tag{2.49a}$$

$$[R_1\text{-}R_2A] = K_{A2}[A][R_1\text{-}R_2], \tag{2.49b}$$

$$[AR_1\text{-}R_2A] = \alpha K_{A2}[A][AR_1\text{-}R_2] = \alpha K_{A1}K_{A2}[A]^2[R_1\text{-}R_2], \tag{2.49c}$$

where $K_{A1} = k_{a_1+}/k_{a_1-}$ and $K_{A2} = k_{a_2+}/k_{a_2-}$ are the equilibrium association constants. The total concentration of dimers can be given as

$$R_{tot} = [R_1\text{-}R_2] + [AR_1\text{-}R_2] + [R_1\text{-}R_2A] + [AR_1\text{-}R_2A] \tag{2.50}$$

$$= [R_1\text{-}R_2](1 + K_{A1}[A] + K_{A2}[A] + \alpha K_{A1}K_{A2}[A]^2), \tag{2.51}$$

which we combine with equations (2.49) to express the concentrations in terms of parameters only, as

$$[R_1\text{-}R_2] = \frac{1}{(1 + K_{A1}[A] + K_{A2}[A] + \alpha K_{A1}K_{A2}[A]^2)} R_{tot}, \tag{2.52a}$$

$$[AR_1\text{-}R_2] = \frac{K_{A1}[A]}{(1 + K_{A1}[A] + K_{A2}[A] + \alpha K_{A1}K_{A2}[A]^2)} R_{tot}, \tag{2.52b}$$

$$[R_1\text{-}R_2A] = \frac{K_{A2}[A]}{(1 + K_{A1}[A] + K_{A2}[A] + \alpha K_{A1}K_{A2}[A]^2)} R_{tot}, \tag{2.52c}$$

$$[AR_1\text{-}R_2A] = \frac{\alpha K_{A1}K_{A2}[A]^2}{(1 + K_{A1}[A] + K_{A2}[A] + \alpha K_{A1}K_{A2}[A]^2)} R_{tot}. \tag{2.52d}$$

From these it is clear that as $[A] \to \infty$ the concentrations $[R_1\text{-}R_2], [AR_1\text{-}R_2], [R_1\text{-}R_2A] \to 0$ and $[AR_1\text{-}R_2A] \to R_{tot}$. As the primary interest is the total amount of ligand bound, due to it generally being a measurable quantity in experiments. At equilibrium this is

$$A_{bound} = [AR_1\text{-}R_2] + [R_1\text{-}R_2A] + 2[AR_1\text{-}R_2A] \tag{2.53}$$

$$= \frac{K_{A1}[A] + K_{A2}[A] + 2\alpha K_{A1}K_{A2}[A]^2}{(1 + K_{A1}[A] + K_{A2}[A] + \alpha K_{A1}K_{A2}[A]^2)} R_{tot}. \tag{2.54}$$

As $[A] \to \infty$, $A_{bound} \to 2R_{tot}$ hence, we calculate

$$A_{50} = \sqrt{\frac{1}{\alpha K_{A1}K_{A2}}}. \tag{2.55}$$

In Figure 2.3.2 we look at the curves of A_{bound} along with the individual species curves to study the effect the equilibrium cooperativity factor α has on the binding relationships. We fix association constants with $K_{A1} > K_{A2}$. It is clear in this figure that the curves are similar in shape to the logDR curves in the homodimer model. As α increases the curves shift towards the left, with a higher concentration of ligand required to reach equilibrium. We also notice we again have inflections in the A_{bound} curve for low cooperativity values, in much the same way as we did in the homodimer model. We also point out that, although the curves of $[AR_1\text{-}R_2]$ and $[R_1\text{-}R_2A]$ are identical in shape, having $K_{A2} < K_{A1}$

results in a much lower concentration of $[R_1\text{-}R_2A]$.

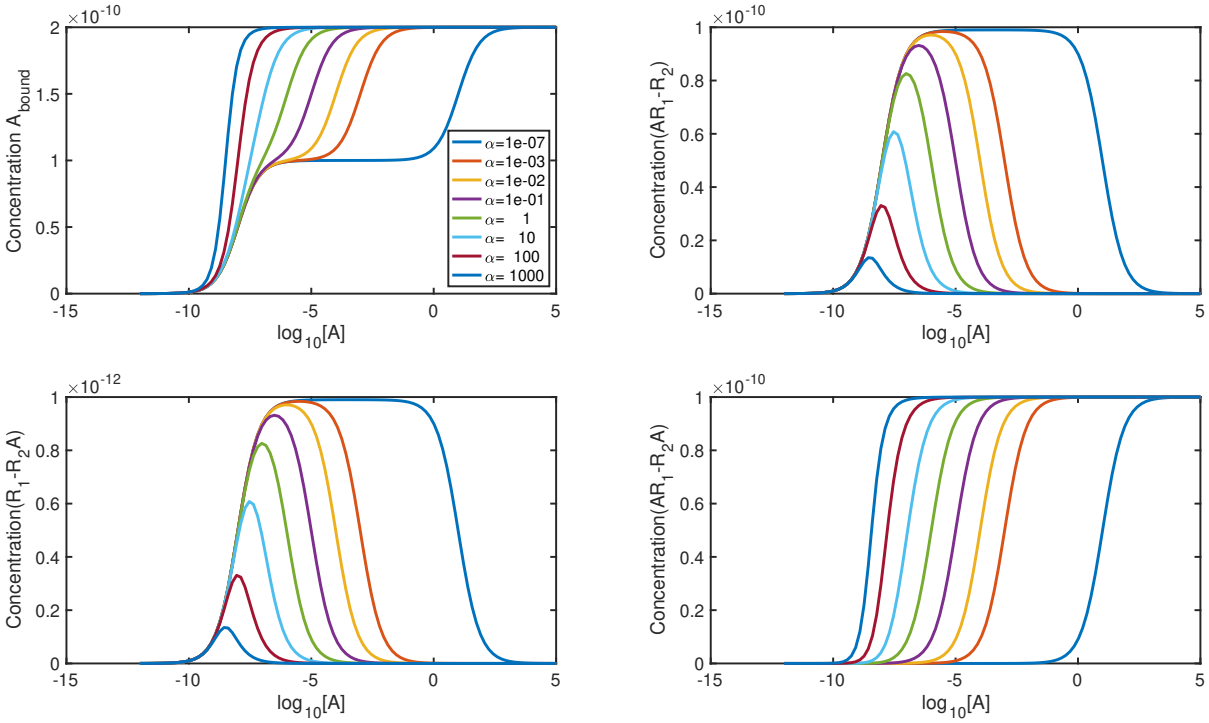

Figure 2.3.2: Plotting logDR curves, using equation (2.54), for varying α show extra inflections appear for heterodimers, similar to those in the homodimer binding curves.

To determine the point of these extra inflections and derive a condition under which these appear we again follow the analysis outlined in Appendix B (with $a = K_{A1} + K_{A2}, b = \alpha K_{A1}K_{A2}, c = 1$ and $X = [A]$). We have an inflection point that is always present at

$$[A] = \sqrt{\frac{1}{\alpha K_{A1}K_{A2}}}. \tag{2.56}$$

We further find that extra inflections appear at the points

$$[A]_{\pm} = \frac{-(8\alpha K_{A1}K_{A2} - (K_{A1} + K_{A2})^2) \pm \sqrt{((K_{A1} + K_{A2})^2 - 16\alpha K_{A1}K_{A2})((K_{A1} + K_{A2})^2 - 4\alpha K_{A1}K_{A2})}}{2\alpha K_{A1}K_{A2}(K_{A1} + K_{A2})}, \tag{2.57}$$

if the following holds

$$(K_{A1} + K_{A2})^2 > 16\alpha K_{A1} K_{A2}. \tag{2.58}$$

Under these conditions we see that, as in the homodimer model, the inflection at the A_{50} value changes from a rising to a falling inflection and we get two extra inflections, one at either side of the A_{50} value.

2.3.2 Single ligand time course results

Here we move on to consider the effects cooperativity has on time course binding dynamics of the system, and observe how having a system consisting of heterodimers differs from that of a homodimer system. In theory, analytical solutions can be found, however these are impractical to compute, so we use numerical methods to create simulations of the results.

In Figure 2.3.3 we fix $k_{a_1+} \gg k_{a_2+}$ and $k_{a_1-} \gg k_{a_2-}$, and cooperativity is fixed to

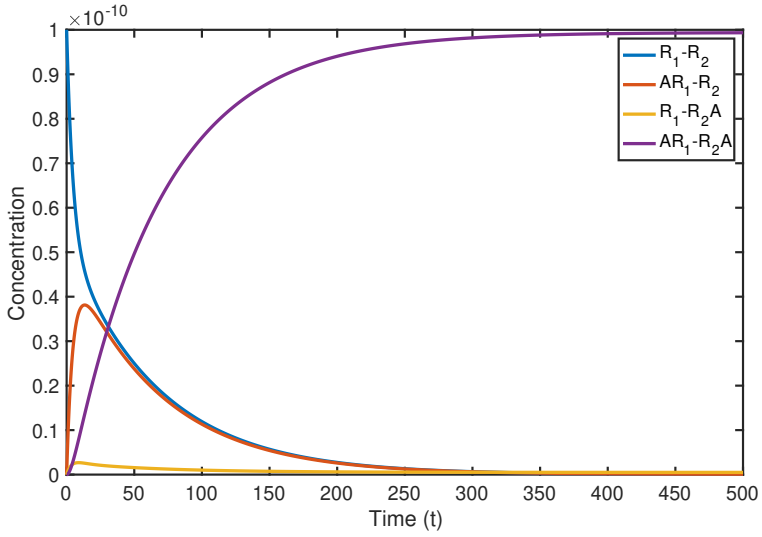

Figure 2.3.3: Plotting the numerical solution to equations (2.44) shows peaks appearing in the time course curves for $[AR_1\text{-}R_2]$ and $[R_1\text{-}R_2A]$. Cooperativity values used for the plot are $\alpha_+ = 2$ and $\alpha_- = 0.01$, while ligand concentration is $[A] = 10^{-8}M$.

give positive cooperativity, though we look at other possibilities later. For the purposes of discussion we describe R_1 receptors as being on the left side of the dimer with R_2

receptors on the right. The curves have clear similarities to the homodimer results (as in Figure 2.1.3), in that we have peaks in the singularly bound receptors followed by an increase in dual bound dimers. Though, as $k_{a_1+} \gg k_{a_2+}$ we initially see a much larger peak in $[AR_1\text{-}R_2]$ while $[R_1\text{-}R_2A]$ remains lower.

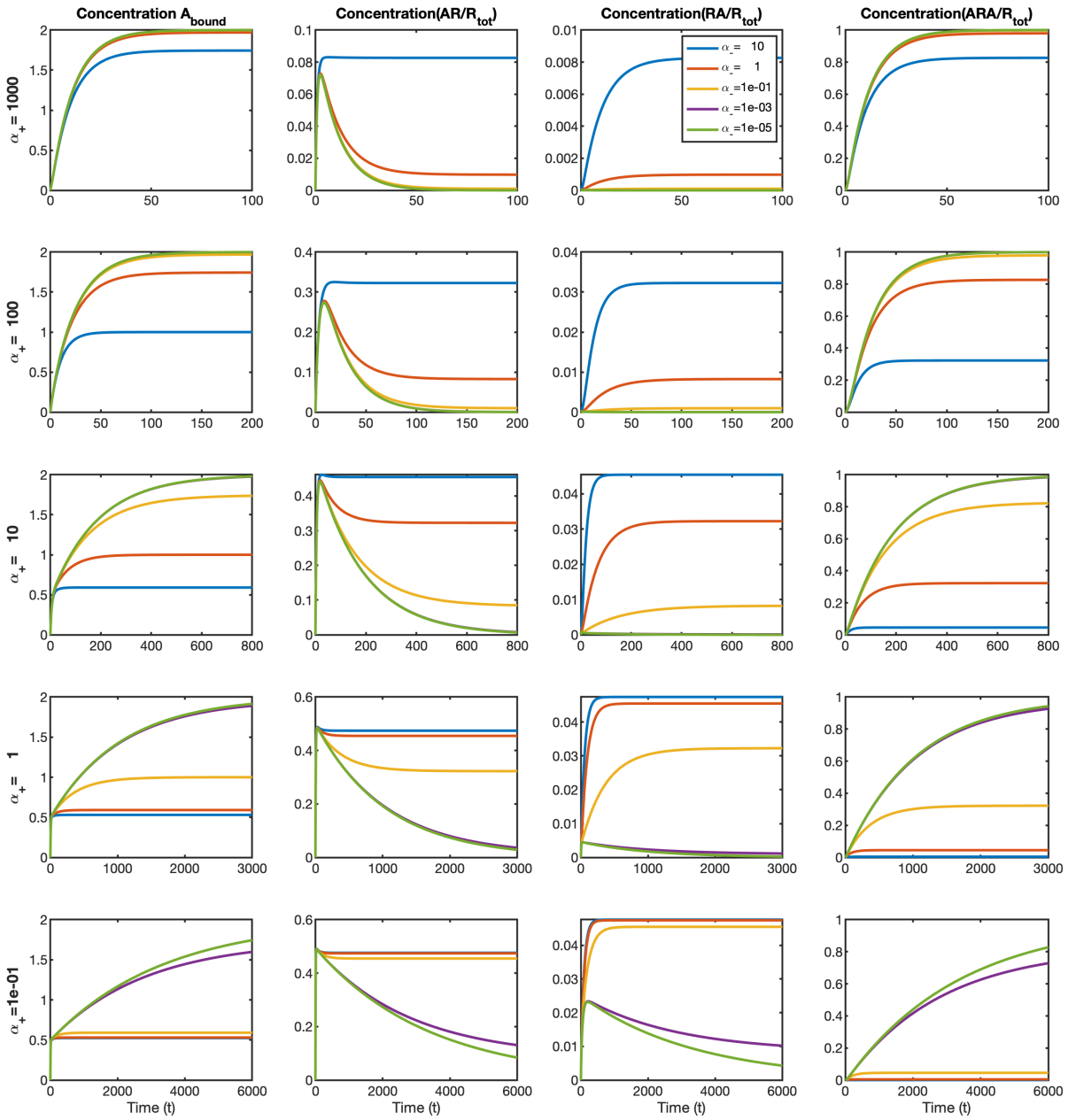

Figure 2.3.4: Plotting the numerical solution to equations (2.44) for a range of both cooperativity factors (α_+ and α_-) result in peaks appearing in some of the $[AR_1\text{-}R_2]$ and $[R_1\text{-}R_2A]$ curves. Ligand concentration used is $[A] = 10^{-8}M$.

In Figure 2.3.4 we see a number effects arising from varying the cooperativity factors

α_+ and α_-. In the A_{bound} curves we see that increasing α_+ results in an increase in the total concentration of ligand bound by equilibrium, while decreasing α_- gives the same effect. Peaks form in the curves of both $[AR_1\text{-}R_2]$ andf $[R_1\text{-}R_2A]$, predominantly when α_- is small, although α_+ also plays a role in the effect. The rate of increase in $[AR_1\text{-}R_2]$ is k_{a_1+}, hence this initially increases the same for all cooperativity values. With an increased α_+ these quickly become dual bound, hence $[AR_1\text{-}R_2]$ falls, and we see an increase in $[AR_1\text{-}R_2A]$. Increasing α_+ results in the peak and then fall happening more quickly and therefore a smaller peak also. This has the added effect that the system then also reaches equilibrium more quickly. The same argument also holds for $[R_1\text{-}R_2A]$, although as $k_{a_1+} > k_{a_2+}$ and $k_{a_1-} > k_{a_2-}$ the overall concentration levels are lower. However, if α_- is increased, when the dimers become dual bound one of the bound ligand molecules quickly become unbound, hence we do not see the fall in $[AR_1\text{-}R_2]$ and $[R_1\text{-}R_2A]$, and instead $[ARA]$ remains low.

2.4 Conclusions

In this Chapter, we have presented dynamic models of ligand binding to pre-dimerised GPCR homodimers, for both a single ligand and two-ligand competition, and single ligand heterodimers. The models are linear ODE systems, allowing analytical solutions for time-dependent and equilibrium responses. The model formulation, solution and results serve as a contribution to the field of pharmacological modelling, and are expected to be of practical use, given the ease with which we can compute solutions. In particular, the bi-exponential single-ligand binding kinetics show a similar model solution structure to the widely-used Motulsky-Mahan model for competition binding at monomers [95]. We therefore propose that these models can be adopted, interpreted and implemented with relative ease by pharmacologists, and as such we have provided a recipe for computational solution.

There are interesting features in the equilibrium logDR curves for the experimental readout (signal) A_{bound}. We have noted the possibility for multiple inflections in logDR

curves, for all models. Multiphasic features in logDR curves are not reproducible by standard Hill functions [32] which only support single-inflection curves. For single-ligand binding, multiphasic logDR curves theoretically rule out monomeric receptor binding as the only ligand binding mechanism; such experimental data therefore would suggest another binding setup, possibly due to dimeric receptors. The existence of multiple inflections depends on the level of cooperativity across a dimerised receptor, and importantly we are able to give conditions on cooperativity factors for a three-inflection curve. The practical use of these conditions is clearly in assessing the sign and magnitude of cooperativity towards quantitative classification of drug-receptor interactions. Given that Hill functions are not suitable for fully characterising multiphasic logDR curves [32], our analysis here goes beyond dimer cooperativity indices which stem in part from cooperativity in the Hill function sense [44]; the present work concerns mechanistic cooperativity which is explicit in the original model schematic, as opposed to empirical measures from a more limited model. In [32], it is noted that multiphasic logDR curves have particular importance in a number of contexts including cancer pharmacology, and effort should be made to move beyond Hill function fitting wherever possible.

Chapter 3

Modelling GPCR binding and activation

Dimeric receptors can theoretically elicit a diverse range of pharmacological effects [91]. With the implications of these on downstream signalling yet to be fully explored, many questions remain, such as, how ligand-receptor crosstalk can be distinguished from downstream crosstalk when interpreting experimental results [92, 93, 22]. The logical next steps in the modelling of dimeric receptors and exploring these possible downstream effects involves extending our GPCR models to include receptor activation [44], G protein binding [93, 6], and ultimately G protein activation and cycling [137, 13].

In this chapter, to begin understanding the downstream effects we extend the GPCR binding models we presented in Chapter 2 to include receptor activation. There are currently few mathematical models of GPCR binding and activation where receptors have formed dimers. Franco *et al.* [40] present a model for the binding and activation of homodimers, where they consider each receptor within the dimer to have its own orthosteric binding site, however, activation occurs for the unit as whole. In [141], Zhou and Giraldo extend the canonical operational model to include homodimers. In their paper they simulate and analyse the functional effect curves seen at equilibrium. Rovira *et al* [109] present a model that describes the binding and activation of homodimers. In their model, they define a dimer with only one protomer active as 'asymmetric', while a dimer with both protomers active as 'symmetric'. They propose that dimers in an asymmetric

state activate a different pathway to those in a symmetric state, hence, compare the signalling outcomes for the two individual pathways. In contrast to the models in the literature, we now present a more complete model, free of any restrictive assumptions.

3.1 The model

The schematic of the model we present is shown in Figure 3.1.1. In this model, we consider a single ligand, A, binding to a pre-dimerised receptor complex. We assume all dimers are homodimers, so the receptors within the complex are identical. Each protomer within the dimer has the ability to bind a ligand molecule and/or activate individually.

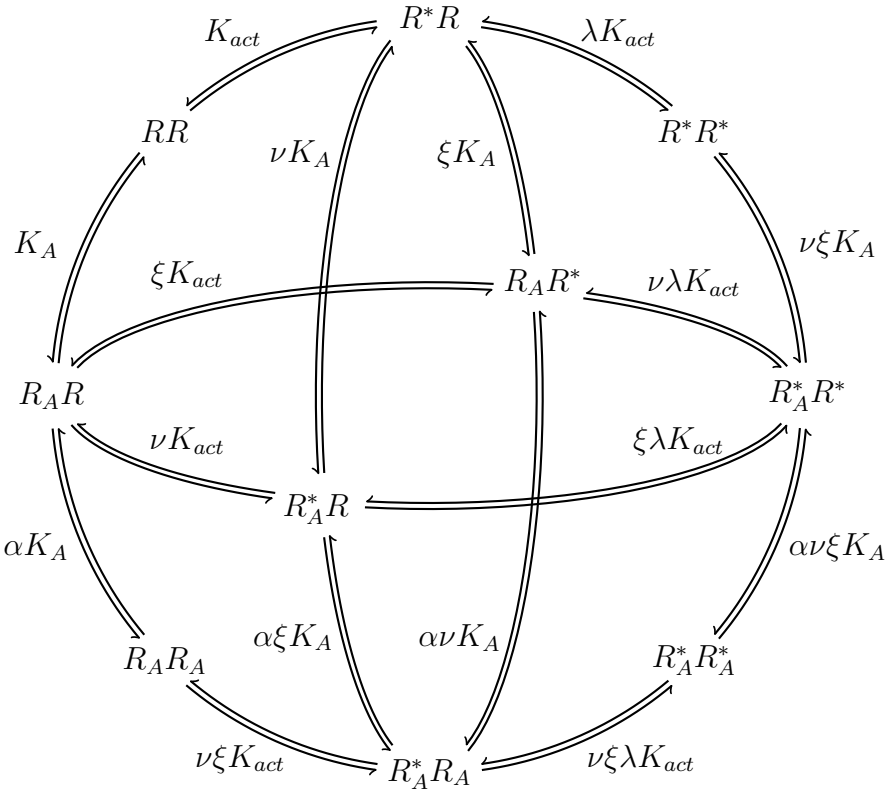

Figure 3.1.1: The binding of a ligand A, and activation of pre-dimerised receptors show a spherical schematic describing interactions with ten dimer complexes. Binding is determined by K_A, while activation K_{act}. Cooperativity factors α, λ, ν and ξ describe crosstalk and the more complex interactions.

3.1.1 Notation

The notation we have used to describe the state of each species is as follows:

- Each protomer is denoted by R, hence an unbound, inactive dimer is RR.
- A subscripted A attached to an R indicates that the receptor is bound by a ligand molecule. As such R_AR implies that one of the protomers in the dimer is bound, while R_AR_A represents a dimer that is dual bound. We assume there is no preference for which protomer the ligand molecule binds to first.
- Ligand association and dissociation kinetic rate constants, for inactive receptors are k_{a+} and k_{a-} respectively, and $K_A = k_{a+}/k_{a-}$ is the equilibrium binding constant.
- An active receptor is represented by an asterisk, hence, R^*R indicates that one receptor within the dimer is active. Similarly, R^*R^* is a dimer with both protomers active. We follow the work of Rovira *et al* and denote these 'asymmetric' and 'symmetric'.
- We have k_{act+} and k_{act-} as the activation and deactivation rates, and $K_{act} = k_{act+}/k_{act-}$ is the equilibrium activation constant. These account for the basal activation of receptors in the absence of ligand, otherwise known as constitutive activity.

As the dimer consists of two identical receptors we assume symmetry throughout, hence R_AR, for example, describes a dimer with a single molecule bound, regardless of which side it is bound to. Furthermore, for simplicity of notation, we order the receptors within the dimer such that a receptor that is bound will be listed first, followed by an active receptor, then finally a free receptor. This means that, a dimer that has one protomer bound by ligand, and the other active, will be written R_AR^*. We highlight however, the reaction

$$A + R^*R \rightleftharpoons R_AR^*$$

to make note of the ordering. This reaction simply binds a ligand molecule to the inactive protomer, giving one active unbound protomer and one inactive bound protomer within the dimer. Whilst before binding ordering is such that the active receptor is on the left, once binding occurs, ordering dictates that the bound receptor is written first, hence the active receptor appears on the right. This ordering is simply a necessity in order to exploit the symmetry of the dimer.

3.1.2 Cooperativity factors

We also require some extra cooperativity factors to fully describe all binding and activation processes:

- The parameter α remains the same as in Figure 2.1.1, in Chapter 2, where its effects were explored extensively. It represents equilibrium binding cooperativity, that is the change in ligand affinity for the dimer when it is already ligand-bound. The value $\alpha = 1$ represents neutral cooperativity, and $\alpha > 1$ and $\alpha < 1$ represent positive and negative cooperativity respectively. With positive cooperativity, less ligand is required for the dimer to become bound by a second molecule, due to the dimer already being bound, and more ligand is needed for negative cooperativity.

- Similarly we have the equilibrium activation cooperativity factor λ which allows for the change in propensity for receptor activation of one receptor within the dimer when the other side is already active. With $\lambda > 1$ there is an increased propensity for activation of the second protomer, while the converse is true if $\lambda < 1$.

- The parameter ν represents the intrinsic efficacy of the ligand and accounts for the change in propensity for receptor activation when ligand is bound as well as change in affinity of the ligand for active protomers over inactive protomers, where detailed balance states that these must be equal. This parameter specifically represents these changes when ligand binding and activation occur on the same receptor within the dimer. Hence, with $\nu > 1$, if a protomer is ligand bound, there is an increased propensity that the same protomer will become active. Similarly, for $\nu > 1$, if a protomer is active it will have a higher affinity for the ligand. The parameter ν

has no effect on the binding and activation rates of the opposite receptor within the dimer. In the two-state model by Leff [76], the parameter K_A^* is defined as an alternate binding rate grouping the cooperativity factor and binding constant together.

– The parameter ξ accounts for the change in propensity for receptor activation when ligand is bound as well as change in affinity of the ligand, where ligand binding and activation occur on opposite receptors. For example, with $\xi > 1$, when one protomer becomes bound, the propensity for activation of the other protomer is increased, and similarly activation of one of the protomers increases the affinity for ligand binding to the opposite protomer.

3.2 Equilibrium analysis

We explore the behaviour of the system, particularly focusing on the effects of the parameters α, λ, ν and ξ, at equilibrium. The equilibrium relationships are

$$\begin{aligned}
[R_A R] &= K_A[A][RR], & [R^* R^*] &= \lambda {K_{act}}^2[RR], \\
[R^* R] &= K_{act}[RR], & [R_A^* R_A] &= \alpha\nu\xi {K_A}^2[A]^2 K_{act}[RR], \\
[R_A R_A] &= \alpha {K_A}^2[A]^2[RR], & [R_A^* R^*] &= \nu\lambda\xi K_A {K_{act}}^2[A][RR], \\
[R_A^* R] &= \nu K_A K_{act}[A][RR], & [R_A^* R_A^*] &= \alpha\nu^2\lambda\xi^2 {K_A}^2 {K_{act}}^2[A]^2[RR]. \\
[R_A R^*] &= \xi K_A K_{act}[A][RR], &&
\end{aligned} \tag{3.1}$$

The total concentration of receptors, a conserved quantity, can be stated as

$$R_{tot} = [RR]+[R_A R]+[R^* R]+[R_A R_A]+[R_A^* R]+[R_A R^*]+[R^* R^*]+[R_A^* R_A]+[R_A^* R^*]+[R_A^* R_A^*], \tag{3.2}$$

which we combine with equations (3.1) to give the equilibrium concentration of each species in terms of model parameters as

$$\begin{aligned}
[RR] &= \frac{R_{tot}}{\sigma}, & [R_AR^*] &= \frac{\xi K_A K_{act}[A]R_{tot}}{\sigma},\\
[R_AR] &= \frac{K_A[A]R_{tot}}{\sigma}, & [R^*R^*] &= \frac{\lambda {K_{act}}^2 R_{tot}}{\sigma},\\
[R^*R] &= \frac{K_{act}R_{tot}}{\sigma}, & [R_A^*R_A] &= \frac{\alpha\nu\xi {K_A}^2[A]^2 K_{act}R_{tot}}{\sigma},\\
[R_AR_A] &= \frac{\alpha {K_A}^2[A]^2 R_{tot}}{\sigma}, & [R_A^*R^*] &= \frac{\nu\lambda\xi K_A {K_{act}}^2[A]R_{tot}}{\sigma},\\
[R_A^*R] &= \frac{\nu K_A K_{act}[A]R_{tot}}{\sigma}, & [R_A^*R_A^*] &= \frac{\alpha\nu^2\lambda\xi^2 {K_A}^2 {K_{act}}^2[A]^2 R_{tot}}{\sigma}.
\end{aligned} \tag{3.3}$$

where

$$\begin{aligned}
\sigma = 1 + K_{act} + \lambda K_{act}^2 + K_A[A] + (\nu+\xi)K_AK_{act}[A] + \nu\xi\lambda K_AK_{act}^2[A]\\
+ \alpha K_A^2[A]^2 + \alpha\nu\xi K_A^2K_{act}[A]^2 + \alpha\nu^2\xi^2\lambda K_A^2K_{act}^2[A]^2.
\end{aligned} \tag{3.4}$$

It is well documented that, receptors in an active state elicit a biochemical response [76]. Although this response will also depend on the downstream signalling pathway, here we define a cursory signal as simply the response coming from an active receptor. However, it is unknown whether a dimer with both protomers active simply gives twice the signal as a single active protomer, or whether other factors may affect this. So in deriving a signal, we introduce an extra parameter ω to account for this. If $\omega = 2$, we assume that the signal arising from a symmetric dimer is exactly twice the signal of an asymmetric dimer, while $\omega = 1$ assumes that, although both protomers within the dimer are active, the signal being transmitted is unchanged by the activation of the second receptor. We therefore state the proposed signal as

$$S = [R^*R] + [R_A^*R] + [R_AR^*] + [R_A^*R_A] + \omega([R^*R^*] + [R_A^*R^*] + [R_A^*R_A^*]). \tag{3.5}$$

This gives the equilibrium signal as

$$S_{eq} = \frac{\begin{array}{r} K_{act} + \nu K_A K_{act}[A] + \xi K_A K_{act}[A] + \alpha\nu\xi {K_A}^2 K_{act}[A]^2 \\ + \omega(\lambda {K_{act}}^2 + \nu\xi\lambda K_A {K_{act}}^2[A] + \alpha\nu^2\xi^2\lambda {K_A}^2 {K_{act}}^2[A]^2) \end{array}}{\begin{array}{r} 1 + K_{act} + \lambda K_{act}^2 + K_A[A] + (\nu+\xi)K_A K_{act}[A] + \nu\xi\lambda K_A K_{act}^2[A] \\ + \alpha K_A^2[A]^2 + \alpha\nu\xi K_A^2 K_{act}[A]^2 + \alpha\nu^2\xi^2\lambda K_A^2 K_{act}^2[A]^2 \end{array}} R_{tot}. \quad (3.6)$$

which we will use to analyse the effect each cooperativity factor has on the dose-response curves.

3.3 Dose-response curves: single parameter effects

We now explore how the parameters α, λ, ν and ξ affect the dose-response relationship. In the following sections, we vary each of these individually while keeping the others fixed, to isolate the effects of a single parameter. The parameters K_A, K_{act} and R_{tot} are fixed in each case with values taken from Woodroffe [138], though we note that these are taken for illustrative computations only, as the values given by Woodroffe are for monomers. All parameter values can be found in Table A.4 in Appendix A. The weighting parameter in the signal is set to $\omega = 2$, so that the signal given from a symmetric dimer is twice that of an asymmetric dimer, although we also later discuss the possible effects of this parameter being other values. All dose-response plots are normalised with respect to R_{tot}, and as such, when referring to the signal, or S_{eq}, we are assuming this to be the normalised signal. All plots in this chapter were created using equation (3.6) for the overall signal and equations (3.3) for the individual species plots.

3.3.1 Varying binding cooperativity parameter α

The first parameter we analyse the effects of is the binding cooperativity factor, α. To do this we fix all other cooperativity factors as $\lambda = \nu = \xi = 1$, to focus on the effects of α alone. Although we note, that these values indicate that the ligand is an antagonist and also constitutive activity is low, and so the signal will also be low. However, fixing the values in this way allows us to isolate the effects of α in order to gain a fuller understand-

ing of the potential consequences of having positive or negative binding cooperativity. We later vary all parameters simultaneously to explore a range of possible signal outcomes.

Before analysing the signal curves, and as α is a binding parameter, we first comment on the amount ligand bound. This allows us to directly compare the results with the previous binding results, in Section 2.1, where we recall the concentration of ligand bound, scaled against R_{tot}, was stated, in equation (2.9), as

$$\frac{A^b_{bound}}{R_{tot}} = \frac{K_A[A] + 2\alpha K_A^2[A]^2}{1 + K_A[A] + \alpha K_A^2[A]^2}. \tag{3.7}$$

The superscript b indicates that the expression is from the binding model. To find the equivalent expression for this model we use the expressions that were given in equations (3.1), while setting $\lambda = \nu = \xi = 1$. Noting that

$$\frac{A^a_{bound}}{R_{tot}} = [R_A R] + [R^*_A R] + [R_A R^*] + [R^*_A R_A] + 2([R_A R_A] + [R^*_A R_A] + [R^*_A R^*_A]), \tag{3.8}$$

where the superscript a indicates the activation model, we calculate

$$\frac{A^a_{bound}}{R_{tot}} = \frac{(1 + 2K_{act} + K^2_{act})K_A[A] + 2\alpha(1 + K_{act} + K^2_{act})K_A^2[A]^2)}{1 + K_{act} + K^2_{act} + (1 + 2K_{act} + K^2_{act})K_A[A] + \alpha(1 + K_{act} + K^2_{act})K_A^2[A]^2}. \tag{3.9}$$

We note that

$$1 + 2K_{act} + K^2_{act} \approx 1 + K_{act} + K^2_{act}, \tag{3.10}$$

as $K_{act} \ll 1$ (although this also holds for $K_{act} \gg 1$), which then gives

$$A^a_{bound} \approx A^b_{bound}. \tag{3.11}$$

Hence, we can conclude that the concentration of ligand bound in the two models is, for all intents and purposes, approximately equal. This is also confirmed in Figure 3.3.1 where we plot dose-response curves showing ligand bound, using the expressions from both the previous binding model and this activation model.

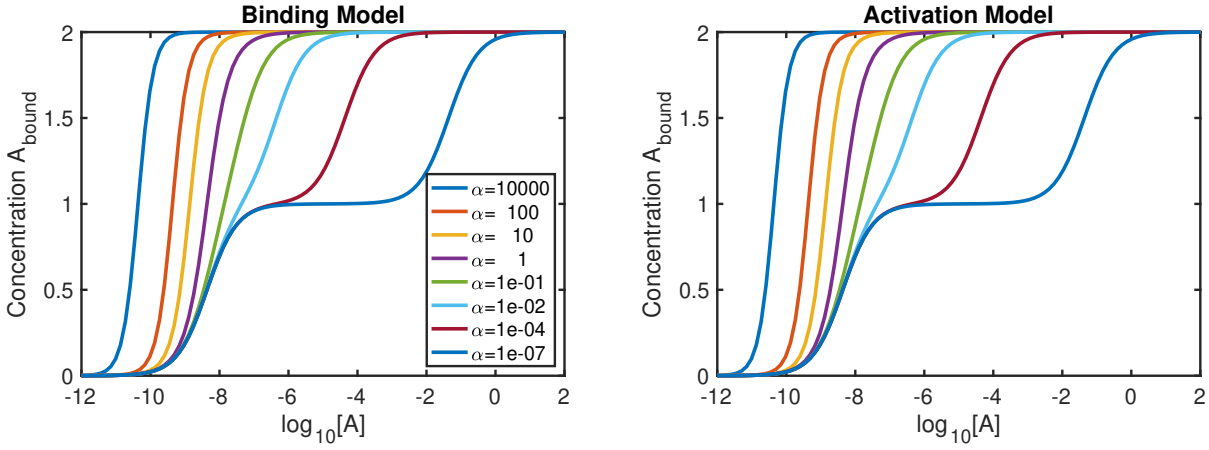

Figure 3.3.1: The logDR A_{bound} curve for this activation model (equation (3.8)) is plotted alongside the logDR A_{bound} curve for the binding model (Section 2.1, equation (3.9)). With equal binding parameters we see that the curves appear to be the equal.

In Figure 3.3.2 we explore the signalling implications of a varied α while fixing $\lambda = \nu = \xi = 1$. On the x axis we have ligand concentration, on a log scale. On the top row, we have the overall signal plotted beside the A_{bound} plot. Underneath these we plot dose-response curves for each individual species in order to assist in understanding the signal plot.

Unlike the sigmoidal shape we see for monomer dose-response curves (as we discussed in Chapter 1), we instead have a bell-shaped signal curve. The signal increases as the ligand concentration increases, and then falls back to the basal level. Research has shown that this can be indicative of non-standard monomer-ligand binding [100]. To understand why this effect occurs we look at the individual species curves.

While the ligand concentration is very low, the signal remains close to the basal level, with most dimers being in the form of either RR, R^*R or R^*R^*. As the ligand in the system increases we first see an increase in singularly bound dimers, that is, $[R_AR]$,

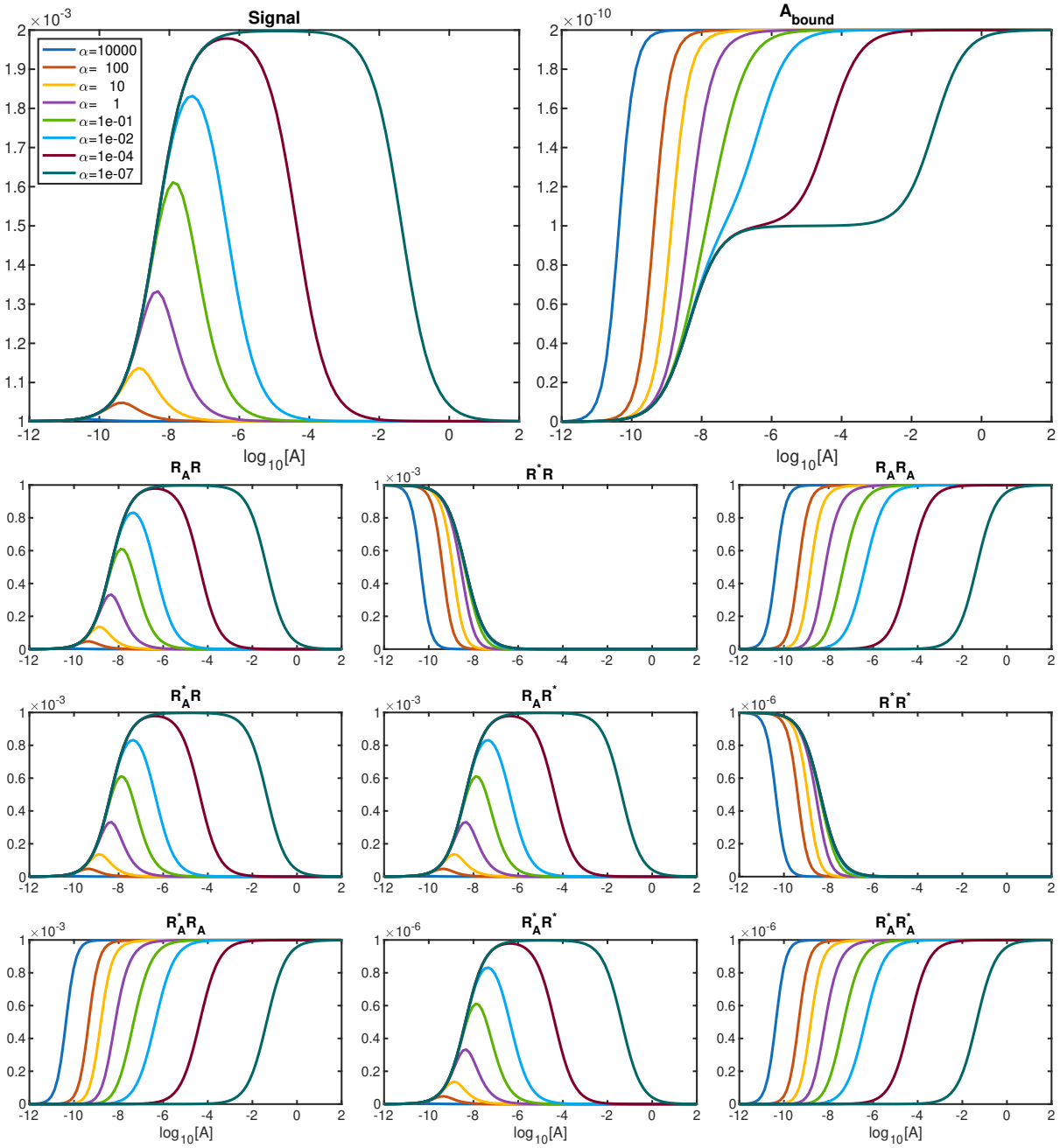

Figure 3.3.2: Plotting a logDR curve (equation (3.12)) for a range of values for binding cooperativity factor α. While all signal curves have a bell shape, increasing α decreases the peak size.

$[R_A^*R]$, $[R_AR^*]$ and $[R_A^*R^*]$. This is followed by a fall in these same concentrations once all dimers become dual bound by the ligand, thus causing a peak in singularly bound species concentrations. This phenomenon was explored extensively in Chapter 2, where we saw that low cooperativity exaggerated the effect, thereby causing a larger peak in singularly bound dimers. Cooperativity values that caused this larger peak also gave an

extra inflection in the total bound ligand curve. In fact, with this A_{bound} plot (top right) alongside the signal (top left), we see how the extra inflection correlates to the largest peaks in the signal curve.

The bell shape in the signal curve comes from the concentrations $[R_A^*R]$ and $[R_AR^*]$. Notice that, when binding cooperativity is taken to be extremely low, the peak value of the signal is exactly double that of the basal level. This low cooperativity, combined with sufficient ligand, means most dimers become trapped in a form with only one receptor in each dimer bound. There are then two ways a singularly bound dimer can be also active (and therefore eliciting a signal), hence we have increases in both $[R_A^*R]$ and $[R_AR^*]$, which leads to the signal doubling in this phase.

This can also be seen mathematically. Taking the signal expression, as in equation (3.6), with $\nu = \xi = \lambda = 1$ and $\omega = 2$ gives

$$S_{eq} = \frac{K_{act}(1 + 2K_{act} + 2(1 + K_{act})K_A[A] + \alpha(1 + 2K_{act})K_A^2[A]^2)}{1 + K_{act} + K_{act}^2 + (1 + K_{act})^2K_A[A] + \alpha(1 + K_{act} + K_{act}^2)K_A^2[A]^2}. \tag{3.12}$$

Differentiating this with respect to $[A]$ gives

$$\frac{dS_{eq}}{d[A]} = \frac{K_AK_{act}(1 - K_{act}^2)(1 - \alpha K_A^2[A]^2)}{(1 + K_{act} + K_{act}^2 + (1 + K_{act})^2K_A[A] + \alpha(1 + K_{act} + K_{act}^2)K_A^2[A]^2)^2}, \tag{3.13}$$

which gives a single stationary point at

$$[A] = \frac{1}{K_A\sqrt{\alpha}}. \tag{3.14}$$

This is clearly identical to the $[A]_{50}$ value and point of inflection in the A_{bound} curve given in Chapter 2 when considering single ligand binding, hence the inflection in the A_{bound} curve directly relates to the peak in the signal curve. Substituting this back into equation

(3.12) gives the signal at this point as

$$S_{eq}\Big|_{[A]=\frac{1}{K_A\sqrt{\alpha}}} = \frac{2K_{act}(1+K_{act}+\sqrt{\alpha}(1+2K_{act})}{(1+K_{act})^2+2\sqrt{\alpha}(1+K_{act}+K_{act}^2)}\ . \tag{3.15}$$

As the signal appears to reach a maximum value when the binding cooperativity factor is small, we let $\alpha \to 0$ to find this maximum peak value as

$$\lim_{\alpha\to 0}\left(S_{eq}\Big|_{[A]=\frac{1}{K_A\sqrt{\alpha}}}\right) = \frac{2K_{act}}{1+K_{act}}\ . \tag{3.16}$$

Furthermore, taking the limit of the signal, as given in equation (3.12), as $[A] \to 0$, and also taking the limit as $[A] \to \infty$ gives

$$\lim_{[A]\to 0} S_{eq} = \lim_{[A]\to\infty} S_{eq} = \frac{K_{act}(1+2K_{act})}{1+K_{act}+K_{act}^2}, \tag{3.17}$$

confirming that at high ligand concentrations the signal falls to the same level as when ligand concentrations are low. One particular point of interest occurs when binding cooperativity is extremely low, enough so that the maximum peak is reached. With the parameters as we fixed them in Figure 3.3.2, the signal does not begin decreasing until the ligand concentration reaches $10^{-4}M$. Typically, concentrations used in experiments are much lower than this. This could potentially result in curves that appear to be monomeric, sigmoidal shaped curves. Not only would this lead to flawed assumptions being inferred, but would also result in the fitting to a model that is not representative of the biological system, and inaccurate binding estimates made.

3.3.2 Varying activation cooperativity parameter λ

We now fix $\alpha = \nu = \xi = 1$ and instead vary λ, to investigate the effects of the activation cooperativity factor. Again, these values mean signal is overall very low, but it is necessary to understand the effects of λ without influence from other cooperativity

values. In Figure 3.3.3 we again plot the overall signal curve, along with all individual species curves. In order to be able to see the details in the curves that are on a lower order concentration, we split the signal plot across two subplots. Clearly the shape of the signal curve is similar to that of the signal curve when we varied α (Figure 3.3.2), that is, a bell-shaped curve. However, we notice that varying λ appears to shift the whole curve, with an increased λ resulting in an increase in concentration at all points.

Again, the individual species curves give explanation for this. It is clear that varying λ has little effect on many of the curves. As the only reactions involving the parameter λ are those that involve the activation of a second protomer, it follows that the concentrations primarily affected by varying λ are the symmetric species, $[R^*R^*]$, $[R_A^*R^*]$ and $[R_A^*R_A^*]$. If we focus on these three concentrations, the reason for the changes in the signal curve become clear. While concentration of A is low, then it follows that $[R^*R^*]$ is high, while $[R_A^*R^*]$ and $[R_A^*R_A^*]$ are close to zero. As $[A]$ increases, we see an increase in $[R_A^*R^*]$ and a decrease in $[R^*R^*]$ as free dimers become bound by a single ligand molecule. As $[A]$ continues to increase, all dimers eventually become dual bound due to an abundance of ligand, hence we see an interchange of $[R_A^*R^*]$ and $[R_A^*R_A^*]$, thus creating the peak in $[R_A^*R^*]$.

As we increase λ, however, we see that this effect becomes more exaggerated, as regardless of the amount of ligand in the system, increasing λ means more asymmetric dimers becoming symmetric. If λ is high, more dimers are in the form $[R^*R^*]$ in the absence of ligand, hence the basal signal is increased with λ. Similarly, once the ligand concentration increases, a high λ leads to an increased peak in $[R_A^*R^*]$, as well as a higher concentration of $[R_A^*R_A^*]$ once there is sufficient ligand for all receptors to become bound.

To determine how much the signal increases at the peak, we consider the equilibrium

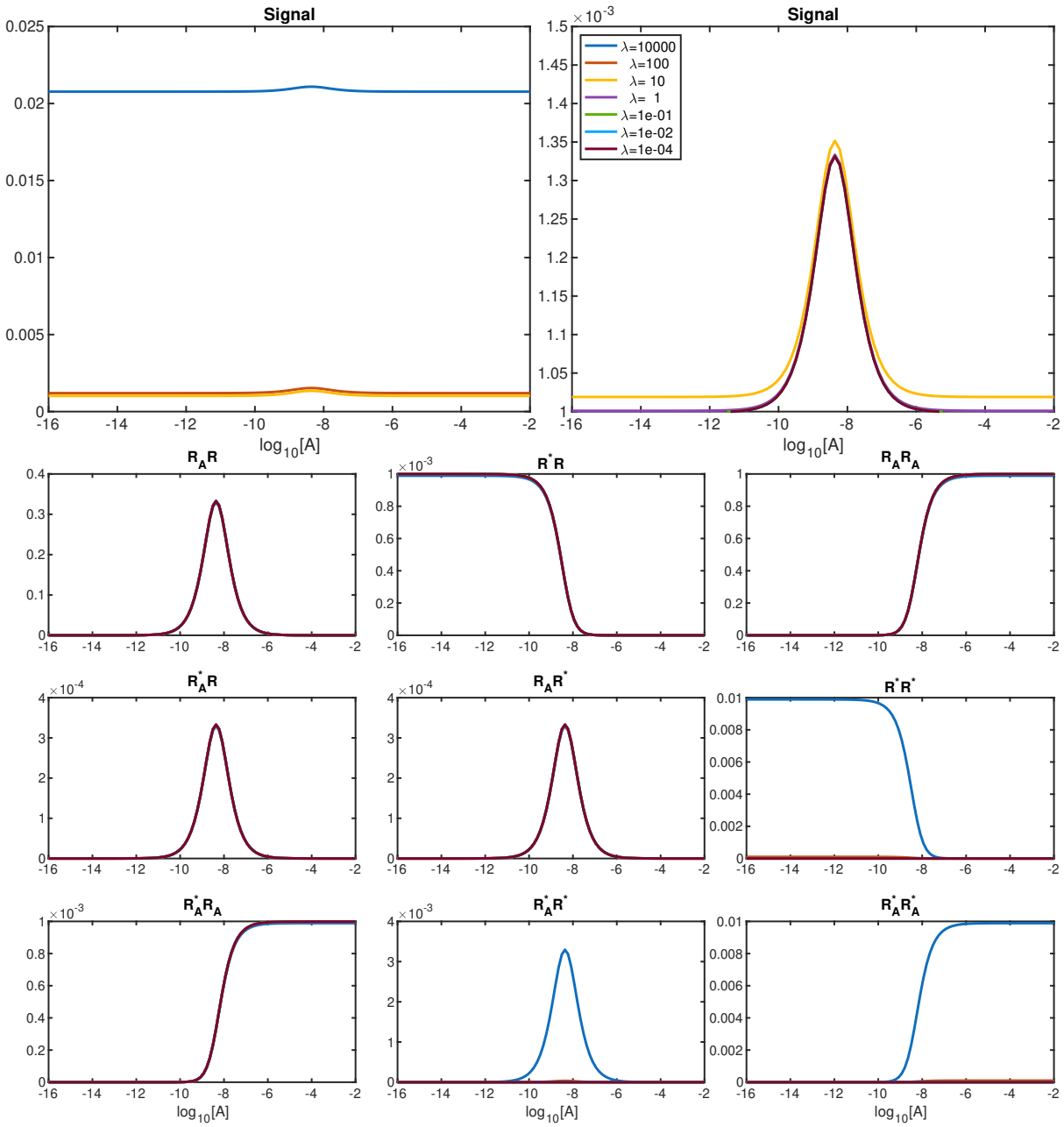

Figure 3.3.3: Plotting a logDR curve (equation (3.18)) for a range of values for activation cooperativity factor λ. Increasing λ increases the signal for all concentration of $[A]$.

signal, with $\alpha = \nu = \xi = 1$ and $\omega = 2$, which gives

$$S_{eq} = \frac{K_{act}(1 + 2\lambda K_{act} + 2(1 + \lambda K_{act})K_A[A] + (1 + 2\lambda K_{act}){K_A}^2[A]^2)}{1 + K_{act} + \lambda K_{act}^2 + (1 + 2K_{act} + \lambda K_{act}^2)K_A[A] + (1 + K_{act} + \lambda K_{act}^2)K_A^2[A]^2}. \tag{3.18}$$

First taking the limit as $[A] \to 0$, and also as $[A] \to \infty$ gives

$$\lim_{[A]\to 0} S_{eq} = \lim_{[A]\to\infty} S_{eq} = \frac{K_{act}(1 + 2\lambda K_{act})}{1 + K_{act} + \lambda K_{act}^2}, \tag{3.19}$$

confirms that the signal is equal at both zero and infinity. Furthermore, if we let $\lambda \to \infty$ in equation (3.18) we have

$$\lim_{\lambda\to\infty} S_{eq} = 2, \tag{3.20}$$

which verifies that, as equilibrium cooperativity increases, the signal tends to a constant, maximal signal.

We also use equation (3.18) to investigate the peak in the signal curve. Differentiating this with respect to $[A]$, gives

$$\frac{dS_{eq}}{d[A]} = \frac{K_A K_{act}(1 - K_A^2[A]^2)(1 - \lambda K_{act}^2)}{(1 + K_{act} + \lambda K_{act}^2 + (1 + 2K_{act} + \lambda K_{act}^2)K_A[A] + (1 + K_{act} + \lambda K_{act}^2)K_A^2[A]^2)^2}. \tag{3.21}$$

which gives a single stationary point at

$$[A] = \frac{1}{K_A}. \tag{3.22}$$

This confirms that the peak occurs at the same point regardless of the value of λ. Substituting equation (3.22) into equation (3.18) gives the signal at the stationary point as

$$S_{eq}\Big|_{[A]=\frac{1}{K_A}} = \frac{2K_{act}(2 + 3\lambda K_{act})}{3 + 4K_{act} + 3\lambda K_{act}^2}\ . \tag{3.23}$$

The difference between the peak and the basal signal can then be calculated as

$$S_p = \frac{K_{act}(1 - \lambda K_{act}^2)}{(1 + K_{act} + \lambda K_{act}^2)(3 + 4K_{act} + 3\lambda K_{act}^2)}. \tag{3.24}$$

As this expression contains the parameter λ, we find that the height of the peak, in relation to the basal signal, does change as the parameter λ changes. Furthermore, if we let $\lambda \to \infty$ we have

$$\lim_{\lambda \to \infty} S_p = 0, \tag{3.25}$$

which confirms that, for a high enough λ, the peak decreases until the curve flattens out, giving a constant maximal signal.

3.3.3 Varying efficacy parameter ν

We now fix $\alpha = \lambda = \xi = 1$ and instead vary the efficacy parameter, ν. We see in Figure 3.3.4 that varying ν results in dose-response curve that are shaped similar to monomeric curves (with the exception of the $\nu = 1$ curve), as opposed to the bell-shaped curves we saw when varying α and λ. As ν is an efficacy parameter, increasing ν by one order of magnitude causes a large increase in signal, until the maximum signal is reached. This, however, means that the signal with $\nu < 1$ are on a much smaller scale than with $\nu > 1$, and so when plotting the signal we have again separated these in order to see the full details of each curve.

Whilst there is little ligand available, varying ν has no effect on any of the concentrations, and therefore the signal. It then follows that the basal level of activity is the same for all values of ν. Once the ligand concentration increases, we then see the effect ν has on the curves. Increasing ν leads to a typical agonist response. The signal curve is sigmoidal in shape, with an increasing ligand concentration leading to an increase in signal. Increasing ν leads to an increased peak in $[R_A^* R]$, which results in increases in

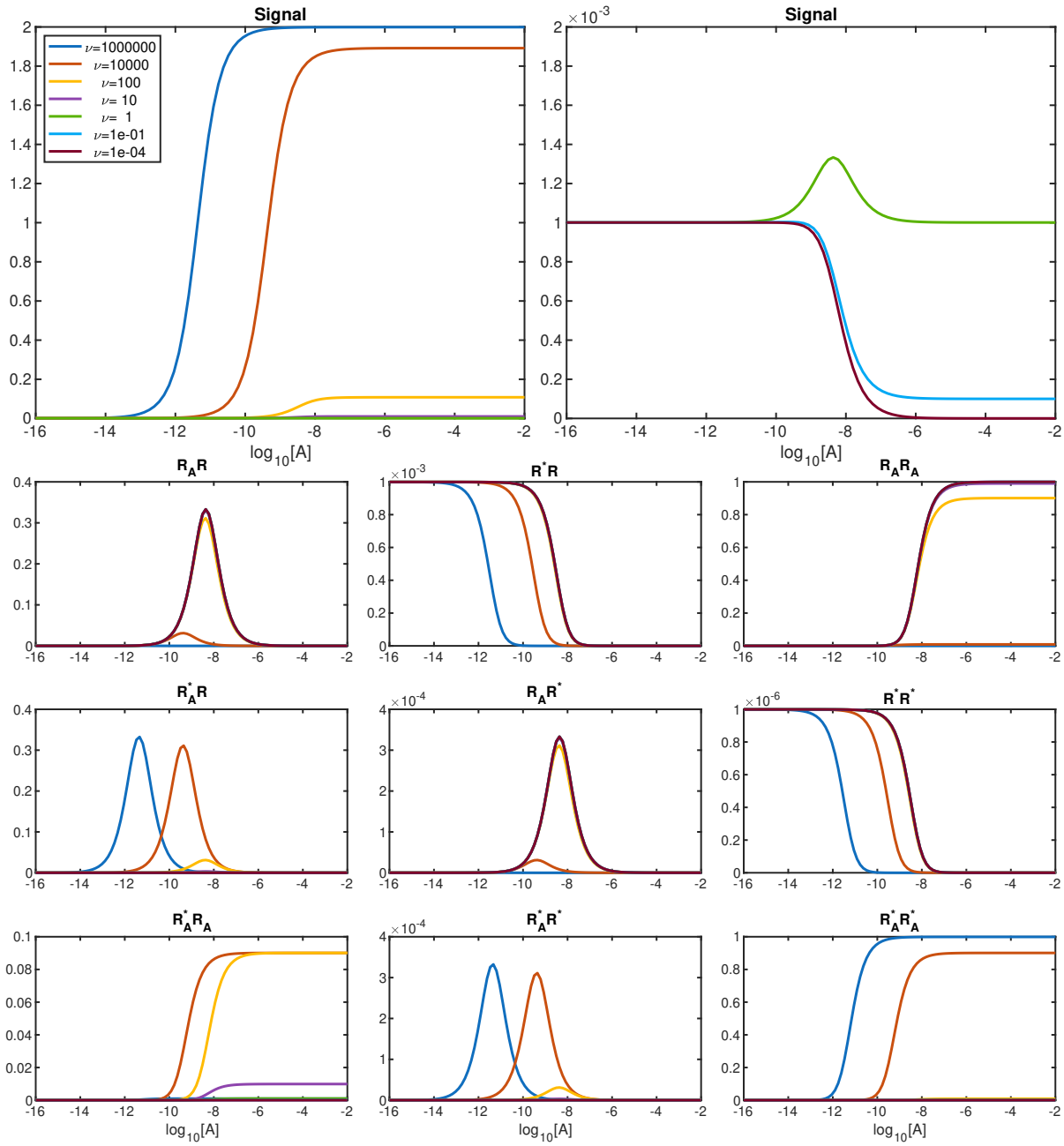

Figure 3.3.4: Plotting a logDR curve (equation (3.26)) for a range of values for efficacy parameter ν. This parameter determines whether the ligand acts as an agonist (when $\nu > 1$), or inverse agonist ($\nu < 1$), while $\nu = 1$ shows a curve with a small peak.

$R_A^* R^*$ and $R_A^* R_A$, and also $[R_A^* R_A^*]$. Whilst increasing ν does mean we also see a decrease in $[R_A R^*]$, we note that this concentration is on a much smaller scale than the increasing concentrations, and as such, the increases far outweigh the decreases.

Conversely, if ν decreases, such that $\nu \ll 1$, then the ligand acts as an inverse agonist

and decreases the signal below the basal level. With a small ν, ability for a protomer to be both bound and active is almost blocked, which leads to an increase in $[R_A R^*]$, and a decrease in $[R_A^* R]$. As the ligand concentration increases the system becomes flooded by the ligand, leading to more dimers becoming bound on both sides. In turn the low value of ν stops either protomer from becoming active, hence we see a drop in active receptors and, therefore, the signal.

To determine how much of an increase (or decrease) in signal is gained from increasing (or decreasing) ν, we set $\alpha = \xi = \lambda = 1$ and $\omega = 2$ in equation (3.6), giving

$$S_{eq} = \frac{K_{act}(1 + 2K_{act} + (1 + \nu + \nu K_{act})K_A[A] + \nu(1 + 2\nu K_{act}){K_A}^2[A]^2)}{1 + K_{act} + K_{act}^2 + (1 + (\nu + 1)K_{act} + \nu K_{act}^2)K_A[A] + (1 + \nu K_{act} + \nu^2 K_{act}^2)K_A^2[A]^2}. \tag{3.26}$$

First taking the limit as $[A] \to 0$ gives

$$\lim_{[A]\to 0} S_{eq} = \frac{K_{act}(1 + 2K_{act})}{1 + K_{act} + K_{act}^2}, \tag{3.27}$$

as the basal signal under these conditions, which we notice depends only on K_{act} and, therefore, does not change when varying ν (which confirms our earlier observations). We also have

$$\lim_{[A]\to \infty} S_{eq} = \frac{\nu K_{act}(1 + 2\nu K_{act})}{1 + \nu K_{act} + \nu^2 K_{act}^2}, \tag{3.28}$$

as the increased or decreased signal level (once saturation is reached). We use this to find the order of magnitude of the signal under a varying ν, noting that this will depend on both ν and K_{act}. First we point out that, if $\nu K_{act} \gg 1$ then it follows that $(\nu K_{act})^2 > \nu K_{act}$. We can then state that, if

$$\nu K_{act} \gg 1, \quad \Rightarrow \quad S_{eq} \approx 1. \tag{3.29}$$

Conversely, if

$$\nu K_{act} \ll 1, \quad \Rightarrow \quad S_{eq} \approx \nu K_{act}. \tag{3.30}$$

Furthermore, we have

$$\lim_{\nu \to \infty} \frac{\nu K_{act}(1 + 2\nu K_{act})}{1 + \nu K_{act} + \nu^2 K_{act}^2} = 2, \tag{3.31}$$

that is, as $\nu \to \infty$ the saturation signal tends to the maximum signal, as expected. Conversely,

$$\lim_{\nu \to 0} \frac{\nu K_{act}(1 + 2\nu K_{act})}{1 + \nu K_{act} + \nu^2 K_{act}^2} = 0, \tag{3.32}$$

concluding that as $\nu \to 0$ the signal is inhibited entirely, with constitutive activity being suppressed, giving the ligand as a full inverse agonist.

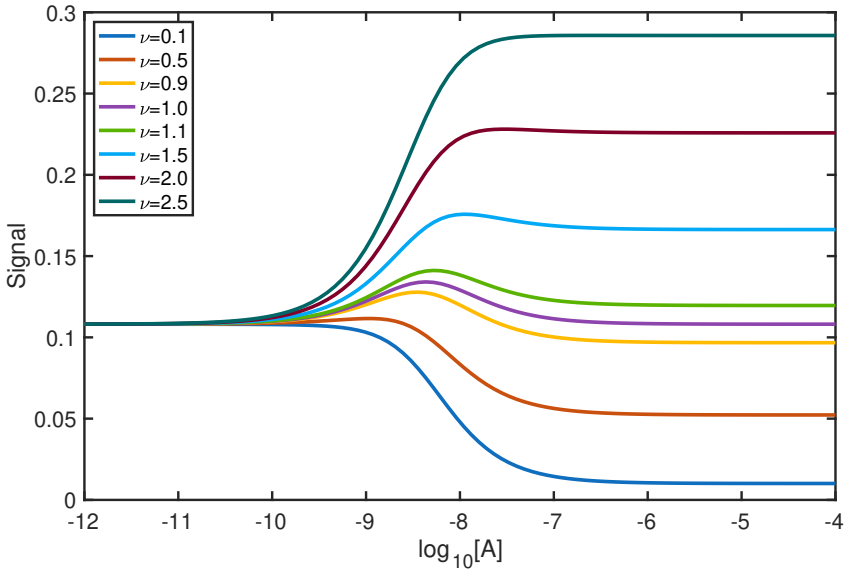

Figure 3.3.5: We plot the signal curve (equation (3.26)) for values of ν close to one, showing the smoothing out of the peak.

We also explore the peak seen in the $\nu = 1$ curve. For a monomeric model, having both $\nu = 1$ and $\xi = 1$ would equate to a neutral agonist, and the signal would be constant. For this dimer model, however, we see a peak appear in the curve. We discussed the reason for the peak appearing when $\alpha = \nu = \xi = \lambda = 1$ in Section 3.3.1. In Figure 3.3.5 we plot the signal for values of ν close to one. In this we see that, as ν increases past one, the peak increases, however, the saturation signal increases by a larger amount, until the curve smooths out. Similarly for ν decreasing, the peak decreases as ν decreases, until there is no longer a peak.

3.3.4 Varying efficacy parameter ξ

In this section, we note that varying ξ gives exactly the same signal as when we varied ν. To show this we take $\nu = x_1$ and $\xi = x_2$, giving the signal as

$$S_{eq} = \frac{\begin{array}{r} K_{act} + x_1 K_A K_{act}[A] + x_2 K_A K_{act}[A] + \alpha x_1 x_2 {K_A}^2 K_{act}[A]^2 \\ + \omega(\lambda {K_{act}}^2 + x_1 x_2 \lambda K_A {K_{act}}^2[A] + \alpha x_1^2 x_2^2 \lambda {K_A}^2 {K_{act}}^2[A]^2) \end{array}}{\begin{array}{r} \alpha x_1^2 x_2^2 \lambda {K_A}^2 {K_{act}}^2[A]^2 + \alpha x_1 x_2 {K_A}^2 K_{act}[A]^2 + \alpha {K_A}^2[A]^2 + x_1 x_2 \lambda K_A {K_{act}}^2[A] \\ + x_1 K_A K_{act}[A] + x_2 K_A K_{act}[A] + K_A[A] + \lambda {K_{act}}^2 + K_{act} + 1 \end{array}}. \tag{3.33}$$

Similarly, taking $\nu = x_2$ and $\xi = x_1$, gives the signal as

$$S_{eq} = \frac{\begin{array}{r} K_{act} + x_1 K_A K_{act}[A] + x_2 K_A K_{act}[A] + \alpha x_1 x_2 {K_A}^2 K_{act}[A]^2 \\ + \omega(\lambda {K_{act}}^2 + x_1 x_2 \lambda K_A {K_{act}}^2[A] + \alpha x_1^2 x_2^2 \lambda {K_A}^2 {K_{act}}^2[A]^2) \end{array}}{\begin{array}{r} \alpha x_1^2 x_2^2 \lambda {K_A}^2 {K_{act}}^2[A]^2 + \alpha x_1 x_2 {K_A}^2 K_{act}[A]^2 + \alpha {K_A}^2[A]^2 + x_1 x_2 \lambda K_A {K_{act}}^2[A] \\ + x_1 K_A K_{act}[A] + x_2 K_A K_{act}[A] + K_A[A] + \lambda {K_{act}}^2 + K_{act} + 1 \end{array}}, \tag{3.34}$$

hence, the signals are equal. This comes from the symmetry in the sphere (Schematic in Figure 3.1.1), and in particular the concentrations of the species $R_A^* R$ and $R_A R^*$. If $\nu > \xi$ then $[R_A^* R]$ increases to a higher concentration than $[R_A R^*]$, while if $\nu < \xi$ then $[R_A R^*]$ increases to the higher concentration. However, the sum $[R_A^* R] + [R_A R^*]$ is the same for both cases. This effect is illustrated in Figure 3.3.6 where we fix all other parameters and take two different values for ν and ξ. Although the curves for $[R_A^* R]$ and $[R_A R^*]$ change as the parameters change, their sum is the same. This results for all other

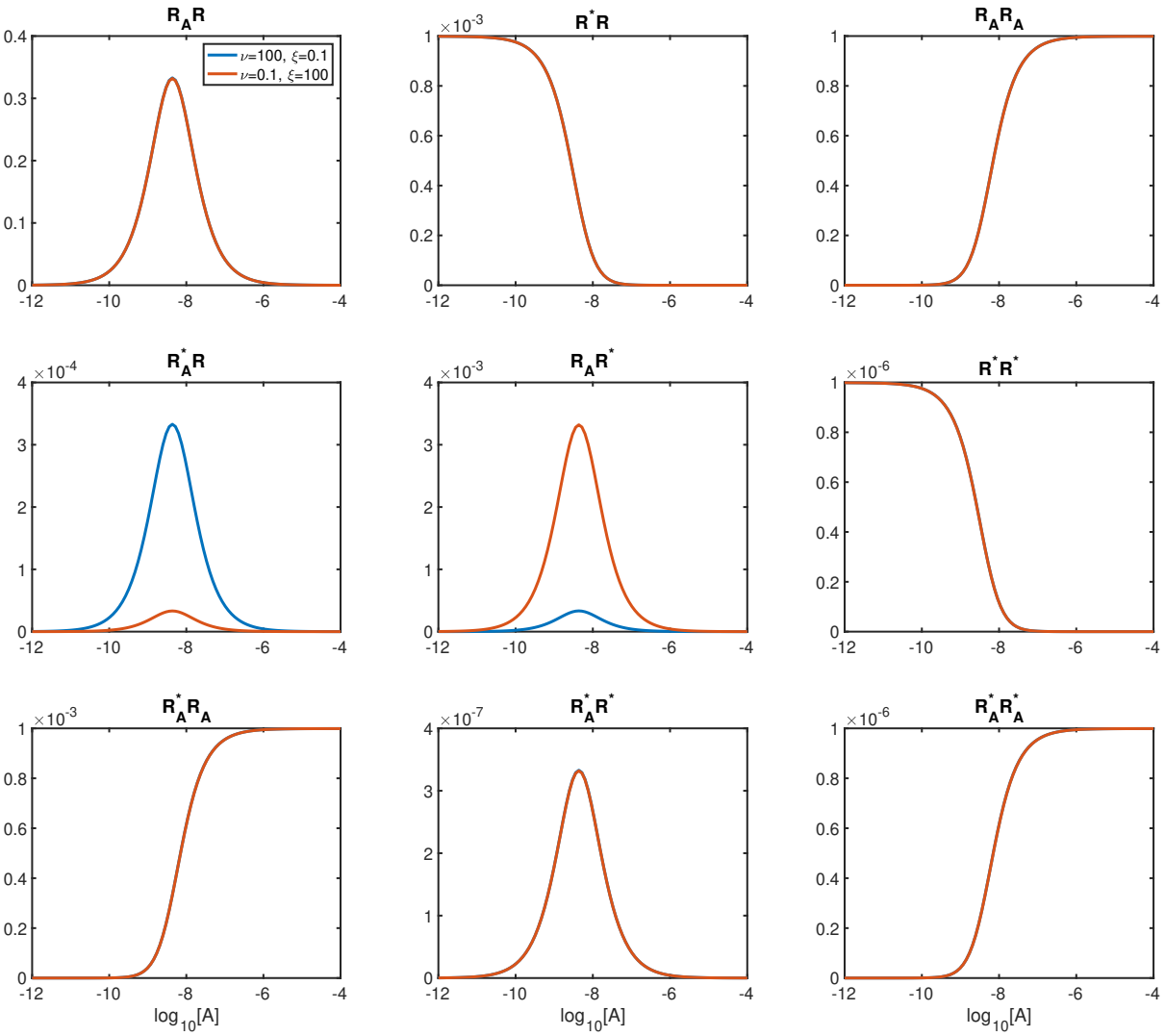

Figure 3.3.6: Plotting two sets of logDR curve (equation (3.33)), with $\nu = 100, \xi = 0.1$ and also $\nu = 0.1, \xi = 100$. The symmetry in the model sphere means that, although the curves of $[R_A^* R]$ and $[R_A R^*]$ change with varying ν and ξ, their sum, as well as all other concentrations, is equal.

species are equal for both cases. We note that, having this symmetry in the parameters ν and ξ will result in identifiability issues if the model is used for data fitting, as it is unable to uniquely identify them individually from the model equations. Identifiability is something we return to in Chapter 5 (though not for this model).

3.3.5 Varying signal weighting parameter ω

The final parameter we consider the effect of is ω. This parameter dictates how much of a signal is elicited from a symmetric dimer in relation to an asymmetric one. Changing this parameter does not affect any of the individual species concentrations, it simply describes how much of a contribution to the overall signal each species has. With $\omega = 1$ the contribution from symmetric dimers is equal to that of asymmetric dimers. Increasing ω means that the signal from symmetric dimers is scaled so that they give a larger contribution to the signal. As such the signal is increased. Similarly, if $\omega < 1$, the contribution from symmetric dimers is less than that of asymmetric ones, and the overall signal is decreased. This can be seen in Figure 3.3.7.

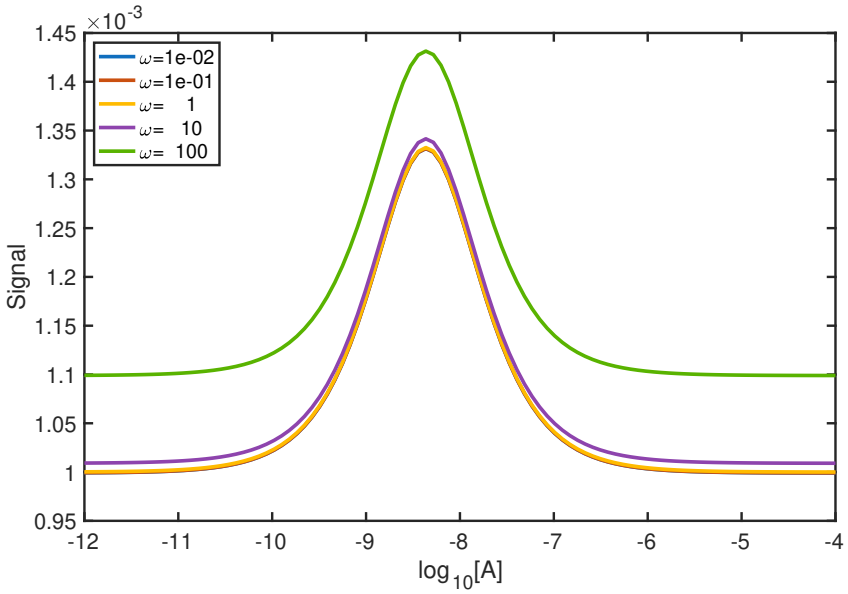

Figure 3.3.7: Plotting a logDR curve (equation (3.6)) for a range of values for weighting parameter ω, with all cooperativity factors fixed at one. Increasing ω increases the contribution symmetric dimers have to the signal, which then increases the signal.

3.4 Dose response curves: varying multiple parameters

Although it was important to vary each of the parameters individually, in order to fully understand the effect they have on the system, it is more likely that these parameters

will all be values other than one. In particular, we are usually interested in the effects of an agonist, that is when efficacy parameters are greater than one. We now examine the possible compound effects arising from a combination of varied parameters.

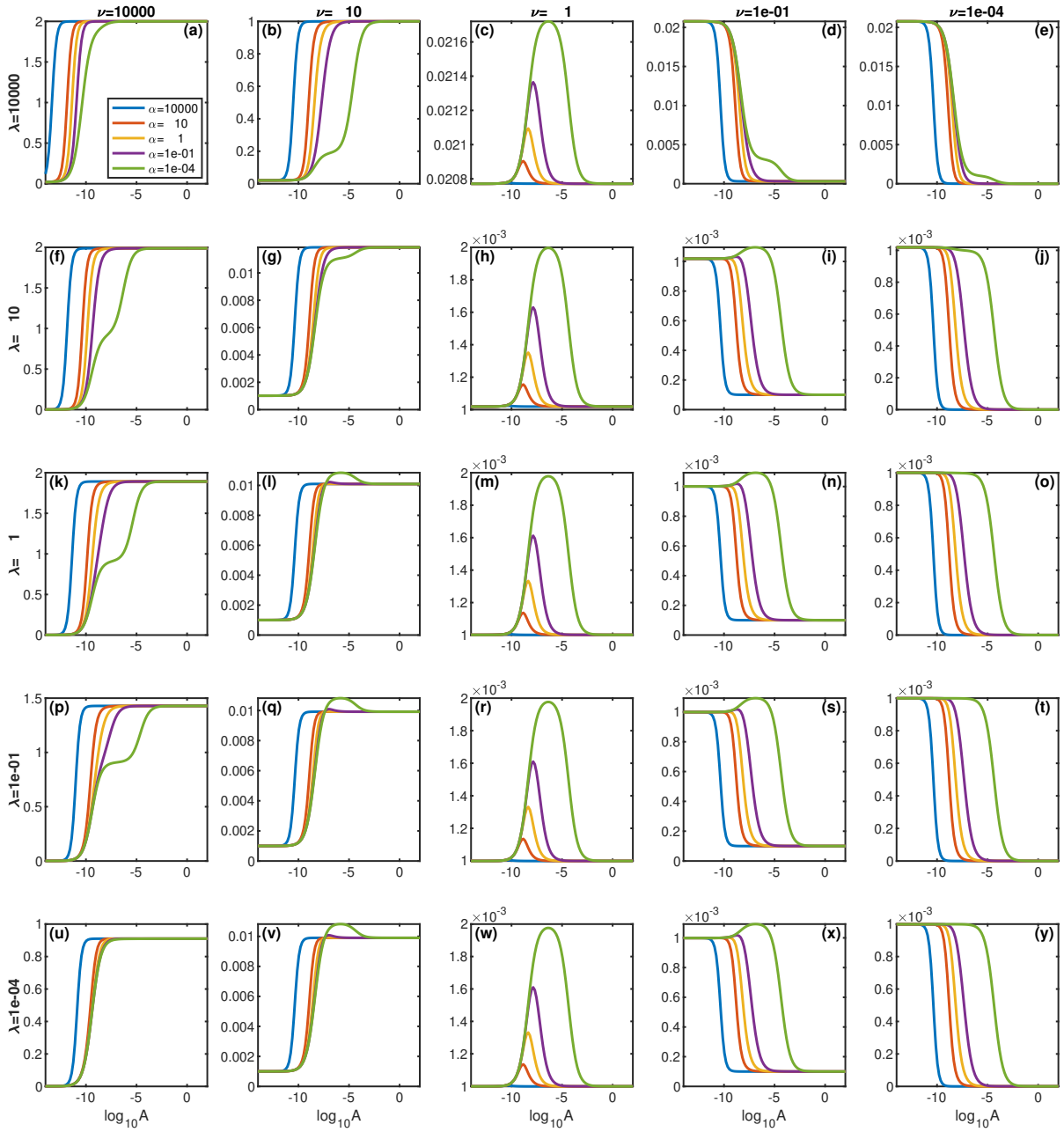

Figure 3.4.1: We fix $\xi = 1$ and vary λ on each row, ν in each column and α within each subplot (using equation (3.6)). We see effects such as inflections and overshoots in some of the curves.

In Figure 3.4.1 we plot only signal curves. We fix $\xi = 1$, then on each row we vary the

parameter λ, each column has a different value of ν, while α is varied within each plot. We note that, although it is unlikely that $\xi = 1$, for now we keep it fixed this way in order to understand the results using findings from the previous sections. In these plots, we observe a number of behaviours that did not appear in the signal curves when varying only a single parameter. The neutral curve (with all cooperativity factors equal to one) appears to be an almost constant signal for most ligand concentrations, though with a small peak for a small range of concentrations. We found that changing the binding cooperativity factor α increases or decreases the size of this peak, while the activation cooperativity factor λ shifts the whole curve. Varying the efficacy parameter ν increases or decreases the signal from the basal level, without affecting the basal level, giving agonist or inverse agonist behaviour. When all parameters are varied simultaneously, we see how these effects combine and result in the emergence of new behaviours, such as extra inflections (for example in subplots **(d)** and **(k)**) or overshoots (subplots **(n)** or **(q)**) in some of the dose-response curves.

If we focus on one particular curve as an example, we will discuss how each parameter contributes to these new effects. We take the green curve in plot **(q)**, that is, with $\alpha = 10^{-4}, \nu = 10, \lambda = 0.1$, and plot this in Figure 3.4.2. Along with this signal curve, we also plot the neutral curve for a point of reference, and the resulting signal for each of these parameter values being applied individually, for example, $\alpha = 10^{-4}, \nu = 1, \lambda = 1$. In this plot, we can clearly see that ν increases the signal from the basal level, and has the biggest contribution in the curve overall.

Whilst it appears that λ has no effect on the curve, we remind the reader of the results in Figure 3.3.3 where we saw that, as λ decreases from $\lambda = 1$ to $\lambda = 10^{-1}$, the signal does decrease but on a much smaller scale, hence it appears to have no effect when plotting on a larger scale, as in Figure 3.4.2. This contribution creates a slight decrease of the saturation signal. Finally, the binding cooperativity factor α is what creates the overshoot in the signal curve. Notice in the curve where $\alpha = 10^{-4}$ and all other cooperativity values are one, the peak increases by approximately 0.001, from the basal level. This appears to be the same as the increase of the overshoot above the saturation signal level, thus

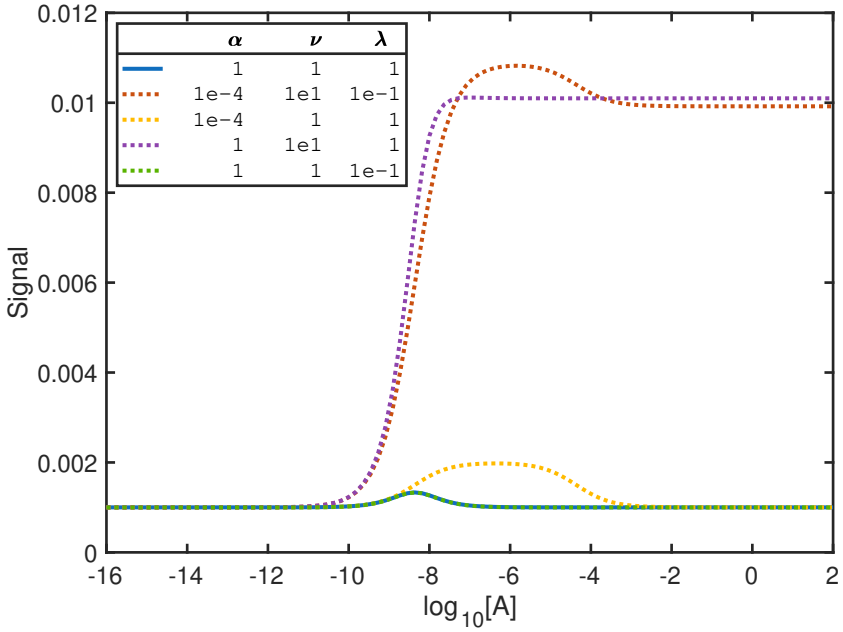

Figure 3.4.2: We plot dose response curves (equation (3.6)) for $\alpha = 10^{-4}, \nu = 10$ and $\lambda = 0.1$, as well as for these each being set individually.

indicating that the large peak arising from having such low binding cooperativity causes the overshoot.

Upon closer inspection, we notice that all curves that have these unusual behaviours when binding cooperativity α is low. This is a clear link to our findings in Chapter 2 where we found low cooperativity to result in extra inflections in dose-response curves for all binding models. However, we also note that we do not see extra inflections or overshoots in all low binding cooperativity signal curves in Figure 3.4.1, and so we find that all cooperativity factors have an effect on whether these occur or not. While it is theoretically possible to derive conditions under which these appear, the calculations are impractical for use and require the solution of high degree polynomials.

Finally, in Figure 3.4.3, instead of fixing $\xi = 1$ we consider the more realistic case, where $\xi = \nu$. That is the efficacy of each protomer is the same. In this, we notice that, while we still see inflections in some of the curves, these appear in fewer plots. Also, there are no overshoots at all that can be seen. It seems that, increasing (or decreasing) ξ together with ν counters some of these effects. We do not explore why this happens at this time, however, this will be something to consider for future work.

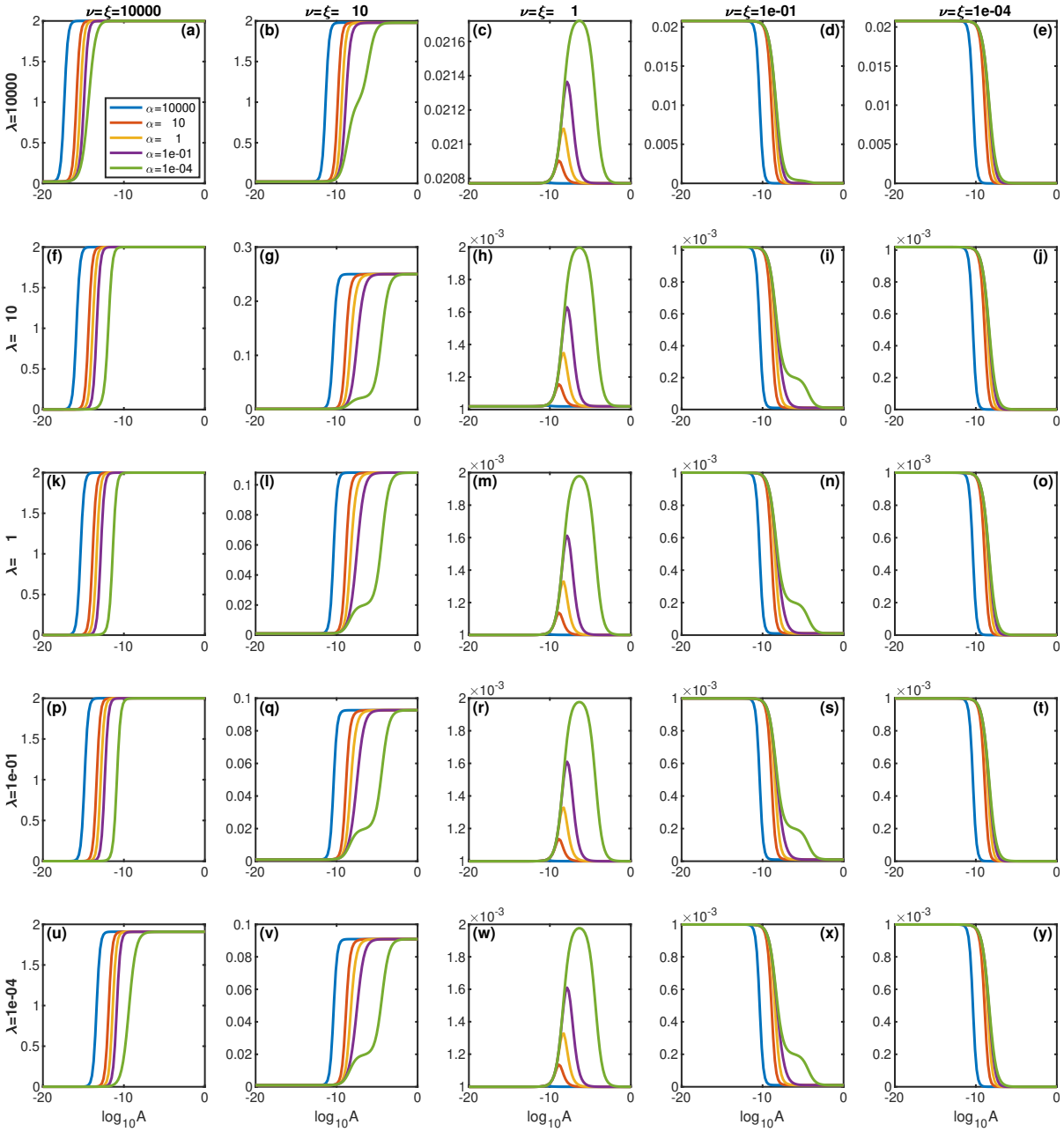

Figure 3.4.3: We fix $\xi = \nu$ and vary λ on each row, ν in each column and α within each subplot. We see some extra inflections but now no overshoots in the curve.

3.5 Conclusions

In this chapter, we have extended the dimer binding model (Chapter 2) to include receptor activation. The increase in complexity of the model from our previous GPCR models is clear, and we also see many new behaviours emerging in the dose response curves. The focus remains on the cooperativity factors that affect secondary binding

and activation. We first considered the effect of each of these parameters individually, observing the effect each had on the dose response curves. With all cooperativity factors being neutral, that is equal to one, the signal curve remains almost constant for most ligand concentrations, but with a small peak for a range of ligand concentrations. This peak is a point of interest as it appears with efficacy parameters equal to one, that is, the ligand acting as a neutral agonist. For monomeric receptors, the signal for a neutral agonist is constant, hence a peak such as this can be used to identify the presence of dimers.

We use this neutral curve as the baseline when comparing the effects of varying the other parameters. The cooperativity factor α causes the peak to increase with an decreasing α. This bell shaped curve is often indicative of dimerised receptors [100]. We also saw how the peak in the signal curve is directly related to the extra inflections seen in the ligand-bound curve. The primary effect of varying λ is a shift in the whole curve. Increasing λ increases the signal for all ligand concentrations and causes subtle changes in the peak size. The symmetry in the model means that we obtain the same results when varying ν or ξ, hence we choose to only explore the effects of ν. Doing so results in a change of the shape of the curve. Increasing ν from neutral leads to the ligand acting as an agonist, while decreasing ν from one sees inverse agonist curve shapes. Furthermore, the amount of increase or decrease in the signal at saturation is directly proportional to the increase or decrease in ν. That is, increasing ν, results in an increase in signal, at saturation.

We also consider how these effects combine, by looking at the results when varying multiple parameters together. In these, we see new behaviours emerge, such as extra inflections and overshoots in the signal curves. We show how each of the parameters contribute to these curves. The model formulation, solution and results serve as a contribution to the field of pharmacological modelling. The expressions given are practical in use and are easily used to simulate results. The findings presented in this chapter can explain experimental data that has a bell-shaped curve or overshoots in the curve, which can, in turn, be used to inform further experiments.

While we have explored some of the possible effects that can be seen at equilibrium, we note that this work is ongoing. Next steps would be to understand why more inflections and overshoots can be seen in the dose response curves when $\xi = 1$ as opposed to when $\xi = \nu$. Following this,we would investigate the time course dynamics of the system, where we expect to see many more interesting behaviours appear.

Chapter 4

The vascular endothelial growth factor system

In Chapter 4 (Section 1.4) we introduced the vascular endothelial growth factor (VEGF) system. VEGF is a key mediator of angiogenesis; a process whereby new blood vessels are formed from the pre-existing vasculature [80, 3]. As angiogenesis is a key factor in many conditions, including cancer and inflammation [80, 111, 102], research into the mechanism of VEGF binding and signalling is essential towards progress in development of new therapies. Although advancements have been made towards understanding the interactions of receptor tyrosine kinase (RTK) ligands and receptors, further research is needed to fully understand their role as therapeutic targets [102, 101, 111].

Although much of existing research is done at equilibrium [80], a recent study by Kilpatrick *et al* [70] has used new technologies to provide a real time quantitative evaluation of VEGF-VEGFR binding. Developments in fluorescent ligand technologies have allowed the complexities of ligand-receptor interactions in living cells to be observed[70, 102]. In this study, synthesized dual poled VEGF ligand molecules, that is, a ligand that can bind two receptors simultaneously, were monitored binding in real time to VEGFR-2 receptors. Results were fitted to a simple association exponential model, however, in many cases the results failed to fit with the standard model for simple mass-action equilibrium. Hence there is need for a model that incorporates the complexities of the VEGF binding dynamics, also taking into account the dimerisation that is induced from ligand binding.

The ligand induced dimerisation process is investigated extensively in [3], where multiple stochastic models are analysed to further understand the role the VEGF receptor system plays in tumour growth, although most models given there include biological processes beyond binding. An interesting point that is also mentioned in [3] is that at high ligand concentrations there is an inhibition of cellular response. This is a phenomenon that is also seen experimentally in [15], although they describe it as a 'window of desensitisation' to VEGF, with a continued increase in signal after a short decrease. Spatial models of the dimerisation process and subsequent signalling are developed in [85], while dynamics of receptor and transducer protein dimerisation are studied using an ordinary differential equation (ODE) model in [127], wherein it is suggested that dimerisation may serve to regulate signalling over multiple time scales. An ODE model of receptor binding and aggregation is presented in [132], although only equilibrium is analysed. Equilibrium models in [136] and [71] also explore the possibility of heterodimers. Mac Gabhann and Popel [80] combine a ligand induced dimerisation model with a 'dynamic pre-dimerisation' model, whereby the dimers are formed before ligand binding, in order to explore the mechanisms of the dimerisation of VEGFR and the possibility of the VEGF receptors having the ability to dimerise in the absence of ligand as well as being induced by ligand. Gabhann and Popel also consider the possibility of VEGF receptors having the ability to form heterodimers, although to what extent this is possible biologically is currently unknown, however their findings suggest that the level of heterodimer formation could be significant as the signalling initiated by these is different and unique.

When formulating our previous GPCR models, we introduced the idea of a cooperativity factor; an extra parameter that alters the binding rate of a second ligand molecule to a pre-formed dimer, due to a first already being bound. Using this cooperativity factor we were able to analyse equilibrium binding results, and contribute to current receptor theory by also considering binding dynamics. Furthermore, analytical expressions for both equilibrium concentration curves and binding kinetics, in terms of this cooperativity value, gave insights into the effects dimerisation has on binding. In this chapter, we continue this work by applying the same framework as we did to the GPCR models in

Chapter 2 to the VEGF-VEGFR system. We again formulate a model in terms of a co-operativity factors, although due to the difference in the dimerisation mechanism of this system, cooperativity in this paper is defined as a factor that alters the binding rate of a second receptor (as opposed to second ligand molecule), once a first receptor is bound to the ligand. The inspirations for this paper lie in the work by Mac Gabhann and Popel [80], where they present a model for the binding and dimerisation of VEGF-VEGFRs.

However, while the model Mac Gabhann and Popel present also includes activation and diffusion, we choose to focus on binding only. Binding dynamics are a current topic of experimental study and a simpler model allows for the study of time course dynamics as well as equilibrium results. Furthermore, this allows us to follow the analysis methods we used in our previous work. However, this ligand induced model that we develop results in a nonlinear system of ODEs (in contrast to the GPCR model in Chapter 2), for which we are unable to give exact analytical solutions. However, the structure of our model is inline with more complex receptor dynamics models ([138], [137], [14]), hence we follow their lead in analysing the system dynamics. We use numerical methods, for a range of control parameters, to explore possible time course behaviours, whilst perturbation analysis gives analytical solutions for reduced problems under interesting parameter regimes. Finally, the model is validated by fitting to recently published data, quantifying the ligand-receptor interactions.

4.1 Model formulation

In formulating the model we assume that all receptors exist constitutively as monomers, represented by R, while ligand A is a two-poled ligand, which has the ability to bind to two receptors simultaneously. Ligand binding and dimerisation is a two step process which can be visualised as in Figure 4.1.1 or given as a reaction scheme in Figure 4.1.2. In the first reversible reaction the ligand, which we fix at a constant concentration, binds to a free monomeric receptor, with association and dissociation rates of k_+ and k_- respectively, thus creating the complex AR. Once a ligand molecule is bound to a receptor the ligand reversibly binds a second receptor monomer, simultaneously dimerising the two receptors.

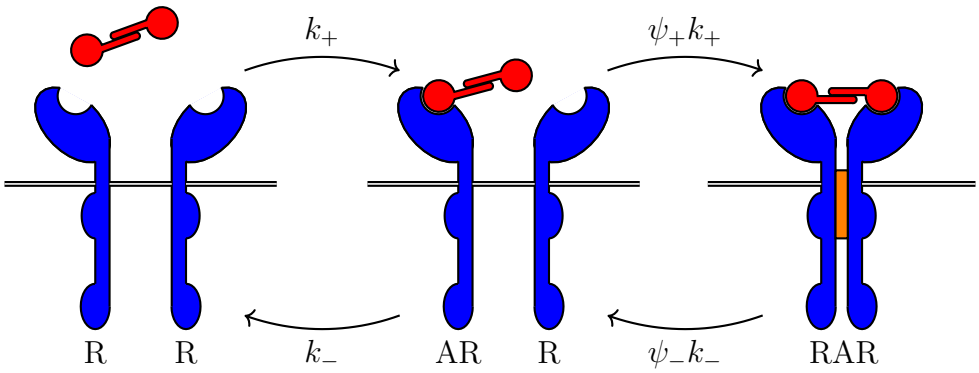

Figure 4.1.1: Ligand binding and dimerisation is a two step process whereby a two poled ligand (in red) first binds a monomeric receptor (blue) before then binding a second and instantaneously dimerising the monomers.

We define the parameter $\psi = \psi_+/\psi_-$ as the cooperativity factor, describing the ligand's increased (or decreased) affinity for the second receptor binding due to the first ligand pole being bound. If, for example, $\psi_+ > 1$ we have positive forward cooperativity, where the rate of a second receptor binding is increased, and negative forward cooperativity if $\psi_+ < 1$.

$$\begin{aligned} \mathrm{A} + \mathrm{R} &\underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{AR} \\ \mathrm{R} + \mathrm{AR} &\underset{\psi_- k_-}{\overset{\psi_+ k_+}{\rightleftharpoons}} \mathrm{RAR} \end{aligned}$$

Figure 4.1.2: Schematic representing the reactions resulting from the binding of a two poled ligand to two monomeric receptors.

In Chapter 2 we also explored the effect cooperativity has on binding, this time in GPCRs. However, there are differences in how cooperativity is described in the earlier chapters and this model. In the previous models we assumed all receptors were predimerised, and so cooperativity was the change in affinity of a second ligand binding, where as in this model it describes the change in affinity of a second receptor binding.

The law of mass action gives rise to a system of ordinary differential equations (ODEs)

that govern the binding kinetics of the reactions, namely:

$$\frac{d[R]}{dt} = -k_+[A][R] + k_-[AR] - \psi_+ k_+[R][AR] + \psi_- k_-[RAR], \tag{4.1a}$$

$$\frac{d[AR]}{dt} = k_+[A][R] - k_-[AR] - \psi_+ k_+[R][AR] + \psi_- k_-[RAR], \tag{4.1b}$$

$$\frac{d[RAR]}{dt} = \psi_+ k_+[R][AR] - \psi_- k_-[RAR]. \tag{4.1c}$$

with initial conditions

$$[R](0) = R_{tot}, \qquad [AR](0) = 0, \qquad [RAR](0) = 0 \tag{4.2}$$

where square brackets denote concentration, and all parameters and measured quantities are restricted to be positive. The signal of interest is proportional to the number of receptors bound [70], so we can therefore state this as

$$S = a([AR] + 2[RAR]), \tag{4.3}$$

where a is a scaling constant. The signal consists of the concentration of AR and double the concentration of RAR. This is because each molecule of RAR is made up of two receptors, hence the signal given off is twice that of the concentration, as is described in [70]. The total concentration of receptors is conserved, with

$$R_{tot} = [R] + [AR] + 2[RAR], \tag{4.4}$$

where R_{tot} is the total concentration of receptors. We use this to reduce the system (4.1)

and eliminate one equation, leaving

$$\frac{d[R]}{dt} = -\left(k_+[A] + \frac{\psi_- k_-}{2}\right)[R] + \left(k_- - \frac{\psi_- k_-}{2}\right)[AR] - \psi_+ k_+[R][AR] + \frac{\psi_- k_- R_{tot}}{2}, \tag{4.5a}$$

$$\frac{d[AR]}{dt} = \left(k_+[A] - \frac{\psi_- k_-}{2}\right)[R] - \left(k_- + \frac{\psi_- k_-}{2}\right)[AR] - \psi_+ k_+[R][AR] + \frac{\psi_- k_- R_{tot}}{2}. \tag{4.5b}$$

The initial value problem has the initial conditions

$$[R](0) = R_{tot}, \quad [AR](0) = 0, \tag{4.5c}$$

since we assume that no bound complex is present to begin. We note here that we assume there exists no constitutive dimerisation, hence dimers are solely ligand-induced. It is worth observing here that the model we have developed is structurally similar to that of 'Model 1' in [3], though moving forward we focus on analysing the binding dynamics, whereas the authors of [3] continue to expand the model to include other biological features of the signalling pathway.

4.2 Equilibrium analysis

In the spirit of classical receptor theory, and in keeping with our previous work, we first investigate the equilibrium behaviour of the system, in particular the effect of the equilibrium cooperativity factor $\psi = \psi_+/\psi_-$. The equilibrium relationships are

$$[AR] = K_A[A][R], \tag{4.6a}$$

$$[RAR] = \psi K_A[R][AR] = \psi K_A^2[A][R]^2, \tag{4.6b}$$

where $K_A = k_+/k_-$ is the equilibrium association constant (that is $K_A = 1/K_D$). Combining these with equation (4.4) gives equilibrium species in terms of parameters as

$$[R] = \frac{-(1 + K_A[A]) + \sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}}}{4\psi K_A^2[A]}, \tag{4.7a}$$

$$[AR] = \frac{-(1 + K_A[A]) + \sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}}}{4\psi K_A}, \tag{4.7b}$$

$$[RAR] = \frac{\left(1 + K_A[A] - \sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}}\right)^2}{16\psi K_A^2[A]}. \tag{4.7c}$$

From this we can state the overall signal at equilibrium to be

$$S_{eq} = \frac{a(1 + K_A[A] + 4\psi K_A^2[A]R_{tot} - \sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}})}{4\psi K_A^2[A]}. \tag{4.8}$$

If we let $[A] \to \infty$ we have $[R], [RAR] \to 0$, while $[AR]$, $S_{eq} \to aR_{tot}$. This signal is a scaled concentration of the total bound ligand, hence is comparable to the total ligand bound expression in our Chapter 2, which we see, although there is similarity in the models, these expressions are very different. Whilst analysing the GPCR model we presented in this earlier chapter, we found a condition under which dimer cooperativity caused there be to extra inflections in the log dose-response (logDR) curve. We check whether it is also possible to get these extra inflections with this ligand induced model. Taking the expression for the (unscaled) signal

$$S_{eq} = \frac{1 + K_A[A] + 4\psi K_A^2[A]R_{tot} - \sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}}}{4\psi K_A^2[A]}, \tag{4.9}$$

we take the second derivative, with respect to $[A]$ (calculated symbolically in Matlab [2]))

$$\frac{d^2 S^{eq}}{d\log_{10}[A]^2} = \frac{\log(10)^2 \left(\begin{array}{l} 3K_A^2[A]^2 + K_A^3[A]^3 + 3K_A[A] + 12\psi K_A^2[A]R_{tot} + 16\psi^2 K_A^4[A]^2 R_{tot}^2 \\ \quad + 8\psi K_A^3[A]^2 R_{tot} + 4\psi K_A^4[A]^3 R_{tot} + 1 - ((1+K_A[A])^2 + 8\psi K_A^2[A]R_{tot})^{3/2} \end{array} \right)}{4\psi K_A^2[A]((1+K_A[A])^2 + 8\psi K_A^2[A]R_{tot})^{3/2}}. \tag{4.10}$$

This is zero when

$$\begin{aligned} 3K_A^2[A]^2 + K_A^3[A]^3 + 3K_A[A] + 12\psi K_A^2[A]R_{tot} + 16\psi^2 K_A^4[A]^2 R_{tot}^2 + 8\psi K_A^3[A]^2 R_{tot} \\ + 4\psi K_A^4[A]^3 R_{tot} + 1 - ((1+K_A[A])^2 + 8\psi K_A^2[A]R_{tot})^{3/2} = 0. \end{aligned} \tag{4.11}$$

That is, when

$$\begin{aligned} (3K_A^2[A]^2 + K_A^3[A]^3 + 3K_A[A] + 12\psi K_A^2[A]R_{tot} + 16\psi^2 K_A^4[A]^2 R_{tot}^2 + 8\psi K_A^3[A]^2 R_{tot} \\ + 4\psi K_A^4[A]^3 R_{tot} + 1)^2 = ((1+K_A[A])^2 + 8\psi K_A^2[A]R_{tot})^3. \end{aligned} \tag{4.12}$$

Expanding and simplifying gives

$$\begin{aligned} -8\psi K_A^3[A]^3 R_{tot}(2\psi K_A R_{tot} + 1)(K_A^4[A]^4 + 2K_A^3(4\psi K_A R_{tot} + 1)[A]^3 \\ + 8\psi K_A^3 R_{tot}(2\psi K_A R_{tot} + 1)[A]^2 - 2K_A(4\psi K_A R_{tot} + 1)[A] - 1) = 0. \end{aligned} \tag{4.13}$$

The left hand side (LHS) is zero when $[A] = 0$, and also when

$$\begin{aligned} K_A^4[A]^4 + 2K_A^3(4\psi K_A R_{tot} + 1)[A]^3 + 8\psi K_A^3 R_{tot}(2\psi K_A R_{tot} + 1)[A]^2 \\ - 2K_A(4\psi K_A R_{tot} + 1)[A] - 1 = 0. \end{aligned} \tag{4.14}$$

For this fourth degree polynomial, we do not find the roots exactly, however we can determine the nature of the roots. We consider Descartes' Rule of Signs [34] which states:

Let $A(X) = \sum_{i=0}^{n} a_i X^i$ *be a polynomial of degree* n *with real coefficients that has exactly* p *positive real roots, counted with multiplicities. Let* $v = var(a0, ..., an)$ *be the number of sign variations in its coefficient sequence. Then* $v \geq p$ *and* $v \equiv p(\mod 2)$. *If all roots of* $A(X)$ *are real, then* $v = p$.

As all parameters are positive, we have $v = 1$, that is one sign change, and so only one positive real root, and therefore only a single inflection point. This concludes that there is no condition under which we get extra inflections. Hence, a logDR curve, for binding, with extra inflections should not be seen for ligand induced dimerisation so may indicate pre-dimerised receptors but not ligand induced ones.

In Figure 4.2.1 we look at the effect the equilibrium cooperativity factor ψ has on the log dose-response (logDR) relationship. Parameter values for the plot can be found in Appendix A. As the maximal ligand bound is Rt, we can calculate the EC_{50} value (the concentration giving half-maximal effect) for the signal, using equation (4.8), as

$$EC_{50} = \frac{a}{K_A(2a-1)(a+(2a-1)\psi K_A R_{tot})}. \tag{4.15}$$

If $a = 1$, that is the signal is unscaled, we have

$$EC_{50} = \frac{1}{K_A(1+\psi K_A R_{tot})}. \tag{4.16}$$

Again this expression is very different from the EC_{50} expression in our previous models, in Chapter 2, with the most noticeable difference being that this now also depends on R_{tot}. We see in Figure 4.2.1 that as ψ increases there is an approximate leftward shift in the overall signal, indicating the smaller EC_{50}. Looking at the individual species curves

explains why this occurs.

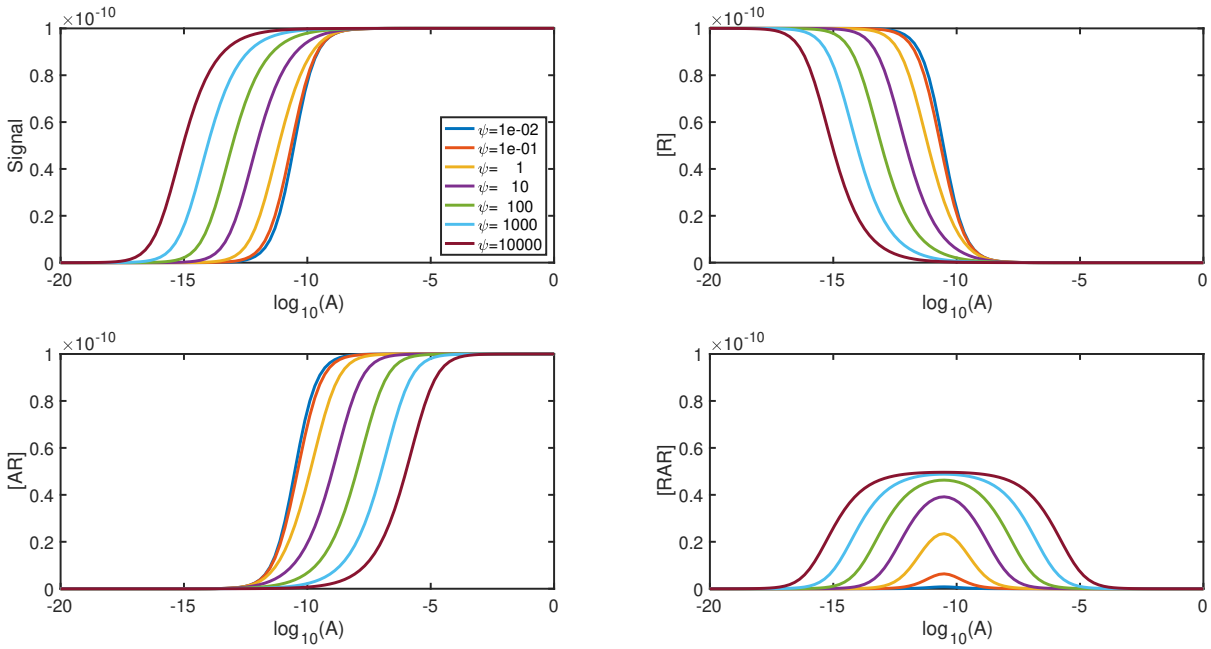

Figure 4.2.1: LogDR curves for varying cooperativity factor, ψ. The plots show the overall signal as well as concentrations of R (equation (4.7a)), AR (equation (4.7b)) and RAR (equation (4.7c)).

When high cooperativity is combined with low concentrations of A, most ligand molecules bind two receptors and thus form dimers, due to the concentration of receptors being much higher than that of the ligand, hence we see $[RAR]$ increase with $[A]$ whilst $[AR]$ remains close to zero. Recalling that the signal is made up of one $[AR]$ and two $[RAR]$ we see the signal is approximately twice the concentration of RAR. This continues until the ligand concentration reaches the point that there are no longer enough monomers to bind both poles of the ligand. From this point we see a plateau in $[RAR]$ and instead $[AR]$ increases. Once the system is saturated with ligand $[RAR]$ falls back towards zero, as the concentration of ligand is so much higher than the receptor concentration, hence there are no available receptors to bind the second pole of the ligand. Interestingly we see that the overall signal reaches saturation at a much lower concentration of A than $[AR]$. This occurs once the system reaches the point where almost all receptors are bound and so as $[A]$ increases from this point there is an interchange of each dimerised bound receptor to two single bound receptors, however, as all receptors are still bound there is

no apparent change in the overall signal.

With low cooperativity, the decreased propensity for dimerisation results in lower concentrations of $[A]$ being required for the same increase in bound monomers than when cooperativity is high. Although a small percentage of ligand bound monomers still bind a second receptor, causing dimerisation, low cooperativity prevents this from increasing substantially. Hence the signal increases largely with monomeric bound receptors. Interestingly we see that the peak in $[RAR]$ appears at the same level of $[A]$ regardless of the cooperativity value. To find the point at which this occurs we begin with expression (4.7c) and differentiate it with respect to $[A]$. This gives

$$\frac{d[RAR]}{d[A]} = \frac{(1 - K_A[A])\left(1 + K_A[A] - \sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}}\right)^2}{16\psi K_A^2[A]^2\sqrt{(1 + K_A[A])^2 + 8\psi K_A^2[A]R_{tot}}}. \tag{4.17}$$

Setting this equal to zero and solving for $[A]$ gives the solution

$$[A] = \frac{1}{K_A}, \tag{4.18}$$

ie, the dissocitation constant, K_D. This is only dependent on the binding rates, hence this peak will appear to some extent regardless of the cooperativity levels. To find the corresponding value of $[RAR]$ we substitute this back into equation (4.7c) to give

$$[RAR]\Big|_{\frac{1}{K_A}} = \frac{\left(1 - \sqrt{1 + 2\psi K_A R_{tot}}\right)^2}{4\psi K_A}. \tag{4.19}$$

Expanding this gives

$$\begin{aligned}&\frac{\left(1+\psi K_A R_{tot}-\sqrt{1+2\psi K_A R_{tot}}\right)}{2\psi K_A},\\ =&\frac{1}{2\psi K_A}+\frac{R_{tot}}{2}-\sqrt{\frac{1+2\psi K_A R_{tot}}{4\psi^2 K_A^2}}.\end{aligned} \tag{4.20}$$

If we take the limit as $\psi \to \infty$ we can clearly see that

$$[RAR]\Big|_{\frac{1}{K_A}} = \frac{R_{tot}}{2}, \tag{4.21}$$

as all other terms tend to zero, we find $R_{tot}/2$.

4.3 Binding dynamics: numerical simulations

Here we present some numerical results to show how cooperativity affects the binding dynamics of the individual species and the resulting overall signal. We fix k_+ and k_-, within the range of reported binding affinities ([80] and references within), to allow us to focus on the effects that cooperativity factors ψ_+ and ψ_- have on the binding dynamics. In Figure 4.3.1 we see how these parameters are a major factor in the behaviour of the individual species dynamics and the overall signal.

First looking to the overall signal, we notice that increasing ψ_+ leads to an increased signal at equilibrium. Increasing ψ_+ gives an increased rate of binding of a second monomer, hence, once a ligand binds a receptor it very quickly binds a second, and so the elevated signal levels come from an high concentration of $[RAR]$. Similarly, decreasing ψ_- gives an increase in the equilibrium signal, as in this case, once a dimer is formed it very quickly returns to bound monomer form, leading to an increased $[AR]$. As ψ_- decreases, the time taken to reach equilibrium is also increased, and in fact, looking further at the signal curves when ψ_- is low is the first indication that there may be further intricacies in the dynamics to be explored, and for this we look at the individual species

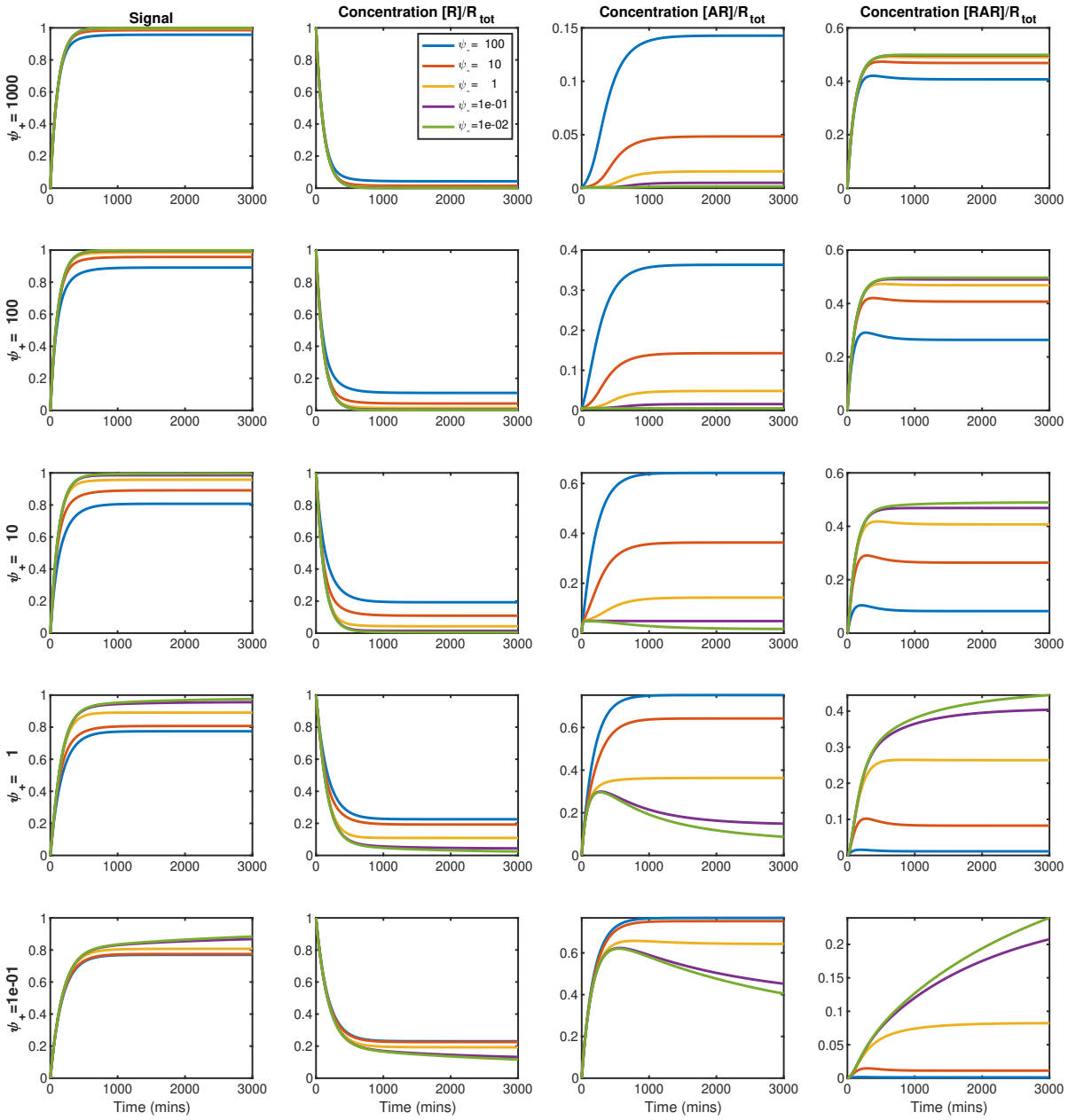

Figure 4.3.1: A numerical investigation into the effects of cooperativity factors ψ_+ and ψ_-. The columns show the overall signal as well as individual species (equations 4.1) while in each row we fix the ψ_+ value. Each plot then shows a varied ψ_-. Plots are created with ligand concentration $[A] = 10^{-10}M$.

curves.

On closer inspection we see three features of interest. Firstly, when ψ_- is small, and therefore a slowed dimer dissociation rate, we see peaks forming in the $[AR]$ curve. Although monomers are initially bound to the ligand, once these form dimers, having a slow dimer dissociation rate reduces the chance of them returning to monomers, so for a time each dimer is essentially stuck in this form, and so the concentration of bound monomers falls. This also explains the extended time for the system to reach equilibrium with small ψ_- as this interchange between bound monomers and dimers increases the time in which it takes the concentrations to come into balance. Conversely, if ψ_- is large, peaks appear instead in the dimer concentration, $[RAR]$. Ligands bind monomers and in turn form dimers. However, with an increased dimer dissociation rate these quickly return back to monomers, and so we see a peak in the curve.

The final feature we notice is again in the $[AR]$ curve, but this time when ψ_+ is large. In this case the $[AR]$ curves in particular appear to have an initial rapid increase in concentration, indicating that we have multiple time scales. The forward cooperativity factor being large drastically increases the rate at which bound monomers form dimers, and so $[RAR]$ increases much faster than when ψ_+ was smaller. This in turn means that there are fewer available monomers to become bound, and so $[AR]$ remains low. Overall, however, the dimerisation happens on a much shorter time scale, hence we see this short transient period in the $[AR]$ concentration curve. We also note that these effects are combined, for example, when we have a large ψ_+ with a small ψ_-, we see the concentration of $[AR]$ increases rapidly to a peak before decreasing towards equilibrium. The analytical tools we used in Chapter 2, for the linear GPCR system, do not apply here, so we instead use asymptotic analysis to explore system (4.1).

4.4 Dimensionless differential equations

To be able to analyse system (4.1) asymptotically we first require it to be nondimensionalised. To rewrite the variables to nondimensional form we first give the units of all

current variables in Table 4.1.

Variable	Units
$[R]$	M
$[AR]$	M
$[RAR]$	M
$[A]$	M
t	s

Parameter	Units
k_+	$M^{-1}s^{-1}$
k_-	s^{-1}
ψ_+	-
ψ_-	-
R_{tot}	M

Table 4.1: The units of the variables and parameters in the system (4.1).

We begin by nondimensionalising $[R]$. As $[R]$ has units M, for balance, we require a scalar that also has units M to rescale with. The natural choice for this is R_{tot}, and so we let dimensionless free receptor r be given by

$$[R] = R_{tot}r \qquad \Rightarrow \; r = \frac{[R]}{R_{tot}}. \tag{4.22}$$

To scale $[AR]$ and $[RAR]$ we first recall the equilibrium concentrations

$$[AR] = K_A[A][R], \qquad [RAR] = \psi K_A[R][AR] = \psi K_A^2[A][R]^2, \tag{4.23}$$

We use these to give us scalings for $[AR]$ and $[RAR]$ as these are the natural scalings to ensure that $[AR]$ and $[ARA]$ are $O(1)$ at equilibrium. Hence, we have

$$[AR] = K_A[A]R_{tot}p \qquad \Rightarrow \; p = \frac{[AR]}{K_A[A]R_{tot}}, \tag{4.24}$$

$$[RAR] = \psi K_A^2[A]R_{tot}^2 q \qquad \Rightarrow \; q = \frac{[RAR]}{\psi K_A^2[A]R_{tot}^2}, \tag{4.25}$$

where r, p and q are the new nondimensional free receptor, single bound receptor and dimerised receptor, respectively. As t has dimension s, we nondimensionalise by a param-

eter with the same dimension. As equilibrium is reached approximately when $t = k_-^{-1}$, we choose to use k_- to scale with, thus

$$t = \frac{\tilde{t}}{k_-} \qquad \Rightarrow \ \tilde{t} = k_- t. \tag{4.26}$$

Substituting these back into system (4.5) gives

$$\frac{dr}{d\tilde{t}} = -\left(\alpha + \frac{\gamma}{2}\right) r + \left(1 - \frac{\gamma}{2}\right) \alpha p - \alpha^2 \beta r p + \frac{\gamma}{2}, \tag{4.27a}$$

$$\frac{dp}{d\tilde{t}} = \left(1 - \frac{\gamma}{2\alpha}\right) r - \left(1 + \frac{\gamma}{2}\right) p - \alpha \beta r p + \frac{\gamma}{2\alpha}. \tag{4.27b}$$

where

$$\alpha = K_A[A], \quad \beta = \frac{\psi_+ R_{tot}}{[A]}, \quad \gamma = \psi_-, \tag{4.28}$$

are the dimensionless parameters for the system. The initial conditions, in equation (4.5c), now become

$$r(0) = 1, \qquad p(0) = 0. \tag{4.29}$$

It is also worth noting that the conservation law, given in equation (4.4), now becomes

$$1 = r + \alpha p + \frac{2\alpha^2 \beta}{\gamma} q, \tag{4.30}$$

which can be used to calculate the concentration of q where necessary. Biologically, α is the ligand concentration scaled by equilibrium dissociation rate, so $\alpha \gg 1$ if either the ligand concentration is very high or has a very high affinity. The grouped parameter β depends on the cooperativity factor ψ_+ as well as the ligand-receptor ratio, so can be

considered a scaled dimerisation rate. So $\beta \gg 1$ if either ψ_+ is large, which leads to an increase in the rate at which dimers are formed, or there is a much higher concentration of receptors than there is ligand. As γ is simply the parameter ψ_-, if $\gamma \gg 1$ then we have fast dimer dissociation. We can also state the equilibrium solutions for this dimensionless system as

$$r^{eq} = p^{eq} = \frac{\gamma\left(-(1+\alpha) + \sqrt{(1+\alpha)^2 + 8\frac{\alpha^2\beta}{\gamma}}\right)}{4\alpha^2\beta}, \qquad q^{eq} = (r^{eq})^2. \tag{4.31}$$

4.5 Asymptotic analysis

In Section 4.3 we noticed that a choice of small ψ_- leads to a peak in the $[AR]$ curve, whilst a large ψ_- leads to a peak instead in $[RAR]$. Also, choosing ψ_+ to be large gives a rapid increase in $[AR]$. This leads to three cases to investigate using asymptotic analysis to further understand these features. As the system is now governed by the three nondimensional parameters α, β and γ, we use these to gain insights into how it will evolve over time. To do this we use the nondimensional system in equations (4.27), though for simplicity we drop the tildes. Binding rates of VEGF are well documented [80] (and references within), and so throughout this section we assume $\alpha = O(1)$, allowing us to focus on the possible effects of cooperativity.

4.5.1 Small γ asymptotics (slow dimer dissociation)

In Figure 4.3.1 we saw that having a slow dimer dissociation rate gave peaks in $[AR]$. As a slow dimer dissociation is equivalent of having a small γ in the dimensionless system, we set $\gamma = \epsilon$, where $\epsilon \ll 1$. Substituting this into system (4.27) gives

$$\frac{dr}{dt} = -\left(\alpha + \frac{\epsilon}{2}\right) r + \left(1 - \frac{\epsilon}{2}\right) \alpha p - \alpha^2 \beta r p + \frac{\epsilon}{2}, \tag{4.32a}$$

$$\frac{dp}{dt} = \left(1 - \frac{\epsilon}{2\alpha}\right) r - \left(1 + \frac{\epsilon}{2}\right) p - \alpha\beta r p + \frac{\epsilon}{2\alpha}, \tag{4.32b}$$

where conservation gives

$$q = \frac{\epsilon}{2\alpha^2\beta}(1 - r - \alpha p) \tag{4.32c}$$

This is a regular perturbed problem. Assuming asymptotic approximations

$$r \approx r_0 + \epsilon r_1 + ..., \quad p \approx p_0 + \epsilon p_1 + ..., \tag{4.33}$$

where each r_i, p_i is $O(1)$, and substituting these into equations (4.32) gives the leading order problem at $O(1)$ as

$$\frac{dr_0}{dt} = -\alpha r_0 + \alpha p_0 - \alpha^2\beta r_0 p_0, \tag{4.34a}$$

$$\frac{dp_0}{dt} = r_0 - p_0 - \alpha\beta r_0 p_0, \tag{4.34b}$$

$$r(0) = 1, \quad p(0) = 0, \tag{4.34c}$$

Although this is still a nonlinear system that we are not able to solve analytically, we can still make some observations. The terms that remain in the ODEs are the linear terms, representing the reversible binding of the ligand, as well as the nonlinear term that dictates the formation of the dimer, that is, the forward dimerisation. With $\gamma = \epsilon$, the terms representing the dissociation of the dimers are approximately zero, hence do not appear in these equations. As such we would expect that as dimers are formed they become stuck in this form, hence we would expect to see the concentration of dimers increasing monotonically. In Figure 4.5.1 we plot the solutions of the full, nondimensional system, as given in equations (4.27), along with the approximated system in equations (4.34). In this we see there is little difference in the two solutions, confirming that negating the terms in the approximated solution has little effect on the overall dynamics.

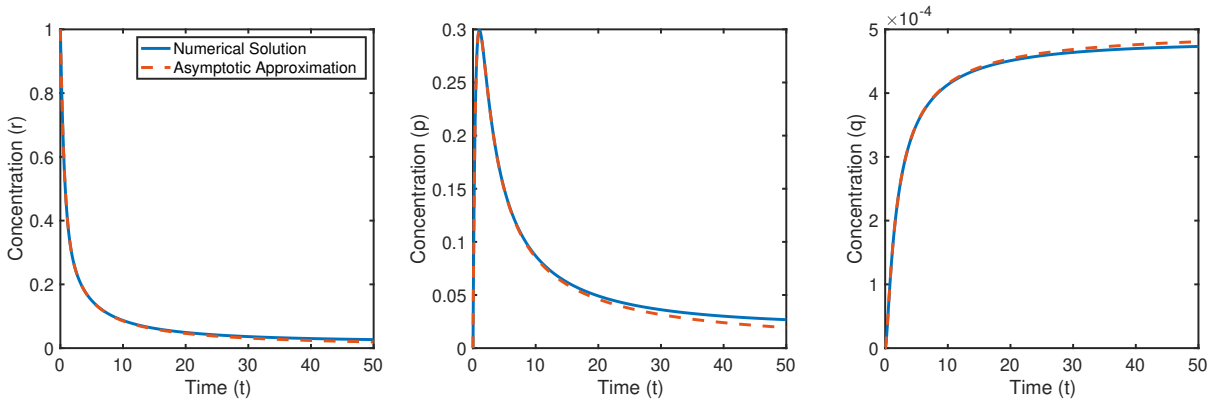

Figure 4.5.1: Both the solution to the full system in equations (4.32) is plotted along side the solution of the reduced, approximated system (equations (4.34)) in red. Conservation of receptors is used to show the approximation for q. Parameters used to create plots are $\alpha = \beta = 1$ and $\gamma = \epsilon = 0.001$.

4.5.2 Large β asymptotics (fast dimerisation)

Another feature appearing in Figure 4.3.1 is that, when ψ_+ is large (particularly in the top row in the figure), we see an initial increase in the concentration of $[AR]$ in a short time period, before the concentration increases towards equilibrium. This indicates that the concentration is evolving on at least two different timescales. This happens when we have a fast dimerisation and ψ_+ large, and since ψ_+ is incorporated in β, we set $\beta = 1/\epsilon$, and hence β is large. Substituting this into system (4.27) and multiplying by ϵ gives

$$\epsilon \frac{dr}{dt} = -\epsilon\left(\alpha + \frac{\gamma}{2}\right) r + \epsilon\left(1 - \frac{\gamma}{2}\right)\alpha p - \alpha^2 rp + \frac{\gamma\epsilon}{2}, \tag{4.35a}$$

$$\epsilon \frac{dp}{dt} = \epsilon\left(1 - \frac{\gamma}{2\alpha}\right) r - \epsilon\left(1 + \frac{\gamma}{2}\right) p - \alpha rp + \frac{\gamma\epsilon}{2\alpha}, \tag{4.35b}$$

$$r(0) = 1, \quad p(0) = 0. \tag{4.35c}$$

where conservation (from equation (4.30)) gives

$$q = \frac{\gamma\epsilon}{2\alpha^2}(1 - r - \alpha p) \tag{4.35d}$$

As ϵ is multiplying the highest derivatives, this is a singular perturbed problem, which supports our observation of multiple timescales in the concentration curves. In Figure 4.5.2 we view the numerical solution to this system on a log-log scale where it also is apparent that the dynamics are evolving on more than one time scale. This plot is used to suggest the time scales needed to asymptotically analyse the system. We find that we require three solutions to accurately approximate all of the intricacies of the curves.

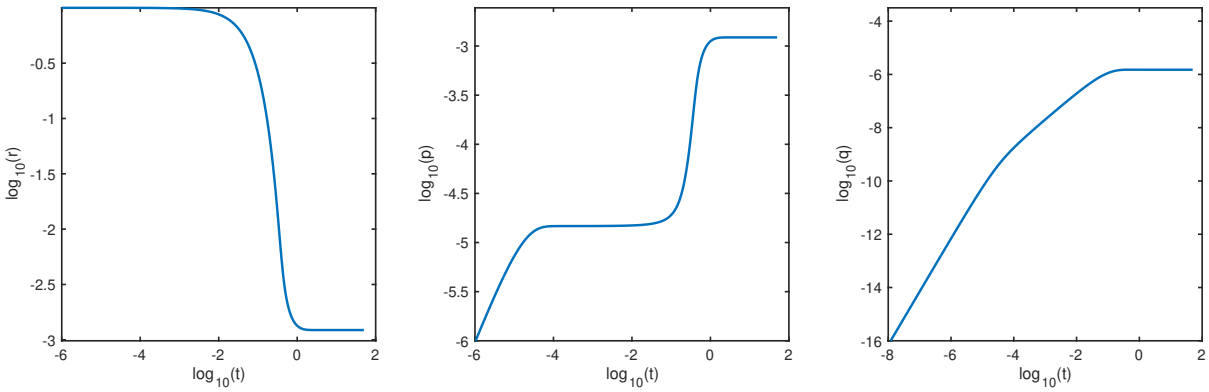

Figure 4.5.2: Numerical solutions of the full system (equation (4.35)), on a log-log scale. Plot created with $\alpha = 6.8$, $\beta = 1/\epsilon$, $\gamma = 1.4$ with $\epsilon = 10^{-4}$.

Inner solution

Looking at Figure 4.5.2, we see two plateaus in the p curve, with the first one appearing when $t = O(\epsilon)$, hence we assume this to be first time scale of interest. We call this region the inner region. We also note that within this region, p is of $O(\epsilon)$, while r remains $O(1)$. Hence we rescale

$$p = \epsilon\tilde{p}, \quad t = \epsilon\tau. \tag{4.36}$$

For matching convenience later, we also use the change of variable $r = \tilde{r}$. This now gives the inner layer problem as

$$\frac{d\tilde{r}}{d\tau} = -\epsilon\left(\alpha + \frac{\gamma}{2}\right)\tilde{r} + \epsilon^2\left(1 - \frac{\gamma}{2}\right)\alpha\tilde{p} - \alpha^2\epsilon\tilde{r}\tilde{p} + \frac{\gamma\epsilon}{2}, \tag{4.37a}$$

$$\frac{d\tilde{p}}{d\tau} = \left(1 - \frac{\gamma}{2\alpha}\right)\tilde{r} - \epsilon\left(1 + \frac{\gamma}{2}\right)\tilde{p} - \alpha\tilde{r}\tilde{p} + \frac{\gamma}{2\alpha}, \tag{4.37b}$$

$$\tilde{r}(0) = 1, \quad \tilde{p}(0) = 0. \tag{4.37c}$$

Substituting asymptotic approximations

$$\tilde{r} \approx \tilde{r}_0 + \epsilon\tilde{r}_1 + ..., \quad \tilde{p} \approx \tilde{p}_0 + \epsilon\tilde{p}_1 + ..., \tag{4.38}$$

gives a leading order problem of

$$\frac{d\tilde{r}_0}{d\tau} = 0, \tag{4.39a}$$

$$\frac{d\tilde{p}_0}{d\tau} = \left(1 - \frac{\gamma}{2\alpha}\right)\tilde{r}_0 - \alpha\tilde{r}_0\tilde{p}_0 + \frac{\gamma}{2\alpha}, \tag{4.39b}$$

$$\tilde{r}_0(0) = 1, \quad \tilde{p}_0(0) = 0. \tag{4.39c}$$

This has the solution

$$\tilde{r}_0(\tau) = 1, \qquad \tilde{p}_0(\tau) = \frac{1 - e^{-\alpha\tau}}{\alpha}, \tag{4.40}$$

which is valid while $t \ll \epsilon$. Interestingly, the solution $p(t)$ depends only on α. Biologically, this implies that, in this initial layer, the concentration of p, and therefore $[AR]$, is dominated by the first reaction, which is the binding and unbinding of the ligand to a monomeric receptor.

Outer solution

Figure 4.5.2 suggests that the long time solution of both r and p are $\mathrm{O}(\epsilon^{\frac{1}{2}})$. We can also confirm this using the equilibrium concentration as stated in equation (4.31). Setting $\beta = 1/\epsilon$ in this, and using a Taylor series expansion to approximate the square root term gives

$$r^{eq} = p^{eq} = \frac{\epsilon\gamma\left(-(1+\alpha) + \sqrt{(1+\alpha)^2 + 8\frac{\alpha^2}{\epsilon\gamma}}\right)}{4\alpha^2} \tag{4.41}$$

$$= \frac{\epsilon\gamma}{4\alpha^2}\left[-(1+\alpha) + \frac{2\alpha\sqrt{2}}{\sqrt{\gamma\epsilon}} + \frac{(1+\alpha)^2\sqrt{\epsilon\gamma}}{4\alpha\sqrt{2}} + O(\epsilon^{\frac{3}{2}})\right] \tag{4.42}$$

$$= \frac{\sqrt{\gamma}}{\alpha\sqrt{2}}\epsilon^{\frac{1}{2}} + \textit{higher order terms} \tag{4.43}$$

hence confirming that the long time solution of both r and p are $\mathrm{O}(\epsilon^{\frac{1}{2}})$. As such we rescale these by substituting

$$r = \epsilon^{\frac{1}{2}}\hat{r}, \qquad p = \epsilon^{\frac{1}{2}}\hat{p}. \tag{4.44}$$

in equations (4.35). After substitution we have

$$\epsilon^{\frac{1}{2}}\frac{d\hat{r}}{dt} = -\epsilon^{\frac{1}{2}}\left(\alpha + \frac{\gamma}{2}\right)\hat{r} + \epsilon^{\frac{1}{2}}\left(1 - \frac{\gamma}{2}\right)\alpha\hat{p} - \alpha^2\hat{r}\hat{p} + \frac{\gamma}{2}, \tag{4.45a}$$

$$\epsilon^{\frac{1}{2}}\frac{d\hat{p}}{dt} = \epsilon^{\frac{1}{2}}\left(1 - \frac{\gamma}{2\alpha}\right)\hat{r} - \epsilon^{\frac{1}{2}}\left(1 + \frac{\gamma}{2}\right)\hat{p} - \alpha\hat{r}\hat{p} + \frac{\gamma}{2\alpha}, \tag{4.45b}$$

We assume asymptotic expansions

$$\hat{r} = \hat{r}_0 + \epsilon^{\frac{1}{2}}\hat{r}_1 + ..., \quad \hat{p} = \hat{p}_0 + \epsilon^{\frac{1}{2}}\hat{p}_1 + ..., \tag{4.46}$$

which give a leading order approximation as

$$0 = \frac{\gamma}{2} - \alpha^2 \hat{r}_0 \hat{p}_0 \qquad \Rightarrow \quad \hat{r}_0 \hat{p}_0 = \frac{\gamma}{2\alpha^2}. \tag{4.47}$$

To find the expressions for $\hat{r}_0$ and $\hat{p}_0$ we also need to consider the $\mathrm{O}(\epsilon^{\frac{1}{2}})$ terms in system (4.45):

$$\frac{d\hat{r}_0}{dt} = -\left(\alpha + \frac{\gamma}{2}\right)\hat{r}_0 + \left(1 - \frac{\gamma}{2}\right)\alpha\hat{p}_0 - \alpha^2(\hat{r}_0\hat{p}_1 + \hat{r}_1\hat{p}_0), \tag{4.48a}$$

$$\frac{d\hat{p}_0}{dt} = \left(1 - \frac{\gamma}{2\alpha}\right)\hat{r}_0 - \left(1 + \frac{\gamma}{2}\right)\hat{p}_0 - \alpha(\hat{r}_0\hat{p}_1 + \hat{r}_1\hat{p}_0). \tag{4.48b}$$

To eliminate the nonlinear terms coming from the higher order terms in the expansion we compute ((4.48a)$-\alpha$(4.48b)) to give

$$\frac{d\hat{r}_0}{dt} - \alpha\frac{d\hat{p}_0}{dt} = 2\alpha(\hat{p}_0 - \hat{r}_0). \tag{4.49}$$

We can use equation (4.47) to eliminate $\hat{p}_0$ by substituting

$$\hat{p}_0 = \frac{\gamma}{2\alpha^2\hat{r}_0}, \quad \text{and} \quad \frac{d\hat{p}_0}{dt} = -\frac{\gamma}{2\alpha^2\hat{p}_0^2}\frac{d\hat{r}_0}{dt}, \tag{4.50}$$

giving the following ODE for the leading order outer solution as

$$\frac{d\hat{r}_0}{dt} = \frac{2\hat{r}_0(\gamma - 2\alpha^2\hat{r}_0^2)}{\gamma + 2\alpha\hat{r}_0^2}. \tag{4.51}$$

This equation can be solved by using separation of variables, with the solution $\hat{r}_0(t)$ given implicitly by

$$2\alpha \log(\hat{r}_0) - (\alpha+1)\log(|2\alpha^2\hat{r}_0^2 - \gamma|) = 2\alpha(2t + c_1), \tag{4.52}$$

where the constant c_1 will be determined by a matching condition. Using equation (4.43) we find that

$$\hat{r}_0 \in \left[\frac{\sqrt{\gamma}}{\alpha\sqrt{2}}, 1\right] \qquad \Rightarrow \qquad \hat{r}_0^2 > \frac{\gamma}{2\alpha^2}, \tag{4.53}$$

hence we can remove the absolute value, giving the solution as

$$2\alpha \log(\hat{r}_0) - (\alpha+1)\log(2\alpha^2\hat{r}_0^2 - \gamma) = 2\alpha(2t + c_1). \tag{4.54}$$

Using equation (4.47) we can also give the implicit solution for $\hat{p}_0$ as

$$2\alpha \log\left(\frac{\gamma}{2\alpha^2\hat{p}_0}\right) - (\alpha+1)\log\left(\frac{\gamma^2}{2\alpha^2\hat{p}_0^2} - \gamma\right) = 2\alpha(2t + c_1). \tag{4.55}$$

Matching

The usual method (as outlined in [96]) at this point would be to match the inner and outer solutions to determine the constant c_1. This would be done by taking the limits of the inner solutions as $\tau \to \infty$ and equating these with the limits of the outer solutions as $t \to 0$. However, we find that these solutions do not match. Taking the inner solution limits, we have

$$\lim_{\tau\to\infty} \tilde{r}_0 = 1, \qquad \lim_{\tau\to\infty} \tilde{p}_0 = \frac{1}{\alpha}. \tag{4.56}$$

The product of these is

$$\tilde{r}_0\tilde{p}_0 = \frac{1}{\alpha}, \tag{4.57}$$

which is a clear contradiction to the product of the outer solutions as found in equation (4.47), hence there is no way that both solutions r and p can match.

Intermediate solution

To overcome the matching problem we consider the solutions in an intermediate layer, between the inner and outer layer. Looking at Figure 4.5.2 we can see that, as t tends to $O(1)$, then r and p begin at order 1 and ϵ respectively before tending to the equilibrium, so for an intermediate solution we rescale

$$r = \bar{r}, \qquad p = \epsilon\bar{p}. \tag{4.58}$$

We note here that these scalings are the same scalings we used for the inner solutions (although we use a separate notation for clarity when matching), in this intermediate layer, time instead remains $O(1)$. Substituting these into equations (4.35) gives

$$\frac{d\bar{r}}{dt} = -\left(\alpha + \frac{\gamma}{2}\right)\bar{r} + \epsilon\left(1 - \frac{\gamma}{2}\right)\alpha\bar{p} - \alpha^2\bar{r}\bar{p} + \frac{\gamma}{2}, \tag{4.59a}$$

$$\epsilon\frac{d\bar{p}}{dt} = \left(1 - \frac{\gamma}{2\alpha}\right)\bar{r} - \epsilon\left(1 + \frac{\gamma}{2}\right)\bar{p} - \alpha\bar{r}\bar{p} + \frac{\gamma}{2\alpha}, \tag{4.59b}$$

which gives the leading order problem as

$$\frac{d\bar{r}_0}{dt} = -\left(\alpha + \frac{\gamma}{2}\right)\bar{r}_0 - \alpha^2\bar{r}_0\bar{p}_0 + \frac{\gamma}{2}, \tag{4.60a}$$

$$0 = \left(1 - \frac{\gamma}{2\alpha}\right)\bar{r}_0 - \alpha\bar{r}_0\bar{p}_0 + \frac{\gamma}{2\alpha}. \tag{4.60b}$$

Solving this system gives

$$\bar{r}_0(t) = c_2 e^{-2\alpha t}, \qquad \bar{p}_0(t) = \frac{(2\alpha - \gamma)c_2 + \gamma e^{2\alpha t}}{2\alpha^2 c_2}. \tag{4.61}$$

Matching the intermediate to inner region

Matching this intermediate solution to the inner solution we require

$$\lim_{t\to 0} \bar{r}_0(t) = \lim_{\tau\to\infty} \tilde{r}_0(\tau). \tag{4.62}$$

Taking these limits gives

$$c_2 = 1. \tag{4.63}$$

We also require p to match, so calculating

$$\lim_{t\to 0} \bar{p}_0(t) = \frac{1}{\alpha}, \qquad \text{and} \qquad \lim_{\tau\to\infty} \bar{p}_0(\tau) = \frac{1}{\alpha}, \tag{4.64}$$

confirms that the intermediate solution matches to the inner solution. These give full intermediate solutions as

$$\bar{r}_0(t) = e^{-2\alpha t}, \qquad \bar{p}_0(t) = \frac{(2\alpha - \gamma) + \gamma e^{2\alpha t}}{2\alpha^2}. \tag{4.65}$$

As the parameter γ appears in this solution we can see that as we move into this region the second reaction, whereby the ligand binds a second monomer, now also plays a role in the dynamics of the system.

Matching the intermediate to outer region

We also match this intermediate solution to that of the outer layer, we determine

$$\lim_{t\to 0}\hat{r}_0(t) = \lim_{t\to\infty}\epsilon^{-\frac{1}{2}}\bar{r}_0(t), \qquad \Rightarrow \qquad \lim_{t\to 0}\hat{r}_0(t) = \epsilon^{-\frac{1}{2}}, \tag{4.66}$$

gives the constant of the outer solution as

$$c_1 = \frac{1}{2\alpha}\log\left(\frac{\epsilon}{(2\alpha^2-\gamma\epsilon)^{\alpha+1}}\right). \tag{4.67}$$

We also need to confirm that the solutions for p match in this region. To do this we use equation (4.47), but first noting that

$$\bar{r}_0(t)\bar{p}_0(t) = \frac{(2\alpha-\gamma)e^{-2\alpha t}+\gamma}{2\alpha^2}, \tag{4.68}$$

we have, with all variables rescaled back to $O(1)$, that

$$\lim_{t\to 0}\hat{r}_0(t)\hat{p}_0(t) = \frac{\gamma}{2\alpha^2\epsilon} \qquad \text{and} \qquad \lim_{t\to\infty}\bar{r}_0(t)\bar{p}_0(t) = \frac{\gamma}{2\alpha^2\epsilon}. \tag{4.69}$$

As we have already determined that the solutions for r match, we can therefore conclude that p must also match. Figure 4.5.3 shows a good match between the intermediate solution and both the inner and outer solutions for all variables, and together the approximations agree with the numerics in each region. These solutions give us insights into the reaction sequences that contribute to the interesting dynamics we see under these conditions. Initially, while time is $O(\epsilon)$, we have solutions depending only on α, indicating that only the first reaction, binding a ligand molecule to a monomer, contributes to an increase in p, whilst there almost no change in r. Once we move to the intermediate region, that is as $t = O(\epsilon^\theta)$ where $\theta > 1$, we see how the second reaction, whereby the

ligand binds a second monomer, begins to have an effect. The solutions indicate that the concentration of r begins to decay exponentially, although this solution is still dependent only on α, so we can assume that this concentration is still governed by the first reaction. However, looking to the solution for p, we see that the concentration increases exponentially, but also the solution contains the parameter γ, hence both reactions play a part in the dynamics in this region. Moving to the outer region, we see how all solutions now depend on both parameters, so we can conclude that by this point, both reactions now affect all concentrations.

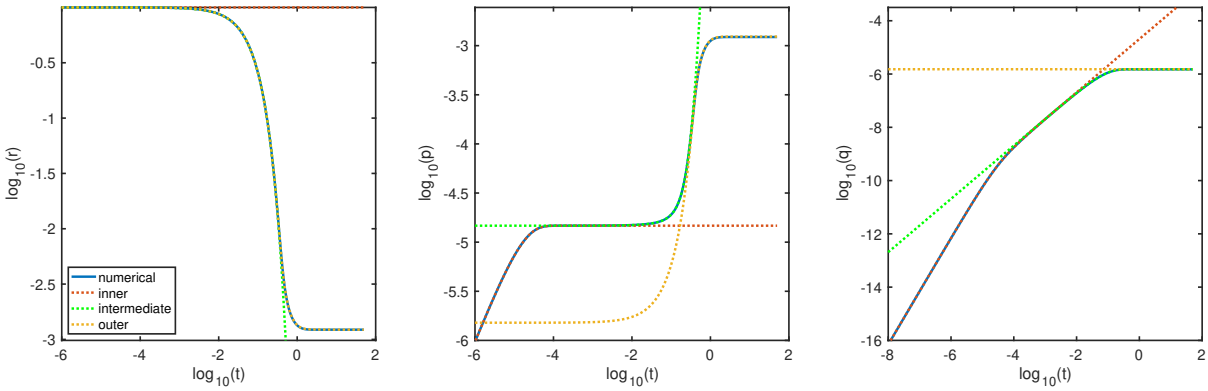

Figure 4.5.3: Plotting the numerical solution of equation (4.35) with inner (equation (4.40))), intermediate (equation (4.65)) and outer (equations (4.52) and (4.55)) solutions. The intermediate solution matches both the inner and outer solution, creating a full approximation to the numerical solution. Conservation of receptors is used to show the approximation for q. Plot created with $\alpha = 6.8$, $\gamma = 1.4$ and $\beta = \epsilon = 10^{-4}$.

4.5.3 Large γ asymptotics

The last case we consider is when we have fast dimer dissociation, that is when γ is large. In Figure 4.3.1 we saw that taking ψ_-, and hence $\gamma = \psi_-$, large caused a peak in the bound dimer, $[RAR]$, curve. As such we set $\gamma = 1/\epsilon$, with $\epsilon \ll 1$, in equations (4.27),

and multiplying with ϵ, this gives

$$\epsilon\frac{dr}{dt} = -\left(\alpha\epsilon + \frac{1}{2}\right) r + \left(\epsilon - \frac{1}{2}\right) \alpha p - \alpha^2\beta\epsilon r p + \frac{1}{2}, \tag{4.70a}$$

$$\epsilon\frac{dp}{dt} = \left(\epsilon - \frac{1}{2\alpha}\right) r - \left(\epsilon + \frac{1}{2}\right) p - \alpha\beta\epsilon r p + \frac{1}{2\alpha}, \tag{4.70b}$$

$$r(0) = 1, \quad p(0) = 0. \tag{4.70c}$$

where conservation (from equation (4.30)) gives

$$q = \frac{1}{2\alpha^2\beta\epsilon}(1 - r - \alpha p) \tag{4.70d}$$

As ϵ is multiplying the highest order terms we find we again have a singular perturbation problem indicating that the solutions evolves on multiple time scales. In Figure 4.5.4 we plot the solutions on a log-log scale, to help determine the required scalings.

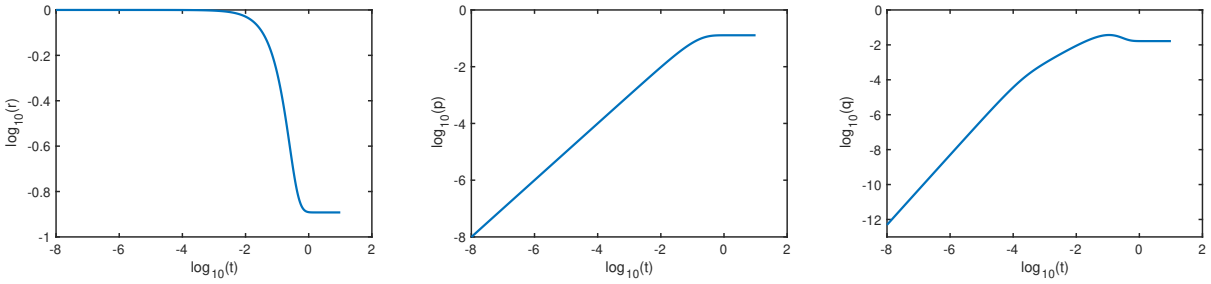

Figure 4.5.4: Numerical solutions of the full system in equation (4.70) on a log-log scale. Plot created with α=6.8, β=1, $\gamma = \epsilon = 10^{-4}$.

Inner Solution

To determine the inner region we rescale time with $t = \epsilon^a \tau$. We also rescale $r = \epsilon^b \tilde{r}$ and $p = \epsilon^c \tilde{p}$ giving the system as

$$\epsilon^{1-a+b}\frac{d\tilde{r}}{d\tau} = -\left(\alpha\epsilon + \frac{1}{2}\right)\epsilon^b\tilde{r} + \left(\epsilon - \frac{1}{2}\right)\alpha\epsilon^c\tilde{p} - \alpha^2\beta\epsilon^{1+b+c}\tilde{r}\tilde{p} + \frac{1}{2}, \tag{4.71a}$$

$$\epsilon^{1-a+c}\frac{d\tilde{p}}{d\tau} = \left(\epsilon - \frac{1}{2\alpha}\right)\epsilon^b\tilde{r} - \left(\epsilon + \frac{1}{2}\right)\epsilon^c\tilde{p} - \alpha\beta\epsilon^{1+b+c}\tilde{r}\tilde{p} + \frac{1}{2\alpha}. \tag{4.71b}$$

To determine a, b and c we use a principle called detailed (or dominant) balance. In Figure 4.5.4 we see that, up to $t = O(1)$ both t and p evolve on the same order, indicating that $a = c$. Furthermore, the relationship between p and t is linear, hence we balance the LHS of (4.71b) with the p term of the equation

$$1 - a + c = c, \qquad \Rightarrow \qquad a = c = 1. \tag{4.72}$$

As this results in ϵ still multiplying the LHS, to have a nontrivial solution, we must have ϵ^0 on the LHS of (4.71a), and so

$$1 - a + b = 0, \qquad \Rightarrow \qquad b = 0. \tag{4.73}$$

Substituting these into (4.71) gives

$$\frac{dr}{d\tau} = -\left(\alpha\epsilon + \frac{1}{2}\right)r + \left(\epsilon - \frac{1}{2}\right)\alpha\epsilon p - \alpha^2\beta\epsilon^2 rp + \frac{1}{2}, \tag{4.74a}$$

$$\epsilon\frac{dp}{d\tau} = \left(\epsilon - \frac{1}{2\alpha}\right)r - \left(\epsilon + \frac{1}{2}\right)\epsilon p - \alpha\beta\epsilon^2 rp + \frac{1}{2\alpha}. \tag{4.74b}$$

$$r(0) = 1, \quad p(0) = 0 \tag{4.74c}$$

Assuming asymptotic approximations

$$\tilde{r} \approx \tilde{r}_0 + \epsilon\tilde{r}_1 + ..., \quad \tilde{p} \approx \tilde{p}_0 + \epsilon\tilde{p}_1 + ..., \tag{4.75}$$

gives a leading order problem of

$$\frac{dr}{d\tau} = -\frac{1}{2}\tilde{r}_0 + \frac{1}{2}, \tag{4.76a}$$

$$0 = -\frac{1}{2\alpha}\tilde{r}_0 + \frac{1}{2\alpha}. \tag{4.76b}$$

which gives the solution

$$\tilde{r}_0(\tau) = 1. \tag{4.77}$$

Considering the $O(\epsilon)$ terms gives

$$\frac{d\tilde{r}_1}{d\tau} = -\alpha\tilde{r}_0 - \frac{1}{2}\tilde{r}_1 - \frac{1}{2}\alpha\tilde{p}_0, \tag{4.78a}$$

$$\frac{d\tilde{p}_0}{d\tau} = \tilde{r}_0 - \frac{1}{2\alpha}\tilde{r}_1 - \frac{1}{2}\tilde{p}_0, \tag{4.78b}$$

$$\tilde{r}_1(0) = 0, \quad \tilde{p}_0(0) = 0. \tag{4.78c}$$

Substituting in $\tilde{r}_0 = 1$ and solving, yields

$$\tilde{r}_1(\tau) = -\alpha\tau, \quad \tilde{p}_0(\tau) = \tau. \tag{4.79}$$

As the features we are particularly interested in occur in $[RAR]$, and therefore q, curve, we also give these solutions. To do this we require the conservation law (equation (4.70d))

which, noting that in the inner region $\tilde{q} = O(\epsilon)$, states that

$$1 = (\tilde{r}_0 + \epsilon\tilde{r}_1 + ...) + \alpha\epsilon(\tilde{p}_0 + \epsilon\tilde{p}_1 + ...) + 2\alpha^2\beta\epsilon^2(\tilde{q}_0 + \epsilon\tilde{q}_1 + ...). \tag{4.80}$$

Equating powers of ϵ, we have

$$1 = \tilde{r}_0, \tag{4.81}$$

$$0 = \tilde{r}_1 + \alpha\tilde{p}_0, \tag{4.82}$$

$$0 = \tilde{r}_2 + \alpha\tilde{p}_1 + 2\alpha^2\beta\tilde{q}_0. \tag{4.83}$$

We see that, to find the leading order solution for q we require further solutions for r and p. To find these we consider $O(\epsilon^2)$ terms in system (4.74a), which, after substituting in $\tilde{r}_0, \tilde{r}_1$ and $\tilde{p}_0$, gives

$$\frac{d\tilde{r}_2}{d\tau} = (\alpha + \alpha^2 - \alpha^2\beta)\tau - \frac{1}{2}\tilde{r}_1 - \frac{1}{2}\alpha\tilde{p}_0, \tag{4.84a}$$

$$\frac{d\tilde{p}_1}{d\tau} = -(1 + \alpha + \alpha\beta)\tau - \frac{1}{2\alpha}\tilde{r}_1 - \frac{1}{2}\tilde{p}_0, \tag{4.84b}$$

$$\tilde{r}_2(0) = 0, \quad \tilde{p}_1(0) = 0. \tag{4.84c}$$

This is again a linear system of ODEs, whose solution is

$$\tilde{r}_2(\tau) = \frac{\alpha}{2}(\alpha + 1)\tau^2 - \alpha^2\beta(\tau + e^{-\tau} - 1), \tag{4.85}$$

$$\tilde{p}_1(\tau) = -\frac{1}{2}(\alpha + 1)\tau^2 - \alpha\beta(\tau + e^{-\tau} - 1). \tag{4.86}$$

Substituting this into equation (4.78c) gives the leading order solution for q as

$$\tilde{q}_0(\tau) = \tau + e^{-\tau} - 1. \tag{4.87}$$

Together, these give the inner solution to the system as

$$\tilde{r}_0(\tau) = 1. \qquad \tilde{p}_0(\tau) = \tau, \qquad \tilde{q}_0(\tau) = \tau + e^{-\tau} - 1. \tag{4.88}$$

The solution for p, in this region, has unbounded, linear growth, while the solution for q is also increasing monotonically. However, these solutions are independent of any model parameters indicating that, under the conditions of large γ, the curves will evolve in this way regardless of the parameters α and β.

Outer Solution

For an outer solution, while time is $O(1)$, r and p are also $O(1)$. This is confirmed mathematically by looking to the equilibrium concentration in equation 4.31, with $\gamma = 1/\epsilon$. Using Taylor series approximation for the square root term, we have

$$r^{eq} = p^{eq} = \frac{\left(-(1+\alpha) + \sqrt{(1+\alpha)^2 + 8\alpha^2\beta\epsilon}\right)}{4\alpha^2\beta\epsilon}, \tag{4.89}$$

$$= \frac{1}{4\alpha^2\beta\epsilon}\left[-(1+\alpha) + (1+\alpha) + \frac{4\alpha^2\beta\epsilon}{1+\alpha} + O(\epsilon^2)\right], \tag{4.90}$$

$$= \frac{1}{1+\alpha} + O(\epsilon) \tag{4.91}$$

hence confirming the long time solutions of r and p as $O(1)$. Assuming asymptotic expansions

$$r \approx r_0 + \epsilon r_1 + ..., \quad p \approx p_0 + \epsilon p_1 + ..., \tag{4.92}$$

and substituting these into the ODEs in (4.70) gives a leading order solution

$$0 = -r_0 - \alpha p_0 + 1 \qquad \Rightarrow \qquad p_0 = \frac{1 - r_0}{\alpha}. \tag{4.93}$$

Looking to the $O(\epsilon)$ terms in (4.70) gives

$$\frac{dr_0}{dt} = -\alpha r_0 - \frac{1}{2}r_1 + \alpha p_0 - \frac{1}{2}\alpha p_1 - \alpha^2\beta r_0 p_0, \tag{4.94a}$$

$$\frac{dp_0}{dt} = r_0 - \frac{1}{2\alpha}r_1 - p_0 - \frac{1}{2}p_1 - \alpha\beta r_0 p_0. \tag{4.94b}$$

We compute ((4.94a)$-\alpha$(4.94b)) to eliminate the nonlinear terms, giving

$$\frac{dr_0}{dt} = 1 - (\alpha + 1)r_0, \tag{4.95}$$

which solves to give

$$r_0 = \frac{c_1 e^{-(\alpha+1)t} + 1}{\alpha + 1}, \qquad p_0 = \frac{1}{\alpha}\left(\frac{\alpha - c_1 e^{-(\alpha+1)t}}{\alpha + 1}\right). \tag{4.96}$$

As the conservation law says

$$1 = (r_0 + \epsilon r_1 + ...) + \alpha(p_0 + \epsilon p_1 + ...) + 2\alpha^2\beta\epsilon(\hat{q}_0 + \epsilon\hat{q}_1 + ...), \tag{4.97}$$

we use this to equate powers of ϵ on the LHS and RHS, giving

$$1 = r_0 + \alpha p_0 \tag{4.98}$$

which, substituting in equations (4.96), confirms this equation is true. Considering the ϵ terms then gives

$$0 = r_1 + \alpha p_1 + 2\alpha^2\beta\hat{q}_0 \qquad \Rightarrow \qquad \hat{q} = \frac{-(r_1 + \alpha p_1)}{2\alpha^2\beta}, \tag{4.99}$$

which, combining with equation (4.94a), gives the time dependent solution for $\hat{q}_0$ as

$$\hat{q}_0 = \frac{-e^{-2(\alpha+1)t}(c_1 - \alpha e^{(\alpha+1)t})(c_1 + e^{(\alpha+1)t})}{\alpha(\alpha+1)^2}. \tag{4.100}$$

Matching

Matching the inner solution to the outer solution we require

$$\lim_{t\to 0} r_0(t) = \lim_{\tau\to\infty} \tilde{r}_0(\tau). \tag{4.101}$$

Taking the limits of $r_0(t)$ in (4.96) and $\tilde{r}_0(\tau)$ in (4.77), gives

$$c_1 = \alpha. \tag{4.102}$$

As the limit of $\tilde{p}_0(\tau)$ in (4.79) as $\tau \to \infty$ is ∞ we instead introduce an intermediate variable to show that p also matches. Recall the inner solution, (rewritten in terms of the original variables) as

$$p_{in}(t) = t, \tag{4.103}$$

where $t = \epsilon\tau$, and outer solution

$$p_{out}(t) = \frac{1 - e^{-(\alpha+1)t}}{\alpha+1}. \tag{4.104}$$

Defining $t = \epsilon^\delta \xi$, where $0 < \delta < 1$ and $\xi = O(1)$. Substituting these into (4.103) and

(4.104) gives

$$p_{in} = \epsilon^{\delta}\xi, \tag{4.105}$$

and

$$p_{out} = \frac{1 - e^{-(\alpha+1)\epsilon^{\delta}\xi}}{\alpha + 1}. \tag{4.106}$$

Expanding the exponential terms in a Taylor series gives

$$\begin{aligned} p_{out} =& \frac{1 - [1 - (\alpha+1))\epsilon^{\delta}\xi + \frac{1}{2}(\alpha+1)^2\epsilon^{2\delta}\xi^2 - ...]}{\alpha+1} \\ =& \epsilon^{\delta}\xi - \textit{higher order terms}, \end{aligned} \tag{4.107}$$

which confirms matching in inner region. The full outer solutions to the system can then be stated as

$$r(t) = \frac{1 + \alpha e^{-(\alpha+1)t}}{\alpha+1}, \quad p(t) = \frac{1 - e^{-(\alpha+1)t}}{\alpha+1}, \quad q(t) = \frac{-e^{-2(\alpha+1)t}(1 - e^{(\alpha+1)t})(\alpha + e^{(\alpha+1)t})}{(\alpha+1)^2}. \tag{4.108}$$

Again, the parameter β does not appear in any of these solutions, suggesting that the first reaction, binding ligand to monomer, dominates the dynamics in this region. Clearly, having such a large γ, and therefore increased rate of dimer dissociation means that any formed dimers quickly return to monomeric state, hence the second reaction, dictating dimer formation has very little effect on any of the concentrations.

The solution for $q(t)$ in (4.108) also governs the peak that appears in the curve.

Differentiating this with respect to t, gives

$$\frac{dq(t)}{dt} = \frac{-e^{-2(\alpha+1)t}(2\alpha + e^{(\alpha+1)t} - \alpha e^{(\alpha+1)t})}{\alpha+1}. \tag{4.109}$$

which, solving for t, gives a stationary point at

$$t_s = \frac{\log\left(\dfrac{2\alpha}{\alpha-1}\right)}{\alpha+1}, \tag{4.110}$$

which has solutions for $\alpha > 1$. To determine the nature of the stationary point, we use the second derivative of $q(t)$ in (4.108), giving

$$\left.\frac{dq(t)}{dt}\right|_{t_s} = -\frac{(\alpha-1)^2}{2\alpha}, \tag{4.111}$$

which is negative, as $\alpha > 0$, thus concluding the stationary point is a maximum. Hence, we conclude that a peak will appear if $\alpha > 1$.

In Figure 4.5.5 we see a good match between the inner and outer solution in all curves. We also see good agreement between the approximate solution to the numerics in both regions, confirming our findings. In the inner region we see very little change in r while p and q begin to increase. Once we move to the outer region, we see exponential decay in the concentration of free receptors, r, while the concentration of bound monomers, p, increases to equilibrium. The concentration of q takes longer to reach equilibrium and is not monotonic, where we see the multiple exponential terms in the solution having an impact on the dynamics.

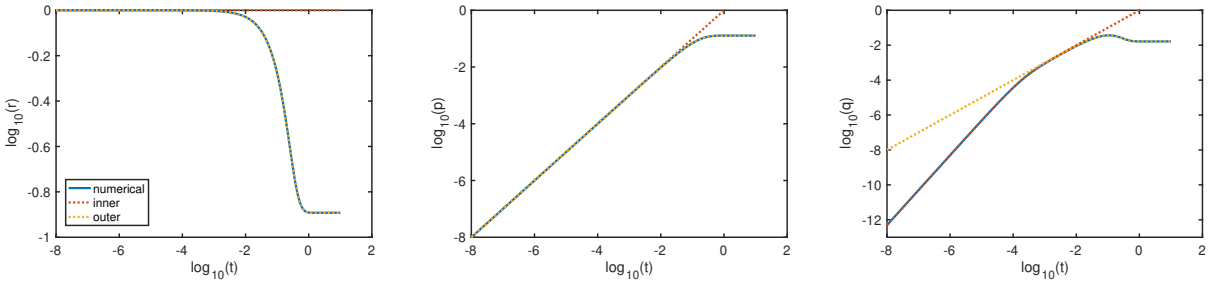

Figure 4.5.5: Plotting the numerical solution of equation (4.70) with asymptotic solutions (4.88) and (4.108) show good agreement in both regions. Parameters used to create plot are α=6.8, β=1, ϵ=1e-4.

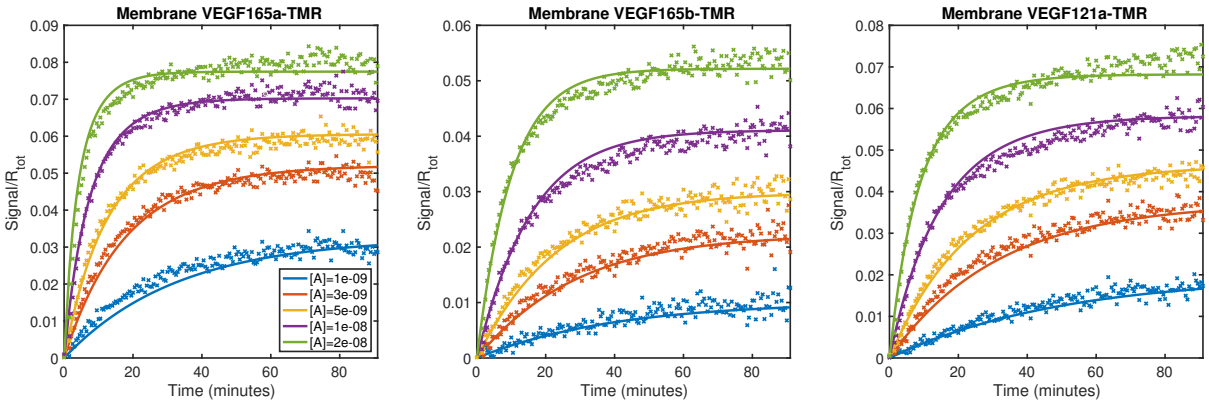

Figure 4.6.1: Data published in [101] is used to estimate the model parameters. Experiments were performed using five concentrations of three different VEGF isoforms (VEGF $_{165}$a-TMR, VEGF $_{165}$b-TMR and VEGF $_{121}$a-TMR) and are scaled with respect to R_{tot}. An excellent fit to the data is seen from fitting to all data sets simultaneously. Parameter values returned can be seen in Table 4.2.

4.6 Model validation

In this section we take the first steps towards validating our model by fitting to published, experimental data, with raw data kindly provided by Dr Chloe Peach [30]. In the paper by Peach *et al* [101], three VEGF isoforms (VEGF $_{165}$a-TMR, VEGF $_{165}$b-TMR and VEGF $_{121}$a-TMR), in five concentrations, are observed binding to VEGFR2. While experiments are performed on both HEK293T cells and membranes, we focus only on membrane results, thus reducing the possibility of data being affected by other processes such as receptor internalisation which may cause a reduction in signal on longer time scales. Including receptor internalisation is something we consider for future work.

The experiments were performed ten times and the average of these was used for fit-

ting. Fitting is performed in Copasi [55] using a particle swarm optimisation method with a swarm size of 100 and 3000 iterations. We note that Copasi is a software application for simulating and analysing biochemical networks and their dynamics. Copasi has many optimisation methods available and fitting was performed using each of these, however, the particle swarm method was chosen as it gave the smallest total error. All data sets are fitted to simultaneously, with a and R_{tot} fixed across all sets, and k_+, k_-, ψ_+ and ψ_- specific to each isoform.

	a	k_+	k_-	ψ_+	ψ_-	R_{tot}
VEGF $_{165}$a-TMR	8.90×10^{-2}	7.57×10^{6}	2.76×10^{-2}	1.87×10^{3}	1.49×10^{3}	2.52×10^{-9}
VEGF $_{165}$b-TMR	8.90×10^{-2}	1.63×10^{6}	7.72×10^{-2}	5.88×10^{3}	1.10×10^{2}	2.52×10^{-9}
VEGF $_{121}$a-TMR	8.90×10^{-2}	2.95×10^{6}	2.64×10^{-2}	1.38×10^{5}	3.88×10^{4}	2.52×10^{-9}

Table 4.2: Estimated parameters returned from fitting to the data, as seen in Figure 4.6.1.

In Table 4.2 we have the parameter estimates returned from fitting, while in Figure 4.6.1 we plot the fitted curve together with the data. Clearly the estimates give an excellent fit to all data sets. Goodness of fit figures can be seen in Table 4.3, where we use the Chi-square test to determine the correlation between each data set and the fit values.

$[A] =$	10^{-9}	3×10^{-9}	5×10^{-9}	10^{-8}	2×10^{-8}
VEGF $_{165}$a-TMR	0.0696	0.0213	0.0133	0.0168	0.0342
VEGF $_{165}$b-TMR	0.0591	0.0276	0.0183	0.0175	0.0125
VEGF $_{165}$a-TMR	0.0324	0.0320	0.0116	0.0183	0.0225

Table 4.3: Goodness of fit values, determined by the Chi-square test.

Looking at the estimated parameter values, we also note that all on and off rates fall into the reported ranges [80] (and references within). In all cases, we have $\psi_+ > \psi_-$ and, therefore, positive cooperativity overall. In Table 4.4 we compare these estimates with those returned in the paper by Chloe Peach *et al* [101], where the data was fitted to a simple monomeric binding model. We see some significant differences in some of

	k_+ estimate	k_+ in [101]	k_- estimate	k_- in [101]
VEGF $_{165}$a-TMR	7.57×10^6	1.27×10^7	2.76×10^{-2}	3×10^{-2}
VEGF $_{165}$b-TMR	1.63×10^6	3.69×10^6	7.72×10^{-2}	2.9×10^{-2}
VEGF $_{165}$a-TMR	2.95×10^6	5.13×10^6	2.64×10^{-2}	1.9×10^{-2}

Table 4.4: Comparing the on and off rates returned from fitting to our model with the values reported in [101].

the values. We see that our association rates are lower than theirs, but in both sets of estimates the on rate for VEGF $_{165}$a-TMR are substantially higher than the other isoforms. This is not, however, the case for the dissociation rates. In our estimates, the value for VEGF $_{165}$b-TMR is significantly higher than the other estimates, but this is not the case for those reported in [101] where all dissociation rates are similar values. This increased value could be explained by looking at the cooperativity values, as we notice that the dimer dissociation rate for the VEGF $_{165}$a-TMR isoform is lower than the others.

	VEGF $_{165}$a-TMR	VEGF $_{165}$b-TMR	VEGF $_{121}$a-TMR
$\alpha = K_A[A]$	2.74E+08×$[A]$	2.11E+07×$[A]$	1.12E+08×$[A]$
$\beta = \psi_+ R_{tot}/[A]$	4.71E-06/$[A]$	1.48E-05/$[A]$	3.49E-04/$[A]$
$\gamma = \psi_-$	1.49E+03	1.10E+02	3.88E+04

Table 4.5: Nondimensional parameter combinations with estimated parameters.

To further understand the implications of these estimations we plot the individual species curves for all results in Figure 4.6.2. The first thing we notice is the appearance of peaks in some of the $[RAR]$ curves. In particular, these appear for the higher ligand concentrations of the VEGF $_{165}$a-TMR and VEGF $_{121}$a-TMR curves. In Table 4.5 we look at the nondimensional parameter combinations for the estimated parameters. Clearly, for all curves we have both large β (to varying degrees) and large γ. However, our asymptotic analysis was based around α being $O(1)$. Only the curves with a higher ligand concentration, for isoforms VEGF $_{165}$a-TMR and VEGF $_{121}$a-TMR meet this requirement, hence this is why we only see the peaks arising from a large γ in these curves.

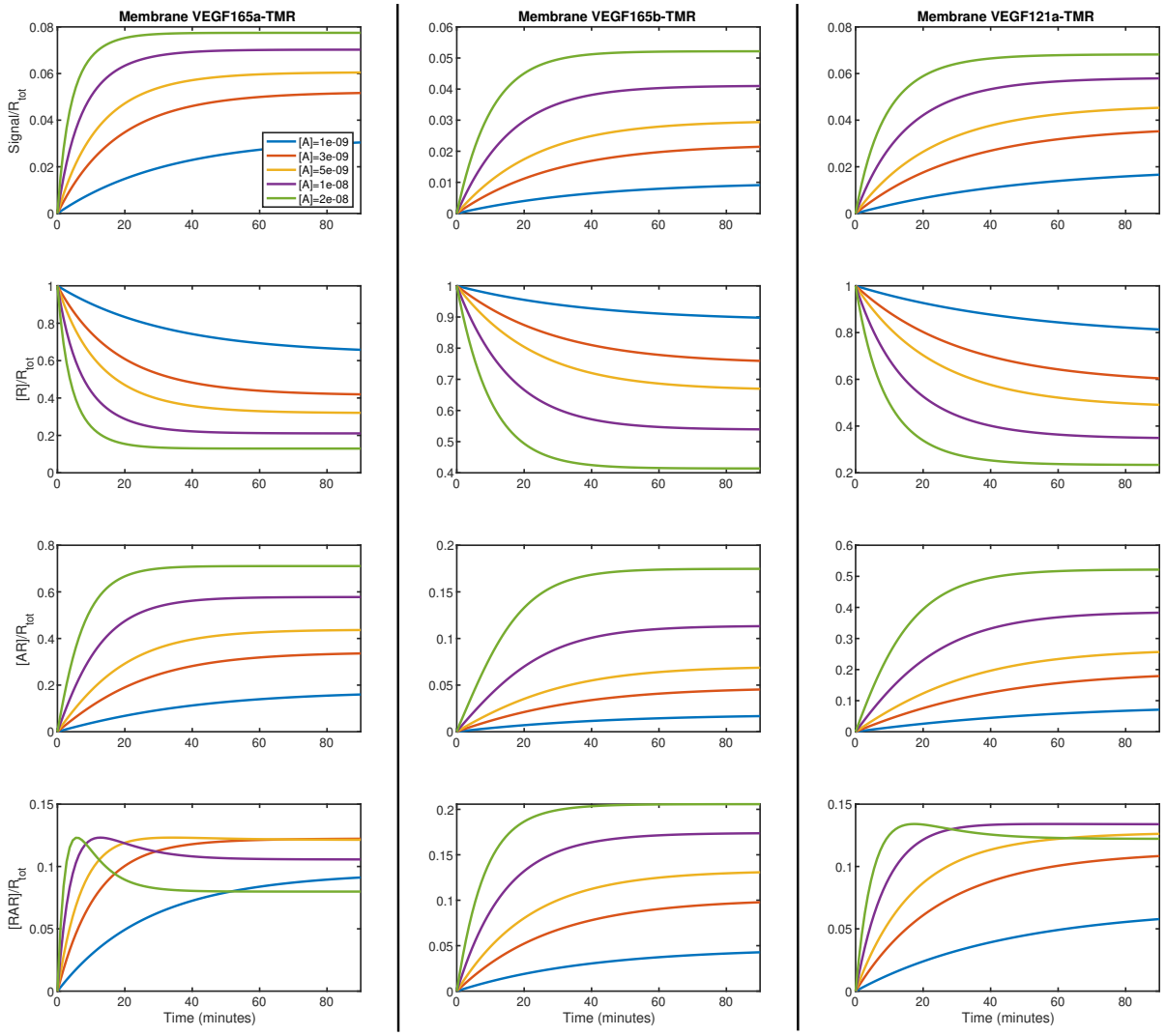

Figure 4.6.2: Individual species curves using the estimated parameters show peaks some of the $[RAR]$ curves.

While the work in this section is the beginnings of model validation, we note that there is still some way to go to fully achieve this. We recommend that next steps towards this goal are further experiments, ideally with differing ligand concentrations, to determine whether the estimates can accurately predict experimental outcomes.

4.7 Conclusions

In this chapter we have presented a model for the binding of VEGF to VEGFRs which leads to the induced dimerisation of these receptors, and is a stripped down version of those models presented in [80] and [3]. Considering a model of such simplicity allows

for in depth analysis of potential responses given various choices of model parameters. Analytical expressions for equilibrium responses provide a user friendly way for pharmacologists to compute solutions, as well as explain how increasing cooperativity leads to a shift in the experiment readout signal. Although in contrast to GPCR dimer binding [135], dose response curves are monophasic, hence there is no parameter values that give extra inflections in the curve. If these are present in the data we can assume that the system may contain pre-dimerised receptors but not ligand induced dimers. Equilibrium results are supported by Gabhann and Popel [80] who also find that increasing the ligand concentration increases the concentration of ligand bound monomers, something that is indicative of ligand induced dimerisation. We extend the previous research in this area by also considering time course dynamics for the model. The model is a non-linear system of ODEs that is unable to be solved analytically, hence numerical methods are used to explore the time dependent behaviour of the curves under different parameter regimes.

Upon investigation of the time course behaviour of the system, a number of emerging effects in the individual species curves can be seen contributing to the overall signal. Depending on the parameters chosen we have the possibility of peaks emerging in the concentration curves of AR and RAR, either individually or in both simultaneously. Also, the numerics show how there are multiple time scales contributing to the dynamics. Asymptotic analysis is used to explain these behaviours and help understand the reaction sequences that create these effects. In summary

- With a small γ, that is a small reverse cooperativity value, giving a reduced rate of dimer dissociation, we obtain a regular perturbation problem where we see, as expected intuitively, that the reversing of the second reaction, binding ligand to a second monomer, can be negated altogether.

- With a large β, that is when either the forward cooperativity value is large or there is a high receptor to ligand ratio, we see multiple time scales emerge. Initially the dynamics are dominated by the first reaction, binding ligand to monomer, however, as time moves past $O(1)$ the second reaction begins to have an effect.

- When γ is instead large, so the dimer dissociation rate is increased, we again have a

singular perturbation problem with multiple time scales contributing to the dynamics. Initially, the concentrations of AR and RAR increase monotonically, regardless of model parameters. As time moves to $O(1)$ we see these increases slow, and a possible drop in the concentration of RAR, however, as β does not appear in any of the solutions we conclude that the second reaction has no effect on the dynamics.

Finally, we take steps towards validating the model by fitting to published, experimental data, where we see the model gives an excellent fit to the data across all curves. This gives some confidence that the model accurately describes the binding and dimerisation processes for the VEGF-VEGFR system. While we have given estimates for the parameter values, we note that, there may be identifiability issues, and as such, other values that give the same fit. This is something we discuss in the next chapter, where we take a theoretical look at which of these parameter values can be relied upon to be accurate.

Chapter 5

Structural identifiability of ligand binding models

5.1 Introduction

In Chapter 1 we introduced the idea of structural identifiability analysis (SIA). Throughout this thesis, we have presented many models representing drug binding scenarios. These have generally resulted in a system of ODEs that govern the system dynamics. While we have mostly looked at possible time course results in a purely theoretical way, the intention is that one of the ways in which these models are used is, by pharmacologists, to fit to experimental data and quantify ligand-receptor interactions. Hence, it is important to understand whether parameter estimates returned from fitting are accurate. We saw an example of this in Chapter 4 (section 4.6) where we fit the ligand induced dimerisation model to real, experimental data. In this, we gave a set of parameter values for the model that gave a 'best fit', however, it is unclear whether there are other values that can possibly give the same fit. The model is mechanistic and the parameters represent biological processes, and so inaccurate parameter estimations may lead to incorrect conclusions being drawn about these.

In this chapter, we first review some of the theoretical foundations of existing SIA techniques, and discuss current work in the field. We cover a variety of methods for

applications to both linear and nonlinear problems, such as those that make use of the Laplace transformation [9] (though this only applies to linear problems), Taylor series expansions [105] and similarity transforms [23, 124]. While the goal is to determine the identifiability of our previous dimer models, we first take tutorial type approach in order to close the gaps in the current literature. To do this, we show the steps taken to determine identifiability of some classical ligand binding models before then also including some of our previous dimer models. For each model we apply a number of the available techniques to determine which parameters can be identified using a single set of time course data. This allows us to give a comparison of the methods and ascertain the ideal method for each scenario. For models that are unidentifiable, we also give, where possible, the identifiable parameter combinations. Following this, for unidentifiable models, we explore alternate ways in which the models can be made to be identifiable. This includes using equilibrium data or washout experiments together with the time course data, or using multiple time course data sets.

5.2 Theoretical foundations

We first discuss some of the theoretical underpinnings of SIA and the methods we will use in the later examples, following the notation as used in [28]. We define $\sum(\mathbf{p})$ as the biological system given by

$$\sum(\mathbf{p}) = \begin{bmatrix} \mathbf{x}'(t,\mathbf{p}) = \mathbf{f}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}) + \mathbf{g}(\mathbf{x}(t,\mathbf{p}),\mathbf{p})\mathbf{u}(t), \\ \mathbf{y}(t,\mathbf{p}) = \mathbf{h}(\mathbf{x}(t,\mathbf{p}),\mathbf{p}), \\ \mathbf{x}(0,\mathbf{p}) = \mathbf{x}_0(\mathbf{p}), \end{bmatrix} \tag{5.1}$$

where $\mathbf{x} = (x_1, ..., x_{n_x})^T \in \mathbb{R}^{n_x}$ is the state vector (the space $\mathbb{R}^{n_x}$ also contains the initial state), $\mathbf{u} = (u_1, ..., u_{n_u})^T \in \mathbb{R}^{n_u}$ is the vector of inputs, and $\mathbf{y} = (y_1, ..., y_{n_y})^T \in \mathbb{R}^{n_y}$ is the output vector of experimentally observed quantities. The functions $\mathbf{f} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ and $\mathbf{g} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$ are analytical function that describe the evolution of the state variable in time, while $\mathbf{h} : \mathbb{R}^{n_x} \to \mathbb{R}^{n_y}$ determines the model outputs. The vector $\mathbf{p} \in \Omega \subset \mathbb{R}^{n_p}$ are the unknown parameters of the system.

Definition 5.2.1 *1. A parameter p_i, $i = 1...n_p$, is structurally globally (or uniquely) identifiable (s.g.i) if for almost any $\widetilde{\boldsymbol{p}} \in \Omega$,*

$$\sum(\boldsymbol{p}) = \sum(\widetilde{\boldsymbol{p}}) \Rightarrow p_i = \widetilde{p}_i.$$

2. A parameter p_i, $i = 1...n_p$, is structurally locally identifiable (s.l.i) if for almost any $\widetilde{\boldsymbol{p}} \in \Omega$, there exists a neighborhood $\boldsymbol{V}(\boldsymbol{p})$ of $\boldsymbol{p}$ such that

$$\widetilde{\boldsymbol{p}} \in \boldsymbol{V}(\boldsymbol{p}) \text{ and } \sum(\boldsymbol{p}) = \sum(\widetilde{\boldsymbol{p}}) \Rightarrow p_i = \widetilde{p}_i.$$

3. A parameter p_i, $i = 1...n_p$, is structurally unidentifiable if for almost any $\widetilde{\boldsymbol{p}} \in \Omega$, there exists no neighborhood $\boldsymbol{V}(\boldsymbol{p})$ of $\boldsymbol{p}$ such that

$$\widetilde{\boldsymbol{p}} \in \boldsymbol{V}(\boldsymbol{p}) \text{ and } \sum(\boldsymbol{p}) = \sum(\widetilde{\boldsymbol{p}}) \Rightarrow p_i = \widetilde{p}_i.$$

The **model** is then said to be s.l.i if all parameters are s.l.i and s.g.i if all parameters are s.g.i [28]. Further theory and the proofs for these definitions can be found in [49].

Although there are now many available methods to determine identifiability, we choose to focus on the transfer function method (based on the Laplace transform), the Taylor series method and the similarity transformation method, for our analysis. While the transfer function method is only applicable to linear systems, its simplicity of application makes it a natural choice for these models. The Taylor series method is one of the most flexible and can be applied to most ODE systems, although does have limitations [24]. The similarity transformation method, although also applicable to linear systems, is one of the most well established for nonlinear systems. Furthermore, these methods are suitable to systems of low dimensionality and have been proven to work well for compartment models that have a similar structure to binding models. We will now

outline the underlying theory for each of these methods before applying them to the models.

5.2.1 Transfer function method

The transfer function method was first proposed by Bellman and Åström [9] and is one that is simple in nature, although is restricted to linear time-invariant systems. Assume a system of the form

$$\sum(\mathbf{p}) = \begin{bmatrix} \mathbf{x}' = F\mathbf{x} + G\mathbf{u}, \\ \mathbf{y} = H\mathbf{x}, \\ \mathbf{x}(0) = \mathbf{x}_0, \end{bmatrix} \tag{5.2}$$

which is a specific case of the system given in equation (5.1). We then have $F \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x}$, $G \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_u}$ and $H \in \mathbb{R}^{n_y} \times \mathbb{R}^{n_x}$. We also note that we do not consider inputs directly to observable function as these are not applicable to the models we use in this chapter.

The Laplace transformation provides us with a transfer function, $Q(s, \mathbf{p})$, describing the input-output relation. Taking the transforms of the system defined in equation (5.2) and applying linearity and time-derivative rules gives

$$s\mathbf{X}(s, \mathbf{p}) - \mathbf{x}_0 = F\mathbf{X}(s, \mathbf{p}) + G\mathbf{U}(s, \mathbf{p}), \tag{5.3a}$$

$$\mathbf{Y}(s, \mathbf{p}) = H\mathbf{X}(s, \mathbf{p}), \tag{5.3b}$$

where $\mathbf{X}(s, \mathbf{p})$, $\mathbf{Y}(s, \mathbf{p})$ and $\mathbf{U}(s, \mathbf{p})$ are the Laplace transforms of $\mathbf{x}$, $\mathbf{y}$ and $\mathbf{u}$ respectively. Using the properties of the Laplace transform, we can then state

$$\mathbf{X}(s, \mathbf{p}) = (sI - F)^{-1}\mathbf{x}_0 + (sI - F)^{-1}G\mathbf{U}(s, \mathbf{p}), \tag{5.4a}$$

$$\mathbf{Y}(s, \mathbf{p}) = H(sI - F)^{-1}\mathbf{x}_0 + H(sI - F)^{-1}G\mathbf{U}(s, \mathbf{p}). \tag{5.4b}$$

Assuming the system has initial conditions $\mathbf{x}_0 = \mathbf{0}$, that is the input-output relation is given entirely by the impulse, then we can define

$$Q(s, \mathbf{p}) = H(sI - F)^{-1}G, \tag{5.5}$$

as the transfer function (when $Q(s, \mathbf{p})$ is a single entry) or matrix of transfer functions of the system. The expanded form of $Q(s, \mathbf{p})$ is the $n_y \times n_u$ matrix, given as

$$Q(s, \mathbf{p}) = \begin{bmatrix} Q_{1,1} & \cdots & Q_{1,n_u} \\ \vdots & \ddots & \vdots \\ Q_{n_y,1} & \cdots & Q_{n_y,n_u} \end{bmatrix}. \tag{5.6}$$

with entries

$$Q_{i,j} = \frac{b_{(i,j),1}s^{n_x-1} + b_{(i,j),2}s^{n_x-2} + ... + b_{(i,j),n_x}}{s^{n_x} + a_1 s^{n_x-1} + ... + a_{n_x}}. \tag{5.7}$$

Each transfer function (and as such the function in each matrix entry) is a unique function relating each input to an output. Thus, the coefficients a_k and $b_{(i,j),k}$ are unique to the transfer function. This uniqueness means we can use these coefficients to determine which parameters are uniquely identifiable. Creating a vector of these coefficients and denoting this vector as $\boldsymbol{\zeta}(\mathbf{p})$, identifiability is then established by setting $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ and applying definition 5.2.1. We outline how this is applied to a system in Figure 5.2.1.

5.2.2 Taylor series method

The Taylor series method was first developed by Pohjanpalo [105], and can be applied to either linear or nonlinear systems, although the algebra involved in applying the method to nonlinear problems can be difficult [24]. We assume a system as defined in equation (5.1) and assume that the functions $\mathbf{f}$ and $\mathbf{g}$ are continuously differentiable, and as such

Applying the transfer function method

1. Ensure the model is in the formulation as given in (5.2), that is, with F, G and H non-zero.
2. Calculate the transfer function for the system, using equation (5.7). This will return a matrix.
3. Create the vector $\boldsymbol{\zeta}(\mathbf{p})$ with the coefficients of s in both the numerator and denominator of each entry of the matrix.
4. Assume a second vector, $\boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ and set $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$.
5. Simplify the relationships in $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ to reduce the expressions.
6. The resulting expressions of $\boldsymbol{\zeta}(\mathbf{p})$ are the identifiable combinations. Expressions that are reduced to a single parameter are then identifiable parameters.

Figure 5.2.1: An algorithm for using the transfer function method to determine identifiability.

the state vectors and observables have infinitely many derivatives with respect to time. We also assume that $\mathbf{h}$ has infinitely many derivatives with respect to the state variables. The Taylor series approach then exploits the fact that, the observations (as given in $\mathbf{y}(t, \mathbf{p})$) are unique analytic functions of time and so all their derivatives with respect to time should also be unique, and should therefore, hold all possible information about the unknown parameters [28]. Thus, the observables can be represented by a Taylor series about the initial state, that can be used to establish identifiability. The Taylor series of $\mathbf{y}$ in a neighbourhood of the initial state is then given by

$$\mathbf{y}(t, \mathbf{p}) = \mathbf{y}(t_0, \mathbf{p}) + (t - t_0)\mathbf{y}'(t_0, \mathbf{p}) + \frac{(t - t_0)^2}{2!}\mathbf{y}''(t_0, \mathbf{p}) + O(t^3). \tag{5.8}$$

We denote

$$\mathbf{a}_k = \frac{d^k}{dt^k}\mathbf{y}(t_0, \mathbf{p}), \qquad k = 0, 1, ..., k_{max}. \tag{5.9}$$

as the Taylor series derivatives. These then create a system of non-linear algebraic equations in the parameters, which we use to check whether the system has a unique solution. That is, the model is evaluated at $\mathbf{a}_k(\mathbf{p})$ and $\mathbf{a}_k(\widetilde{\mathbf{p}})$, and the model is structurally identifiable if $\mathbf{a}_k(\mathbf{p}) = \mathbf{a}_k(\widetilde{\mathbf{p}}) \Rightarrow \mathbf{p} = \widetilde{\mathbf{p}}$ ([28]).

For linear problems the maximum number of coefficients, that is k_{max}, needed to determine identifiability is

$$k_{max} = 2n_x - 1, \tag{5.10}$$

where n_x is the number of state variables [28]. The number of coefficients required is unbounded for nonlinear problems, and as such is one of the disadvantages of the Taylor series method for these type of problems. Furthermore, for nonlinear problems, the coefficients can be algebraically complex and hence, identifiability cannot always be determined using this method. In Figure 5.2.2 we outline the steps needed to use this method to determine identifiability of the system.

5.2.3 Similarity transformation method

The similarity transformation method (or exhaustive modelling approach) was first proposed by Walter and Lecourtier [131], though originally was only applicable to linear problems. This was later extended to include nonlinear problems [123]. Before we discuss the method itself and its implementation, we first detail some requirements on the system for the method to be applicable.

Controllability and observability

In order to use the similarity transformation method, the system must be a minimal representation, that is it must be both controllable and observable, concepts introduced by Kalman in 1960 [62]. A system is said to be controllable if the system states, in $\mathbf{x}$, are changed by changing the input, and observable if the initial state $\mathbf{x}_0$ can be uniquely

Applying the Taylor series method

1. Calculate the first Taylor series coefficient by substituting the initial condition into the observable function, that is, calculate $\mathbf{y}(t_0)$. This is the first element of $\boldsymbol{\zeta}(\mathbf{p})$.
2. Differentiate $\mathbf{y}$ to give the next Taylor coefficient.
3. Substitute the expressions in $\mathbf{f}$, into this coefficient, to reduce the differential terms to algebraic expressions.
4. Substitute the initial conditions of the system in place of the state variables. This should result in an expression containing only model parameters.
5. Add the resulting expression to the vector $\boldsymbol{\zeta}(\mathbf{p})$.
6. Repeat steps 2-5, repeatedly differentiating $\mathbf{y}$ to calculate higher order derivatives and repeatedly substituting in $\mathbf{f}$ in place of state variables.
7. Assume a second vector, $\boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ and set $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$.
8. Simplify the relationships in $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ to reduce the expressions.
9. The resulting expressions of $\boldsymbol{\zeta}(\mathbf{p})$ are the identifiable combinations. Expressions that are reduced to a single parameter are then identifiable parameters.

Figure 5.2.2: An algorithm for using the Taylor series method to determine identifiability.

determined from a set of input-output measurements. For a linear system, in the form of equation (5.2), we define the controllability matrix as

$$\mathcal{C} = [G \vdots FG \vdots F^2G \vdots \cdots \vdots F^{n_x-1}G], \tag{5.11}$$

and the observability matrix

$$\mathcal{O} = \begin{bmatrix} H \\ HF \\ HF^2 \\ \vdots \\ HF^{n_x-1} \end{bmatrix}, \tag{5.12}$$

The system is then said to be controllable if $\text{rank}(\mathcal{C}) = n_x$, and observable if $\text{rank}(\mathcal{O}) = n_x$ [123]. The derivation of these matrices, and the theorems detailing these rank tests, along with the theory they are built upon can be found in [113].

For a nonlinear system, as in equation (5.1), the controllability matrix makes use of the Lie bracket, which can be stated as

$$(ad^1_{\mathbf{f}}, \mathbf{g}) = [\mathbf{f}, \mathbf{g}] = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}) \tag{5.13}$$

with higher order derivatives calculated as

$$(ad^k_{\mathbf{f}}, \mathbf{g}) = [f, (ad^{k-1}_{\mathbf{f}}, \mathbf{g})]. \tag{5.14}$$

The full definition of the Lie bracket can be found in Section C.1, while further explanation of the physical meaning of the Lie bracket and how this relates to controllability can be found in [134]. The controllability matrix is then given by

$$\mathcal{C} = [\mathbf{g}, (ad^1_{\mathbf{f}}, \mathbf{g}), \cdots, (ad^{n-1}_{\mathbf{f}}, \mathbf{g})], \tag{5.15}$$

Following this, calculating the observability matrix involves using the *Lie derivative*, which is defined as

$$L^1_{\mathbf{f}}\mathbf{y}(\mathbf{x}) = \frac{\partial \mathbf{y}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}), \tag{5.16}$$

and higher order derivatives being calculated recursively as

$$L^i_{\mathbf{f}}\mathbf{y}(x) = \frac{\partial L^{i-1}_{\mathbf{f}}\mathbf{y}(x)}{\partial x}\mathbf{f}(x). \tag{5.17}$$

The full definition of the Lie derivative can be found in Section C.2. The observability matrix is then calculated as

$$\mathcal{O} = \begin{bmatrix} \dfrac{\partial L_{\mathbf{f}}^{0}\mathbf{y}(\mathbf{x})}{\partial \mathbf{x}} \\ \dfrac{\partial L_{\mathbf{f}}^{1}\mathbf{y}(\mathbf{x})}{\partial \mathbf{x}} \\ \vdots \\ \dfrac{\partial L_{\mathbf{f}}^{n-1}\mathbf{y}(\mathbf{x})}{\partial \mathbf{x}} \end{bmatrix}. \tag{5.18}$$

Again, the system is controllable and observable if $\text{rank}(\mathcal{C}) = n_x$ and $\text{rank}(\mathcal{O}) = n_x$ respectively.

The method

The objective of this method is, given the model (5.1), with $\mathbf{p} \in \Omega$, find all parameter values $\widetilde{\mathbf{p}} \in \Omega$ and corresponding models of the form

$$\sum(\widetilde{\mathbf{p}}) = \begin{bmatrix} \widetilde{\mathbf{x}}'(t,\widetilde{\mathbf{p}}) = \mathbf{f}(\widetilde{\mathbf{x}}(t,\widetilde{\mathbf{p}}),\widetilde{\mathbf{p}}) + \mathbf{g}(\widetilde{\mathbf{x}}(t,\widetilde{\mathbf{p}}),\widetilde{\mathbf{p}})\mathbf{u}(t), \\ \mathbf{y}(t,\widetilde{\mathbf{p}}) = \mathbf{h}(\widetilde{\mathbf{x}}(t,\widetilde{\mathbf{p}}),\widetilde{\mathbf{p}}), \\ \widetilde{\mathbf{x}}(0,\widetilde{\mathbf{p}}) = \widetilde{\mathbf{x}}_0(\widetilde{\mathbf{p}}). \end{bmatrix} \tag{5.19}$$

that have the same input-output map. This involves assuming the existence of an alternate system, $\sum(\widetilde{\mathbf{p}})$, and applying a set of conditions on this alternate system to ensure that it is equivalent to the original system, $\sum(\mathbf{p})$. Once equivalence holds, the system is identifiable if

$$\sum(\mathbf{p}) = \sum(\widetilde{\mathbf{p}}) \Rightarrow \mathbf{p} = \widetilde{\mathbf{p}}. \tag{5.20}$$

Linear systems

The linear equivalence of two linear systems is given by the algebraic equivalence theorem [46, 112] for ODE systems. We consider a system in the form of equations 5.2, and assume an alternate system with $\widetilde{F} = F(\widetilde{\mathbf{p}})$, $\widetilde{G} = G(\widetilde{\mathbf{p}})$ and $\widetilde{H} = H(\widetilde{\mathbf{p}})$. These are then equivalent if there exists a matrix $T \in \mathbb{R}^{n_x} \times \mathbb{R}^{n_x}$ such that

$$\det T \neq 0, \tag{5.21a}$$

$$T\widetilde{\mathbf{x}}_0 = \mathbf{x}_0, \tag{5.21b}$$

$$T\widetilde{F} = FT, \tag{5.21c}$$

$$T\widetilde{G} = G, \tag{5.21d}$$

$$\widetilde{H} = HT, \tag{5.21e}$$

The full equivalence theorem can be seen in Section C.3. The process is then to apply the constraints on F, G and H to determine the parameters $\mathbf{p}$. The system is globally identifiable if it follows from applying these conditions that $\mathbf{p} = \widetilde{\mathbf{p}}$ and locally identifiable if there is a finite number of possibilities for each p_i.

Nonlinear systems

The equivalence theorem has also been extended to include nonlinear systems [24, 28, 123]). With a system in the form of equations (5.1), the conditions to ensure equivalence

(see Section C.4 for the full details) become

$$\operatorname{rank}\left(\frac{\partial\lambda(\widetilde{\mathbf{x}})}{\partial\widetilde{\mathbf{x}}}\right) = n_x, \quad \forall\widetilde{\mathbf{x}} \in V, \tag{5.22a}$$

$$\lambda(\widetilde{\mathbf{x}}_0(\mathbf{p})) = \mathbf{x}_0(\widetilde{\mathbf{p}}), \tag{5.22b}$$

$$f(\lambda(\widetilde{\mathbf{x}}), \mathbf{p}) = \frac{\partial\lambda(\widetilde{\mathbf{x}})}{\partial\widetilde{\mathbf{x}}} f(\widetilde{\mathbf{x}}, \widetilde{\mathbf{p}}), \tag{5.22c}$$

$$g(\lambda(\widetilde{\mathbf{x}}), \mathbf{p}) = \frac{\partial\lambda(\widetilde{\mathbf{x}})}{\partial\widetilde{\mathbf{x}}} g(\widetilde{\mathbf{x}}, \widetilde{\mathbf{p}}), \tag{5.22d}$$

$$h(\lambda(\widetilde{\mathbf{x}}), \mathbf{p}) = h(\widetilde{\mathbf{x}}, \widetilde{\mathbf{p}}), \tag{5.22e}$$

for some linear transformation $\lambda(\widetilde{\mathbf{x}})$. Then the system is globally identifiable if applying these conditions implies $\widetilde{\mathbf{p}} = \mathbf{p}$.

The steps needed to apply this method can be seen in Figure 5.2.3.

Applying the similarity transformation method

1. Calculate the rank of the controllability and observability matrices. These must both be equal to the number of state variables to proceed.
2. If the system is linear, assume a square transformation matrix (denoted T) of equal size to the number of state variables. For nonlinear systems, assume the existence of a linear transformation, denoted $\lambda(\mathbf{x})$, equal in size to the state vector.
3. Assume all conditions hold. For linear systems these conditions are given in (5.21) while the conditions for nonlinear systems can be found in (5.22). The conditions can be applied in any order, depending on the individual system they are being applied to.
4. While applying the conditions, it may be found that $p_i = \tilde{p}_i$ for any parameter. This parameter is then considered identifiable. Identifiable parameter combinations may also be found.
5. The system is fully identifiable if it is found that T is the identity matrix (for linear systems), or $\lambda(\mathbf{x}) = \mathbf{x}$ (for nonlinear systems).

Figure 5.2.3: An algorithm for using the similarity transformation method to determine identifiability.

5.2.4 Current literature

The origins of SIA lie in the works of Kalman [62] for linear systems, and Hermann and Krener [53] for nonlinear models. Following these, many methods have emerged to assess the identifiability of a model. The methods previously discussed in this chapter (transfer function, Taylor series and similarity transformation methods) are still some the most commonly used methods and have formed a basis for much of the SIA work over the past decades. Often compartment models are used as an example of application [24, 123, 45]. In [25], Chappell and Gunn explore the possibility of reparameterising unidentifiable non-linear systems in order to force locally identifiability. This work is continued by Evans and Chappell in [35].

In [140], the authors develop a method that builds upon the similarity transformation method to determine the identifiability of nonlinear models. This method also has the potential to also be applicable to partial differential equations (PDEs). Cheung and Yates [26] consider whether models are identifiable if data from multiple experiments is used. The paper looks at linear compartment models and uses the similarity transformation method to determine identifiability. In [129], Villaverde, Barreiro and Papachristodoulou develop a novel method where they assume local identifiability is a generalised form of observability, if the parameters are viewed as constant state variables. With this work, the authors also present a software implementation of their algorithm, named STRIKE-GOLDD, which is given as a Matlab [2] toolbox. In this the user is able to input their own model and the software determines parameter identifiability and, in some cases, returns identifiable parameter combinations. Anguelova *et al* [4] take a different approach and consider the effect conservation laws have on identifiability.

Further methods have been also developed, and are reviewed in [28] and [27]. One such method is the differential algebra approach which can be applied to linear or nonlinear models, providing the functions are polynomial or rational in form [81]. The method involves replacing the input-output relationship of the model by a polynomial and testing for injectivity [28]. The software tool, DAISY [10], implements the differential algebra

method using the computer algebra system, REDUCE version 3.8. Lecourtier and Raksanyi [75] also use REDUCE to look at algorithms for reducing lengthy expressions. The generating function approach [131] follows the Taylor series method but makes use of the Lie derivative, although a major limitation of this method is that the number of coefficients needed to be calculated is unknown [27]. Lecourtier, Lamnabhi-Lagarrigue and Walter extend this method together with Volterra [74] to develop a new method that overcomes the limitations of the generating series approach. This new method that they propose creates a transformation that is a generalisation of the Laplace transform, though the method is still applicable to nonlinear systems. Identifiability tableaus [7] are frequently used with the generating series approach. Calculating the Jacobian of the power series coefficients, with respect to the parameters, the tableau shows the non-zero elements. The rank of this matrix is then used to determine identifiability.

Despite the importance of SIA when fitting ODE models to experimental data, it is frequently overlooked for binding models such as those presented throughout this thesis, although Janzén *et al* [59] take some steps towards showing the benefits of SIA for pharmacodynamics models. Also, Middendorf and Aldrich, in [89] and [90] develop a method to determine the identifiability of equilibrium binding parameters for models that describe ligand binding to proteins that have multiple binding sites. Their method aims to overcome the limitations of existing methods such as those that use likelihood intervals or brute-force methods.

5.3 Structural identifiability analysis of ligand binding models with a single time course

In this section we show how the techniques can be applied to some ligand binding models. We perform SIA for a range of models, beginning with some classical binding models before then analysing our new models for dimer binding. For each of these, we consider a single set of time course data and apply three methods (transfer function, Taylor series

and similarity transformation methods). Not only does this provide a tutorial of how each of the methods work in practice but also allows to compare the different techniques and explore which is best in each situation.

5.3.1 Monomeric receptor binding with a single drug

The simple model of a ligand binding with a monomeric receptor has been the basis of much of the theoretical foundations of receptor theory of the past few decades and is still widely used in drug development research [66, 69]. For this reason (and because of its simplicity), we begin our analysis with this model. Assuming the ligand, A, is a constant concentration that binds to the monomeric receptor $[R]$ with association and dissociation rates k_{a+} and k_{a-} respectively, to create the complex $[AR]$. This can be described as a chemical reaction as:

$$\mathrm{A} + \mathrm{R} \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} \mathrm{AR}$$

The law of mass action gives us the system of ODEs

$$[R]' = -k_{a+}[A][R] + k_{a-}[AR], \tag{5.23a}$$

$$[AR]' = k_{a+}[A][R] - k_{a-}[AR], \tag{5.23b}$$

which, together with the initial conditions

$$[R](0) = R_{tot}, \qquad [AR](0) = 0, \tag{5.23c}$$

where R_{tot} is the total receptor concentration, form the initial value problem describing the kinetics of the system. The concentration of $[AR]$ is measured experimentally, hence the observed output is

$$y = [AR]. \tag{5.23d}$$

Assuming the only known parameter is the drug concentration $[A]$, then we have $\mathbf{p} = (k_{a+}, k_{a-}, R_{tot})$. Although this is a very simple system and can be solved analytically, we will still perform each of the SIA methods for the model. This will allow us to illustrate the methods for the simple case before increasing the complexity.

The transfer function method

We first outline the transfer function method. In its current form, the input to the system is in the initial conditions. This means that we have $G = 0$ in equation (5.2), and so the transfer function will also be zero. To overcome this we need to reformulate the system equations. Either we use conservation laws to reduce the system, or we use the Dirac δ function so that the input appears as a forcing term. As throughout this thesis we have used conservation laws to reduce the system dimensions we will continue to use this method in this chapter also. In this model, the conservation law can be stated as

$$R_{tot} = [R] + [AR], \tag{5.24}$$

to reduce the system by removing $[R]$. This gives the single differential equation

$$\frac{d[AR]}{dt} = -(k_{a+}[A] + k_{a-})[AR] + k_{a+}[A]R_{tot}, \tag{5.25a}$$

with initial condition

$$[AR](0) = 0, \tag{5.25b}$$

and output

$$y = [AR]. \tag{5.25c}$$

Calculating the transfer function of a single ODE is trivial as each component is one-dimensional. We have

$$F = -(k_{a+}[A] + k_{a-}), \qquad G = k_{a+}[A]R_{tot}, \qquad H = 1, \tag{5.26}$$

which gives a transfer function of

$$Q(s, \mathbf{p}) = \frac{k_{a+}[A]R_{tot}}{s + k_{a+}[A] + k_{a-}}. \tag{5.27}$$

The unique parameter combinations are the coefficients of s, hence in this case we have

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{a+}[A] + k_{a-} \end{bmatrix}. \tag{5.28}$$

Clearly we have three unknowns and only two expressions so it is not possible to identify all parameters, in fact none of the parameters are identifiable individually, and so we conclude that only the grouped parameters

$$k_{a+}R_{tot} \qquad \text{and} \qquad k_{a+}[A] + k_{a-}, \tag{5.29}$$

are identifiable.

Non-identifiability can be seen in Figure 5.3.1, where we see how three different parameter sets can result in the same measured output curve but have different concentration magnitudes of the free receptor concentrations (not observed), showing the obviously different binding parameters. However, in Table 5.1 where we calculate these grouped parameters for these three parameter sets, we see that these are equal for all sets, and therefore identifiable.

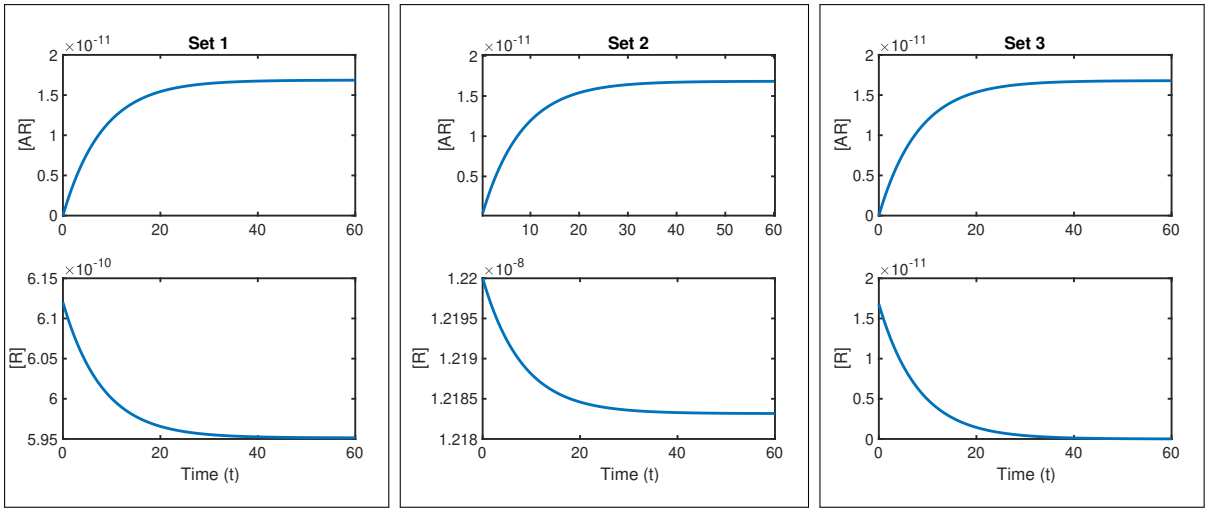

Figure 5.3.1: Three sets of parameters are used to plot the system given in equations (5.23). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.1 together with $[A] = 10^{-8}M$.

	Set 1	Set 2	Set 3
k_{a+}	$3.4 \times 10^{5} M^{-1}s^{-1}$	$1.7 \times 10^{4} M^{-1}s^{-1}$	$1.23 \times 10^{7} M^{-1}s^{-1}$
k_{a-}	$1.2 \times 10^{-1} s^{-1}$	$1.23 \times 10^{-1} s^{-1}$	$10^{-5} s^{-1}$
R_{tot}	$6.12 \times 10^{-10} M$	$1.22 \times 10^{-8} M$	$1.68 \times 10^{-11} M$
$k_{a+}R_{tot}$	$2.07 \times 10^{-4} M$	$2.07 \times 10^{-4} M$	$2.07 \times 10^{-4} M$
$k_{a+}[A] + k_{a-}$	$1.23 \times 10^{-1} s^{-1}$	$1.23 \times 10^{-1} s^{-1}$	$1.23 \times 10^{-1} s^{-1}$

Table 5.1: The parameters for three different parameter sets are used to plot Figure 5.3.1. The parameter combinations are equal in each case, so we can confirm these as identifiable.

Taylor series method

For this method we use the system in the form as given in (5.23), although results would be the identical if using the reduced system. We recall equation (5.10), which states that the maximum number of Taylor series coefficients (we refer to these simply as coefficients throughout) is $2n_x - 1$, where n_x is the number of state variables. As in this case we have $n_x = 2$ we will need to calculate a maximum of 3 coefficients. The first coefficient

is simply

$$y_0 = y(\mathbf{x}_0) = [AR](0) = 0. \tag{5.30}$$

We note that a subscript zero indicates that we are evaluating at $t = 0$. As this contains no parameters it gives no information about parameter identifiability. As we have only two remaining coefficients to calculate but three unknown parameters, we determine that the system will not be globally identifiable. However, we will still continue to determine which, if any, parameters are identifiable and also any identifiable parameter combinations. The first derivative of the output function is simply

$$y^{(1)} = [AR]^{(1)} = k_{a+}[A][R] - k_{a-}[AR]. \tag{5.31}$$

Substitution of the initial conditions gives the first coefficient as

$$y_0^{(1)} = k_{a+}[A]R_{tot}. \tag{5.32}$$

This can be simplified to

$$c_1 = k_{a+}R_{tot}, \tag{5.33}$$

where we denote c_i as the coefficient in its minimal form. A coefficient is considered to be in a minimal form once it is not possible to simplify it using already known quantities or expressions. Further derivatives are calculated with recursive substitution of the

equations in system (5.23). This gives the second derivative as

$$\begin{aligned} y^{(2)} = [AR]^{(2)} &= k_{a+}[A][R]^{(1)} - k_{a-}[AR]^{(1)} \\ &= -k_{a+}[A](k_{a+}[A][R] - k_{a-}[AR]) - k_{a-}(k_{a+}[A][R] - k_{a-}[AR]) \\ &= -k_{a+}[A](k_{a+}[A] + k_{a-})[R] + k_{a-}(k_{a+}[A] + k_{a-})[AR], \end{aligned} \tag{5.34}$$

leading to, after substitution of the initial conditions, the second coefficient as

$$y_0^{(2)} = -k_{a+}[A]R_{tot}(k_{a+}[A] + k_{a-}). \tag{5.35}$$

As we have

$$y_0^{(2)} = -c_1[A](k_{a+}[A] + k_{a-}), \tag{5.36}$$

we are able to simplify and reduce this, leaving

$$c_2 = k_{a+}[A] + k_{a-}, \tag{5.37}$$

as the second identifiable parameter combination. Hence, we have two identifiable parameter combinations which agree with the parameter combinations found when using the transfer function method. As the complexity of the calculation for each is similar, the only real difference is in the requirements of the methods. For completeness we will show that calculating further derivatives gives no unique expressions. The third derivative is calculated in the same way, which, after some simplification gives

$$y^{(3)} = k_{a+}[A](k_{a+}[A] + k_{a-})^2[R] + k_{a-}(k_{a+}[A] + k_{a-})^2[AR], \tag{5.38}$$

which gives the next coefficient as

$$y_0^{(3)} = k_{a+}[A]R_{tot}(k_{a+}[A] + k_{a-})^2. \tag{5.39}$$

As

$$y_0^{(3)} = c_1[A]c_2^2 \tag{5.40}$$

this gives no further information, hence, we have only the same two identifiable parameter combinations as we found in the transfer function method.

Similarity transformation method

We first note that, with the system in its original form, we have $G = (0, 0)^T$ in equation (5.2), hence we are unable to determine identifiability. Similar to the transfer function method, this is due to the inputs being in the initial conditions as opposed to the model equations, hence a reformulation is required. We again use the reduced system, using conservation of receptors, to overcome this problem, giving

$$F = -(k_{a+}[A] + k_{a-}), \quad G = k_{a+}[A]R_{tot}, \quad H = 1. \tag{5.41}$$

As $n = 1$, the system must be both controllable and observable, hence we now move forward with checking identifiability. To do this we assume a transformation matrix T and apply conditions (5.21). As $n_x = 1$, then T is a single entry matrix, and so we first note that $\det T \neq 0$, providing $T \neq 0$, and so condition (5.21a) holds if this remains true. From condition (5.21e) we have

$$\widetilde{H} = HT, \quad \Rightarrow \quad 1 = 1 \cdot T, \quad \Rightarrow \quad T = 1, \tag{5.42}$$

hence conditions (5.21a) and (5.21b) both hold. Applying condition (5.21d), we have

$$T\widetilde{G} = G, \quad \Rightarrow \quad \widetilde{k_{a+}}[A]\widetilde{R_{tot}} = k_{a+}[A]R_{tot}, \quad \Rightarrow \quad \widetilde{k_{a+}}\widetilde{R_{tot}} = k_{a+}R_{tot}. \tag{5.43}$$

Finally, we have, from condition (5.21c)

$$T\widetilde{F} = FT, \quad \Rightarrow \quad -(\widetilde{k_{a+}}[A]+\widetilde{k_{a-}}) = -(k_{a+}[A]+k_{a-}), \quad \Rightarrow \quad \widetilde{k_{a+}}[A]+\widetilde{k_{a-}} = k_{a+}[A]+k_{a-}. \tag{5.44}$$

Hence, we have the same identifiable parameter combinations as in the previous methods. This confirms that all methods give the same results. Comparing the methods is futile in this case, however, as the model is of such low dimension.

5.3.2 Monomeric receptor two drug competition binding

The next model we consider is a competitive binding scenario as in Motulsky-Mahan [95], where two ligands, A and B, are each able to bind to a monomeric receptor. The ligands are assumed to be constant in concentration and are similar in formation, hence both bind to the same receptor type. The reactions describing the interactions are:

$$\mathrm{A} + \mathrm{R} \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} \mathrm{AR} \qquad \mathrm{B} + \mathrm{R} \underset{k_{b-}}{\overset{k_{b+}}{\rightleftharpoons}} \mathrm{BR}$$

The system of ODEs governing the dynamics of the system is given as

$$[R]' = -(k_{a+}[A] + k_{b+}[B])[R] + k_{a-}[AR] + k_{b-}[BR], \tag{5.45a}$$

$$[AR]' = k_{a+}[A][R] - k_{a-}[AR], \tag{5.45b}$$

$$[BR]' = k_{b+}[B][R] - k_{b-}[BR], \tag{5.45c}$$

together with the initial conditions

$$[R](0) = R_{tot}, \qquad [AR](0) = 0, \qquad [BR](0) = 0, \tag{5.45d}$$

where R_{tot} is the total receptor concentration, form the initial value problem describing the kinetics of the system. Only the concentration of $[AR]$ is measured experimentally, hence the output is

$$y = [AR]. \tag{5.45e}$$

We assume the only known parameters are the drug concentrations $[A]$ and $[B]$, hence we have $\mathbf{p} = (k_{a+}, k_{a-}, k_{b+}, k_{b-}, R_{tot})$. We again apply the transfer function method, the Taylor series method and the similarity transformation method to determine identifiability.

Transfer function method

We first apply the transfer function method to determine identifiability of the parameters. Reducing the system (5.45) using the conservation law, $R_{tot} = [R] + [AR] + [BR]$ gives

$$[AR]' = -(k_{a+}[A] + k_{a-})[AR] - k_{a+}[A][BR] + k_{a+}[A]R_{tot}, \tag{5.46a}$$
$$[BR]' = -k_{b+}[B][AR] - (k_{b+}[B] + k_{b-})[BR] + k_{b+}[B]R_{tot}. \tag{5.46b}$$

Hence, we have the matrices, as in equation (5.2), as

$$F = \begin{bmatrix} -(k_{a+}[A] + k_{a-}) & -k_{a+}[A] \\ -k_{b+}[B] & -(k_{b+}[B] + k_{b-}) \end{bmatrix}, \quad G = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{b+}[B]R_{tot} \end{bmatrix}, \quad H = [1, 0]. \tag{5.47}$$

To check the identifiability of the system, we calculate the transfer function as

$$Q(s,\mathbf{p}) = \frac{k_{a+}[A]R_{tot}s + k_{a+}k_{b-}[A]R_{tot}}{s^2 + (k_{a+}[A] + k_{b+}[B] + k_{a-} + k_{b-})s + k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B] + k_{a-}k_{b-}}, \quad (5.48)$$

which gives the vector of coefficients as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{a+}k_{b-}[A]R_{tot} \\ k_{a+}[A] + k_{b+}[B] + k_{a-} + k_{b-} \\ k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B] + k_{a-}k_{b-} \end{bmatrix}. \quad (5.49)$$

Setting $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ allows us to determine identifiability. Using the first two entries we find k_{b-} to be identifiable, however, this is the only parameter to be so. This leaves the remaining identifiable parameter combinations as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}R_{tot} \\ k_{a+}[A] + k_{b+}[B] + k_{a-} \\ k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B] + k_{a-}k_{b-} \end{bmatrix}. \quad (5.50)$$

In Figure 5.3.2 the non-identifiability of the model is clear, as three parameter sets (as given in Table 5.2) result in vastly different species curves, yet the measured output curve of $[AR]$ is the same for all sets. In particular, we notice that the $[BR]$ curves have peaks in some of the curves, depending on the parameters used.

Taylor series method

To apply the Taylor series method we again use the full system, as stated in equation (5.45). As we now have three state variables, we have $n_x = 3$ and so require a maximum

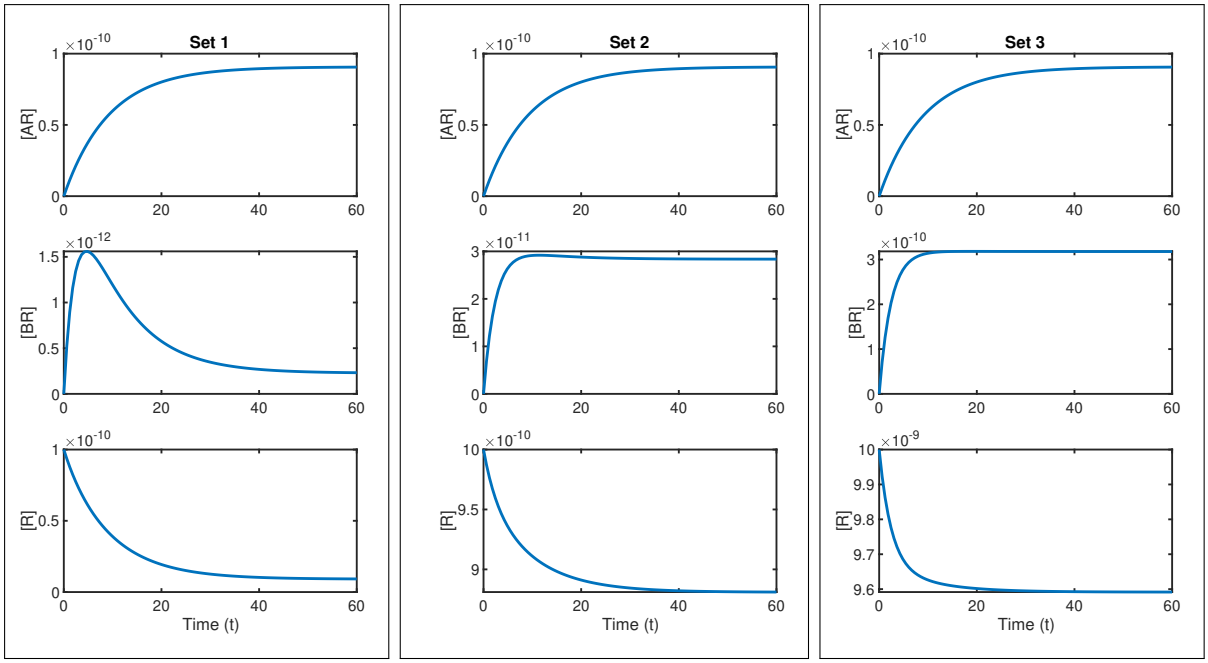

Figure 5.3.2: Three sets of parameters are used to plot the system given in equations (5.45). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.2 together with $[A] = 10^{-8}M$.

	Set 1	Set 2	Set 3
k_{a+}	$10^7 M^{-1}s^{-1}$	$10^6 M^{-1}s^{-1}$	$10^5 M^{-1}s^{-1}$
k_{a-}	$0.01s^{-1}$	$0.097s^{-1}$	$0.105s^{-1}$
k_{b+}	$10^5 M^{-1}s^{-1}$	$1.29 \times 10^5 M^{-1}s^{-1}$	$1.33 \times 10^5 M^{-1}s^{-1}$
k_{b-}	$0.4s^{-1}$	$0.4s^{-1}$	$0.4s^{-1}$
R_{tot}	$10^{-10}M$	$10^{-9}M$	$10^{-8}M$
$k_{a+}R_{tot}$	10^{-3M}	$10^{-3}M$	$10^{-3}M$
$k_{a+}[A] + k_{b+}[B] + k_{a-}$	$0.12s^{-1}$	$0.12s^{-1}$	$0.12s^{-1}$
$k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B] + k_{a-}k_{b-}$	$0.044s^{-2}$	$0.044s^{-2}$	$0.044s^{-2}$

Table 5.2: The parameters for three different parameter sets are used to plot Figure 5.3.2. The parameter k_{b-}, as well as the parameter combinations are equal in each case, so we can confirm these as identifiable.

of five coefficients to determine identifiability. Again, the first coefficient is

$$y_0 = 0, \tag{5.51}$$

hence gives no information regarding identifiability. As we have five unknown parameters, we can conclude that the system is not globally identifiable. The first derivative is

$$y^{(1)} = k_{a+}[A][R] - k_{a-}[AR], \tag{5.52}$$

giving the first Taylor coefficient as

$$y_0^{(1)} = k_{a+}[A]R_{tot}, \tag{5.53}$$

and as such, the first identifiable parameter combination as

$$c_1 = k_{a+}R_{tot}. \tag{5.54}$$

Using recursive substitution of equations (5.45), we have the second derivative as

$$\begin{aligned} y^{(2)} &= -k_{a+}[A]((k_{a+}[A] + k_{b+}[B])[R] - k_{a-}[AR] - k_{b-}[BR]) - k_{a-}(k_{a+}[A][R] - k_{a-}[AR]) \\ &= -k_{a+}[A](k_{a+}[A] + k_{b+}[B] + k_{a-})[R] + k_{a-}(k_{a-} + k_{a+}[A])[AR] + k_{a+}k_{b-}[A][BR], \end{aligned} \tag{5.55}$$

which gives the coefficient

$$y_0^{(2)} = k_{a+}[A]R_{tot}(k_{a+}[A] + k_{b+}[B] + k_{a-}). \tag{5.56}$$

This gives

$$y_0^{(2)} = k_{a+}[A]R_{tot}(k_{a+}[A] + k_{b+}[B] + k_{a-}), \tag{5.57}$$

$$= c_1[A](k_{a+}[A] + k_{b+}[B] + k_{a-}) \tag{5.58}$$

and therefore, we have a second identifiable combination as

$$c_2 = k_{a+}[A] + k_{b+}[B] + k_{a-}. \tag{5.59}$$

Further coefficients are calculated in the same way, using recursive substitution of the system variables to calculate higher order derivatives followed by substitution of the initial conditions. Using this method we obtain the third coefficient as

$$\begin{aligned} y_0^{(3)} &= k_{a+}[A]R_{tot}(k_{a+}^2[A]^2 + 2k_{a+}k_{b+}[A][B] + 2k_{a+}k_{a-}[A] \\ &\qquad + k_{b+}^2[B]^2 + k_{a-}k_{b+}[B] + k_{b+}k_{b-}[B] + k_{a-}^2) \\ &= c_1[A](c_2^2 + k_{b+}(k_{a-} - k_{b-})) \end{aligned} \tag{5.60}$$

which gives the third identifiable combination as

$$c_3 = k_{b+}(k_{a-} - k_{b-}). \tag{5.61}$$

We also obtain,

$$\begin{aligned} y_0^{(4)} &= k_{a+}[A]R_{tot}(k_{a+}^4[A]^4 + 4k_{a+}^3k_{b+}[A]^3[B] + 4k_{a+}^3k_{a-}[A]^3 + 6k_{a+}^2k_{b+}^2[A]^2[B]^2 \\ &+ 9k_{a+}^2k_{a-}k_{b+}[A]^2[B] + 3k_{a+}^2k_{b+}k_{b-}[A]^2[B] + 6k_{a+}^2k_{a-}^2[A]^2 + 4k_{a+}k_{b+}^3[A][B]^3 \\ &+ 6k_{a+}k_{a-}k_{b+}^2[A][B]^2 + 6k_{a+}k_{b+}^2k_{b-}[A][B]^2 + 6k_{a+}k_{a-}^2k_{b+}[A][B] + 4k_{a+}k_{a-}k_{b+}k_{b-}[A][B] \\ &+ 2k_{a+}k_{b+}k_{b-}^2[A][B] + 4k_{a+}k_{a-}^3[A] + k_{b+}^4[B]^4 + k_{a-}k_{b+}^3[B]^3 + 3k_{b+}^3k_{b-}[B]^3 + k_{a-}^2k_{b+}^2[B]^2 \\ &+ 2k_{a-}k_{b+}^2k_{b-}[B]^2 + 3k_{b+}^2k_{b-}^2[B]^2 + k_{a-}^3k_{b+}[B] + k_{a-}^2k_{b+}k_{b-}[B] + k_{a-}k_{b+}k_{b-}^2[B] \\ &+ k_{b+}k_{b-}^2[B] + k_{b+}k_{b-}^3[B] + k_{a-}^4), \end{aligned} \tag{5.62}$$

which, we find via some trial and error, is equal to

$$y_0^{(4)} = c_1(c_2^3 + c_3(2c_2 + k_{b-})) \tag{5.63}$$

which gives the identifiable parameter

$$c_4 = k_{b-}. \tag{5.64}$$

We note that calculations are all performed using Matlab's symbolic toolbox [2]. To conclude, we have k_{b-} is identifiable, and also the identifiable combinations

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}R_{tot}, \\ k_{b+}(k_{a-} - k_{b-}), \\ k_{a+}[A] + k_{b+}[B] + k_{a-}. \end{bmatrix} \tag{5.65}$$

Although it appears that these parameter combinations are not identical to the ones found from the transfer function method (equation (5.50)), we note that

$$k_{b+}(k_{a-} - k_{b-}) + k_{b-}(k_{a+}[A] + k_{b+}[B] + k_{a-}) = k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B] + k_{a-}k_{b-}, \quad (5.66)$$

hence, they are equivalent for identifiability purposes.

Similarity transformation method

To apply the similarity transformation method we require the system to be controllable and observable, however, in its original form the system has a zero input vector, so we are unable to determine controllability. Therefore, we instead reformulate and use the reduced form. Recall, in this form we have

$$F = \begin{bmatrix} -(k_{a+}[A] + k_{a-}) & -k_{a+}[A] \\ -k_{b+}[B] & -(k_{b+}[B] + k_{b-}) \end{bmatrix}, \quad G = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{b+}[B]R_{tot} \end{bmatrix}, \quad H = \begin{bmatrix} 1 & 0 \end{bmatrix}. \quad (5.67)$$

In this case, we have $n_x = 2$, and so the controllability matrix is given as

$$\mathcal{C} = \begin{bmatrix} k_{a+}[A]R_{tot} & -k_{a+}[A]R_{tot}(k_{a+}[A] + k_{a-} + k_{b+}[B]) \\ k_{b+}[B]R_{tot} & -k_{b+}[B]R_{tot}(k_{a+}[A] + k_{b+}[B] + k_{b-}) \end{bmatrix}, \quad (5.68)$$

which has $rank(\mathcal{C}) = 2$, and so the system is controllable. The observability matrix is

$$\mathcal{O} = \begin{bmatrix} 1 & 0 \\ -(k_{a+}[A] + k_{a-}) & -k_{a+}[A] \end{bmatrix}, \quad (5.69)$$

which also has $rank(\mathcal{O}) = 2$, therefore, both controllability and observability conditions are met. To check the identifiability, we assume a linear transformation

$$T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}, \tag{5.70}$$

and first noting, that (5.21b) holds for any T as we have $\mathbf{x}_0 = \widetilde{\mathbf{x}}_0 = (0,0)^T$, and therefore $T\widetilde{\mathbf{x}}_0 = (0,0)^T$. We now check condition (5.21e), which gives

$$\begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix} \quad \Rightarrow \quad t_{11} = 1, t_{12} = 0, \tag{5.71}$$

so we now have

$$T = \begin{bmatrix} 1 & 0 \\ t_{21} & t_{22} \end{bmatrix} \tag{5.72}$$

Condition (5.21d) states that

$$\begin{bmatrix} 1 & 0 \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} \widetilde{k_{a+}}[A]\widetilde{R_{tot}} \\ \widetilde{k_{b+}}[B]\widetilde{R_{tot}} \end{bmatrix} = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{b+}[B]R_{tot} \end{bmatrix}. \tag{5.73}$$

From the top row, we have

$$\widetilde{k_{a+}}[A]\widetilde{R_{tot}} = k_{a+}[A]R_{tot} \qquad \Rightarrow \qquad \widetilde{k_{a+}}\widetilde{R_{tot}} = k_{a+}R_{tot}, \tag{5.74}$$

which is our first identifiable parameter combination. The second row gives

$$\begin{aligned} & t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}} + t_{22}\widetilde{k_{b+}}[B]\widetilde{R_{tot}} = k_{b+}[B]R_{tot} \\ \Rightarrow \quad & t_{22} = \frac{k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}}}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}, \end{aligned} \tag{5.75}$$

and so T now becomes

$$T = \begin{bmatrix} 1 & 0 \\ t_{21} & \dfrac{k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}}}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}} \end{bmatrix}. \tag{5.76}$$

We next check condition (5.21c), giving

$$\begin{aligned} & \begin{bmatrix} 1 & 0 \\ t_{21} & \dfrac{k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}}}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}} \end{bmatrix} \begin{bmatrix} -(\widetilde{k_{a+}}[A] + \widetilde{k_{a-}}) & -\widetilde{k_{a+}}[A] \\ -\widetilde{k_{b+}}[B] & -(\widetilde{k_{b+}}[B] + \widetilde{k_{b-}}) \end{bmatrix} \\ & = \begin{bmatrix} -(k_{a+}[A] + k_{a-}) & -k_{a+}[A] \\ -k_{b+}[B] & -(k_{b+}[B] + k_{b-}) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t_{21} & \dfrac{k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}}}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}} \end{bmatrix}. \end{aligned} \tag{5.77}$$

Expanding gives

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}, \tag{5.78}$$

where

$$M_{11} = \widetilde{k_{a+}}[A] + \widetilde{k_{a-}}, \tag{5.79a}$$

$$M_{12} = \widetilde{k_{a+}}[A], \tag{5.79b}$$

$$M_{21} = t_{21}\widetilde{k_{a-}} + \frac{k_{b+}[B]R_{tot}}{\widetilde{R_{tot}}}, \tag{5.79c}$$

$$M_{22} = \frac{\widetilde{k_{b+}}k_{b+}[B]^2\widetilde{R_{tot}} + k_{b+}\widetilde{k_{b-}}[B]R_{tot} - t_{21}\widetilde{k_{a+}}\widetilde{k_{b-}}[A]\widetilde{R_{tot}}}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}, \tag{5.79d}$$

and

$$N_{11} = k_{a+}[A] + k_{a-} + t_{21}k_{a+}[A], \tag{5.80a}$$

$$N_{12} = \frac{k_{a+}[A](k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}})}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}, \tag{5.80b}$$

$$N_{21} = k_{b+}[B] + t_{21}k_{b+}[B] + t_{21}k_{b-}, \tag{5.80c}$$

$$N_{22} = \frac{(k_{b+}[B] + k_{b-})(k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}})}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}. \tag{5.80d}$$

We now equate the entries, M and N. First we solve $M_{12} = N_{12}$ for t_{21}, giving

$$\begin{aligned}\widetilde{k_{a+}}[A] &= \frac{k_{a+}[A](k_{b+}[B]R_{tot} - t_{21}\widetilde{k_{a+}}[A]\widetilde{R_{tot}})}{\widetilde{k_{b+}}[B]\widetilde{R_{tot}}} \\ t_{21} &= \frac{k_{a+}k_{b+}[B]R_{tot} - \widetilde{k_{a+}}\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}{k_{a+}\widetilde{k_{a+}}[A]\widetilde{R_{tot}}}.\end{aligned} \tag{5.81}$$

We note that, with this the transformation matrix becomes

$$T = \begin{bmatrix} 1 & 0 \\ \dfrac{k_{a+}k_{b+}[B]R_{tot} - \widetilde{k_{a+}}\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}{k_{a+}\widetilde{k_{a+}}[A]\widetilde{R_{tot}}} & \dfrac{\widetilde{k_{a+}}}{k_{a+}} \end{bmatrix}, \tag{5.82}$$

which has determinant

$$det(T) = \frac{\widetilde{k_{a+}}}{k_{a+}}. \tag{5.83}$$

As $det(T) \neq 0$, we find that condition (5.21a) holds. After substituting in the expression for t_{21}, we equate $M_{11} = N_{11}$, giving

$$\widetilde{k_{a+}}[A] + \widetilde{k_{a-}} = \frac{k_{a+}k_{b+}[B]R_{tot} + \widetilde{k_{a+}}k_{a-}\widetilde{R_{tot}} + k_{a+}\widetilde{k_{a+}}[A]\widetilde{R_{tot}} - \widetilde{k_{a+}}\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}{\widetilde{k_{a+}}\widetilde{R_{tot}}}. \tag{5.84}$$

However, as we have already determined that $\widetilde{k_{a+}}\widetilde{R_{tot}} = k_{a+}R_{tot}$, we have

$$\begin{aligned}
& \widetilde{k_{a+}}[A] + \widetilde{k_{a-}} = \frac{\widetilde{k_{a+}}k_{a-}\widetilde{R_{tot}} + k_{a+}\widetilde{k_{a+}}[A]\widetilde{R_{tot}} + \widetilde{k_{a+}}k_{b+}[B]\widetilde{R_{tot}} - \widetilde{k_{a+}}\widetilde{k_{b+}}[B]\widetilde{R_{tot}}}{\widetilde{k_{a+}}\widetilde{R_{tot}}} \\
\Rightarrow \quad & \widetilde{k_{a+}}[A] + \widetilde{k_{a-}} = k_{a-} + k_{a+}[A] + k_{b+}[B] - \widetilde{k_{b+}}[B] \\
\Rightarrow \quad & \widetilde{k_{a+}}[A] + \widetilde{k_{a-}} + \widetilde{k_{b+}}[B] = k_{a+}[A] + k_{a-} + k_{b+}[B].
\end{aligned} \tag{5.85}$$

Equating $M_{22} = N_{22}$ gives

$$\frac{k_{a+}k_{b+}[B]R_{tot} + \widetilde{k_{a+}}\widetilde{k_{b-}}\widetilde{R_{tot}}}{k_{a+}\widetilde{R_{tot}}} = \frac{\widetilde{k_{a+}}(k_{b+}[B] + k_{b-})}{k_{a+}}. \tag{5.86}$$

With $\widetilde{k_{a+}}\widetilde{R_{tot}} = k_{a+}R_{tot}$, this can simplified to give

$$\widetilde{k_{b-}} = k_{b-}. \tag{5.87}$$

Finally, equating $M_{21} = N_{21}$, after some simplification, gives

$$\widetilde{k_{b+}}(\widetilde{k_{a-}} - \widetilde{k_{b-}}) = k_{b+}(k_{a-} - k_{b-}). \tag{5.88}$$

Hence, we again have k_{b-} as identifiable, and identifiable parameter combinations

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}R_{tot}, \\ k_{b+}(k_{a-} - k_{b-}), \\ k_{a+}[A] + k_{b+}[B] + k_{a-}, \end{bmatrix} \tag{5.89}$$

as in the previous case. Comparing the methods, it is clear that, although all methods give the same identifiable parameter and parameter combinations, the transfer function method is by far the simplest in terms of ease of use. The Taylor series method, in particular, results in expressions that are algebraically complex and require much simplification in order to obtain reduced expressions.

5.3.3 Pre-formed homodimer binding with a single drug

The next model, and final linear model, we consider is the GPCR homodimer model we presented and analysed in Chapter 2, for singe ligand binding. Recall the schematic for the model as:

$$\mathrm{A} + \mathrm{R} \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} \mathrm{AR} \qquad \mathrm{A} + \mathrm{AR} \underset{\alpha_- k_{a-}}{\overset{\alpha_+ k_{a+}}{\rightleftharpoons}} \mathrm{ARA}$$

The ODE system describing the model dynamics is given as

$$\frac{d[R]}{dt} = -k_{a+}[A][R] + k_{a-}[AR], \tag{5.90a}$$

$$\frac{d[AR]}{dt} = k_{a+}[A][R] - (k_{a-} + \alpha_+ k_{a+}[A])[AR] + \alpha_- k_{a-}[ARA], \tag{5.90b}$$

$$\frac{d[ARA]}{dt} = \alpha_+ k_{a+}[A][AR] - \alpha_- k_{a-}[ARA], \tag{5.90c}$$

with initial conditions

$$[R](0) = R_{tot}, \qquad [AR](0) = 0, \qquad [ARA](0) = 0, \tag{5.90d}$$

The measured quantity is bound ligand, hence the output can be stated as

$$y = [AR] + 2[ARA]. \tag{5.90e}$$

We assume the only known parameter is the drug concentration, $[A]$, and so we have the vector of unknown parameters as $\mathbf{p} = (\alpha_+, \alpha_-, k_{a+}, k_{a-}, R_{tot})$. Whilst at first glance this system looks similar to the previous system, for monomer competition binding, in terms of complexity, we note that the output function is now a combination of two states, as opposed to just one, adding a slight but significant difference to the proceeding computations.

Transfer function method

Again, we first consider the transfer function method to determine identifiability. We first use conservation of receptors, which is stated in this case as

$$R_{tot} = [R] + [AR] + [ARA], \tag{5.91}$$

to reduce the system, giving

$$\frac{d[AR]}{dt} = -(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A])[AR] + (\alpha_- k_{a-} - k_{a+}[A])[ARA] + k_{a+}[A]R_{tot}, \tag{5.92a}$$

$$\frac{d[ARA]}{dt} = \alpha_+ k_{a+}[A][AR] - \alpha_- k_{a-}[ARA]. \tag{5.92b}$$

This gives the components, as stated in equation (5.2), as

$$F = \begin{bmatrix} -(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A]) & \alpha_- k_{a-} - k_{a+}[A] \\ \alpha_+ k_{a+}[A] & -\alpha_- k_{a-} \end{bmatrix},$$
$$G = \begin{bmatrix} k_{a+}[A]R_{tot} \\ 0 \end{bmatrix}, \qquad H = \begin{bmatrix} 1 & 2 \end{bmatrix}. \tag{5.93}$$

We give the transfer function of the system as

$$L(s, \mathbf{p}) = \frac{k_{a+}[A]R_{tot}(s + 2\alpha_+ k_{a+}[A] + \alpha_- k_{a-})}{s^2 + (k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A] + \alpha_- k_{a-})s + (\alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+}k_{a-}[A] + \alpha_- k_{a-}^2)}, \tag{5.94}$$

which gives the vector of coefficients as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}R_{tot} \\ 2\alpha_+ k_{a+}[A] + \alpha_- k_{a-} \\ k_{a+}[A] + k_{a-} - \alpha_+ k_{a+}[A] \\ \alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+}k_{a-}[A] + \alpha_- k_{a-}^2 \end{bmatrix}. \tag{5.95}$$

Hence we have no identifiable parameters but do have four identifiable parameter combinations. Again this can be seen in Figure 5.3.3 where we show how three sets of different parameter values result in vastly different individual species curves, yet all give the same A_{bound}, measured output curve. While the curves of $[R]$ and $[ARA]$ are similar in shape, they have different magnitudes of concentration. However, the largest differences are seen in the $[AR]$ curves, where the different parameter sets result in curves that have distinctly different evolution patterns, with some curves have a peak and fall while others don't. This highlights how naive parameter estimation performed without this knowledge of identifiability could lead to incorrect conclusions being drawn about the underlying qualitative dynamics. The parameter values used for the plots can be seen in Table 5.3.

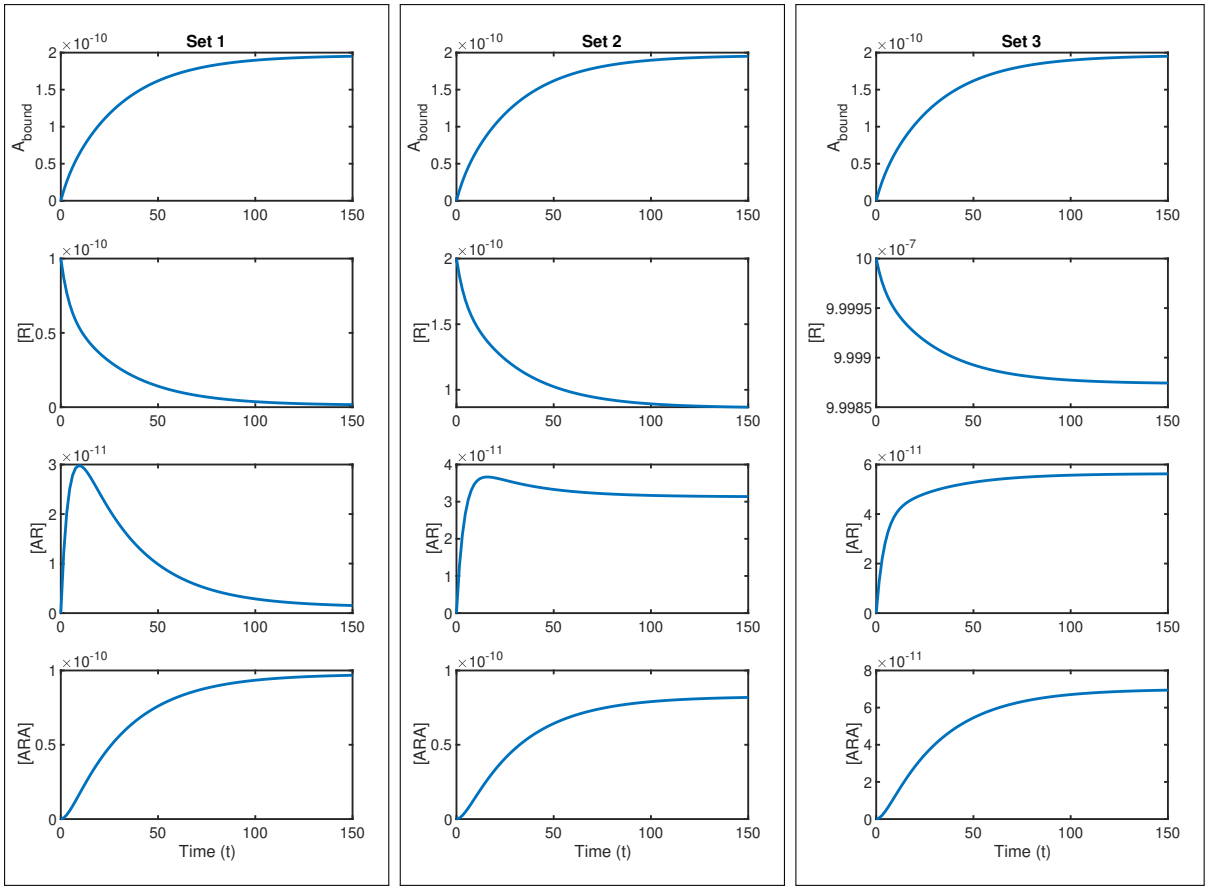

Figure 5.3.3: Three sets of parameters are used to plot the system given in equations (5.90). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.3 together with $[A] = 10^{-8}M$.

Taylor series method

We proceed with the Taylor series method to determine identifiability, with repeated substitution of the ODEs and initial conditions. While the process is the same as when we applied this method to the competition binding model, the output function being a combination of two state variables adds an extra complexity to the calculations. As in all previous sections, the first coefficient is trivial, that is

$$y_0 = 0. \tag{5.96}$$

	Set 1	Set 2	Set 3
k_{a+}	$10^7 M^{-1}s^{-1}$	$5 \times 10^6 M^{-1}s^{-1}$	$10^3 M^{-1}s^{-1}$
k_{a-}	$0.1s^{-1}$	$0.14s^{-1}$	$0.18s^{-1}$
α_+	0.8	1.35	5.74×10^3
α_-	0.01	0.19	0.26
R_{tot}	$10^{-10}M$	$2 \times 10^{-10}M$	$10^{-6}M$
$k_{a+}R_{tot}$	$10^{-3}s^{-1}$	$10^{-3}s^{-1}$	$10^{-3}s^{-1}$
$2\alpha_+ k_{a+}[A] + \alpha_- k_{a-}$	$0.161s^{-1}$	$0.161s^{-1}$	$0.161s^{-1}$
$k_{a+}[A] + k_{a-} - \alpha_+ k_{a+}[A]$	$0.12s^{-1}$	$0.12s^{-1}$	$0.12s^{-1}$
$\alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+}k_{a-}[A] + \alpha_- k_{a-}^2$	$0.0082s^{-2}$	$0.0082s^{-2}$	$0.0082s^{-2}$

Table 5.3: The parameters for three different parameter sets are used to plot Figure 5.3.3. The parameter combinations are equal in each case, so we can confirm these as identifiable.

The first derivative of the output function is given as

$$\begin{aligned} y^{(1)} &= k_{a+}[A][R] - (k_{a-} + \alpha_+ k_{a+}[A])[AR] + \alpha_- k_{a-}[ARA] + 2(\alpha_+ k_{a+}[A][AR] - \alpha_- k_{a-}[ARA]) \\ &= k_{a+}[A][R] - (k_{a-} - \alpha_+ k_{a+}[A])[AR] - \alpha_- k_{a-}[ARA], \end{aligned} \tag{5.97}$$

giving the first unique coefficient as

$$y_0^{(1)} = c_1 = k_{a+}[A]R_{tot}. \tag{5.98}$$

The remaining three coefficents are calculated in much the same way, repeatedly differentiating the output expression and substituting in the dynamic equations (equations (5.90)). This gives

$$y_0^{(2)} = k_{a+}[A]R_{tot}(\alpha_+ k_{a+}[A] - k_{a+}[A] - k_{a-}), \tag{5.99}$$

$$y_0^{(3)} = k_{a+}[A]R_{tot}(k_{a+}^2[A]^2 - \alpha_+^2 k_{a+}^2[A]^2 - \alpha_+ k_{a+}^2[A]^2 + 2k_{a+}k_{a-}[A] - \alpha_- k_{a+}k_{a-}[A] + k_{a-}^2), \tag{5.100}$$

$$\begin{aligned} y_0^{(4)} = {} & k_{a+}[A]R_{tot}(\alpha_+^3 k_{a+}^3[A]^2 + \alpha_+^2 k_{a+}^3[A]^3 + \alpha_+ k_{a+}^3[A]^3 - k_{a+}^3[A]^3 \\ & + 2\alpha_+^2\alpha_- k_{a+}^2 k_{a-}[A]^2 + \alpha_+\alpha_- k_{a+}^2 k_{a-}[A]^2 + \alpha_+^2 k_{a+}^2 k_{a-}[A]^2 - 3k_{a+}^2 k_{a-}[A]^2 \\ & + \alpha_+\alpha_-^2 k_{a+}k_{a-}^2[A] - \alpha_+ k_{a+}k_{a-}^2[A] - 3k_{a+}k_{a-}^2[A] - k_{a-}^3). \end{aligned} \tag{5.101}$$

These can again be reduced by division and subtraction of previous coefficients, giving the vector of coefficients, and, therefore, identifiable parameter combinations as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}R_{tot} \\ 2\alpha_+ k_{a+}[A] + \alpha_- k_{a-} \\ \alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+}k_{a-}[A] + \alpha_- k_{a-}^2 \\ k_{a+}[A] + k_{a-} - \alpha_+ k_{a+}[A] \end{bmatrix}, \tag{5.102}$$

hence, the results are the same as when using the transfer function method (that is, equal to equation (5.95)).

Similarity transformation method

Recall the reduced form of the system as

$$\begin{aligned} F &= \begin{bmatrix} -(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A]) & \alpha_- k_{a-} - k_{a+}[A] \\ \alpha_+ k_{a+}[A] & -\alpha_- k_{a-} \end{bmatrix}, \\ G &= \begin{bmatrix} k_{a+}[A]R_{tot} \\ 0 \end{bmatrix}, \qquad H = \begin{bmatrix} 1 & 2 \end{bmatrix}. \end{aligned} \tag{5.103}$$

Before we determine identifiability for the system, we first check whether the controllability and observability conditions are satisfied. The controllability matrix for this system is given as

$$\mathcal{C} = \begin{bmatrix} k_{a+}[A]R_{tot} & -k_{a+}[A]R_{tot}(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A]) \\ 0 & \alpha_+ k_{a+}^2[A]^2 R_{tot} \end{bmatrix}, \tag{5.104}$$

and the observability matrix is

$$\mathcal{O} = \begin{bmatrix} 1 & 2 \\ \alpha_+ k_{a+}[A] - k_{a+}[A] - k_{a-} & -(\alpha_- k_{a-} + k_{a+}[A]) \end{bmatrix}, \tag{5.105}$$

As $rank(\mathcal{C}) = rank(\mathcal{O}) = 2$ we can conclude that the system is both controllable and observable, hence we continue with identifiability. Assuming a linear transformation, with the matrix

$$T = \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}, \tag{5.106}$$

we first note that again having all initial conditions being zero means that condition (5.21b) automatically applies. We now apply condition (5.21e), giving

$$\begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{bmatrix}. \tag{5.107}$$

From this we determine

$$t_{11} = 1 - 2t_{21} \tag{5.108}$$

$$t_{12} = 2 - 2t_{22}, \tag{5.109}$$

which gives the transformation matrix as

$$T = \begin{bmatrix} 1 - 2t_{21} & 2 - 2t_{22} \\ t_{21} & t_{22} \end{bmatrix}. \tag{5.110}$$

We next apply condition (5.21d), giving

$$\begin{bmatrix} 1 - 2t_{21} & 2 - 2t_{22} \\ t_{21} & t_{22} \end{bmatrix} \begin{bmatrix} \widetilde{k_{a+}}[A]\widetilde{R_{tot}} \\ 0 \end{bmatrix} = \begin{bmatrix} k_{a+}[A]R_{tot} \\ 0 \end{bmatrix}. \tag{5.111}$$

From the bottom row, we find

$$t_{21} = 0. \tag{5.112}$$

With this, the top row then gives the identifiable parameter combination

$$\widetilde{k_{a+}}\widetilde{R_{tot}} = k_{a+}R_{tot}. \tag{5.113}$$

The transformation matrix now becomes

$$T = \begin{bmatrix} 1 & 2 - 2t_{22} \\ 0 & t_{22} \end{bmatrix}. \tag{5.114}$$

Applying condition (5.21c), gives

$$\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}, \tag{5.115}$$

where

$$M_{11} = \widetilde{\alpha_+}\widetilde{k_{a+}}[A] - \widetilde{k_{a+}}[A] - \widetilde{k_{a-}} - 2t_{22}\widetilde{\alpha_+}\widetilde{k_{a+}}[A], \tag{5.116a}$$

$$M_{12} = 2t_{22}\widetilde{\alpha_-}\widetilde{k_{a-}} - \widetilde{k_{a+}}[A] - \widetilde{\alpha_-}\widetilde{k_{a-}}, \tag{5.116b}$$

$$M_{21} = t_{22}\widetilde{\alpha_+}\widetilde{k_{a+}}[A], \tag{5.116c}$$

$$M_{22} = -t_{22}\widetilde{\alpha_-}\widetilde{k_{a-}}, \tag{5.116d}$$

and

$$N_{11} = -(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A]), \tag{5.117a}$$

$$N_{12} = t_{22}(k_{a+}[A] + 2k_{a-} + 2\alpha_+ k_{a+}[A] + \alpha_- k_{a-}) - 2(k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A]), \tag{5.117b}$$

$$N_{21} = \alpha_+ k_{a+}[A], \tag{5.117c}$$

$$N_{22} = 2\alpha_+ k_{a+}[A] - 2t_{22}\alpha_+ k_{a+}[A] - t_{22}\alpha_- k_{a-}. \tag{5.117d}$$

We solve $M_{21} = N_{21}$ for T_{22}, which gives

$$t_{22} = \frac{\alpha_+ k_{a+}}{\widetilde{\alpha_+}\widetilde{k_{a+}}}. \tag{5.118}$$

We note that, while it is possible to begin with a different matrix entry, and therefore have a different expression for t_{22}, this should still give the same identifiability results. With this expression, the transformation matrix is given by

$$T = \begin{bmatrix} 1 & 2 - 2\dfrac{\alpha_+ k_{a+}}{\widetilde{\alpha_+}\widetilde{k_{a+}}} \\ 0 & \dfrac{\alpha_+ k_{a+}}{\widetilde{\alpha_+}\widetilde{k_{a+}}} \end{bmatrix}, \tag{5.119}$$

which has determinant

$$det(T) = \frac{\alpha_+ k_{a+}}{\widetilde{\alpha_+}\widetilde{k_{a+}}} \neq 0, \tag{5.120}$$

hence, condition (5.21a) holds. We substitute the expression for t_{22} into the remaining matrix entries. Solving $M_{11} = N_{11}$, we have, after some simplification

$$\widetilde{k_{a+}}[A] + \widetilde{k_{a-}} - \widetilde{\alpha_+}\widetilde{k_{a+}}[A] = k_{a+}[A] + k_{a-} - \alpha_+ k_{a+}[A]. \tag{5.121}$$

Equating $M_{22} = N_{22}$ and $M_{12} = N_{12}$, we obtain

$$2\widetilde{\alpha_+}\widetilde{k_{a+}}[A] + \widetilde{\alpha_-}\widetilde{k_{a-}} = 2\alpha_+ k_{a+}[A] + \alpha_- k_{a-}, \tag{5.122}$$

and

$$\widetilde{\alpha_+}\widetilde{k_{a+}}^2[A]^2 + \widetilde{\alpha_-}\widetilde{k_{a+}}\widetilde{k_{a-}}[A] + \widetilde{\alpha_-}\widetilde{k_{a-}}^2 = \alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+} k_{a-}[A] + \alpha_- k_{a-}^2, \tag{5.123}$$

respectively. Hence, once again we have the identifiable parameter combinations for this model as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+} R_{tot} \\ 2\alpha_+ k_{a+}[A] + \alpha_- k_{a-} \\ \alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+} k_{a-}[A] + \alpha_- k_{a-}^2 \\ k_{a+}[A] + k_{a-} - \alpha_+ k_{a+}[A] \end{bmatrix}, \tag{5.124}$$

confirming that all methods give the same results. Again, if we compare the three methods applied to this system, we find that the transfer function method to be decidedly

simpler to use, whereas the Taylor series method results in expressions that require much simplification.

5.3.4 Ligand induced dimerisation with a single drug

The final model we consider is the VEGF, ligand induced dimerisation model that we presented in Chapter 4. Recall the schematic for the model is given by

$$\mathrm{A} + \mathrm{R} \underset{k_-}{\overset{k_+}{\rightleftharpoons}} \mathrm{AR} \qquad \mathrm{R} + \mathrm{AR} \underset{\psi_- k_-}{\overset{\psi_+ k_+}{\rightleftharpoons}} \mathrm{RAR}.$$

The ODE system that governs the system dynamics can then be stated as

$$\frac{d[R]}{dt} = -k_+[A][R] + k_-[AR] - \psi_+ k_+[R][AR] + \psi_- k_-[RAR], \tag{5.125a}$$

$$\frac{d[AR]}{dt} = k_+[A][R] - k_-[AR] - \psi_+ k_+[R][AR] + \psi_- k_-[RAR], \tag{5.125b}$$

$$\frac{d[RAR]}{dt} = \psi_+ k_+[R][AR] - \psi_- k_-[RAR]. \tag{5.125c}$$

with initial conditions

$$[R](0) = R_{tot}, \qquad [AR](0) = 0, \qquad [RAR](0) = 0 \tag{5.125d}$$

The signal of interest, and measured output for the system, is proportional to the number of receptors bound, that is

$$y = a([AR] + 2[RAR]), \tag{5.125e}$$

for a scalar a. Again the only known parameter is the drug concentration, $[A]$, and so we have $\mathbf{p} = (a, k_+, k_-, \psi_+, \psi_-, R_{tot})$. We note that this system is nonlinear, which affects the

available methods to determine identifiability. The transfer function method is applicable only to linear systems, and so we are unable to use this method.

Taylor series method

We first apply the Taylor series method to the VEGF system. While we have $n = 3$, the system is nonlinear and so it is unknown how many coefficients will need to be calculated for identifiability to be established. As in all previous sections, the first coefficient is trivial, that is

$$y_0 = 0. \tag{5.126}$$

and so gives no information. The first derivative of the output function is given as

$$\begin{aligned} y^{(1)} &= a([AR] + 2[RAR]) && (5.127)\\ &= a(k_+[A][R] - k_-[AR] + \psi_+ k_+[R][AR] - \psi_- k_-[RAR]), && (5.128) \end{aligned}$$

giving the first unique Taylor coefficient as

$$y_0^{(1)} = c_1 = ak_+[A]R_{tot}. \tag{5.129}$$

The second coefficient we calculate as

$$\begin{aligned} y_0^{(2)} &= ak_+[A]R_{tot}(k_+[A] + k_- - \psi_+ k_+ R_{tot}) \\ \Rightarrow \quad c_2 &= k_+[A] + k_- - \psi_+ k_+ R_{tot}. \end{aligned} \tag{5.130}$$

We continue calculating coefficents using repeated differentiation and substitution, giving

$$y_0^{(3)} = ak_+[A]R_{tot}(k_+^2[A]^2 + 2k_+k_-[A] + k_-^2 - 4\psi_+k_+^2[A]R_{tot} - \psi_+^2k_+^2R_{tot}^2 - \psi_+\psi_-k_+k_-R_{tot}), \tag{5.131}$$

$$y_0^{(4)} = ak_+[A]R_{tot}(\psi_+^3k_+^3R_{tot}^3 + 2\psi_+^2\psi_-k_+^2k_-R_{tot}^2 + 3\psi_+^2k_+^3[A]R_{tot}^2 + \psi_+^2k_+^2k_-R_{tot}^2 + \psi_+\psi_-^2k_+k_-^2R_{tot} + 4\psi_+\psi_-k_+^2k_-[A]R_{tot} + 11\psi_+k_+^3[A]^2R_{tot} + 6\psi_+k_+^2k_-[A]R_{tot} - \psi_+k_+k_-^2R_{tot} - k_+^3[A]^3 - 3k_+^2k_-[A]^2 - 3k_+k_-^2[A] - k_-^3), \tag{5.132}$$

$$y_0^{(5)} = -ak_+[A]R_{tot}(\psi_+^4k_+^4R_{tot}^4 + 3\psi_+^3\psi_-k_+^3k_-R_{tot}^3 - 4\psi_+^3k_+^4[A]R_{tot}^3 + 2\psi_+^3k_+^3k_-R_{tot}^3 + 3\psi_+^2\psi_-^2k_+^2k_-^2R_{tot}^2 + 6\psi_+^2\psi_-k_+^3k_-[A]R_{tot}^2 + 2\psi_+^2\psi_-k_+^2k_-^2R_{tot}^2 - 2\psi_+^2k_+^4[A]^2R_{tot}^2 + 18\psi_+^2k_+^3k_-[A]R_{tot}^2 + \psi_+\psi_-^3k_+k_-^3R_{tot} + 4\psi_+\psi_-^2k_+^2k_-^2[A]R_{tot}[A]^2R_{tot} + 11\psi_+\psi_-k_+^3k_- + 6\psi_+\psi_-k_+^2k_-^2[A]R_{tot} - \psi_+\psi_-k_+k_-^3R_{tot} + 26\psi_+k_+^4[A]^3R_{tot} + 34\psi_+k_+^3k_-[A]^2R_{tot} + 6\psi_+k_+^2k_-^2[A]R_{tot} - 2\psi_+k_+k_-^3R_{tot} - k_+^4[A]^4 - 4k_+^3k_-[A]^3 - 6k_+^2k_-^2[A]^2 - 4k_+k_-^3[A] - k_-^4). \tag{5.133}$$

Although theoretically it may be possible to reduce these expressions, as we did with the linear models, it is not obvious as to how to do this. Using these coefficients directly to create the vector of identifiable parameter combinations is also impractical, as setting $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ would involve solving a fourth order polynomial. If further coefficients were also needed to determine identifiability these would be of higher degree still. These issues were also noted by Chappell, Godfrey and Vajda [24] where they compared the Taylor series method and similarity transformation method for nonlinear compartment models. Hence, we conclude that the Taylor series method is impractical to use for nonlinear problems.

Similarity transformation method

Although we have used this method to establish identifiability for the previous models, these were all linear models. There are a number of key differences in the steps taken to apply the method to a nonlinear system. We first reduce the system (equation (5.125)),

using the conservation law

$$R_{tot} = [R] + [AR] + [RAR], \tag{5.134}$$

which gives the system components as

$$\mathbf{f}(\mathbf{x}, \mathbf{p}) = \begin{bmatrix} \psi_+ k_+ [AR]^2 + 2\psi_+ k_+ [AR][RAR] - (k_+[A] + k_- \\ \qquad + \psi_+ k_+ R_{tot})[AR] + (\psi_- k_- - 2k_+[A])[RAR] \\ -\psi_+ k_+ [AR]^2 - 2\psi_+ k_+ [AR][RAR] + \psi_+ k_+ R_{tot}[AR] - \psi_- k_- [RAR] \end{bmatrix}, \tag{5.135}$$

$$\mathbf{g}(\mathbf{x}, \mathbf{p}) = \begin{bmatrix} k_+[A]R_{tot} \\ 0 \end{bmatrix}, \tag{5.136}$$

$$\mathbf{h}(\mathbf{x}, \mathbf{p}) = a([AR] + 2[RAR]). \tag{5.137}$$

In order to use this method, we require the system to be observable and controllable. Recall, the definitions of controllability and observability as given in Section 5.2.3. The observability condition makes use of the Lie derivative, so we calculate

$$L^0_{\mathbf{f}}\mathbf{y}(\mathbf{x}) = \mathbf{y}(\mathbf{x}), \tag{5.138}$$

and

$$\begin{aligned} L^1_{\mathbf{f}}\mathbf{y}(\mathbf{x}) &= \frac{\partial \mathbf{y}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) \\ &= -a(\psi_+ k_+([AR] + 2[RAR] - R_{tot})[AR] \\ &\qquad + (k_+[A] + k_-)[AR] + (\psi_- k_- + 2k_+[A])[RAR]), \end{aligned} \tag{5.139}$$

giving the observability matrix as

$$\mathcal{O} = \begin{bmatrix} a & 2a \\ -a(\psi_+ k_+(2[AR] + 2[RAR] - R_{tot}) + (k_+[A] + k_-)) & -a(2\psi_+ k_+[AR][AR] + (\psi_- k_- + 2k_+[A])) \end{bmatrix}, \tag{5.140}$$

which has $rank(\mathcal{O}) = 2$. We calculate the controllability matrix, making use of the Lie bracket. Noting that

$$\begin{aligned} [f, g] &= \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{f}(\mathbf{x}) - \frac{\partial \mathbf{f}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{g}(\mathbf{x}) \\ &= \begin{bmatrix} k_+[A]R_{tot}(k_+[A] + k_- - \psi_+ k_+(2[AR] + 2[RAR] - R_{tot})) \\ \psi_+ k_+^2[A]R_{tot}(2[AR] + 2[RAR] - R_{tot}) \end{bmatrix}, \end{aligned} \tag{5.141}$$

giving the controllability matrix as

$$\mathcal{C} = \begin{bmatrix} k_+[A]R_{tot} & k_+[A]R_{tot}(k_+[A] + k_- - \psi_+ k_+(2[AR] + 2[RAR] - R_{tot})) \\ 0 & \psi_+ k_+^2[A]R_{tot}(2[AR] + 2[RAR] - R_{tot}) \end{bmatrix}, \tag{5.142}$$

which also has $rank(\mathcal{C}) = 2$. Hence, we can confirm that the system is both observable and controllable. We now assume a transformation $\boldsymbol{\lambda}(\mathbf{x}) = (\lambda_1(\mathbf{x}), \lambda_2(\mathbf{x}))$ and assume that this transformation preserves the system output, $\mathbf{y}(\mathbf{x})$, as well as the algebraic struc-

ture of the system. This is done by showing that all conditions in equations (5.22) hold.

The first condition we check is (5.22e), that is, $\mathbf{h}(\boldsymbol{\lambda}(\widetilde{\mathbf{x}}), \mathbf{p}) = \mathbf{h}(\widetilde{\mathbf{x}}, \widetilde{\mathbf{p}})$. This gives

$$a(\lambda_1 + 2\lambda_2) = \widetilde{a}([AR] + 2[RAR]), \qquad \Rightarrow \qquad \lambda_1 = \frac{\widetilde{a}}{a}([AR] + 2[RAR]) - 2\lambda_2. \quad (5.143)$$

Differentiating this with respect to $[AR]$ gives

$$\frac{\partial \lambda_1}{\partial [AR]} = \frac{\widetilde{a}}{a} - 2\frac{\partial \lambda_2}{\partial [AR]}. \quad (5.144)$$

Similarly, differentiating with respect to $[RAR]$ gives

$$\frac{\partial \lambda_1}{\partial [RAR]} = 2\left(\frac{\widetilde{a}}{a} - \frac{\partial \lambda_2}{\partial [RAR]}\right). \quad (5.145)$$

Using these, we have the Jacobian of $\boldsymbol{\lambda}(\mathbf{x})$ as

$$\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{x}} = \begin{bmatrix} \dfrac{\widetilde{a}}{a} - 2\dfrac{\partial \lambda_2}{\partial [AR]} & 2\left(\dfrac{\widetilde{a}}{a} - \dfrac{\partial \lambda_2}{\partial [RAR]}\right) \\ \dfrac{\partial \lambda_2}{\partial [AR]} & \dfrac{\partial \lambda_2}{\partial [RAR]} \end{bmatrix}. \quad (5.146)$$

We now consider condition (5.22d) which says we must have $\mathbf{g}(\boldsymbol{\lambda}(\widetilde{\mathbf{x}}), \mathbf{p}) = \dfrac{\partial \boldsymbol{\lambda}(\widetilde{\mathbf{x}})}{\partial \widetilde{\mathbf{x}}} \mathbf{g}(\widetilde{\mathbf{x}}, \widetilde{\mathbf{p}})$.

This gives

$$\begin{bmatrix} k_+[A]R_{tot} \\ \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{\widetilde{a}}{a} - 2\frac{\partial\lambda_2}{\partial[AR]} & 2\left(\frac{\widetilde{a}}{a} - \frac{\partial\lambda_2}{\partial[RAR]}\right) \\ \frac{\partial\lambda_2}{\partial[AR]} & \frac{\partial\lambda_2}{\partial[RAR]} \end{bmatrix} \begin{bmatrix} \widetilde{k_+}[A]\widetilde{R_{tot}} \\ \\ 0 \end{bmatrix} \tag{5.147}$$

$$= \begin{bmatrix} \left(\frac{\widetilde{a}}{a} - 2\frac{\partial\lambda_2}{\partial[AR]}\right)\widetilde{k_+}[A]\widetilde{R_{tot}} \\ \\ \widetilde{k_+}[A]\widetilde{R_{tot}}\frac{\partial\lambda_2}{\partial[AR]} \end{bmatrix}. \tag{5.148}$$

The second row of these gives

$$\frac{\partial\lambda_2}{\partial[AR]} = 0. \tag{5.149}$$

Substituting this into the first row gives our first identifiable parameter combination as

$$ak_+R_{tot} = \widetilde{a}\widetilde{k_+}\widetilde{R_{tot}}. \tag{5.150}$$

The Jacobian matrix now becomes

$$\frac{\partial\boldsymbol{\lambda}}{\partial\mathbf{x}} = \begin{bmatrix} \frac{\widetilde{a}}{a} & 2\left(\frac{\widetilde{a}}{a} - \frac{\partial\lambda_2}{\partial[RAR]}\right) \\ \\ 0 & \frac{\partial\lambda_2}{\partial[RAR]} \end{bmatrix}. \tag{5.151}$$

We now check condition (5.22c) which we recall says, $\mathbf{f}(\boldsymbol{\lambda}(\widetilde{\mathbf{x}}),\mathbf{p}) = \dfrac{\partial\boldsymbol{\lambda}(\widetilde{\mathbf{x}})}{\partial\widetilde{\mathbf{x}}}\mathbf{f}(\widetilde{\mathbf{x}},\widetilde{\mathbf{p}})$, giving

$$\frac{1}{a^2}\begin{bmatrix} a(ak_+(2+\psi_- + 2\psi_+k_+R_{tot}) - 2\widetilde{a}\psi_+k_+([AR]+2[RAR]))\lambda_2 \\ \quad -\widetilde{a}([AR]+2[RAR])(a(k_+[A]+k_- + \psi_+k_+R_{tot}) - \widetilde{a}([AR]+2[RAR])) \\ \\ -a(a\psi_-k_- + \psi_+k_+(aR_{tot} - \widetilde{a}([AR]+2[RAR])))\lambda_2 \\ \quad +\widetilde{a}\psi_+k_+([AR]+2[RAR])(aR_{tot} - \widetilde{a}([AR]+2[RAR])) \end{bmatrix} =$$

$$\begin{bmatrix} \dfrac{\widetilde{a}}{a} & 2\left(\dfrac{\widetilde{a}}{a} - \dfrac{\partial\lambda_2}{\partial[RAR]}\right) \\ \\ 0 & \dfrac{\partial\lambda_2}{\partial[RAR]} \end{bmatrix} \begin{bmatrix} \widetilde{\psi_+}\widetilde{k_+}([AR]+2[RAR])[AR] \\ \quad -(\widetilde{k_+}[A] + \widetilde{k_-} + \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}})[AR] + (\widetilde{\psi_-}\widetilde{k_-} - 2\widetilde{k_+}[A])[RAR] \\ \\ -\widetilde{\psi_+}\widetilde{k_+}[AR]^2 - 2\widetilde{\psi_+}\widetilde{k_+}[AR][RAR] + \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}}[AR] - \widetilde{\psi_-}\widetilde{k_-}[RAR] \end{bmatrix}. \tag{5.152}$$

Solving these for λ_2 and $\partial\lambda_2/\partial[RAR]$ gives

$$\lambda_2 = \frac{\widetilde{a}(c_1[AR]^2 + c_2[AR][RAR] + c_3[AR] + c_4[RAR]^2 + c_5[RAR])}{a(ak_-(2-\psi_-) + 2\psi_+k_+(\widetilde{a}([AR]+2[RAR]) - aR_{tot}))}, \tag{5.153}$$

where

$$\begin{aligned} c_1 &= \widetilde{a}\psi_+k_+ - a\widetilde{\psi_+}\widetilde{k_+}, \\ c_2 &= 4\widetilde{a}\psi_+k_+ - 2a\widetilde{\psi_+}\widetilde{k_+}, \\ c_3 &= a(k_+[A] + k_- - \psi_+k_+R_{tot} - (\widetilde{k_+}[A] + \widetilde{k_-} - \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}})) \\ c_4 &= 4\widetilde{a}\psi_+k_+, \\ c_5 &= a(2(k_+[A] + k_- - \psi_+k_+R_{tot}) - 2\widetilde{k_+}[A] - \widetilde{\psi_-}\widetilde{k_-}), \end{aligned}$$

and

$$\frac{\partial\lambda_2}{\partial[RAR]} = \frac{\begin{array}{r} d_1([AR]^3 + 4[AR]^2[RAR] + 4[AR][RAR]^2) \\ + d_2[AR]^2 + d_3[AR][RAR] + d_4[AR] + d_5[RAR]^2 + d_6[RAR] \end{array}}{\begin{array}{r} a(\widetilde{\psi_+}\widetilde{k_+}([AR]+2[RAR])[AR] - \widetilde{R_{tot}}) + \widetilde{\psi_-}\widetilde{k_-}[RAR]) \\ \times (ak_-(2-\psi_-) + 2\psi_+k_+(\widetilde{a}([AR]+2[RAR]) - aR_{tot})) \end{array}}, \tag{5.154}$$

where

$$
\begin{aligned}
d_1 &= 2\widetilde{a}\psi_+\widetilde{\psi_+}k_+\widetilde{k_+},\\
d_2 &= 2\widetilde{a}\psi_+k_+(\widetilde{k_-} - k_+[A] + \widetilde{k_+}[A] - \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}}) - a\widetilde{\psi_+}\widetilde{k_+}(\psi_-k_- + 2\psi_+k_+R_{tot}),\\
d_3 &= 2\widetilde{a}\psi_+k_+(2\widetilde{k_-} - 4k_+[A] + 4\widetilde{k_+}[A] - 2\widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}}) - 2a\widetilde{\psi_+}\widetilde{k_+}(\psi_-k_- + 2\psi_+k_+R_{tot}),\\
d_4 &= a(\psi_-k_-(k_- + k_+[A] - \widetilde{k_-} - \widetilde{k_+}[A] + \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}})\\
&\qquad + 2\psi_+k_+R_{tot}(k_+[A] - \widetilde{k_+}[A] - \widetilde{k_-} + \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}})),\\
d_5 &= 8\widetilde{a}\psi_+k_+(2\widetilde{k_+}[A] - 2k_+[A] + \widetilde{\psi_-}\widetilde{k_-}),\\
d_6 &= a(\psi_-k_-(2k_- + 2k_+[A] - \widetilde{\psi_-}\widetilde{k_-} - 2\widetilde{k_+}[A]) + 2\psi_+k_+R_{tot}(2k_+[A] - 2\widetilde{k_+}[A] - \widetilde{\psi_-}\widetilde{k_-})).
\end{aligned}
$$

Differentiating λ_2, that is equation (5.153), with respect to $[AR]$ gives

$$
\frac{\partial\lambda_2}{\partial[AR]} = \frac{\widetilde{a}(m_1([AR]^2 + 4[AR][RAR] + 4[RAR]^2) + m_2[AR] + m_3[RAR] + m_4)}{a(ak_-(2-\psi_-) + 2\psi_+k_+(\widetilde{a}([AR] + 2[RAR]) - aR_{tot}))^2}, \tag{5.155}
$$

where

$$
\begin{aligned}
m_1 &= 2\widetilde{a}\psi_+k_+(\widetilde{a}\psi_+k_+ - a\widetilde{\psi_+}\widetilde{k_+}),\\
m_2 &= -2a(\widetilde{a}\psi_+k_+ - a\widetilde{\psi_+}\widetilde{k_+})(\psi_-k_- - 2k_- + 2\psi_+k_+R_{tot}),\\
m_3 &= 2a(\widetilde{a}\psi_+k_+(4k_- - 2\widetilde{k_-} - 2\psi_-k_- + \widetilde{\psi_-}\widetilde{k_-} - 4\psi_+k_+R_{tot}\\
&\qquad + 2\widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}}) + a\widetilde{\psi_+}\widetilde{k_+}(\psi_-k_- - 2k_- + 2\psi_+k_+R_{tot})),\\
m_4 &= -a^2(\psi_-k_- - 2k_- + 2\psi_+k_+R_{tot})(k_+[A] + k_- - \psi_+k_+R_{tot} - (\widetilde{k_+}[A] + \widetilde{k_-} - \widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}})).
\end{aligned}
$$

However, we found earlier, in equation (5.149), that $\partial\lambda_2/\partial[AR] = 0$, hence each of these coefficients (m_1-m_4) must equal zero. We find

$$
m_1 = 0, \quad \Rightarrow \quad \widetilde{a}\psi_+k_+ = a\widetilde{\psi_+}\widetilde{k_+}, \tag{5.156}
$$

as a second identifiable parameter combination. Also

$$m_3 = 0, \quad \Rightarrow \quad 2k_- - \psi_- k_- - 2\psi_+ k_+ R_{tot} = 2\widetilde{k_-} - \widetilde{\psi_-}\widetilde{k_-} - 2\widetilde{\psi_+}\widetilde{k_+}\widetilde{R_{tot}}. \tag{5.157}$$

Differentiating λ_2 with respect to $[RAR]$, and simplifying using the combinations identified in equations (5.150), (5.156) and (5.157), gives

$$\frac{\partial \lambda_2}{\partial [RAR]} = \frac{4\widetilde{a}^4 \psi_+^2 k_+^2 ([AR]^2 + 4[AR] + 4[RAR]^2)}{a(ak_-(2-\psi_-) + 2\psi_+ k_+ (\widetilde{a}([AR] + 2[RAR]) - aR_{tot}))^2}. \tag{5.158}$$

This must be equal to equation (5.154). Equating coefficients, while using previously identified parameter combinations to simplify where possible, we find

$$aR_{tot} = \widetilde{a}\widetilde{R_{tot}}, \quad k_+ = \widetilde{k_+}, \quad k_- = \widetilde{k_-}, \quad \psi_+ R_{tot} = \widetilde{\psi_+}\widetilde{R_{tot}}, \quad \psi_- = \widetilde{\psi_-}. \tag{5.159}$$

Substituting these into equation (5.153) gives the λ_2 solution as

$$\lambda_2(\mathbf{x}) = \frac{\widetilde{a}}{a}[RAR], \tag{5.160}$$

and it follows that the λ_1 solution is therefore

$$\lambda_1(\mathbf{x}) = \frac{\widetilde{a}}{a}[AR]. \tag{5.161}$$

Clearly we have

$$\lambda_1(\mathbf{0}) = 0, \quad \text{and} \quad \lambda_2(\mathbf{0}) = 0, \tag{5.162}$$

confirming that, condition (5.22b) holds. Finally, we have

$$\frac{\partial \lambda_2}{\partial [RAR]} = \frac{\widetilde{a}}{a}, \tag{5.163}$$

and so the Jacobian becomes

$$\frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\widetilde{a}}{a} & 0 \\ \\ 0 & \frac{\widetilde{a}}{a} \end{bmatrix}. \tag{5.164}$$

This has $rank = 2$, hence we have confirmed that the final condition (5.22a) also holds, concluding that the algebraic structure is preserved under this linear transformation. With this we can successfully identify parameters

$$k_+, \quad k_-, \quad \psi_-, \tag{5.165}$$

and have the identifiable parameter combinations

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} aR_{tot} \\ \psi_+ R_{tot} \end{bmatrix} \tag{5.166}$$

Although the system in not fully identifiable, we note that this is due only to the output being scaled. If this is not the case, or if the value of this scaler is known, the system becomes fully identifiable. This is clear if we look at Figure 5.3.4 together with parameter values in Table 5.4. In these we see that, although we have nonidentifiability, the differences in the curves are simply a change in magnitude of concentrations. That is, they are all scaled to a different magnitude.

The results are also confirmed using the STRIKE-GOLDD toolbox [129], in Matlab

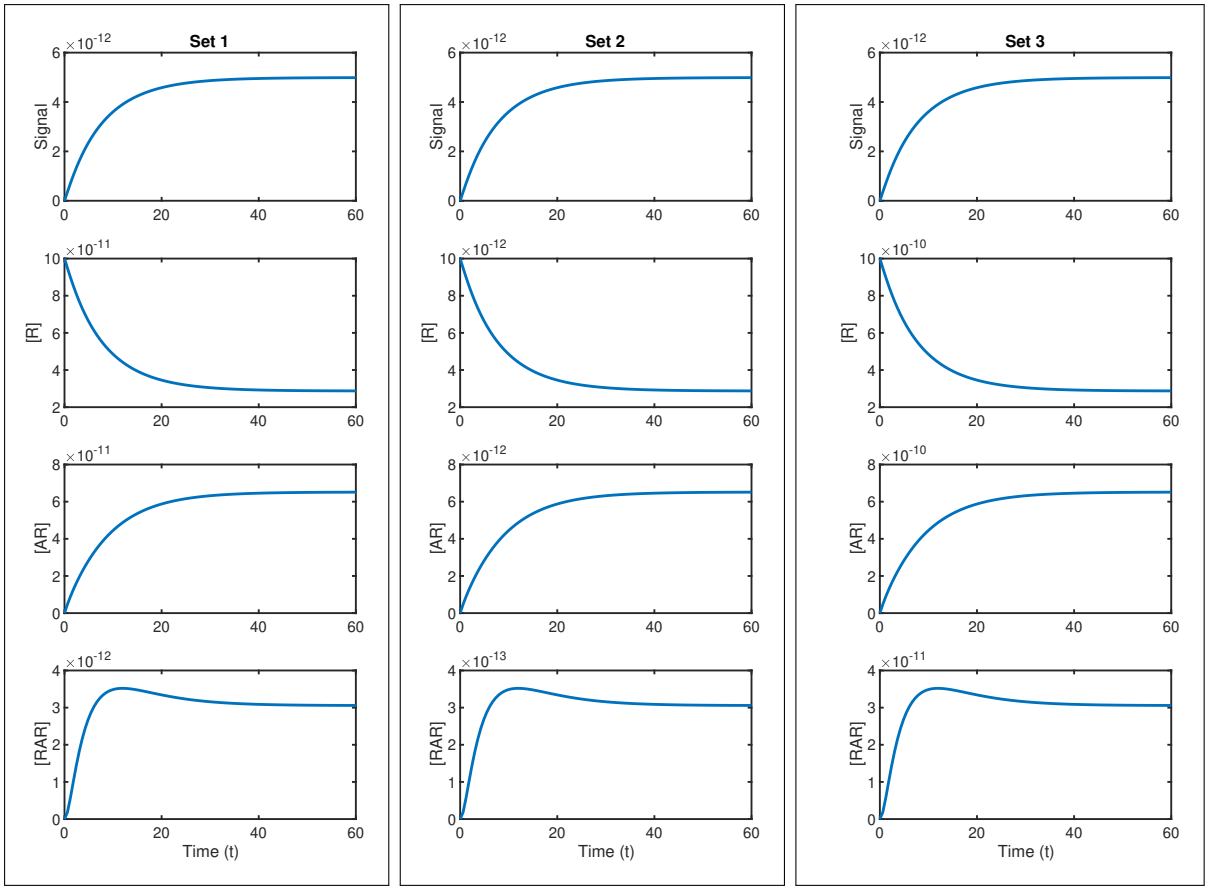

Figure 5.3.4: Three sets of parameters are used to plot the system given in equations (5.125). All three parameter sets give the same measured output curve, A_{bound}. However, non-identifiability can be seen in the individual species curves. Each set of plots is created using the values in Table 5.4 together with $[A] = 10^{-8}M$.

[2] toolbox, the results of which can be seen in Figure 5.3.5.

5.3.5 A comparison of methods

For linear systems, we can conclude that all methods give the same results for the models studied. In all considered models, we find that it is not possible to identify all parameters from a single set of time course data. In fact, very few parameters can be successfully recovered, hence further information is required to ensure identifiability. In all cases, we find the transfer function method requires the least computation, with the identifiable combinations returned directly, that is, without the need for simplification. The Taylor series the returned parameter combinations are substantially more complex than those found in the transfer function method, requiring much more simplification. The Taylor

	Set 1	Set 2	Set 3
a	0.07	0.7	0.007
k_+	$8.4 \times 10^6 M^{-1} s^{-1}$	$8.4 \times 10^6 M^{-1} s^{-1}$	$8.4 \times 10^6 M^{-1} s^{-1}$
k_-	$0.037 s^{-1}$	$0.037 s^{-1}$	$0.037 s^{-1}$
ψ_+	187	1.87	18.7
ψ_-	26	26	26
R_{tot}	$10^{-10} M$	$10^{-11} M$	$10^{-9} M$
$k_+ R_{tot}$	$7 \times 10^{-12} s^{-1}$	$7 \times 10^{-12} s^{-1}$	$7 \times 10^{-12} s^{-1}$
$\psi_+ R_{tot}$	$1.87 \times 10^{-8} M$	$1.87 \times 10^{-8} M$	$1.87 \times 10^{-8} M$

Table 5.4: The parameters for three different parameter sets are used to plot Figure 5.3.4. The parameter combinations are equal in each case, so we can confirm these as identifiable.

series method does have the advantage, however, that the system can remain in its full form without any reformulation needed, and unknowns in the initial conditions can also be determined. The main disadvantage of the similarity transformation method is in the inability to algorithmically compute the solutions. That is, the results of each step need to be considered before deciding on how to proceed with the next. This is in contrast to the transfer function method and Taylor series method which both follow a structured algorithm to obtain solutions.

For nonlinear systems we apply only the Taylor series and similarity transformation methods, as the transfer function method is applicable to only linear systems. From the Taylor series method, we are unable to determine what parameters are identifiable or unidentifiable as the parameter combinations are long and algebraically complex. From the similarity transformation method, we find some parameters to be identifiable, and in fact if the observable function scalar can be predetermined, all parameters can be theoretically identified. Although, we once again note that this is assuming noise-free data. This is unlikely to be the case so issues of practical identifiability also need to be taken into account. Though the method is not simple, and will be more difficult for larger systems, it gives clear results. The Matlab [2] toolbox, STRIKE-GOLDD confirms the

```
--------------------------------
>>> STRIKE-GOLDD toolbox 3.0
--------------------------------

Analyzing the lid_s model...

>>> The model contains:
2 states:
matrix([[x1], [x2]])
1 outputs:
a*(x1 + 2*x2)
0 known inputs:

0 unknown inputs:

6 parameters:
matrix([[a], [yp], [ym], [kp], [km], [Rt]])

>>> Building the observability-identifiability matrix requires at least 7 Lie derivatives
    Calculating derivatives: 1 2 3 4 5 6 7
>>> Observability-Identifiability matrix built with 7 Lie derivatives
    (calculated in 2.851094e+00 seconds)
>>> Calculating rank...
    Rank = 7 (calculated in 1.101564e+01 seconds)
>>> Observability-Identifiability matrix built with 8 Lie derivatives
    (calculated in 1.140776e+01 seconds)
>>> Calculating rank...
    Rank = 7 (calculated in 6.496877e+01 seconds)
   The model is structurally unidentifiable as a whole
   => Parameter a is structurally unidentifiable
   => Parameter yp is structurally unidentifiable
   => Parameter ym is structurally identifiable
   => Parameter kp is structurally identifiable
   => Parameter km is structurally identifiable
   => Parameter Rt is structurally unidentifiable

------------------------
>>> RESULTS SUMMARY:
------------------------

>>> The model is structurally unidentifiable.
>>> These parameters are identifiable:
     matrix([[km, kp, ym]])
>>> These parameters are unidentifiable:
     matrix([[Rt, a, yp]])
```

Figure 5.3.5: The STRIKE-GOLDD toolbox [129] confirms the identifiability of the parameters k_+, k_- and ψ_-.

results of identifiable parameters but fails to return any parameter combinations.

5.4 Identifiability with equilibrium, washout and multiple time courses

The results thus far have shown none of the models to be fully identifiable from a single set of time course data. In this section we consider alternative ways in which full identifiability can be established. As we discussed in Chapter 1, commonly performed experiments for drug binding are equilibrium (or saturation) experiments. In these, for each ligand concentration, the experiments are run until equilibrium is reached, then the amount of ligand bound is observed. These experiments are often used to estimate equilibrium constants $K_D = 1/K_A$ and R_{tot}, for monomeric receptors [102]. For each model, we aim to identify the model equilibrium parameters and then using time course data to identify the remaining binding parameters.

Washout experiments can also be used to gain further insights into the binding kinetics of ligands and provide identifiability information. In these experiments the free ligand is removed by repeated washing, ensuring that no further drug associates with the receptors and the rate of decrease of effect is then observed. We use this type of experimental data do determine the identifiability of the dissociation parameters before then using association time course data to also determine identifiability of association parameters.

Finally, we also consider multiple binding experiments, whereby each data set is collected from experiments performed using different ligand concentrations. These data sets are then used simultaneously, with a view to determine the minimum number of data sets required to make the model globally identifiable. In each case, we choose the appropriate identifiability method to apply.

5.4.1 Monomer receptor binding with a single drug

Recall the model for a monomer binding a single drug, as given by the schematic

$$\mathrm{A} + \mathrm{R} \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} \mathrm{AR}$$

Although we previously found no identifiable parameters, we have identifiable parameter combinations as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{a+}[A] + k_{a-} \end{bmatrix}. \tag{5.167}$$

Saturation curves

We first attempt to identify equilibrium parameters. Note that, at equilibrium, we have

$$[AR] = K_A[A][R], \tag{5.168}$$

where $K_A = k_{a+}/k_{a-}$. This gives the expression for the concentration of ligand bound at equilibrium as

$$A_{bound} = \frac{K_A[A]R_{tot}}{1 + K_A[A]}. \tag{5.169}$$

Taking two ligand concentrations, $[A]_1$ and $[A]_2$, and the corresponding output measurements, x_1 and x_2 gives

$$\frac{K_A[A]_1R_{tot}}{1 + K_A[A]_1} = x_1 \qquad \text{and} \qquad \frac{K_A[A]_2R_{tot}}{1 + K_A[A]_2} = x_2. \tag{5.170}$$

Solving these for R_{tot} and $K_A = k_{a+}/k_{a-}$ shows there is the single solutions

$$R_{tot} = \frac{x_1 x_2([A]_1 - [A]_2)}{[A]_1 x_2 - [A]_2 x_1} \qquad \text{and} \qquad K_A = \frac{[A]_1 x_2 - [A]_2, x_1}{[A]_1[A]_2(x_1 - x_2)} \tag{5.171}$$

and hence these are identifiable from a single dose response curve (in fact, only two points from the curve). With these known, only a single parameter remains to be found from time course data. This can be achieved using only one of the parameter combinations as stated in equation (5.167), giving

$$k_{a-} K_A [A] R_{tot} = \widetilde{k_{a-}} K_A [A] R_{tot} \qquad \Rightarrow \qquad k_{a-} = \widetilde{k_{a-}}. \tag{5.172}$$

Hence, we conclude that, using equilibrium data together with a single set of time course data, it is possible to identify all model parameters.

Washout experiments

We also consider using washout experiment data to identify dissociation parameters. In a washout experiment, the drug is removed from the system (usually once equilibrium has been reached), hence, we set $[A] = 0$ in the model given in equations (5.23). This gives

$$[R]' = k_{a-}[AR], \tag{5.173a}$$

$$[AR]' = -k_{a-}[AR]. \tag{5.173b}$$

We note that, it is clearly possible to solve the ODE for $[AR]$, so it is possible to determine identifiability directly from this. However, we will continue to use a method that is applicable to any system in order to highlight the process.

As the concentrations of each species is unknown at the start point of washout, the

initial conditions are therefore

$$[R](0) = R_{tot} - x_1, \qquad [AR](0) = x_1, \tag{5.173c}$$

while the output remains unchanged

$$y = [AR]. \tag{5.173d}$$

The unknown parameters in the model in this form are k_{a-}, R_{tot} and x_1, although as we are only interested in identifying the parameters in the overall system (as stated in equation 5.23), it is not necessary to determine identifibility of x_1. Reducing the system using conservation still leads to a zero force term, so we instead use the Taylor series method. We have $n_x = 2$ so calculate a maximum of three Taylor coefficients. Calculating the first of these coefficient gives

$$y = AR, \qquad \Rightarrow \qquad y_0 = x_1, \tag{5.174}$$

which clearly gives no information about k_{a-}, but does (unnecessarily in this case) show x_1 to be identifiable. The second coefficient is

$$y^{(1)} = AR' = -k_{a-}[AR], \qquad \Rightarrow \qquad y_0^{(1)} = a_1 = -k_{a-}x_1, \tag{5.175}$$

and so we find k_{a-} to be identifiable. The final coefficient is calculated as

$$y_0^{(2)} = a_2 = k_{a-}^2 x_1, \tag{5.176}$$

and so gives no further information, hence, only k_{a-} is identifiable from washout, dissociation data.

Returning to the association coefficients, as k_{a-} is now known we have $\mathbf{p} = (k_{a+}, R_{tot})$ as the remaining unknown parameters. Taking the coefficients in equation (5.167) and solving $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ for $\mathbf{p}$ confirms these parameters as also now identifiable.

Multiple time courses

We consider multiple sets of time course data, each with a different drug concentration. Each of these will give the identifiable parameters, as stated in equation (5.50) for their corresponding concentration of $[A]$. That is

$$\boldsymbol{\zeta}(\mathbf{p})_i = \begin{bmatrix} k_{a+}[A]_i R_{tot} \\ k_{a+}[A]_i + k_{a-} \end{bmatrix}, \tag{5.177}$$

for $i = 1, 2, ...k$, where k is the number of time courses being considered. As fitting is performed to all sets simultaneously, we solve the system as one also. For example, for two time courses we have

$$\boldsymbol{\zeta}(\mathbf{p})_2 = \begin{bmatrix} k_{a+}[A]_1 R_{tot} \\ k_{a+}[A]_1 + k_{a-} \\ k_{a+}[A]_2 R_{tot} \\ k_{a+}[A]_2 + k_{a-} \end{bmatrix}. \tag{5.178}$$

Setting $\boldsymbol{\zeta}(\mathbf{p})_2 = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})_2$ results in all parameters being successfully identified. Hence we conclude that the model is fully identifiable from just two time courses.

5.4.2 Monomeric receptor two drug competition binding

The model for a monomer binding with two ligands in a competition binding scenario is described by the schematic

$$\mathrm{A} + \mathrm{R} \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} \mathrm{AR} \qquad \mathrm{B} + \mathrm{R} \underset{k_{b-}}{\overset{k_{b+}}{\rightleftharpoons}} \mathrm{BR}$$

Identifiability analysis from a single time course resulted in the parameter k_{b-} being successfully identifiable, along with parameter combinations

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}R_{tot}, \\ k_{a+}[A] + k_{b+}[B] + k_{a-}, \\ k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B] + k_{a-}k_{b-}. \end{bmatrix} \tag{5.179}$$

Saturation curves

At equilibrium, we have the relations

$$[AR] = K_A[A][R], \qquad [BR] = K_B[B][R], \tag{5.180}$$

where $K_A = k_{a+}/k_{a-}$ and $K_B = k_{b+}/k_{b-}$. Taking this, together with the conservation law

$$[R] + [AR] + [BR] = R_{tot}, \tag{5.181}$$

we can state the equilibrium concentration of the measured ligand bound ($[AR]$) as

$$[A]_{bound} = \frac{K_A[A]R_{tot}}{1 + K_A[A] + K_B[B]}. \tag{5.182}$$

Taking two points from the dose response curve (with $[A]$ being the varied parameter and a single dose of $[B]$), we have

$$\frac{K_A[A]_i R_{tot}}{1 + K_A[A]_i + K_B[B]} = x_i, \tag{5.183}$$

for $i = 1, 2$, and solving for K_A and R_{tot} gives the solutions

$$K_A = \frac{(1 + K_B[B])([A]_1 x_2 - [A]_2 x_1)}{[A]_1[A]_2(x_1 - x_2)}, \qquad R_{tot} = \frac{x_1 x_2([A]_1 - [A]_2)}{[A]_1 x_2 - [A]_2 x_1}. \tag{5.184}$$

Considering further points on the curve gives no extra information, hence from a single dose response curve the only identifiable parameter is R_{tot}, as the solution for K_A still depends the unknown parameter K_B.

We then look again to time course data to identify remaining parameters. Assuming R_{tot} is known, we now use equation (5.179) and set $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$. From this we find the two solution sets

$$\begin{pmatrix} k_{a+} \\ k_{a-} \\ k_{b+} \\ k_{b-} \end{pmatrix} = \begin{pmatrix} \widetilde{k_{a+}} \\ \widetilde{k_{b+}}[B] + \widetilde{k_{b-}} \\ \frac{\widetilde{k_{a-}} - \widetilde{k_{b-}}}{[B]} \\ \widetilde{k_{b-}} \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} k_{a+} \\ k_{a-} \\ k_{b+} \\ k_{b-} \end{pmatrix} = \begin{pmatrix} \widetilde{k_{a+}} \\ \widetilde{k_{a-}} \\ \widetilde{k_{b+}} \\ \widetilde{k_{b-}} \end{pmatrix}, \tag{5.185}$$

and so we find that only k_{a+} and k_{b-} are identifiable. As we have two possible solutions for k_{a-} and k_{b+} these can be described as locally identifiable and as such there is no way of knowing whether the estimated parameters are accurate or not. However, if $\widetilde{k_{b-}} > \widetilde{k_{a-}}$, then the first solution set contains a negative value which makes this solution set invalid, and therefore, all the parameters are identifiable.

We also consider using a second dose response curve, with an alternative concentration

of the competitor drug B (as in the work done by Motulsky and Mahan [95]), giving the equations to be solved as

$$\frac{K_A[A]_1R_{tot}}{1+K_A[A]_1+K_B[B]_1}=x_1,\quad \frac{K_A[A]_2R_{tot}}{1+K_A[A]_2+K_B[B]_1}=x_2,\quad \frac{K_A[A]_1R_{tot}}{1+K_A[A]_1+K_B[B]_2}=x_3. \tag{5.186}$$

These give solutions

$$\begin{aligned}
R_{tot} &= \frac{x_1x_2([A]_1-[A]_2)}{[A]_1x_2-[A]_2x_1},\\
K_A &= \frac{x_3([B]_1-[B]_2)([A]_1x_2-[A]_2x_1)}{[A]_1([A]_1[B]_1x_2(x_1-x_3)-[A]_2[B]_1x_1(x_2-x_3)-[A]_2[B]_2x_3(x_1-x_2))},\\
K_B &= \frac{x_3([A]_1-[A]_2)(x_1-x_3)}{[A]_1[B]_1x_2(x_1-x_3)-[A]_2[B]_1x_1(x_2-x_3)-[A]_2[B]_2x_3(x_1-x_2)},
\end{aligned} \tag{5.187}$$

and so all equilibrium parameters are determined to be identifiable. Remaining parameters are then all identifiable from time course data, hence, the model is fully identifiable.

Washout experiments

To determine identifiability using washout experiment data we set $[A]=0$ in equation 5.45 to give the model with only dissociation terms as

$$[R]' = k_{a-}[AR]+k_{b-}[BR], \tag{5.188a}$$

$$[AR]' = -k_{a-}[AR], \tag{5.188b}$$

$$[BR]' = -k_{b-}[BR], \tag{5.188c}$$

initial conditions are unknown so we have

$$[R](0)=R_{tot}-x_1-x_2,\qquad [AR](0)=x_1,\qquad [BR](0)=x_2. \tag{5.188d}$$

The only measured quantity is $[AR]$ and so the output is

$$y = [AR]. \tag{5.188e}$$

We have unknown parameters $\mathbf{p} = (k_{a-}, k_{b-}, R_{tot}, x_1, x_2)$, although we note that we will be unable to identify R_{tot} as $[R]$ does not appear in any of the model equations, and it is not necessary to find estimates for x_1 and x_2 for the overall system to be identifiable. We again use the Taylor series method as reducing the system results in a zero force term (that is $G = 0$). We calculate the first Taylor coefficient as

$$y = AR, \qquad \Rightarrow \qquad y_0 = x_1, \tag{5.189}$$

and so x_1 is identifiable. The second coefficient is calculated as

$$y' = AR' = -k_{a-}[AR], \qquad \Rightarrow \qquad y_0' = -k_{a-}x_1, \tag{5.190}$$

giving k_{a-} as also being identifiable. Further coefficients give no new information, and so only k_{a-} is identifiable from dissociation data. This leaves parameters $\mathbf{p} = (k_{a+}, k_{b+}, k_{b-}, R_{tot})$ to be identified from association data. Using the coefficients in equation (5.179), we find all remaining parameters to be identifiable.

Multiple time courses

There are two ways in which we can consider using multiple time course data to determine identifiability of this model. Taking the coefficients as stated in equation (5.179) we can look at time courses with either multiple concentrations of A or multiple concentrations

of B, and so giving either

$$\boldsymbol{\zeta}(\mathbf{p})_i = \begin{bmatrix} k_{a+}R_{tot} \\ k_{a+}[A]_i + k_{b+}[B] + k_{a-} \\ k_{a+}k_{b-}[A]_i + k_{a-}k_{b+}[B] + k_{a-}k_{b-} \end{bmatrix} \text{ or } \quad \boldsymbol{\zeta}(\mathbf{p})_i = \begin{bmatrix} k_{a+}R_{tot} \\ k_{a+}[A] + k_{b+}[B]_i + k_{a-} \\ k_{a+}k_{b-}[A] + k_{a-}k_{b+}[B]_i + k_{a-}k_{b-} \end{bmatrix} \tag{5.191}$$

for $i = 1, 2....$ If we first look to the first case, having multiple A concentrations, and solve $\boldsymbol{\zeta}(\mathbf{p})_i = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})_i$, we have

$$\begin{pmatrix} k_{a+} \\ k_{a-} \\ k_{b+} \\ k_{b-} \end{pmatrix} = \begin{pmatrix} \widetilde{k_{a+}} \\ \widetilde{k_{b+}}[B] + \widetilde{k_{b-}} \\ \frac{\widetilde{k_{a-}} - \widetilde{k_{b-}}}{[B]} \\ \widetilde{k_{b-}} \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} k_{a+} \\ k_{a-} \\ k_{b+} \\ k_{b-} \end{pmatrix} = \begin{pmatrix} \widetilde{k_{a+}} \\ \widetilde{k_{a-}} \\ \widetilde{k_{b+}} \\ \widetilde{k_{b-}} \end{pmatrix}, \tag{5.192}$$

from two sets of time course data, and so we once again have k_{a+}, k_{b-} and R_{tot} identifiable, but two possible solutions for k_{a-} and k_{b+}, unless $\widetilde{k_{b-}} > \widetilde{k_{a-}}$. Further data sets with a third concentration of A yields no further information. If we instead consider having time courses for multiple $[B]$, solving $\boldsymbol{\zeta}(\mathbf{p})_i = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})_i$ gives all parameters as being identifiable from just two time courses.

5.4.3 Pre-formed homodimer binding with a single drug

The schematic for the dimer binding with a single drug model is given by

$$\mathrm{A} + \mathrm{R} \underset{k_{a-}}{\overset{k_{a+}}{\rightleftharpoons}} \mathrm{AR} \qquad \mathrm{A} + \mathrm{AR} \underset{\alpha_- k_{a-}}{\overset{\alpha_+ k_{a+}}{\rightleftharpoons}} \mathrm{ARA}$$

and we saw in the previous section (section 5.3.3) that we have no identifiable parameters, however, we do have the following identifiable parameter combinations

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} k_{a+}[A]R_{tot} \\ k_{a+}[A]R_{tot}(2\alpha_+ k_{a+}[A] + \alpha_- k_{a-}) \\ k_{a+}[A] + k_{a-} + \alpha_+ k_{a+}[A] + \alpha_- k_{a-} \\ \alpha_+ k_{a+}^2[A]^2 + \alpha_- k_{a+}k_{a-}[A] + \alpha_- k_{a-}^2 \end{bmatrix}. \tag{5.193}$$

Saturation curves

We take the expression for the concentration of ligand bound at equilibrium, as stated in Chapter 2 (equation (2.9)), as

$$A_{bound} = \frac{(K_A[A] + 2\alpha K_A^2[A]^2)}{1 + K_A[A] + \alpha K_A^2[A]^2} R_{tot}. \tag{5.194}$$

where $K_A = k_{a+}/k_{a-}$ and $\alpha = \alpha_+/\alpha_-$. Taking three drug concentrations $[A]_i$ and the corresponding concentration of ligand bound at equilibrium as x_i, we solve

$$\frac{(K_A[A]_i + 2\alpha K_A^2[A]_i^2)}{1 + K_A[A]_i + \alpha K_A^2[A]_i^2} R_{tot} = x_i, \qquad i = 1, 2, 3, \tag{5.195}$$

for $\mathbf{p} = (K_A, \alpha, R_{tot})$. This is simple to solve in a symbolic equation solver, using software such as Matlab [2] or Mathematica [58], where we find that all three equilibrium parameters are identifiable from a single dose-response curve. However, the expressions are extremely lengthy and so impractical to write down, and so we refrain from doing so here. The dynamic parameters can be then identified from a single time course, and can be seen by solving $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$, where $\boldsymbol{\zeta}(\mathbf{p})$ is defined as in equation (5.193).

Washout experiments

Again we consider washout experimental data to aid in identifiability, which is represented by setting $[A] = 0$ in equations (5.90), giving

$$\frac{d[R]}{dt} = k_{a-}[AR], \tag{5.196a}$$
$$\frac{d[AR]}{dt} = -k_{a-}[AR] + \alpha_- k_{a-}[ARA], \tag{5.196b}$$
$$\frac{d[ARA]}{dt} = -\alpha_- k_{a-}[ARA], \tag{5.196c}$$

The initial conditions for each species are now unknown and will instead have to be determined through fitting. Also using conservation of receptors, we have the initial conditions as

$$[R](0) = R_{tot} - x_1 - x_2, \qquad [AR](0) = x_1, \qquad [ARA](0) = x_2, \tag{5.196d}$$

while the output remains as

$$y = [AR] + 2[ARA] \tag{5.196e}$$

We have $\mathbf{p} = (\alpha_-, k_{a-}, R_{tot}, x_1, x_2)$ as the vector of unknown parameters, though again it is not necessary to identify x_1 and x_2 for the system to be identifiable. We again use the Taylor series method to determine identifiability in this section. Through repeated differentiation of y and substitution of the initial conditions we obtain the vector of

coefficients as

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} x_1 + 2x_2, \\ -k_{a-}(x_1 + \alpha_- x_2), \\ k_{a-}^2(x_1 + (\alpha_- - 1)\alpha_- x_2), \\ -k_{a-}^3(x_1 + (\alpha_-^2 - \alpha_- - 1)\alpha_- x_2), \end{bmatrix} \tag{5.197}$$

Setting $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ and solving for $\mathbf{p} = (k_{a-}, \alpha_-, x_1, x_2)$ gives the two sets of solutions

$$\begin{bmatrix} k_{a-} \\ \alpha_- \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \widetilde{k_{a-}} \\ \widetilde{\alpha_-} \\ \widetilde{x_1} \\ \widetilde{x_2} \end{bmatrix}, \qquad \text{and} \qquad \begin{bmatrix} k_{a-} \\ \alpha_- \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \widetilde{\alpha_-}\widetilde{k_{a-}} \\ \dfrac{1}{\widetilde{\alpha_-}} \\ \dfrac{\widetilde{x_1} + 2(\widetilde{\alpha_-} - 1)\widetilde{x_2}}{2\widetilde{\alpha_-} - 1} \\ \dfrac{(\widetilde{\alpha_-} - 1)\widetilde{x_1} + \widetilde{\alpha_-}\widetilde{x_2}}{2\widetilde{\alpha_-} - 1} \end{bmatrix}, \tag{5.198}$$

and so we find that all parameters are locally identifiable. In terms of fitting, this means estimates will tend to one of these solution sets. However, when running fitting we impose the restriction that all parameters can only take positive values. Hence with this we find that the second solution set can only be obtained if

$$\frac{\widetilde{x_1} + 2(\widetilde{\alpha_-} - 1)\widetilde{x_2}}{2\widetilde{\alpha_-} - 1} > 0, \quad \text{and} \quad \frac{(\widetilde{\alpha_-} - 1)\widetilde{x_1} + \widetilde{\alpha_-}\widetilde{x_2}}{2\widetilde{\alpha_-} - 1} > 0. \tag{5.199}$$

If the estimated parameters do not fall within these ranges then the only solution is that the parameters are identifiable and as such the estimates are accurate for all parameters. We use time course data to then identify the remaining parameters. Setting $\boldsymbol{\zeta}(\mathbf{p}) = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})$ in equations (5.95) and solving for $\mathbf{p} = (k_{a+}, \alpha_+, R_{tot})$ confirms these as identifiable.

Multiple experiments

To determine the minimum number of concentrations needed to ensure full identifiability, we assume multiple experiments, each with a different ligand concentration. Each of these have identifiable parameter combinations, as defined in (5.193), with $[A]_i$ as the concentration for experiment y_i. This gives the identifiable parameter combinations as

$$\boldsymbol{\zeta}(\mathbf{p})_i = \begin{bmatrix} k_{a+}[A]_i R_{tot} \\ k_{a+}[A]_i R_{tot}(2\alpha_+ k_{a+}[A]_i + \alpha_- k_{a-}) \\ k_{a+}[A]_i + k_{a-} + \alpha_+ k_{a+}[A]_i + \alpha_- k_{a-} \\ \alpha_+ k_{a+}^2 [A]_i^2 + \alpha_- k_{a+} k_{a-} [A]_i + \alpha_- k_{a-}^2 \end{bmatrix}. \tag{5.200}$$

To determine identifiability we again set $\boldsymbol{\zeta}(\mathbf{p})_i = \boldsymbol{\zeta}(\widetilde{\mathbf{p}})_i$, however, when solving we solve the systems for $i = 1, 2...$ simultaneously. From this we can conclude that only two data sets are required to ensure all parameters are identifiable.

5.4.4 Ligand induced dimerisation with a single drug

In section 5.3.4 we saw that, from a single time course, the parameters

$$k_+, \quad k_-, \quad \psi_-, \tag{5.201}$$

are identifiable. Furthermore, we have identifiable parameter combinations

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} aR_{tot} \\ \psi_+ R_{tot} \end{bmatrix}. \tag{5.202}$$

In Chapter 4, we used this model to fit to experimental data and estimate the model parameters. In this, we had three sets of experiments, each with a different VEGF

isoform. Within each of these sets, the experiments are repeated with five concentrations of VEGF. The quantity R_{tot} and scaler a are the same in all experiment sets (as the cell membranes and measuring techniques are the same for each), while all other parameters are unique. If we consider $\boldsymbol{\zeta}(\mathbf{p})$ as above, we note that $[A]$ does not appear in any of the identifiable combinations, hence, multiple concentrations of drug for the same experiment does not make further any parameters identifiable. However, as ψ_- is different for each VEGF isoform, with two sets of experiments we have

$$\boldsymbol{\zeta}(\mathbf{p}) = \begin{bmatrix} aR_{tot} \\ \psi_+ R_{tot} \\ \widehat{\psi_+} R_{tot} \end{bmatrix}, \tag{5.203}$$

where $\widehat{\psi_+}$ indicates that it is a different value to ψ_+. In this, we see that there are no further identifiable parameters. This will also be the case for more than two sets of experiments. Thus we can conclude that, although the model gives a good fit to the data, the parameters a, R_{tot} and ψ_+ are not identifiable and cannot be taken as accurate estimates. For the model to be fully identifiable, prior knowledge of either a or R_{tot} is needed. The consequence of these parameters not being identifiable is that we cannot evaluate the cooperativity.

5.5 Conclusions

In this chapter, we have used well-established SIA techniques to determine model identifiability for a number of standard ligand-receptor binding models, followed then by our own dimer models. The models we consider are: monomer binding with a single ligand, monomer competition binding, GPCR dimer binding and VEGF dimer binding. Analysis such as this is not typically undertaken for pharmacodynamics models, and parameter estimates obtained from fitting to real, experimental data are accepted without knowledge of their accuracy or uniqueness. As a consequence, incorrect characteristics of ligands are assumed. Furthermore, when we consider that data is rarely noise-free, practical identi-

fiability issues compound these inaccuracies which can lead to unsuitable models being used and further experiment designs may be flawed. As such, SIA should be undertaken at an early stage. To aid with this, we have taken a tutorial style approach to approaching the SIA for each of the models.

We used three methods to determine identifiability for each model (where applicable), namely, the transfer function method based on the Laplace transform, the Taylor series method and the similarity tranform method. In each case, we demonstrate the necessary steps taken to perform identifiability, for each method. We compare the methods, highlighting the method that is best suited to each model. In this, we find that the transfer function is best for linear models due to its simplicity. The Taylor series and similarity transformation methods, while give the same end results, require more calculation and are more algebraically complex. For nonlinear problems we again find that the Taylor series method results in expressions that are too lengthy and complex to be practically useful. We instead use the similarity transformation method to determine identifiability for the nonlinear, VEGF model. For all models, we find them to be unidentifiable overall, though we point out that the VEGF model would be identifiable if the output scaler is known. In each case, we highlight which parameters are identifiable and also give identifiable parameter combinations.

Following this we consider the possibilities for making the models fully identifiable. For this we also look at dose-response data, washout experiments and multiple sets of time course data. For the monomer, single ligand binding model, we find that the model is fully identifiable from either dose-response or washout data, together with a single set of time course data, or just two sets of time course data. For the monomer competition binding model, we find that using dose response data, with a single concentration of the unlabelled ligand, together with the time course data gives the system as locally identifiable. That is, we have two possible solutions, either the estimates will be accurate (theoretically), or they will tend to a second solution. To ensure full identifiability, a second dose-response curve, with a different concentration of unlabelled ligand can be used. Using washout data will provide full identifiability also. When using multiple time

courses, only local identifiability can be established when using differing concentrations of labelled ligand, regardless of how many time courses are used, however, the system is fully identifiable when the time courses have different concentration of the unlabelled ligand for only two data sets. For the GPCR dimer binding model, full identifiability can be obtained from using dose-response data, or from two sets of time course data, however, only local identifiability can be established from washout data.

Chapter 6

Conclusions and future work

6.1 Results summary

In this thesis, we have developed a number of models that describe ligand binding and receptor activation to dimerised and dimerising. The central theme of the thesis is *co-operativity* across the dimer, which we focus on analysing the effects of, for each model. We have used a range of mathematical methods and techniques to analyse the different problems. Here we detail the main results of the thesis.

In Chapter 2, we develop models for ligand binding to pre-dimerised GPCRs. We consider both a single ligand and two ligand competition binding to a homodimer and also a single ligand binding to a heterodimer. In these, we study the effects of binding cooperativity, which is defined as the effect one protomer being bound by a ligand has on the binding of a second ligand molecule to the other protomer in the dimer. For single ligand models cooperativity is a single parameter, while the competition binding model requires three cooperativity factors to fully describe all ligand-receptor interactions. In each case, we first look at equilibrium results where we see extra inflections in the logDR curves. LogDR curves for monomeric receptors are typically Hill curves and have a sigmoidal shape. We find that, biphasic curves such as the ones seen in Chapter 2 are indicative of dimerised receptors and were seen for all models presented. For each model, we use analytical expressions to analyse the curves and give conditions upon the parameters for exactly when the extra inflections occur. For the single ligand homodimer

binding model, this condition depends only on the cooperativity factor, where we find that the extra inflections appear when cooperativity is particularly low. Also in Chapter 2, explore time course binding. For the single ligand homodimer binding model we give analytical solutions to the ODE system. These are of practical use to pharmacologists due to the ease at which they can be implemented and interpreted.

In Chapter 3, we see how extending the GPCR dimer model to include receptor activation leads to a wide range of equilibrium signal logDR curves (we note that, unlike the other chapters, we only look at equilibrium results). There are four cooperativity factors to explore. The parameter α represents ligand binding cooperativity, and is defined the same as in Chapter 2. We also have activation cooperativity factor λ, which describes the effect one protomer being active has on the activation of the other protomer. Finally, ν and ξ are efficacy parameters and represent the change in propensity for activation when ligand bound and the change in affinity for active protomers over inactive protomers. Specifically, ν accounts for when binding and activation occur on the same protomer, while ξ accounts for when these occur on opposite protomers. A neutral curve (where all cooperativity factors are equal to one) shows a small peak in an otherwise constant signal. Varying each cooperativity factor individually, shows how each alters the curve from the neutral. Overall, we have three potential effects. Varying α alters the peak height of the curve. We also see how an increased peak relates to the extra inflections in the logDR binding curves. Increasing λ causes and upward shift in the whole curve, although there is little change for lower values of λ. Varying either ν or ξ individually cause the curves to show agonist or antagonist behaviour. When varying multiple cooperativity factors these effects combine and can result in signal curves that show a biphasic shape (symmetrical as in Chapter 2 or non-symmetrical) or curves that overshoot the saturation signal. This work is still ongoing and requires both further exploration into possible equilibrium effects as well as time course dynamics.

In Chapter 4 we follow the framework that we outlined in Chapter 2, by first analysing equilibrium logDR curves and then considering time course dynamics, but we instead apply it to the growth factor, VEGF system (although the model can be applied to most

growth factor systems). As the growth factor family is known to dimerise in response to ligand binding, we develop a ligand induced model. This results in a nonlinear system which is a key difference from all previous models. We again give analytical expressions for equilibrium concentrations, and use these to explain the shifts in the signal curve, although an important point to note is that there are no extra inflections in this case. When investigating the time course dynamics of the system, we notice some interesting behaviours such as peaks in the curves and curves which evolve on multiple time scales. Using asymptotic analysis we find that, when dissociation cooperativity is low, that is a slow dimer dissociation, we have peaks in the concentrations of bound monomers. Conversely, when this is instead high, we have multiple time scales contributing the concentrations of bound dimers. If instead, we have either high forward cooperativity or a high receptor to ligand ratio, we observe multiple time scales in the bound monomer curves. The model is validated by fitting to published data, where we see an excellent fit across multiple time courses, when fitting to datasets with five concentrations of three VEGF isoforms simultaneously.

In Chapter 5 we use structural identifiability analysis to determine whether parameters returned from data fitting can be relied upon to be accurate. This is something that is often overlooked in systems biology. Thus, we present this chapter as a tutorial for determining identifiability to some classical monomeric binding models, before then analysing the dimer binding models presented throughout this thesis. When applying the techniques, we find that, using only a single set of time course data results in a non-identifiable model, for all models. We are, however, able to find parameter combinations that can be identified successfully. For linear models, we apply three methods, using transfer functions, Taylor series and similarity transformations. We find that all methods give the same results, though there are advantages and disadvantages to each. For nonlinear problems both the Taylor series method, and similarity transformation methods are applicable, however, we show how, even for a small problem, the Taylor series method is problematic. An important point to note on this problem, is that, the model is only unidentifiable as the measured output is scaled. If this scalar is known then the model is fully identifiable. We also consider ways in which the models can be made fully

identifiable. We use equilibrium data or washout experiment data together with the set of time course data to ensure identifiability, as well as multiple sets of time course data.

6.2 Future work

6.2.1 G proteins

In Chapter 3, we highlighted the importance of looking beyond ligand binding to dimerised receptors, and considering the signalling implications of including the full G protein cycle. We took first steps towards this goal by developing a model that included receptor activation. So far, we looked at the equilibrium effects of binding and activation. However, we are yet to explore the possible time course dynamics of the system. As time course data is becoming more readily available, considering these effects is the logical next thing to examine.

Following this, extending the model to include G protein binding, and eventually downstream pathways, is necessary and essential to fully understand the signalling implications for drug discovery and development. To achieve this, a second layer would be added to the current, spherical schematic. The current schematic would sit inside, and link to, a larger sphere, that contains reactions involving the attachment of G proteins. This will result in a large model involving many species and reactions, however, would provide a full insight into the consequences of dimerisation for GPCRs.

6.2.2 Receptor internalisation

Another potential area way in which the models we have developed thus far, would be to include receptor internalisation (or endocytosis), which plays a key role in cell signaling. Once a receptor becomes activated, and a signal has been transmitted, the receptor moves from the plasma membrane towards the endosomal compartment where they are recycled to the cell membrane or degraded [16, 114]. This process is known as receptor internalisation. Initially it was thought that the primary implication of receptor internalisation was signal desensitisation, that is, the decreased responsiveness to a ligand

[16]. However, while the full implications for internalisation remain unknown, recent research is showing that an internalised receptor can continue to signal [18, 114] from the endosome. Furthermore, a signal initiated from the cell surface may produce a different signal from an intracellular one [17, 18, 16, 114]. The ligand induced model in-particular would benefit from including receptor internalisation. In the work by Peach *et al* [101], it was found that internalisation has an effect on the binding of VEGF on longer time scales.

6.2.3 Allosteric modulators

While GPCRs are targets for approximately 50% of all current drugs [13], only a small fraction of these receptors have been exploited [121]. One reason for this seems to be an issue with traditional agonists and antagonists, that bind to the orthosteric site, producing unwanted side effects [98]. The configuration of the orthosteric site is similar for several receptors, which lowers the specificity of the drug for the required receptor [50]. This means ligands must have a high affinity for the receptor, in order to produce an effect with minimal side effects. Allosteric ligands bind at a site other than the orthosteric site and can modulate a receptor's signalling pathways, and the ligands' effects [84]. These ligands work by causing a conformation change in the receptor, and often a change in the orthosteric site. This in turn may cause an increase in efficacy of the ligand bound to the orthosteric site, thus reducing the required dose, and in turn, the level of side effects.

A receptor can (and often does) have multiple allosteric sites, hence, in developing new drugs, it is possible to find a site that it more unique, causing the ligand to act with an increased specificity, thus making it more safe [50]. Furthermore, many orthosteric drugs work as antagonists, blocking the natural ligand from binding. Allosteric modulators can reduce the activity of a ligand without blocking it entirely, and so is less disruptive [98]. Although there is already a rise in drugs targeting the allosteric sites of GPCRs [50], models are needed to fully understand and exploit the potential benefits of these ligands.

Appendix A

Parameter values

In this section, we give details of the parameters used for plots in Chapters 2 - 4. There are three sections that make up Chapter 2 (with a model in each section), and we give the parameters for each of these sections individually, although we note that some parameters appear in all of the models. In Table A.1, we have parameters for the single ligand, GPCR binding model given in Section 2.1, while parameters for the competition binding model presented in Section 2.2 are given in Table A.2, and the heterodimer binding model, as given in Section 2.3, are detailed in Table A.3. All parameter values in this chapter are taken from [83] and [138], and references within.

Parameter	R_{tot}	k_{a+}	k_{a-}	K_A
Value	$10^{-10}M$	$10^7M^{-1}s^{-1}$	$0.1s^{-1}$	10^8M^{-1}

Table A.1: Values used for the single ligand, GPCR homodimer binding model, as presented in Section 2.1.

Parameter	R_{tot}	k_{a+}	k_{a-}	K_A	k_{b+}	k_{b-}	K_B
Value	$10^{-10}M$	$10^7M^{-1}s^{-1}$	$0.1s^{-1}$	10^8M^{-1}	$10^6M^{-1}s^{-1}$	$0.5s^{-1}$	10^7M^{-1}

Table A.2: Values used for the two ligand, GPCR homodimer, competition binding model, as presented in Section 2.2.

In Table A.4 we have the parameter values used for all plots in Chapter 3, describing the binding and activation of a single ligand with GPCRs. These values are taken from

Parameter	R_{tot}	k_{a_1+}	k_{a_1-}	K_{A1}	k_{a_2+}	k_{a_2-}	K_{A2}
Value	$10^{-10}M$	$10^{7}M^{-1}s^{-1}$	$0.1s^{-1}$	$10^{8}M^{-1}$	$10^{6}M^{-1}s^{-1}$	$0.001s^{-1}$	$10^{6}M^{-1}$

Table A.3: Values used for the single ligand, GPCR heterodimer binding model, as presented in Section 2.3.

[137] and [138]. We note that, analysis was undertaken for the model at equilibrium only, hence we have only equilibrium parameters.

Parameter	R_{tot}	K_A	K_{act}
Value	$10^{-10}M$	$2.3 \times 10^{8}M^{-1}$	10^{-3}

Table A.4: Values used for the single ligand, GPCR homodimer, binding and activation model, as presented in Chapter 3.

Finally, in Table A.5, we give figures used in the VEGF-VEGFR dimer binding model presented in Chapter 4. The parameter values in this chapter are taken from [80] (and references within), though we note that we have taken a slightly higher receptor total concentration for simplicity and continuity in the analysis. The value taken still falls into reported ranges.

Parameter	R_{tot}	k_+	k_-	K_A
Value	$2 \times 10^{-10}M$	$4.4 \times 10^{7}M^{-1}s^{-1}$	$1.32 \times 10^{-3}s^{-1}$	$3.3 \times 10^{10}M^{-1}$

Table A.5: Values used for the single ligand, VEGFR homodimer binding model, as presented in Chapter 4.

Appendix B

Inflections in the GPCR logDR curves

In Chapter 2 we saw how extra inflections appear in many of the logDR curves. In this appendix, we investigate these inflections and derive conditions upon when they appear. We begin by stating a general form, that applies to all models presented, of the concentration of ligand bound as

$$A_{bound} = \frac{aX + 2bX^2}{c + aX + bX^2} R_{tot}. \tag{B.1}$$

Calculating the second differential, with respect to $\log_{10} X$, gives

$$\frac{d^2 A_{bound}}{d \log_{10} X^2} = \frac{X(\log 10)^2 (c - bX^2)(abX^2 + (8bc - a^2)X + ac)}{(c + aX + bX^2)^3} R_{tot} \tag{B.2}$$

To find the inflection points, we set this equal to zero and solve for X. Clearly we have the trivial root at $X = 0$ as well as a root when

$$c - bX^2 = 0. \tag{B.3}$$

Discarding the negative root, we conclude that we have a possible inflection at

$$X_c = \sqrt{\frac{c}{b}}. \tag{B.4}$$

We also get extra roots when

$$abX^2 + (8bc - a^2)X + ac = 0, \tag{B.5}$$

which using the quadratic formula gives us

$$\begin{aligned} X_\pm &= \frac{-(8bc - a^2) \pm \sqrt{(8bc - a^2)^2 - 4a^2bc}}{2ab} \\ &= \frac{-(8bc - a^2) \pm \sqrt{(a^2 - 16bc)(a^2 - 4bc)}}{2ab}. \end{aligned} \tag{B.6}$$

We see that we get a single, real root when

$$(a^2 - 16bc)(a^2 - 4bc) = 0, \tag{B.7}$$

that is when $a^2 = 16bc$ or $a^2 = 4bc$. However we see that, if $a^2 = 4bc$ then $X = -2c/a$, hence a negative root. Whereas if $a^2 = 16bc$ then $X = 4c/a$, so we have a positive root. Thus we can conclude that we have a possible inflection at $X = 4c/a$ under the condition of $a^2 = 16bc$.

Finally we have two real distinct roots when

$$(a^2 - 16bc)(a^2 - 4bc) > 0. \tag{B.8}$$

As $a^2 - 16bc < a^2 - 4bc$ then this holds if either $a^2 > 16bc$ or $a^2 < 4bc$. To confirm whether these conditions give positive or negative roots we first note that

$$\begin{aligned}(a^2 - 8bc)^2 - 4a^2bc &> 0 \\ \Rightarrow (a^2 - 8bc)^2 &> 4a^2bc \\ \Rightarrow \sqrt{(a^2 - 8bc)^2 - 4a^2bc} &< a^2 - 8bc. \end{aligned} \tag{B.9}$$

Using this we can state that, if $a^2 > 16bc$ then $a^2 - 8bc > 16bc - 8bc = 8bc > 0$, hence we have two positive roots. Conversely, if $a^2 < 4bc$, then $a^2 - 8bc < 4bc - 8bc = -4bc < 0$, hence both roots are negative. Thus we can conclude that we get two positive roots, and possible inflection points, under the condition

$$a^2 > 16bc. \tag{B.10}$$

To confirm that these are in fact inflection points, we first confirm that $X_- < X_c < X_+$. First noting that $a^2 - 16bc < a^2 - 4bc \Rightarrow (a^2 - 16bc)^2 < (a^2 - 16bc)(a^2 - 4bc) \Rightarrow a^2 - 16bc < \sqrt{(a^2 - 16bc)(a^2 - 4bc)}$, then we can state that

$$\begin{aligned} X_- =& \frac{a^2 - 8bc - \sqrt{(a^2 - 16bc)(a^2 - 4bc)}}{2ab} \\ <& \frac{a^2 - 8bc - (a^2 - 16bc)}{2ab} \\ =& \frac{4c}{a} \\ =& \sqrt{\frac{16c^2}{a^2}} \\ <& \sqrt{\frac{c}{b}} = X_c. \end{aligned} \tag{B.11}$$

Thus we can confirm that $X_- < X_c$. Also

$$\begin{aligned} X_+ =& \frac{a^2 - 8bc + \sqrt{(a^2 - 16bc)(a^2 - 4bc)}}{2ab} \\ >& \frac{a^2 - 8bc}{2bc} \\ =& \sqrt{\frac{(a^2 - 8bc)^2}{4a^2b^2 2}} \\ >& \sqrt{\frac{4a^2bc}{4a^2b^2}} \\ =& \sqrt{\frac{c}{b}} = X_c, \end{aligned} \tag{B.12}$$

hence confirming that $X_c < X_+$. We now use the second derivative to confirm there is a sign change at each of the points. We can see as $X \to 0^+$, we have $d^2A_{bound}/d\log_{10} X^2 > 0$. Taking $\hat{X}_- = 4c/a$ as a point between X_- and X_c, we evaluate

$$\left.\frac{d^2A_{bound}}{d\log_{10} X^2}\right|_{X=\hat{X}_-} = -\frac{12c^3(a^2 - 16bc)^2}{a^4} < 0, \tag{B.13}$$

thus we can confirm there is an inflection point at X_- with the curve changing from convex to concave. Taking $\hat{X}_+ = (a^2 - 8bc)/2ab$ as a point between X_c and X_+, we evaluate

$$\left.\frac{d^2A_{bound}}{d\log_{10} X^2}\right|_{X=\hat{X}_-} = \frac{(a^2 - 8bc)((a^2 - 8bc)^2 - 4a^2bc)^2}{32a^4b^3} > 0, \tag{B.14}$$

hence there is a sign change at the point X_c with the curve changing from convex to concave. Finally taking the limit as $X \to \infty$, we see that $d^2A_{bound}/d\log_{10} X^2 < 0$, thus we have a third inflection point with the curve changing from convex to concave. So to conclude, under the conditions $16bc < a^2$, we have three inflection points.

Appendix C

Definitions and theorems for SIA

C.1 The Lie bracket

The controllability matrix, as discussed in Section 5.2.3, is constructed by use of the Lie bracket. We define this, as in [134]

Definition C.1.1 *The **Lie bracket** is an algebraic operation on two vector fields $\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{g}(\boldsymbol{x}) \in \mathbb{R}^{n_x}$ that creates a third vector field $\mathcal{F}(\boldsymbol{x})$, which when taken with $\boldsymbol{g}$ as the input control vector together with $\boldsymbol{u} \in \mathbb{R}^{n_u}$ defines an embedding in $\mathbb{R}^{n_x}$ that maps the input to states. It is given as*

$$(ad^1_{\boldsymbol{f}}, \boldsymbol{g}) = [\boldsymbol{f}, \boldsymbol{g}] = \frac{\partial \boldsymbol{g}(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}) - \frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{g}(\boldsymbol{x}), \tag{C.1}$$

with higher order brackets being calculated recursively as

$$\begin{aligned} (ad^2_{\boldsymbol{f}}, \boldsymbol{g}) &= [\boldsymbol{f}, [\boldsymbol{f}, \boldsymbol{g}]] = \frac{\partial (ad^1_{\boldsymbol{f}}, \boldsymbol{g})}{\partial \boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}) - \frac{\partial \boldsymbol{f}(\boldsymbol{x})}{\partial x} (ad^1_{\boldsymbol{f}}, \boldsymbol{g}), \\ &\vdots \\ (ad^k_{\boldsymbol{f}}, \boldsymbol{g}) &= [\boldsymbol{f}, (ad^{k-1}_{\boldsymbol{f}}, \boldsymbol{g})]. \end{aligned} \tag{C.2}$$

C.2 The Lie derivative

To construct the observability matrix, as discussed in Section 5.2.3, we use the Lie bracket, which we define as in [129]:

Definition C.2.1 *The **Lie derivative** of $\boldsymbol{f}(\boldsymbol{x})$ with respect to $\boldsymbol{y}(\boldsymbol{x})$ is given as*

$$L_{\boldsymbol{f}}^{1}\boldsymbol{y}(\boldsymbol{x}) = \frac{\partial \boldsymbol{y}(\boldsymbol{x})}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}), \tag{C.3}$$

with higher order derivatives being calculated recursively as

$$\begin{aligned} L_{\boldsymbol{f}}^{2}\boldsymbol{y}(\boldsymbol{x}) &= \frac{\partial L_{\boldsymbol{f}}^{1}\boldsymbol{y}(\boldsymbol{x})}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}), \\ &\vdots \\ L_{\boldsymbol{f}}^{i}\boldsymbol{y}(\boldsymbol{x}) &= \frac{\partial L_{\boldsymbol{f}}^{i-1}\boldsymbol{y}(\boldsymbol{x})}{\partial \boldsymbol{x}}\boldsymbol{f}(\boldsymbol{x}), \end{aligned} \tag{C.4}$$

C.3 Linear equivalence

To determine identifiability of a linear ODE system, we use the algebraic equivalence theorem [46, 112], which states that

Theorem C.3.1 *(F, G, H) is algebraically equivalent to $(\widetilde{F}, \widetilde{G}, \widetilde{H})$ if and only if there exists a continuously differentiable matrix $T : \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$, such that*

(i)
$$\det T \neq 0, \tag{C.5a}$$

(ii)
$$T\widetilde{\boldsymbol{x}}_0 = \boldsymbol{x}_0, \tag{C.5b}$$

(iii)
$$T\widetilde{F} = FT, \tag{C.5c}$$
$$T\widetilde{G} = G, \tag{C.5d}$$
$$\widetilde{H} = HT, \tag{C.5e}$$

where $\widetilde{F} = F(\widetilde{\boldsymbol{p}})$, $\widetilde{G} = G(\widetilde{\boldsymbol{p}})$ *and* $\widetilde{H} = H(\widetilde{\boldsymbol{p}})$.

C.4 Nonlinear equivalence

Identifiability of a nonlinear ODE system is determined via the local state isomorphism theorem, given as (as in [24, 28, 123]):

Theorem C.4.1 *Assume that the model of* (5.1) *is locally reduced at* $\boldsymbol{x}(\boldsymbol{p})$ *for all* $\boldsymbol{p} \in \Omega$. *Consider the parameter values* $\boldsymbol{p}, \widetilde{\boldsymbol{p}} \in \Omega$, *an open neighbourhood* V *of* $\boldsymbol{x}_0(\widetilde{\boldsymbol{p}})$ *in* M, *and any analytical mapping* $\lambda : V \to \mathbb{R}, V \subset \mathbb{R}^{n_x}$ *such that*

$$(i) \qquad rank\left(\frac{\partial \lambda(\widetilde{\boldsymbol{x}})}{\partial \widetilde{\boldsymbol{x}}}\right) = n_x, \quad \forall \widetilde{\boldsymbol{x}} \in V, \tag{C.6a}$$

$$(ii) \qquad \lambda(\widetilde{\boldsymbol{x}}_0(\boldsymbol{p})) = \boldsymbol{x}_0(\widetilde{\boldsymbol{p}}), \tag{C.6b}$$

$$(iii) \qquad f(\lambda(\widetilde{\boldsymbol{x}}), \boldsymbol{p}) = \frac{\partial \lambda(\widetilde{\boldsymbol{x}})}{\partial \widetilde{\boldsymbol{x}}} f(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{p}}), \tag{C.6c}$$

$$g(\lambda(\widetilde{\boldsymbol{x}}), \boldsymbol{p}) = \frac{\partial \lambda(\widetilde{\boldsymbol{x}})}{\partial \widetilde{\boldsymbol{x}}} g(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{p}}), \tag{C.6d}$$

$$h(\lambda(\widetilde{\boldsymbol{x}}), \boldsymbol{p}) = h(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{p}}), \tag{C.6e}$$

for all $\widetilde{\boldsymbol{x}} \in V$. *Then the system in* (5.1) *is globally identifiable at* $\boldsymbol{p}$ *if and only if* (5.22a)-(5.22e) *implies* $\widetilde{\boldsymbol{p}} = \boldsymbol{p}$.

Bibliography

[1] Graphpad prism. Graphpad Software www.graphpad.com.

[2] Matlab. The MathWorks Inc.

[3] Tomás Alarcón and Karen M Page. Mathematical models of the vegf receptor and its role in cancer therapy. *Journal of The Royal Society Interface*, 4(13):283–304, 2007.

[4] Milena Anguelova, Gunnar Cedersund, M Johansson, CJ Franzen, and Bernt Wennberg. Conservation laws and unidentifiability of rate expressions in biochemical models. *IET systems biology*, 1(4):230–237, 2007.

[5] Everhardus Jacobus Ariens et al. Affinity and intrinsic activity in the theory of competitive inhibition. 1. problems and theory. *Archives internationales de pharmacodynamie et de thérapie*, 99:32–49, 1954.

[6] Mei Bai. Dimerization of g-protein-coupled receptors: roles in signal transduction. *Cellular signalling*, 16(2):175–186, 2004.

[7] Eva Balsa-Canto, Antonio A Alonso, and Julio R Banga. An iterative identification procedure for dynamic modeling of biochemical networks. *BMC systems biology*, 4(1):11, 2010.

[8] Daniel J Bearup, Neil D Evans, and Michael J Chappell. The input-output relationship approach to structural identifiability analysis. 2010.

[9] Ror Bellman and Karl Johan Åström. On structural identifiability. *Mathematical biosciences*, 7(3-4):329–339, 1970.

[10] Giuseppina Bellu, Maria Pia Saccomani, Stefania Audoly, and Leontina D'Angiò. Daisy: A new software tool to test global identifiability of biological and physiological systems. *Computer methods and programs in biomedicine*, 88(1):52–61, 2007.

[11] James Whyte Black and Paul Leff. Operational models of pharmacological agonism. *Proceedings of the Royal society of London. Series B. Biological sciences*, 220(1219):141–162, 1983.

[12] JW Black, P Leff, NP Shankley, and J Wood. An operational model of pharmacological agonism: the effect of e/[a] curve shape on agonist dissociation constant estimation. *British journal of pharmacology*, 84(2):561, 1985.

[13] LJ Bridge. Modeling and simulation of inverse agonism dynamics. *Methods in enzymology*, 485:559–582, 2009.

[14] LJ Bridge, JR King, SJ Hill, and MR Owen. Mathematical modelling of signalling in a two-ligand g-protein coupled receptor system: Agonist–antagonist competition. *Mathematical biosciences*, 223(2):115–132, 2010.

[15] Jun Cai, Wen G Jiang, Asif Ahmed, and Mike Boulton. Vascular endothelial growth factor-induced endothelial cell proliferation is regulated by interaction between vegfr-2, sh-ptp1 and enos. *Microvascular research*, 71(1):20–31, 2006.

[16] Davide Calebiro and Amod Godbole. Internalization of g-protein-coupled receptors: Implication in receptor function, physiology and diseases. *Best Practice & Research Clinical Endocrinology & Metabolism*, 32(2):83–91, 2018.

[17] Davide Calebiro, Viacheslav O Nikolaev, Maria Cristina Gagliani, Tiziana De Filippis, Christian Dees, Carlo Tacchetti, Luca Persani, and Martin J Lohse. Persistent camp-signals triggered by internalized g-protein–coupled receptors. *PLoS Biol*, 7(8):e1000172, 2009.

[18] Davide Calebiro, Viacheslav O Nikolaev, Luca Persani, and Martin J Lohse. Signaling by internalized g-protein-coupled receptors. *Trends in pharmacological sciences*, 31(5):221–228, 2010.

[19] Vicent Casadó, Carla Ferrada, Jordi Bonaventura, Eduard Gracia, Josefa Mallol, Enric I Canela, Carmen Lluís, Antoni Cortés, and Rafael Franco. Useful pharmacological parameters for g-protein-coupled receptor homodimers obtained from competition experiments. agonist–antagonist binding modulation. *Biochemical pharmacology*, 78(12):1456–1463, 2009.

[20] Vincent Casadó, Antoni Cortés, Francisco Ciruela, Josefa Mallol, Sergi Ferré, Carmen Lluis, Enric I Canela, and Rafael Franco. Old and new ways to calculate the affinity of agonists and antagonists interacting with g-protein-coupled monomeric and dimeric receptors: The receptor–dimer cooperativity index. *Pharmacology & therapeutics*, 116(3):343–354, 2007.

[21] J del Castillo and Bernard Katz. Interaction at end-plate receptors between different choline derivatives. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, 146(924):369–381, 1957.

[22] Marc Chabre, Philippe Deterre, and Bruno Antonny. The apparent cooperativity of some gpcrs does not necessarily imply dimerization. *Trends in pharmacological sciences*, 30(4):182–187, 2009.

[23] MJ Chapman and KR Godfrey. On structural equivalence and identifiability constraint ordering. *IFAC Proceedings Volumes*, 18(5):523–529, 1985.

[24] Michael J Chappell, Keith R Godfrey, and Sandor Vajda. Global identifiability of the parameters of nonlinear systems with specified inputs: a comparison of methods. *Mathematical Biosciences*, 102(1):41–73, 1990.

[25] Michael J Chappell and Roger N Gunn. A procedure for generating locally identifiable reparameterisations of unidentifiable non-linear systems by the similarity transformation approach. *Mathematical biosciences*, 148(1):21–41, 1998.

[26] SY Amy Cheung, James WT Yates, and Leon Aarons. The design and analysis of parallel experiments to produce structurally identifiable models. *Journal of pharmacokinetics and pharmacodynamics*, 40(1):93–100, 2013.

[27] Oana Chis, Julio R Banga, and Eva Balsa-Canto. Methods for checking structural identifiability of nonlinear biosystems: A critical comparison. *IFAC Proceedings Volumes*, 44(1):10585–10590, 2011.

[28] Oana-Teodora Chis, Julio R Banga, and Eva Balsa-Canto. Structural identifiability of systems biology models: a critical comparison of methods. *PloS one*, 6(11):e27755, 2011.

[29] Alfred Joseph Clark. The mode of action of drugs on cells. 1933.

[30] Personal communication with Lloyd Bridge and Dr Chloe Peach. Excel data files for time courses published in [101] provided for data fitting purposes, courtesy of dr chloe peach and the institute of cell signalling, university of nottingham.

[31] Andre De Lean, JMand Stadel, and RJ Lefkowitz. A ternary complex model explains the agonist-specific binding properties of the adenylate cyclase-coupled beta-adrenergic receptor. *Journal of Biological Chemistry*, 255(15):7108–7117, 1980.

[32] Giovanni Y Di Veroli, Chiara Fornari, Ian Goldlust, Graham Mills, Siang Boon Koh, Jo L Bramhall, Frances M Richards, and Duncan I Jodrell. An automated fitting procedure and software for dose-response curves with multiphasic features. *Scientific reports*, 5:14701, 2015.

[33] Thierry Durroux. Principles: a model for the allosteric interactions between ligand binding sites within a dimeric gpcr. *Trends in pharmacological sciences*, 26(7):376–384, 2005.

[34] Arno Eigenwillig. Real root isolation for exact and approximate polynomials using descartes' rule of signs. 2008.

[35] Neil D Evans and Michael J Chappell. Extensions to a procedure for generating locally identifiable reparameterisations of unidentifiable systems. *Mathematical Biosciences*, 168(2):137–159, 2000.

[36] Marcello Farina, Rolf Findeisen, Eric Bullinger, Sergio Bittanti, Frank Allgower, and Peter Wellstead. Results towards identifiability properties of biochemical re-

action networks. In *Proceedings of the 45th IEEE Conference on Decision and Control*, pages 2104–2109. IEEE, 2006.

[37] Sergi Ferré, Vicent Casadó, Lakshmi A Devi, Marta Filizola, Ralf Jockers, Martin J Lohse, Graeme Milligan, Jean-Philippe Pin, and Xavier Guitart. G protein–coupled receptor oligomerization revisited: functional and pharmacological perspectives. *Pharmacological reviews*, 66(2):413–434, 2014.

[38] R Franco, V Casadó, A Cortés, J Mallol, F Ciruela, S Ferré, C Lluis, and EI Canela. G-protein-coupled receptor heteromers: function and ligand pharmacology. *British journal of pharmacology*, 153(S1), 2008.

[39] Rafael Franco, Vicent Casadó, Antoni Cortés, Carla Ferrada, Josefa Mallol, Amina Woods, Carme Lluis, Enric I Canela, and Sergi Ferré. Basic concepts in g-protein-coupled receptor homo-and heterodimerization. *The Scientific World Journal*, 7:48–57, 2007.

[40] Rafael Franco, Vicent Casadó, Josefa Mallol, Carla Ferrada, Sergi Ferré, Kjell Fuxe, Antoni Cortés, Francisco Ciruela, Carmen Lluis, and Enric I Canela. The two-state dimer receptor model: a general model for receptor dimers. *Molecular pharmacology*, 69(6):1905–1912, 2006.

[41] Rafael Franco, Vicent Casadó, Josefa Mallol, Sergi Ferré, Kjell Fuxe, Antonio Cortés, Francisco Ciruela, Carmen Lluis, and Enric I Canela. Dimer-based model for heptaspanning membrane receptors. *Trends in biochemical sciences*, 30(7):360–366, 2005.

[42] ROBERT F FURCHGOTT. The pharmacology of vascular smooth muscle. *Pharmacological reviews*, 7(2):183–265, 1955.

[43] Rudolf Gesztelyi, Judit Zsuga, Adam Kemeny-Beke, Balazs Varga, Bela Juhasz, and Arpad Tosaki. The hill equation and the origin of quantitative pharmacology. *Archive for history of exact sciences*, 66(4):427–438, 2012.

[44] J Giraldo. On the fitting of binding data when receptor dimerization is suspected. *British journal of pharmacology*, 155(1):17–23, 2008.

[45] Keith R Godfrey and Michael J Chapman. Identifiability and indistinguishability of linear compartmental models. *Mathematics and Computers in Simulation*, 32(3):273–295, 1990.

[46] KR Godfrey and JJ DiStefano III. Identifiability of model parameter. *IFAC Proceedings Volumes*, 18(5):89–114, 1985.

[47] Manoj Gopalakrishnan, Kimberly Forsten-Williams, and Uwe C Täuber. Ligand-induced coupling versus receptor pre-association: cellular automaton simulations of fgf-2 binding. *Journal of theoretical biology*, 227(2):239–251, 2004.

[48] AN Gorban. Detailed balance in micro-and macrokinetics and micro-distinguishability of macro-processes. *Results in Physics*, 4:142–147, 2014.

[49] M Grewal and Keith Glover. Identifiability of linear and nonlinear dynamical systems. *IEEE Transactions on automatic control*, 21(6):833–837, 1976.

[50] Ashok Kumar Grover. Use of allosteric targets in the discovery of safer drugs. *Medical Principles and Practice*, 22(5):418–426, 2013.

[51] Dong Guo, Julia M Hillger, Adriaan P IJzerman, and Laura H Heitman. Drug-target residence time—a case for g protein-coupled receptors. *Medicinal research reviews*, 34(4):856–892, 2014.

[52] Caitlin D Hanlon and Deborah J Andrew. Outside-in signaling–a brief review of gpcr signaling with a focus on the drosophila gpcr family. *Journal of cell science*, 128(19):3533–3542, 2015.

[53] Robert Hermann and Arthur Krener. Nonlinear controllability and observability. *IEEE Transactions on automatic control*, 22(5):728–740, 1977.

[54] Samuel RJ Hoare, Nicolas Pierre, Arturo Gonzalez Moya, and Brad Larson. Kinetic operational models of agonism for g-protein-coupled receptors. *Journal of theoretical biology*, 446:168–204, 2018.

[55] Stefan Hoops, Sven Sahle, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. Copasi—a complex pathway simulator. *Bioinformatics*, 22(24):3067–3074, 2006.

[56] Xiang-Qun Hu and Robert W Peoples. The 5-ht3b subunit confers spontaneous channel opening and altered ligand properties of the 5-ht3 receptor. *Journal of Biological Chemistry*, 283(11):6826–6831, 2008.

[57] Edward C Hulme and Mike A Trevethick. Ligand binding assays at equilibrium: validation and interpretation. *British journal of pharmacology*, 161(6):1219–1237, 2010.

[58] Wolfram Research, Inc. Mathematica, Version 11.1. Champaign, IL, 2017.

[59] David LI Janzén, Linnéa Bergenholm, Mats Jirstrand, Joanna Parkinson, James Yates, Neil D Evans, and Michael J Chappell. Parameter identifiability of fundamental pharmacodynamic models. *Frontiers in physiology*, 7:590, 2016.

[60] Kenneth A Johnson and Roger S Goody. The original michaelis constant: translation of the 1913 michaelis–menten paper. *Biochemistry*, 50(39):8264–8269, 2011.

[61] Alfonsas Juška. Minimal models of multi-site ligand-binding kinetics. *Journal of theoretical biology*, 255(4):396–403, 2008.

[62] Rudolf Emil Kalman et al. Contributions to the theory of optimal control. *Bol. soc. mat. mexicana*, 5(2):102–119, 1960.

[63] Bernard Katz and S Thesleff. A study of thedesensitization'produced by acetylcholine at the motor end-plate. *The Journal of physiology*, 138(1):63, 1957.

[64] Bertram G Katzung. *Basic and clinical pharmacology*. Mc Graw Hill, 2012.

[65] James P Keener and James Sneyd. *Mathematical physiology*, volume 1. Springer, 1998.

[66] T. Kenakin. *Pharmacology in Drug Discovery and Development: Understanding Drug Response*. Elsevier Science, 2016.

[67] Terry Kenakin. Principles: receptor theory in pharmacology. *Trends in pharmacological sciences*, 25(4):186–192, 2004.

[68] Terry Kenakin. Receptor theory. *Current Protocols in Pharmacology*, 41(1):1–2, 2008.

[69] Terry Kenakin. *A pharmacology primer: theory, application and methods*. Academic Press, 2009.

[70] Laura E Kilpatrick, Rachel Friedman-Ohana, Diana C Alcobia, Kristin Riching, Chloe J Peach, Amanda J Wheal, Stephen J Briddon, Matthew B Robers, Kris Zimmerman, Thomas Machleidt, et al. Real-time analysis of the binding of fluorescent vegf165a to vegfr2 in living cells: effect of receptor tyrosine kinase inhibitors and fate of internalized agonist-receptor complexes. *Biochemical pharmacology*, 136:62–75, 2017.

[71] Peter Klein, Dawn Mattoon, Mark A Lemmon, and Joseph Schlessinger. A structure-based model for ligand binding and dimerization of egf receptors. *Proceedings of the National Academy of Sciences*, 101(4):929–934, 2004.

[72] Wesley K Kroeze, Douglas J Sheffler, and Bryan L Roth. G-protein-coupled receptors at a glance. *Journal of cell science*, 116(24):4867–4869, 2003.

[73] Irving Langmuir. The adsorption of gases on plane surfaces of glass, mica and platinum. *Journal of the American Chemical society*, 40(9):1361–1403, 1918.

[74] Y Lecourtier, F Lamnabhi-Lagarrigue, and E Walter. Volterra and generating power series approaches to identifiability testing. *Identifiability of parametric models*, pages 50–66, 1987.

[75] Y Lecourtier and A Raksanyi. The testing of structural properties through symbolic computation. In *Identifiability of parametric models*, pages 75–84. Pergamon Press Oxford, 1987.

[76] Paul Leff. The two-state model of receptor activation. *Trends in pharmacological sciences*, 16(3):89–97, 1995.

[77] Gilbert N Lewis. A new principle of equilibrium. *Proceedings of the National Academy of Sciences of the United States of America*, 11(3):179, 1925.

[78] Lee E Limbird. *Cell surface receptors: a short course on theory and methods: a short course on theory and methods*. Springer Science & Business Media, 2012.

[79] E Wang Lund. Guldberg and waage and the law of mass action. *Journal of Chemical Education*, 42(10):548, 1965.

[80] Feilim Mac Gabhann and Aleksander S Popel. Dimerization of vegf receptors and implications for signal transduction: a computational study. *Biophysical chemistry*, 128(2):125–139, 2007.

[81] Gabriella Margaria, Eva Riccomagno, Michael J Chappell, and Henry P Wynn. Differential algebra methods for the study of the structural identifiability of rational function state-space models in the biosciences. *Mathematical biosciences*, 174(1):1–26, 2001.

[82] Ichiro Maruyama. Mechanisms of activation of receptor tyrosine kinases: monomers or dimers. *Cells*, 3(2):304–330, 2014.

[83] Lauren T May, Lloyd J Bridge, Leigh A Stoddart, Stephen J Briddon, and Stephen J Hill. Allosteric interactions across native adenosine-a3 receptor homodimers: quantification using single-cell ligand-binding kinetics. *The FASEB Journal*, 25(10):3465–3476, 2011.

[84] Lauren T May, Katie Leach, Patrick M Sexton, and Arthur Christopoulos. Allosteric modulation of g protein–coupled receptors. *Annu. Rev. Pharmacol. Toxicol.*, 47:1–51, 2007.

[85] Kapil Mayawala, Dionisios G Vlachos, and Jeremy S Edwards. Spatial modeling of dimerization reaction dynamics in the plasma membrane: Monte carlo vs. continuum differential equations. *Biophysical chemistry*, 121(3):194–208, 2006.

[86] Charlene McQueen. *Comprehensive toxicology*. Elsevier, 2017.

[87] Leonor Menten and MI Michaelis. Die kinetik der invertinwirkung. *Biochem Z*, 49(333-369):5, 1913.

[88] Microsoft Corporation. Microsoft excel.

[89] Thomas R Middendorf and Richard W Aldrich. Structural identifiability of equilibrium ligand-binding parameters. *Journal of General Physiology*, 149(1):105–119, 2017.

[90] Thomas R Middendorf and Richard W Aldrich. The structure of binding curves and practical identifiability of equilibrium ligand-binding parameters. *Journal of General Physiology*, 149(1):121–147, 2017.

[91] Graeme Milligan. G protein-coupled receptor dimerization: function and ligand pharmacology. *Molecular pharmacology*, 66(1):1–7, 2004.

[92] Graeme Milligan. G-protein-coupled receptor heterodimers: pharmacology, function and relevance to drug discovery. *Drug discovery today*, 11(11):541–549, 2006.

[93] Graeme Milligan. G protein-coupled receptor hetero-dimerization: contribution to pharmacology and function. *British journal of pharmacology*, 158(1):5–14, 2009.

[94] Jacque Monod, Jeffries Wyman, and Jean-Pierre Changeux. On the nature of allosteric transitions: a plausible model. *J Mol Biol*, 12(1):88–118, 1965.

[95] Harvey J Motulsky and LC Mahan. The kinetics of competitive radioligand binding predicted by the law of mass action. *Molecular pharmacology*, 25(1):1–9, 1984.

[96] James D Murray. *Mathematical biology: I. An introduction*, volume 17. Springer Science & Business Media, 2007.

[97] Mark Nickerson. Receptor occupancy and tissue response. *Nature*, 178(4535):697–698, 1956.

[98] Ruth Nussinov and Chung-Jung Tsai. The different ways through which specificity works in orthosteric and allosteric drugs. *Current pharmaceutical design*, 18(9):1311–1316, 2012.

[99] Anna-Karin Olsson, Anna Dimberg, Johan Kreuger, and Lena Claesson-Welsh. Vegf receptor signalling? in control of vascular function. *Nature reviews Molecular cell biology*, 7(5):359, 2006.

[100] Shawn C Owen, Allison K Doak, Ahil N Ganesh, Lyudmila Nedyalkova, Christopher K McLaughlin, Brian K Shoichet, and Molly S Shoichet. Colloidal drug formulations can explain ̈bell-shapedc̈oncentration–response curves. *ACS chemical biology*, 9(3):777–784, 2014.

[101] Chloe J Peach, Laura E Kilpatrick, Jeanette Woolard, and Stephen J Hill. Comparison of the ligand-binding properties of fluorescent vegf-a isoforms to vegf receptor 2 in living cells and membrane preparations using nanobret. *British journal of pharmacology*, 176(17):3220–3235, 2019.

[102] Chloe J Peach, Viviane W Mignone, Maria Augusta Arruda, Diana C Alcobia, Stephen J Hill, Laura E Kilpatrick, and Jeanette Woolard. Molecular pharmacology of vegf-a isoforms: binding and signalling at vegfr2. *International journal of molecular sciences*, 19(4):1264, 2018.

[103] Lambertus A Peletier and Johan Gabrielsson. Impact of mathematical pharmacology on practice and theory: four case studies. *Journal of pharmacokinetics and pharmacodynamics*, 45(1):3–21, 2018.

[104] MC Peterson and MM Riggs. Fda advisory meeting clinical pharmacology review utilizes a quantitative systems pharmacology (qsp) model: a watershed moment? *CPT: Pharmacometrics & Systems Pharmacology*, 4(3):189–192, 2015.

[105] Hannu Pohjanpalo. System identifiability based on the power series expansion of the solution. *Mathematical biosciences*, 41(1-2):21–33, 1978.

[106] Heino Prinz. Hill coefficients, dose–response curves and allosteric mechanisms. *Journal of chemical biology*, 3(1):37–44, 2010.

[107] Humphrey P Rang, James M Ritter, Rod J Flower, and Graeme Henderson. *Rang & Dale's Pharmacology E-Book: with STUDENT CONSULT Online Access*. Elsevier Health Sciences, 2014.

[108] X Rovira, M Vivó, J Serra, D Roche, PG Strange, and J Giraldo. Modelling the interdependence between the stoichiometry of receptor oligomerization and ligand binding for a coexisting dimer/tetramer receptor system. *British journal of pharmacology*, 156(1):28–35, 2009.

[109] Xavier Rovira, Jean-Philippe Pin, and Jesús Giraldo. The asymmetric/symmetric activation of gpcr dimers as a possible mechanistic rationale for multiple signalling pathways. *Trends in pharmacological sciences*, 31(1):15–21, 2010.

[110] RR Ruffolo Jr. Important concepts of receptor theory. *Journal of autonomic pharmacology*, 2(4):277–295, 1982.

[111] Masabumi Shibuya. Vascular endothelial growth factor (vegf) and its receptor (vegfr) signaling in angiogenesis: a crucial target for anti-and pro-angiogenic therapies. *Genes & cancer*, 2(12):1097–1105, 2011.

[112] L Silverman. Realization of linear dynamical systems. *IEEE Transactions on Automatic Control*, 16(6):554–567, 1971.

[113] Leonard M Silverman and HE Meadows. Controllability and observability in time-variable linear systems. *SIAM Journal on Control*, 5(1):64–73, 1967.

[114] Michael Simons. An inside view: Vegf receptor trafficking and signaling. *Physiology*, 27(4):213–222, 2012.

[115] Nicola J Smith and Graeme Milligan. Allostery at g protein-coupled receptor homo- and heteromers: uncharted pharmacological landscapes. *Pharmacological reviews*, 62(4):701–725, 2010.

[116] Thomas J Snowden, Piet H van der Graaf, and Marcus J Tindall. Model reduction in mathematical pharmacology. *Journal of pharmacokinetics and pharmacodynamics*, 45(4):537–555, 2018.

[117] RP Stephenson. A modification of receptor theory. *British journal of pharmacology and chemotherapy*, 11(4):379, 1956.

[118] Janet L Stringer. *Basic concepts in pharmacology*. McGraw Hill Professional, 2006.

[119] Edward Stuttfeld and Kurt Ballmer-Hofer. Structure and function of vegf receptors. *IUBMB life*, 61(9):915–922, 2009.

[120] David A Sykes, Leigh A Stoddart, Laura E Kilpatrick, and Stephen J Hill. Binding kinetics of ligands acting at gpcrs. *Molecular and cellular endocrinology*, 485:9–19, 2019.

[121] David M Thal, Alisa Glukhova, Patrick M Sexton, and Arthur Christopoulos. Structural insights into g-protein-coupled receptor allostery. *Nature*, 559(7712):45–53, 2018.

[122] Narendra Tuteja. Signaling through g protein coupled receptors. *Plant signaling & behavior*, 4(10):942–947, 2009.

[123] Sandor Vajda, Keith R Godfrey, and Herschel Rabitz. Similarity transformation approach to identifiability analysis of nonlinear compartmental models. *Mathematical biosciences*, 93(2):217–248, 1989.

[124] JM Van den Hof. Structural identifiability of linear compartmental systems. *IEEE Transactions on Automatic Control*, 43(6):800–818, 1998.

[125] Piet H van der Graaf, Neil Benson, and Lambertus A Peletier. Topics in mathematical pharmacology. *Journal of Dynamics and Differential Equations*, 28(3-4):1337–1356, 2016.

[126] AJ Venkatakrishnan, Xavier Deupi, Guillaume Lebon, Christopher G Tate, Gebhard F Schertler, and M Madan Babu. Molecular signatures of g-protein-coupled receptors. *Nature*, 494(7436):185–194, 2013.

[127] Julio Vera, Thomas Millat, Walter Kolch, and Olaf Wolkenhauer. Dynamics of receptor and protein transducer homodimerisation. *BMC systems biology*, 2(1):92, 2008.

[128] Alejandro F Villaverde. Observability and structural identifiability of nonlinear biological systems. *Complexity*, 2019, 2019.

[129] Alejandro F Villaverde, Antonio Barreiro, and Antonis Papachristodoulou. Structural identifiability of dynamic systems biology models. *PLoS computational biology*, 12(10):e1005153, 2016.

[130] JG Wagner. Kinetics of pharmacologic response i. proposed relationships between response and drug concentration in the intact animal and man. *Journal of Theoretical Biology*, 20(2):173–201, 1968.

[131] Eric Walter and Yves Lecourtier. Unidentifiable compartmental models: what to do? *Mathematical biosciences*, 56(1-2):1–25, 1981.

[132] Sumanas Wanant and Michael J Quon. Insulin receptor binding kinetics: modeling and simulation studies. *Journal of theoretical Biology*, 205(3):355–364, 2000.

[133] Jack M Weiss, Paul H Morgan, Michael W Lutz, and Terry P Kenakin. The cubic ternary complex receptor–occupancy model i. model description. *Journal of theoretical biology*, 178(2):151–167, 1996.

[134] Andrew J Whalen, Sean N Brennan, Timothy D Sauer, and Steven J Schiff. Observability and controllability of nonlinear networks: The role of symmetry. *Physical Review X*, 5(1):011005, 2015.

[135] Carla White and Lloyd J Bridge. Ligand binding dynamics for pre-dimerised g protein-coupled receptor homodimers: Linear models and analytical solutions. *Bulletin of mathematical biology*, pages 1–33.

[136] C Wofsy, B Goldstein, K Lund, and HS Wiley. Implications of epidermal growth factor (egf) induced egf receptor aggregation. *Biophysical journal*, 63(1):98–110, 1992.

[137] PJ Woodroffe, LJ Bridge, JR King, CY Chen, and SJ Hill. Modelling of the activation of g-protein coupled receptors: drug free constitutive receptor activity. *Journal of mathematical biology*, 60(3):313–346, 2010.

[138] PJ Woodroffe, LJ Bridge, JR King, and SJ Hill. Modelling the activation of g-protein coupled receptors by a single drug. *Mathematical biosciences*, 219(1):32–55, 2009.

[139] Peter J Woolf and Jennifer J Linderman. An algebra of dimerization and its implications for g-protein coupled receptor signaling. *Journal of theoretical biology*, 229(2):157–168, 2004.

[140] James WT Yates, Neil D Evans, and Michael J Chappell. Structural identifiability analysis via symmetries of differential equations. *Automatica*, 45(11):2585–2591, 2009.

[141] Bin Zhou and Jesús Giraldo. An operational model for gpcr homodimers and its application in the analysis of biased signaling. *Drug discovery today*, 23(9):1591–1595, 2018.