University of Reading
Department of Mathematics and Statistics

Mathematical Modelling of Bacterial Chemotaxis Signalling Pathways

by

Matthew Paul Edgington

May 2015

Thesis submitted for the degree of Doctor of Philosophy

Abstract

Chemotactic bacteria are able to swim toward areas with favourable conditions. In order to control this movement cells utilise intracellular signalling pathways which pass externally detected signals from chemoreceptor clusters, through the cytoplasm, to the flagellar motors. Mathematical modelling has been key in the understanding of such systems, however the models used are often rather large and complex.

In this thesis a four-dimensional nonlinear ordinary differential equation model of the *Escherichia coli* signalling pathway is investigated mathematically and used to understand specific features of the chemotactic response. We investigate fold-change detection (FCD) and demonstrate that an extension to previous theoretical results in which we account for two receptor types produces good agreement with experimentally observed regimes of FCD behaviour. The model is then used to provide insight into the overshoot response. In particular, stability analyses allow conditions to be found describing when overshoot will occur. Simplified models are used to identify a possible mechanism for the emergence of overshoot and subsequently show that cell protein concentration, signalling timescales and negative feedback are closely linked to this phenomenon.

Single cell results are then extrapolated to the population scale using agent-based modelling. In particular, we show the effects of single cell phenomena on the behaviour of a cell population. Additionally, we investigate how the presence of two different chemoattractant gradients affects the chemotactic behaviour of a population of cells. This shows that the response is determined by the sensitivity of chemoreceptors to the precise ligand concentrations detected.

Finally, we utilise both our knowledge of chemotaxis in *E. coli* and new experimental data in order to elucidate the chemosensory pathways of *Rhodobacter sphaeroides.* We begin by identifying the main weaknesses of previous modelling efforts before using simplified models to further understand key system features. This leads to the proposal of a new mathematical model of chemotaxis in *R. sphaeroides* that provides improved agreement with experimental data.

Acknowledgements

Firstly, I would like to express my sincerest thanks to my supervisor Dr. Marcus Tindall. Throughout this project his enthusiasm for the subject matter was a great source of inspiration. In addition to this his excellent supervision kept me on track yet still allowed me the freedom to produce my best work.

I am also extremely grateful to Prof. Judy Armitage and Dr. George Wadhams in the Department of Biochemistry at the University of Oxford who each provided a number of really useful suggestions that made much of the work in this thesis possible. Also, thanks go to Dr. Jennifer de Beyer who provided the experimental data upon which the entirety of the *R. sphaeroides* theoretical work in this thesis is based.

I must also thank all those in the Department of Mathematics and Statistics who have made this time so enjoyable. To the staff, I thank you for being there whenever I needed help. This really did make the whole process so much less stressful than it could otherwise have been. Also to all of the PhD students, past and present, especially those with whom I have shared offices, I really would have gone mad if it weren't for you. You were always available when I wanted to talk maths as well as those times where I would have preferred to speak about anything else. I must also thank my office mates for the many games of darts that provided a much needed break when things weren't going to plan and (no matter how bad we all are) it won't quite be the same without you.

To my close friends and family I will be eternally grateful. Without your encouragement I would never have returned to academia. You will never know how much your love, kindness and support has meant to me these past years and I just hope that I can pay it all back in the years to come.

Finally, I am extremely thankful for the funding provided by the EPSRC. Without their financial support this research project would not have been possible.

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Signed .... Date ....

Publications Arising

Edgington, M. and Tindall, M. (2014)

Fold-Change Detection in a Whole-Pathway Model *Escherichia coli* Chemotaxis

Bulletin of Mathematical Biology 76:1376-1395.

Edgington, M. and Tindall, M. (2015)

Understanding the Link Between Single Cell and Population Scale Responses of *Escherichia coli* in Differing Ligand Gradients

Computational and Structural Biotechnology Journal 13:528-538.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The aim of this thesis is to formulate and analyse mathematical models of intracellular signalling pathways describing chemotactic cell movement in *Escherichia coli* and *Rhodobacter sphaeroides* cells. In particular, this allows us to link single cell behaviour to the population scale by using well informed and analysed single cell models. We also seek to utilise knowledge gained from *E. coli* studies to inform models of the more complex bacterial species *R. sphaeroides.*

In order to make sense of the term chemotaxis we may break it down into the prefix 'chemo' meaning chemical and a suffix 'taxis' which is from the Greek for order or arrangement. In a literal sense then, chemotaxis could be considered as the arrangement of bacteria in relation to chemicals. This relates well to the modern definition which is the directed movement of cells or organisms in relation to a concentration gradient of some particular chemical substance. The property allows cells to move around in search of regions containing chemicals that are beneficial to both their development and survival (chemoattractants), or away from those which could prove harmful (chemorepellents).

It is quite likely that most people have never heard of chemotaxis. Why, then, do we study the process in such depth within this thesis? To answer this, we consider a number of examples demonstrating its vital role from conception all the way through to death. For example, without the chemotaxis of sperm, humans would be unable to reproduce [1]. Within the embryonic phase of pregnancy, cells use chemotaxis in order to migrate from their source to target locations where they become important components of various bodily organs. Neurons inside the developing brain also display chemotaxis which allows them to move from the place they are created to find suitable sites where they form connections that allow functions such as learning and memory [2]. Once born, chemotaxis continues to play a key role in the survival of the human race. In particular, white blood cells display chemotactic behaviour in order to search for sites of infection [3]. This allows them to eradicate foreign and harmful organisms, helping to maintain the healthy functioning of the body. However, in spite of this obvious health benefit, chemotaxis is not always good news. For example, during metastasis (the spreading of cancer throughout the body), cancer cells use chemotaxis to direct their motion toward lymphatic vessels and then use the lymphatic system (a component of the circulatory system) in order to spread to different parts of the human body [4].

Clearly chemotaxis plays a key role in the survival (and sometimes death) of animal species, but does it have any further uses? Bacterial chemotaxis has been shown to be important in the formation of biofilms - groups of microorganisms that stick to one another on or close to a surface [5]. It has been suggested that this could help to make bioremediation a more widely used technique for neutralising difficult to clear sources of contamination such as oil spills in the ocean [6].

Whilst it is known that chemotaxis plays a role in each of the processes described above, the exact mechanisms behind the sensing of the environment and cell migration are not fully understood. The study of relatively simple, model bacteria that display chemotaxis can provide key insights into the mechanisms controlling it. So, what are the model species used in the study of bacterial chemotaxis? A number of bacterial species have been identified as chemotactic, however by far the best studied example is that of *Escherichia coli.* More recently though other bacterial species such as *Bacillus subtilis* and *Rhodobacter sphaeroides* have increasingly been examined in the literature.

Why has *E. coli* been studied so thoroughly when other species have not? One key reason is that, in general, *E. coli* is fairly harmless [7]. Most strains of *E. coli* are part of the normal flora of the gut and can actually be of benefit to their host by producing vitamin K_2 as well as protecting the intestine against colonisation by pathogenic bacteria [8]. However, certain strains may cause food poisoning and under certain circumstances are capable of causing harmful liver and kidney infections. Whilst it is clear that working with a relatively harmless bacteria is desirable, it is perhaps more important that this species may be grown both easily and inexpensively in a laboratory setting [9]. This, coupled with the fact that *E. coli* cells are relatively simple in their genetics compared to other species [9], are perhaps the most likely reasons for *E. coli* being the model organism for studies on bacterial chemotaxis.

Utilising model systems such as that of *E. coli* yields interesting questions and insights on its own. However, the knowledge of how these relatively simple model systems function should form a knowledge base that can be used to help inform further investigations of both more complex systems and future applications. In this thesis in particular, we study a number of phenomena present within the *E. coli* model system and subsequently seek to use the knowledge obtained in an attempt to gain insight into the function of *R. spaheroides* which is a far more complex chemotactic species of bacteria.

1.1 Thesis Outline

Within Chapter 2 of this thesis we review the biology behind the process of bacterial chemotaxis. Firstly, the history of bacterial chemotaxis research shall be reviewed. This begins with the initial discovery of different bacterial species as well as their ability to move (i.e. their motility). Following this, the experimental biochemical studies that have shaped the modern view of how chemotaxis functions within bacterial cells are reviewed. In addition to this, a summary of the work that is beginning to shape an understanding of *R. sphaeroides* chemotaxis is given. Finally, throughout this chapter mathematical models that have been used to help elucidate the features of single cell and population level chemotaxis are reviewed with a particular emphasis placed on the relative strengths and weaknesses of each approach.

The majority of experimental work that has been instrumental in forming the modern understanding of the intracellular signalling cascades controlling chemotactic behaviour has focused on the chemotactic function of individual cells. As such, Chapter 3 analyses a recent model of the *E. coli* intracellular chemotaxis signalling pathway. This model includes a representation of the key biological processes without the issues associated with the large and complicated models common in the literature. In the context of this mathematical model, different offset energy functions are compared to experimental data in order to analyse which produces the most effective fit. This model is subjected to analytics including steady-state analysis. The asymptotic stability of the biologically feasible, unique steady-state is then investigated as is the stiffness of the model system. This study underlies much of the work included within subsequent chapters.

One feature that has been demonstrated experimentally in *E. coli* cells as well as being analysed theoretically is that of fold-change detection (FCD) - a response depending on fold-changes, rather than absolute changes in input. Within Chapter 4 the feature of FCD is analysed in the context of the signalling pathway model from Chapter 3, using sufficient conditions stated in the theoretical literature. Further to demonstrating the theoretical existence of FCD within the model system, we investigate the robustness of the phenomenon to changes in model parameters and determine under what conditions FCD fails. Finally, theoretical predictions are given relating to the number of regimes of FCD behaviour that should be expected when chemoreceptor clusters meet certain conditions.

Within the experimental literature it has been shown that the *E. coli* intracellular signalling cascade does not always display monotonic relaxation to steady-state following the response to a ligand stimulus. Chapter 5 examines conditions under which the mathematical model from

Chapter 3 is able to display an overshoot response to ligand stimuli. In doing so a number of key processes in the signalling cascade are identified as responsible for the emergence of the response. Results are compared with experimental data and discussed with reference to their biological feasibility.

An agent-based model (ABM) is developed in Chapter 6 to understand how individual behaviour identified in Chapters 3 to 5 maps to the population scale. Using this ABM we further these single cell investigations by studying phenomena in the context of chemotactic *E. coli* cell populations. In particular, the effects of variation in intracellular signalling protein concentration, such as those required to exhibit overshoot, are discussed in terms of how efficiently cells respond when placed in chemoattractant gradients of differing steepness. Also considered here is the ability of cells to perform chemotaxis when two separate ligand gradients are present simultaneously.

Within Chapter 7 we seek to utilise the knowledge gained from the study of *E. coli* chemotaxis in order to help further the study of chemotaxis signalling pathways in *R. sphaeroides*. In a similar manner to Chapter 3, this chapter will present a model of the *R. sphaeroides* chemotaxis signalling pathways from the literature. This mathematical model is then analytically treated in order to ascertain that there is just one biologically feasible steady-state, which is subsequently shown to be asymptotically stable and also belongs to a very stiff system of ordinary differential equations (ODEs). Considering this model in the context of recent experimental results by developing separate models of chemotaxis mutant cells reveals that the adaptation processes assigned to the polar and cytoplasmic chemotaxis clusters within the model used here are inadequate to represent the experimental data, thus invalidating this model.

Within Chapter 8 we consider simplified mathematical models of *R. sphaeroides* chemotactic signalling. Firstly, a three ODE model is utilised in order to propose roles for each of the processes occurring at the cytoplasmic cluster. Then, we utilise a non-adapting version of the model in Chapter 7 in order to investigate two outstanding questions. In particular, using a number of slightly altered versions of this model it may be found that the signalling reactions (i.e. not those relating to adaptation) appear to produce behaviour in line with the experimental data. A specific case considered is the reverse phosphotransfer reaction from CheB$_2$-P onto CheA$_2$ that has been demonstrated *in vitro* but has not yet been shown to exist *in vivo*. This chapter provides theoretical support for the idea that this process does in fact occur *in vivo*. The same non-adapting model is then used to prove that proteins CheY$_3$ and CheY$_4$ must play a role in setting the flagellar rotation behaviour of *R. sphaeroides* cells. It is also shown that a certain ratio of the three CheY proteins of *R. sphaeroides* produces a correlation with the experimental fraction of time that the flagella spend in a stopped state. We end this chapter by postulating a new *R. sphaeroides* signalling model and show that this helps to remove some of the issues associated with previous modelling efforts.

Finally, in Chapter 9 a summary of the work contained within this thesis shall be given. Within the context of this summary, the implications of this work both for the future study of chemotaxis signalling pathways as well as other biological systems will be discussed.

In addition to the work discussed in Chapters 1-9, Appendices A-E provide further details on mathematical results and biological concepts referred to within this thesis for the interested reader. To this end a glossary of important biological terms appearing within this thesis is also included.

Chapter 2

Background

In this chapter we begin with a summary of the key biological processes underlying bacterial chemotaxis in *E. coli*. Following this, we review the main mathematical modelling approaches that have been used to elucidate various aspects of the intracellular chemotaxis signalling pathways of individual cells. We then progress onto a discussion of approaches used to model cell populations. This chapter concludes with a summary of multi-scale approaches that seek to elucidate the link between the single cell and population scales. In particular, throughout this chapter we draw attention to the issues of model complexity and the extrapolation of single cell behaviour onto the population scale.

2.1 Biology Underlying Chemotaxis in *E. coli*

Within this section we summarise the key biological processes that allow cells to exhibit bacterial chemotaxis. We begin by outlining how cells physically produce such a response before explaining how this is controlled by an intracellular signalling pathway.

Using a three-dimensional tracking microscope [10], Berg showed that unstimulated *E. coli* cells will move about their environment by executing a random walk [11]. These cells will move in (approximately) a straight line (runs), however these runs are interspersed with abrupt changes in direction (tumbles). This is often referred to as the chemotactic run and tumble swimming pattern (see Figure 2.1).

Within this run and tumble swimming pattern the direction of movement is altered at least once every few seconds [5]. In order to display chemotaxis, cells increase the length of runs when moving up an attractant gradient [12] (see Figure 2.1). *E. coli* cells utilise an intracellular signalling cascade (see Figure 2.2) to control the balance between runs and tumbles, thus allowing them to find areas containing beneficial substances.

This intracellular signalling cascade begins with the chemoreceptors of the cell (also called

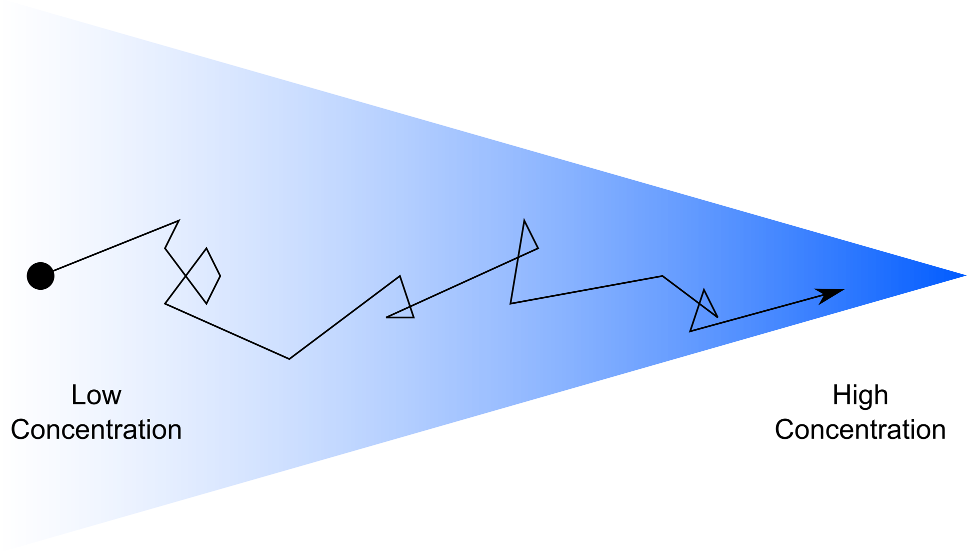


Figure 2.1: Chemotactic cells utilise a run and tumble swimming pattern in order to find regions containing beneficial nutrients. Runs act to propel the cell forward whereas tumbles act to randomly reorient the cell. When unstimulated, cells execute a three-dimensional random walk, exploring their environment. Upon sensing a beneficial attractant gradient, cells elongate their runs, biasing the random walk in the beneficial direction. This differs from the sensing of a negative gradient after which cells will increase the frequency of tumbles.

methyl-accepting chemotaxis proteins, MCPs). Each of these MCPs spans the cytoplasmic membrane of the cell, with the extracellular part able to bind molecules of certain attractants. Within *E. coli* cells there have been five types of chemoreceptor identified, namely

- MCPI (Tsr) [13] - identified as the serine chemoreceptor [14];
- MCPII (Tar) [14] - identified as the aspartate chemoreceptor [15];
- MCPIII (Trg) - identified as the chemoreceptor for ribose and galactose [16];
- MCPIV (Tap) [17] - later discovered to be the dipeptide chemoreceptor [18]; and
- MCP-like protein (Aer) [19, 20] - identified as the receptor for oxygen and has little or none of the extracellular part associated with full MCPs.

Each of the chemoreceptors identified are essentially very sensitive devices capable of detecting attractant concentrations as low as the nanomolar (nM, 10^{-9} molar) range [21, 22, 23].

Since bacterial cells are too small to sense a spatial change along the length of their bodies, the change in time of chemoreceptor's ligand occupancy is detected [24]. Each receptor type produces a signal dependent upon the sensing of their respective attractants which is then passed into the cell. In order to translate these signals into a signalling pathway response, the intracellular domains of chemoreceptors associate with a linker protein CheW. A histidine protein kinase (CheA) is then able to bind CheW, thus forming functional chemotaxis complexes that are localised to the poles of the cell [5]. Once these complexes are formed, CheA

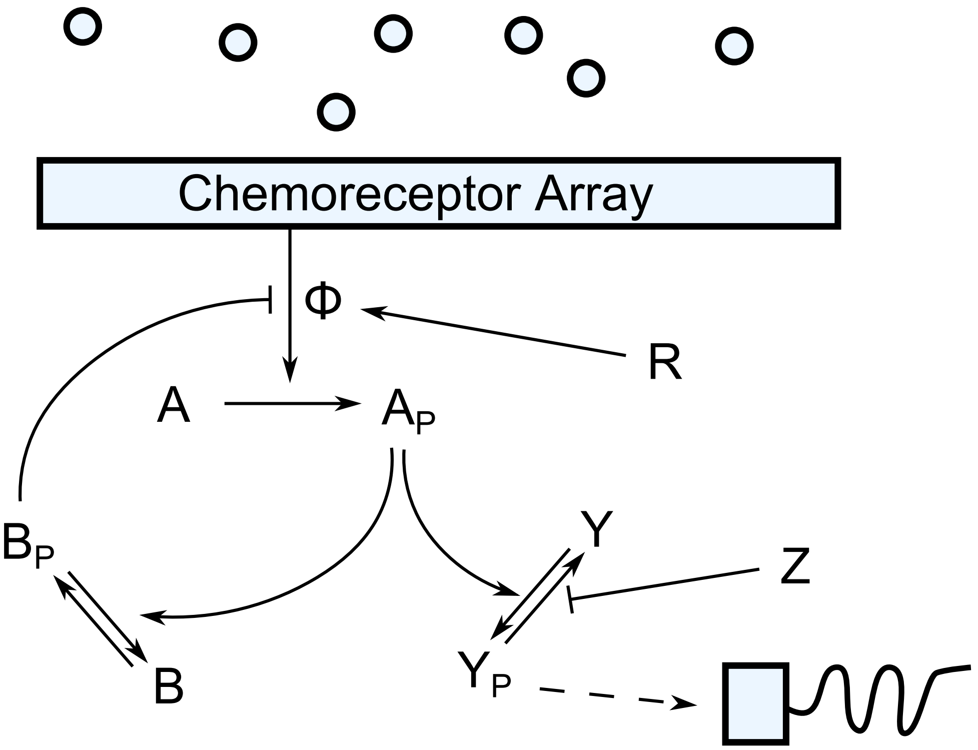


Figure 2.2: A schematic of the *E. coli* chemotaxis signalling pathway. Ligand molecules (blue circles) bind to receptors that together form an array of chemoreceptors. This affects the activity of the chemoreceptor array (Φ) and in turn, the rate of CheA (A) autophosphorylation, forming CheA-P (A_p). The phosphoryl groups produced are passed on to either CheB (B) or CheY (Y). Phosphorylated CheY (CheY-P, Y_p) diffuses through the cell toward the flagellar motor, binding to the motor driving protein FliM, altering the rotational behaviour. CheZ (Z) acts to dephosphorylate CheY-P, thus helping to terminate the signal. CheR (R) and CheB (B) form an adaptation module that acts to reset the cell to its pre-stimulus state. CheR adds methyl groups onto the chemoreceptors at a constant rate while CheB-P (B_p) acts to demethylate them. Note that CheW is neglected in this schematic due to the mathematical modelling assumption that there is always a sufficient quantity to allow for the formation of chemoreceptor, CheW and CheA complexes.

is able to autophosphorylate (forming CheA-P) at a rate dependent upon the activity of the chemoreceptors.

Once the CheA has been phosphorylated, the phosphoryl groups may be passed from CheA-P onto either the response regulator protein CheY or the methylesterase CheB [25]. Phosphorylated CheY (CheY-P) may diffuse within the cell cytoplasm toward the flagellar motors [26, 27]. Once the CheY-P has reached the flagellar motor, it is able to bind the motor-switching protein FliM causing a greater bias towards tumbling [28, 29]. In addition to autodephosphorylation of CheY-P, a phosphatase CheZ increases the rate at which this dephosphorylation occurs, thus giving more rapid termination of the signal [30].

Phosphorylated CheB (CheB-P) alongside the methyltransferase CheR forms the adaptation module of the signalling cascade, re-setting the chemoreceptors to their pre-stimulus state. Within this adaptation module the protein CheR constantly methylates chemoreceptors, thus increasing their activity [31] while CheB-P demethylates chemoreceptors, decreasing their ac-

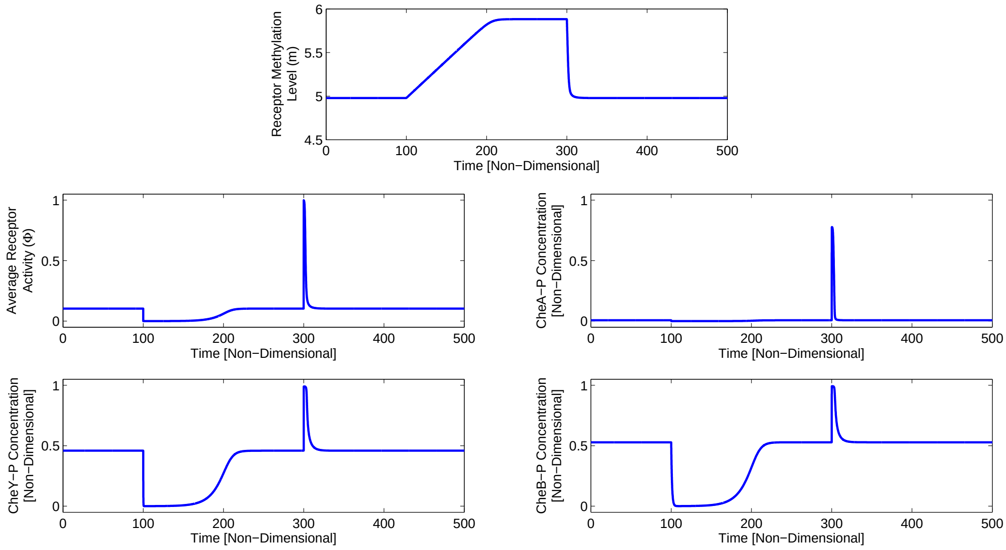


Figure 2.3: A graphical representation of adaptation in an *E. coli* cell. Shown here is the behaviour of receptor methylation, receptor activity (Φ) and phosphorylated fractions of CheA, CheY and CheB (i.e. CheA-P, CheY-P and CheB-P, respectively). Within this diagram the cell begins at a steady-state before the ligand concentration is increased at $\tau = 100$ and the decreased at $\tau = 300$.

tivity [32].

Upon detection of a positive change in the extracellular chemoattractant concentration the activity of the chemoreceptors falls and so does the rate of CheA autophosphorylation. This causes a subsequent drop in the CheY-P and CheB-P concentrations. The reduction in CheY-P concentration within the cell causes the flagellar motors to rotate counterclockwise (CCW) more frequently, inducing a chemotactic run. The reduction in CheB-P concentration allows the receptors to become methylated due to the action of CheR. This causes the receptor activity to rise, leading to a return to pre-stimulus levels of CheA autophosphorylation and subsequently those of CheY-P and CheB-P.

For a negative change, receptor activity and thus the rate of CheA autophosphorylation rises. This leads to an increase in the concentrations of both CheY-P and CheB-P. The flagellar motors will in this case rotate clockwise (CW) causing the cell to tumble more often. This is caused by the increase in the CheY-P concentration. The increased CheB-P concentration results in the demethylation of receptors, thus causing their activity to fall. This results in a return to pre-stimulus levels of CheA autophosphorylation as well as the CheB-P and CheY-P concentrations.

It is this ability of *E. coli* cells to return to their pre-stimulus state (via adaptation) that is one of the key factors enabling them to remain responsive over the observed five orders of magnitude in ligand concentration [33].

2.2 Mathematical Modelling of the Single Cell

From the early 1970s onward there has been a large amount of theoretical work aimed at understanding bacterial chemotaxis. The majority of this work has focused on elucidating key features of the signalling pathway (discussed in Section 2.1) in individual *E. coli* cells. Within this section we summarise the main approaches used to model ligand binding and adaptation; receptor sensitivity and gain; and the phosphorylation cascade. Throughout this section we pay particular attention to model complexity which has been an important issue in the analysis of past mathematical models. It is worth noting here that much of the theoretical work discussed was conducted prior to some of the key experimental work upon which current understanding is based.

2.2.1 Ligand Binding and Adaptation

Much of the early work on the mathematical modelling of individual chemotactic cells focused on representing the processes of ligand binding and adaptation. Within this section we summarise work that sought to postulate mechanisms associated with these processes prior to the existence of a full experimental understanding.

In 1972 Macnab & Koshland [34] proposed a three-variable model in which chemotactic bacteria may utilise some type of internal "memory" in order to respond to extracellular stimuli. The mechanism they hypothesised involved a compound W that could be rapidly converted to an alternative state X before the system is returned to its pre-stimulus level through a much slower degradation of X. Koshland [35] extended this work by observing that the control of enzymatic activity of compound X was most likely due to methylation of receptors.

Block et al. [36] examined the responses of bacteria to pulses of an attractant or repellent substance. In doing so they found that the timescales of excitation and adaptation are sufficiently separate as to conclude that each was controlled by a different process. In addition, they noted that cells must integrate attractant/repellent stimuli over a number of seconds. Using these results Block et al. [36] went on to propose a two-state model of flagellar switching (between counter-clockwise and clockwise states). In fact this was extended in order to consider different forms of stimulus and subsequently showed that cells must make temporal comparisons of the receptor occupancy [37]. In order to consider this, a model of light adaptation in *Phycomedes* [38] was adapted using an expression of the form

$$\frac{dA}{dt} = \frac{1}{\tau}\left(P - A\right), \tag{2.1}$$

where τ is the adaptation time and A is the adapting variable which changes according to receptor occupancy P. Block et al. [37] observed that this model did not account for the observed response thresholds and as such, they hypothesised that the difference between A and P could be accounted for by variation in receptor occupancy and methylation.

Goldbeter & Koshland [39] were the first to put forward a model that explicitly included receptor methylation. This took the form of a model in which receptors may exist in one of four states, representing the possible combinations of ligand bound and unbound receptors that may be either methylated or unmethylated. This model was subsequently extended by Asakura & Honda [40] to consider multiple methylation states which proceed in a preferred order and assumed that only unbound receptors may be methylated whilst only bound receptors may be demethylated. This results in an model consisting of two ODEs for each methylation state. Results obtained from this work showed that the model displayed excitation and adaptation and allowed them to conclude that methylation must have a dual role - allowing adaptation and increasing sensitivity to small changes in excitation.

Segel & Goldbeter [41] produced a similar four-state model to Goldbeter & Koshland [39] that displays exact adaptation for an appropriate parameter set, however the robustness of this was not examined. Hauri & Ross [42] added a description of CheB and CheY phosphorylation and dephosphorylation to a similar model (resulting in a ten-state model). However this failed to display observed levels of sensitivity and gain, nor did it display robust adaptation. Arocena & Acerenza [43] considered receptor modification via attractant binding rather than methylation or phosphorylation. This displayed a response over the experimentally observed five orders of magnitude in ligand concentration but did not match experimentally observed levels of sensitivity. They go on to propose that a wider response range could be caused by clustering of chemoreceptors.

More recently, Barkai & Leibler [44] constructed an adaptation model for a single receptor species that is based upon methylation of receptors. This was initially considered in isolation and as such did not consider the phosphorylation pathway. Here receptor methylation was assumed to be caused by CheR acting to constantly methylate inactive receptors while CheB demethylated active receptors at a rate dependent on the receptor activity. Parameter sensitivity analysis showed that this model was robust for a wide range of parameter values. Further to this, the model represents an important step since it was able to produce similar behaviour to previous models without the large quantity of ODEs.

It is worth noting that whilst multiple methylation state models such as those discussed above (and more recently Xin & Othmer [45]) may produce a reasonable comparison with experimental data, they often produce large systems of ODEs (especially when coupled with a signalling cascade model) from which it may be difficult to obtain an intuitive understanding. In addition to this, high order mathematical models are significantly more difficult to work with due to the large number of parameters and variables and can often make analytical approaches essentially impossible. The work of Barkai & Leibler [44] went some way to addressing this issue however their model did not include a representation of the signalling pathway, thus it was of limited use without further work.

2.2.2 Chemoreceptor Sensitivity and Gain

Whilst the models of Section 2.2.1 displayed adaptation and responses to ligand changes, they did not capture the experimentally observed levels of receptor sensitivity and signal amplification (gain) [46]. This feature allows cells to produce a chemotactic response in cases where only a very small change in the extracellular chemoattractant concentration is sensed. In particular, Sourjik & Berg [47] showed that the fractional change in kinase activity is amplified by a factor of ∼35 relative to the fractional change in receptor occupancy. Furthermore they state that this signal amplification occurs at the beginning of the intracellular signalling pathway and is capable of explaining the significant degree of gain observed in *E. coli* cells when combined with the existence of an ultrasensitive flagellar motor such as that described by Cluzel et al. [29]. As such, within this section we summarise some of the most common (and successful) approaches used to study this problem.

Bray et al. [48] were perhaps responsible for one of the most significant contributions to the study of receptor sensitivity and gain. Similar to the idea of Arocena & Acerenza [43], it was suggested that rather than working individually, receptors would function together (as a cluster). They began by investigating the effects of allowing neighbouring receptors to influence each other's state. In a mechanism such as this the inactivation of one receptor will result in one or more nearby receptors becoming inactivated due to nearby neighbour interactions. In particular, they noted that sensitivity increased as a result of this mechanism but receptors must adapt in order to display an appropriate range of ligand response.

In 2004, Bray & Duke [49] summarised the experimental data that indicated conformational changes (changes in state) could spread. Earlier, they had considered a statistical mechanics model (analogous to an Ising model, discussed in Section 2.2.2) in order to consider propagation of conformational changes in a ring of proteins [50]. Results here showed that rather than automatically switching between states, increasing numbers of proteins changed state as a signal propagated via nearest neighbour interactions.

It is clear from the work discussed here that a large amount of the theoretical work studying chemoreceptor sensitivity and gain has been centred on the idea of receptor clustering. The central idea behind this is that external signals will be amplified since ligand bound receptors will be able to alter the activity of their nearby neighbours (known as conformational spread) when they are part of the same cluster. Whilst the work summarised above made some progress there are two main mathematical models that have been key to the understanding of chemoreceptor sensitivity and gain. These are the Ising and Monod-Wyman-Changeux models discussed in the following sections.

Ising Models

One model successfully used in the early 2000s to study receptor sensitivity and gain is the Ising model. Such models were initially used to consider electron behaviour in the presence of a magnetic field. It has subsequently been considered an appropriate model for the study of receptor coupling since it allowed for the effects of particle coupling, i.e. the ability of particles to change the state of their near neighbours.

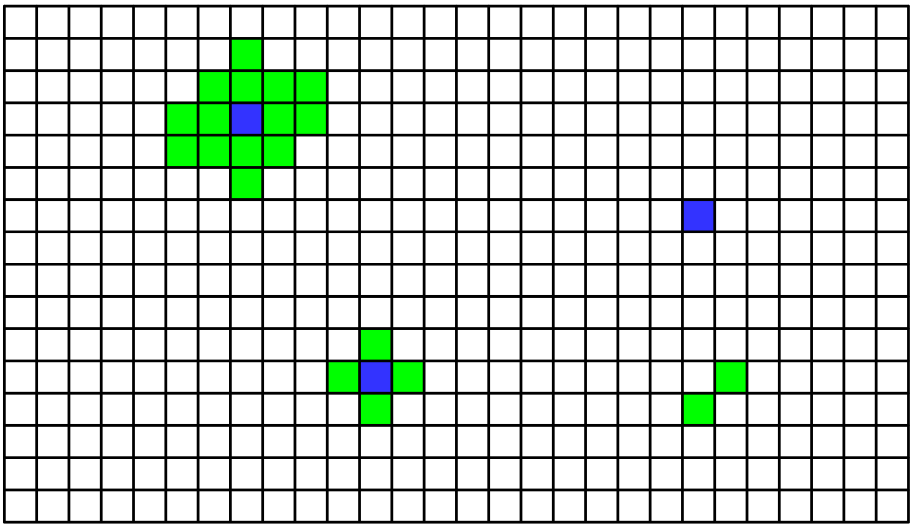


Figure 2.4: A graphical representation of the Ising model. Each square represents a single receptor that may exist in either active or inactive states. A non-ligand bound receptor (white square) will be active whilst a ligand bound receptor (blue square) will be inactive. Depending upon the coupling strength a number of close neighbours will be inactivated (green squares). Displayed are examples with large (top left), small (bottom left) and zero (top right) coupling strengths. Also shown is an example of the thermal fluctuation (noise) that exists in this type of model (bottom right).

Within an Ising model receptors are allowed to exist in two states ($i = \pm 1$) between which they may switch. The total energy of an array of receptors is calculated using a Hamiltonian energy functional (H) of the form

$$H(t) = -\sum_{\langle ij \rangle} J_{ij} S_i S_j - \sum_i B_i(t) S_i, \tag{2.2}$$

where S_i and S_j represent the “spin” (so called from the original physics definition) of the i-th and j-th receptors, respectively, J_{ij} denotes the coupling strength between receptors and $B_i(t)$ is the attractant concentration. When no attractant binding occurs, each receptor may freely switch between their two states.

The first application of this type of model to chemoreceptor coupling was due to Shi & Duke [51]. Using mean-field theory they demonstrated that the coupling strength of neighbouring receptors had a large effect upon the sensitivity of a receptor array, however adaptation was not studied in great detail within this work. Duke & Bray [52] later conducted Monte Carlo simulations of this model in which each receptor was coupled to its four closest neighbours. These simulations were able to show that changes in attractant concentration could be detected over the experimentally observed five orders of magnitude.

Shimizu et al. [53] incorporated an Ising type model into a spatially extended version of the algorithm used by the StochSim (software for stochastic simulation of biochemical systems) program alongside a description of methylation and adaptation similar to that of Barkai & Leibler [44]. This was used to examine the effects of different receptor array geometries, leading to the conclusion that the largest gain was obtained where there were more nearest neighbours.

Shi [54] later incorporated the effects of CheB-P and CheR (adaptation) into the earlier model [51]. In order to do this, a differential equation of the form

$$\frac{dS_i}{dt} = -\sigma S_i\left(t - t_r\right), \tag{2.3}$$

was combined with equation (2.2). Within this, σ is a rate constant and t_r is a delay time representing the sum of time taken for demethylation and phosphotransfer from CheA-P to CheB. The vastly different timescales of ligand binding, phosphorylation and adaptation allow the Hamiltonian to be coarse grained and as such Shi [54] showed that the effects of CheB-P and CheR are sufficient to return receptor activity to a pre-stimulus state (perfect adaptation). Finally, this work noted that the adaptation mechanism allowed receptor coupling strength conditions to be relaxed, hence it actually improves the robustness of the model. Mello and colleagues [55, 56] considered an alternative extension including adaptation similar to that of Barkai & Leibler [44]. Mean-field analysis of this model found it produced good agreement with experimental data for a number of mutant cell types.

Shi [57] later compared their earlier model results to experimental data and in doing so found good agreement with a number of experimental data sources. It was also noted here that whilst modelling approaches used previously had been informative, it was not always the case that the physical meaning of each component was clear. This has been pointed out as one of the main strengths of Ising type models. Whilst such models clearly produce good agreement with much of the experimental data, their main weakness is in their computational nature. This can make it difficult to intuit how the model will function in different situations.

Allosteric Models

The Ising models discussed in Section 2.2.2 clearly had some success in terms of comparison with experimental data. Around the mid-2000s however, theoretical modelling of receptor sensitivity and gain began to utilise Monod-Wyman-Changeux (MWC) [58] models. Such models assume an array of receptors is divided into several smaller clusters. Within these smaller clusters all receptors are assumed to be in the same (active or inactive) state. It is also assumed that separate clusters do not interact with others. Tu [59] was able to show that this represents a special case of the Ising model within which the interaction strength, J_{ij}, between receptors is infinite for those of the same small cluster but zero for those of different clusters. In a model of this type the average receptor activity is described by

$$\Phi = \frac{L\left(1 + [L]/K_i\right)^N}{L\left(1 + [L]/K_i\right)^N + \left(1 + [L]/K_a\right)^N}, \tag{2.4}$$

where $[L]$ is the attractant concentration, $K_{a/i}$ are the ligand dissociation constants of active and inactive receptors, respectively, N is the number of receptors in the cluster and L is an equilibrium constant. Further discussion on the relative strengths and weaknesses of this type of model can be found in the review of Tu [59].

Keymer et al. [60] were able to use an MWC model to investigate changes in receptor activity due to the addition of an attractant (aspartate). In doing so they were able to show good agreement with experimentally obtained fluorescence resonance energy transfer (FRET) data for various different receptor methylation states. Within this work the receptor methylation level was considered as a change in the offset energy (a component of the equilibrium constant). Mello & Tu [61] took a similar approach, however they generalised the MWC model in order to consider heterogeneity in receptor clusters, i.e. they considered different receptor types within clusters as well as the effects of binding two different attractants.

Endres et al. [62] later used an MWC model in order to investigate the idea that the level of receptor cooperativity is dynamic (i.e. cluster size varies with receptor modification). In order to do this they produced model fits of experimental dose-response data by varying the number, N, of receptors within a cluster.

A similar model was also used by Mello & Tu [33] in order to investigate the sensitivity of receptor clusters in different background attractant concentrations. They took a model of the form of equation (2.4) and defined the sensitivity as the change in receptor cluster activity with

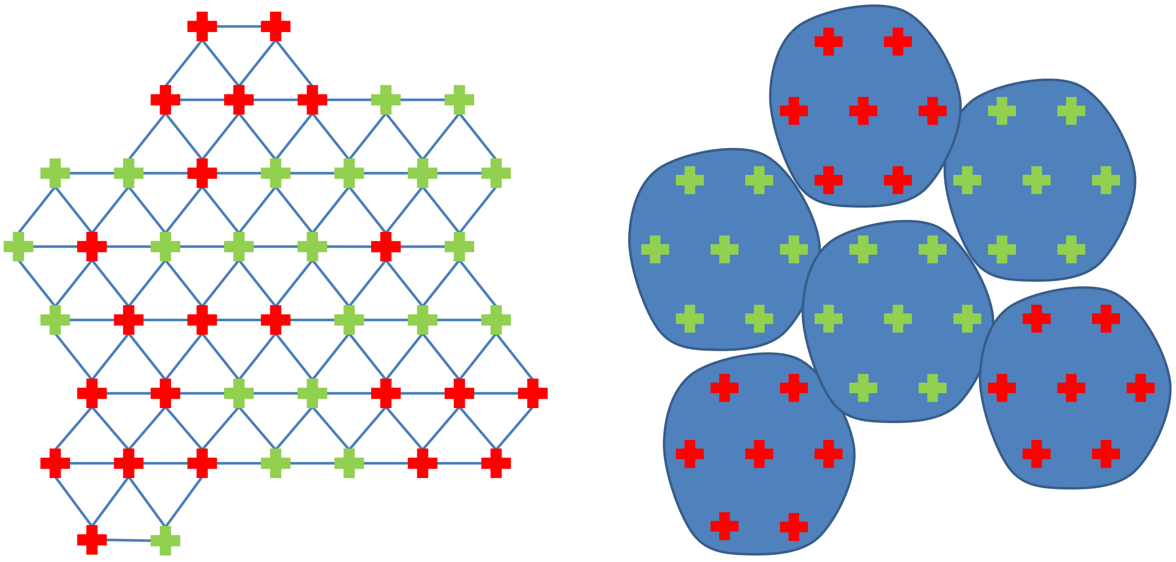

(a) Ising Model (b) MWC Model

Figure 2.5: A graphical representation of the differences between the Ising and MWC modelling approaches. (a) Within Ising models receptors may exist in either active (red) or inactive (green) states. Depending on the structure of the lattice (trigonal in this example), each receptor may affect the states of a number of near neighbours. (b) In MWC type models receptors exist in small "teams" (blue blobs) within which they are all either active or inactive simultaneously. Each MWC team is equivalent to a small Ising model lattice with an infinite coupling strength. This diagram is adapted from the work of Tu [59].

respect to a change in the attractant concentration,

$$S \equiv -\frac{\partial \ln \Phi}{\partial \ln [L]}. \tag{2.5}$$

Using this they were able to show that for low to intermediate aspartate concentrations the effect of low affinity binding of aspartate to Tsr receptors is negligible. However, for high concentrations of aspartate the low affinity binding of aspartate to Tsr receptors begins to have an effect.

Tu et al. [63] continued the theoretical study of this model type by considering the effects of different time varying forms of the attractant stimulus. In particular, they considered exponential ramp, exponentiated sine wave and large step stimuli by utilising a variety of analytical techniques. Kalinin et al. [64] experimentally studied the swimming responses of cells in different (well controlled) concentration gradients. In doing so they also considered the contribution of methylation to the free-energy of a receptor (offset energy). It was concluded that this must be a linearly decreasing function of receptor methylation level.

Endres et al. [65] studied the effects of methylation on kinase (CheA) activity. This work followed a similar approach to previous literature [60, 61]. As with this previous work, different values were used for the offset energy in order to study the effects of different methylation states.

The adaptation model of Barkai & Leibler [44] was extended by Meir et al. [66] to allow the incorporation of an MWC description of receptor clustering. In this work they discuss the existence of cell-to-cell variation in methylation and demethylation rates which is proposed to be a survival/bet hedging strategy allowing for population survival in varied environments. In addition they note the ability of their model to display the experimentally observed asymmetry in adaptation time for addition or removal of attractant.

The experiments of Li & Hazelbauer [67] were the focus of work by Endres & Wingreen [68]. Li & Hazelbauer [67] observed that the adaptation proteins CheB and CheR can bind an approximately 35 amino acid tether allowing them to act upon groups of between five and seven receptors, termed an "assistance neighbourhood". The work of Endres & Wingreen [68] then sought to examine this effect by incorporating an adaptation model, similar to Barkai & Leibler [44], into an MWC one. In doing so they found two different responses for high attractant concentrations. In particular, it was noted that the response could either be terminated by receptor saturation (i.e. no further ligand binding is possible) or for large stimuli, receptors may become fully methylated. At this point, low affinity binding of aspartate to Tsr receptors can allow the cell to respond to further stimuli. Hansen et al. [69] extended this by explicitly considering the actions of CheR and CheB in the form of an ODE for the average methylation level of receptors within a cluster.

More recently, Clausznitzer et al. [70] considered an MWC modelling approach combining much of the MWC modelling work discussed above. Within this work the kinetics of the average receptor methylation level (m) were examined by considering temporal changes in receptor cluster activity

$$\frac{\partial \Phi}{dt} = \frac{\partial \Phi}{\partial m}\frac{dm}{dt} + \frac{\partial \Phi}{\partial [L]}\frac{d[L]}{dt}. \tag{2.6}$$

In considering the time immediately after a step change in the attractant concentration, $[L]$, this reduces to

$$\frac{\partial \Phi}{dt} = \frac{\partial \Phi}{\partial m}\frac{dm}{dt}, \tag{2.7}$$

since $d[L]/dt = 0$. It was then possible to compare potential mechanisms for temporal changes in the methylation levels. Within this work they were able to produce a functional form for the offset energy as well as an ODE describing the average methylation level of chemoreceptors. As such, the model produced here was able to produce good agreement with experimental data.

Skoge et al. [71] compared results of both spatially one- and two-dimensional Ising models to those obtained from MWC type models and experimental data. They were able to establish which produced the best results for a range of mutant cell types. This analysis showed that the MWC model produced better agreement with experimental data for wild-type and a number of mutant cells.

It is clear from the work summarised here that MWC models are able to produce good agreement with the experimental literature in terms of receptor sensitivity and gain. In particular, they are able to display a response over the experimentally observed five-orders of magnitude in ligand concentration (i.e. they can detect MeAsp concentrations in the range from 10^{-3}mM to 10^{2}mM [47]). Such models have been highly popular in recent years due to their relative simplicity in comparison with alternative models. This has allowed researchers to produce theoretical analyses that were previously unmanageable using large ODE models. Thus, MWC models are proving an essential tool in the study of chemotaxis in single *E. coli* cells.

2.2.3 The Phosphorylation Cascade

Within Section 2.1 we summarised the workings of the *E. coli* intracellular signalling pathway. In doing so we referred to a large amount of experimental literature that has helped to elucidate mechanisms in this cascade. The availability of this experimental literature inspired an increasing amount of theoretical work that sought to mathematically model this aspect of the chemotactic response. In this section we summarise some of the key work in this area, however the review of Tindall et al. [72] provides a much more comprehensive overview.

In 1993 Bray et al. [73] became the first to produce a mathematical model of all known reactions in the phosphorylation cascade. This model, built in to a computer program called BCT (Bacterial ChemoTaxis), is essentially an extension to the work of Block et al. [36] in that interactions between flagellar motors and CheY-P are incorporated. However, it failed to include the effects of receptor methylation. In spite of using an extensive set of experimentally determined parameter values, this model failed to display adaptation or observed levels of gain. In contrast to this Spiro [74] and Spiro et al. [75] produced a model that includes attractant binding, methylation, phosphorylation and motor bias. This model was able to display adaptation but also failed to produce appropriate levels of sensitivity or gain.

Alon et al. [76] later considered an extension to the work of Barkai & Leibler [44] which included a phosphorylation pathway description. This work demonstrated that variation in protein concentrations causes changes in steady-state behaviour and adaptation time but that perfect adaptation remains robust. Yi et al. [77] later showed that these models are examples of integral feedback control systems, for which robustness is known to be an inherent property.

As in Alon et al. [76], the work of Levin et al. [78] studied the implications of variation in protein concentration upon the phosphorylation pathway. They developed a “fine-tuned model” that considers an adaptation mechanism similar to that of Segel & Goldbeter [41] and a “robust model” merging the receptor kinetics of Asakura & Honda [40] and methylation of Barkai & Leibler [44]. Varying the concentration of individual proteins displayed relatively similar results in terms of phosphorylation levels, however the “robust model” maintained the ability to adapt where the “fine-tuned model” failed, suggesting robustness is a desirable feature.

It was discussed in Section 2.2.2 that Clausznitzer et al. [70] formulated an MWC model and adaptation mechanism that produced a good fit with experimental data. In addition to this, they incorporated a description of the phosphorylation pathway. As such, this mathematical model was able to produce a good fit to FRET experimental results, contained all of the key processes and is simple enough as to be amenable to mathematical analysis.

Within the modelling of the intracellular signalling cascade there has been relatively little work aimed at studying the effects of spatial organisation of chemotaxis signalling network components. Lipkow et al. [26] utilised three-dimensional stochastic simulations of individual protein molecules (using Smoldyn, based on Smoluchowski dynamics [79]) to study the effects of CheZ localisation on the distribution of CheY-P. This showed that polar localisation reduces the CheY-P gradient observed along the length of the cell, a result also shown by Rao et al. [80] using a partial differential equation (PDE) approach. This was later extended to study dynamic protein localisation [30] whereby CheZ localisation depends upon the attractant/repellent concentration. For a positive attractant gradient, CheY-P concentration will fall allowing CheZ to diffuse freely within the cell. Upon detection of a repellent, the concentration of CheY-P rises causing CheY-P, CheZ and CheA$_s$ to oligomerise at cell poles. The close proximity of CheY-P and CheZ leads to enhanced CheY-P dephosphorylation and as such a negative feedback upon CheY-P concentration has been introduced. Lipkow [30] goes on to state that this mechanism represents a robust second tier of adaptation for chemotactic cells.

2.3 Mathematical Modelling of Cell Populations

It is clear from Section 2.2 that a great deal of both experimental and theoretical work has sought to explain chemotaxis in single *E. coli* cells. In addition to this work, there has also been a significant body of work focused on describing chemotaxis on the scale of cell populations - see for example the review of Tindall et al. [81].

In the 1960s, Julius Adler began his work on chemotaxis by studying the migration of bacterial cell populations [82, 83]. This was done using agar plates (often referred to as swarm plates) similar to those used in pioneering work by Beijerinck & Sherris and coworkers [84, 85, 86]. These consist of a petri dish containing a jelly like substance (agar) to which different attractant chemicals are added. In Adler’s experiments he placed *E. coli* cells in the centre of one of these plates. As these cells consume the attractant chemicals (in this case serine, aspartate, threonine and oxygen) they form bands of cells depending upon the consumption of the attractants.

Firstly, some cells will move outward using up serine until they reach the outer ring of the serine. Then, of those which remain, some will move outward in search of aspartate, forming a second ring. Finally, some of the remaining cells will search for threonine, forming a third ring. The observation of these chemotactic bands was a key factor in sparking an interest in the modelling of population scale behaviour.

2.3.1 Keller-Segel Models

Throughout the literature it is most common, when attempting to model the behaviour of chemotactic bacterial populations (such as chemotactic band formation), to consider a mathematical model such as that devised by Keller & Segel [87]. Commonly referred to as the Keller-Segel (or K-S) model, this was originally created for slime moulds but has been successfully applied to bacterial chemotaxis [88]. This model takes a continuum approach and uses two partial differential equations (PDEs) to represent a population density and concentration of some attractant substance across a spatial domain. A generalised version of this model may be written as:

$$\frac{\partial b}{\partial t} = \nabla \cdot (\mu(s)\nabla b) - \nabla \cdot (\chi(s) b \nabla s) + g(b,s) - h(b,s), \tag{2.8}$$

$$\frac{\partial s}{\partial t} = D\nabla^2 s - f(b,s), \tag{2.9}$$

within which $b(\mathbf{x},t)$ represents the population density, $s(\mathbf{x},t)$ the attractant concentration, $\mathbf{x}$ is the spatial position, t denotes time, χ is the chemotactic coefficient, g indicates cell growth, h cell death, f indicates degradation of attractant whilst μ and D are the diffusion coefficients of the bacteria and attractant, respectively. For more detailed information on the impact of the K-S model in this and other applications we recommend the reviews of Horstmann [89, 90].

It is clear upon examination of the literature that the K-S model has been useful within the study of bacterial chemotaxis of cell populations. One of the most notable areas in which this model has been of benefit is in explaining the formation of various patterns observed in petri-dish experiments (see Erban & Othmer [91] for a more comprehensive review). Much of this work seeks to find parameter limits that lead to diffusion-driven (Turing) instabilities, for example Zhu & Murray [92] and Maini et al. [93]. Using the K-S model, Keller & Segel [94] were able to determine that in order to produce chemotactic bands, the chemotactic coefficient must be larger than the bacterial diffusion coefficient (i.e. $\chi/\mu > 1$). In addition to this Brenner et al. [95] considered more complex patterning in a novel case whereby cells secreted the chemoattractant to which they respond.

Segel subsequently sought to understand how microscopic (individual) behaviour acts to influence the behaviour observed on the macroscale (population) level by incorporating receptor dynamics into a generic population model [96]. Later, by assuming the turning probability of the bacteria depends upon the temporal rate of change of receptor occupancy a better fit to experimental data was obtained [97].

Unfortunately, whilst K-S models have been useful in certain settings, they do not tell the whole story. The main issue associated with such models is that they do not explain how population scale phenomena are affected by the intracellular features of single cells. As such, in spite of the work of Segel [96, 97] and others [98, 99, 100] that have begun to address the issue, further work would be required in order to bridge the gap between single and population scale modelling of chemotactic *E. coli* cells.

2.3.2 Stochastic Models

Also considered within the literature are stochastic approaches for explaining the behaviour of chemotactic cell populations. A key piece of work in the field is that of Patlak [101] who studied random walks in which there exists a correlation between the orientation in consecutive time steps (i.e. a correlated random walk, CRW). This type of random walk model tends to be applicable in biology since most organisms have a tendency to move forward [102]. Among the most cited examples of work extending that of Patlak [101] is that due to Alt [103] who considered a biased and correlated random walk (BCRW) that is applicable to cases in which the bias is provided by a response to a spatially varying chemical gradient. This gave a governing stochastic equation of the form:

$$\frac{\partial \sigma}{\partial t} + \frac{\partial \sigma}{\partial \tau} + \theta \cdot \nabla (c\sigma) = -\beta\sigma, \tag{2.10}$$

for $\tau > 0$. Within this expression $\sigma(t, \mathbf{x}, \theta, \tau)$ is the bacterial density distribution which turns an angle θ, and has a run time of τ. Also defined are $c(\mathbf{x}, t)$ which represents the mean run speed, $\beta(\mathbf{x}, t, \theta, \tau)$ the tumble frequency distribution, the turning angle η which is derived from a turn angle distribution $k(\mathbf{x}, t, \theta; \eta)$, s is the unit sphere in n-dimensional space and σ is defined for a new run in a new direction by:

$$\sigma(\mathbf{x}, t, \eta, 0) = \int_0^\infty \int_s (\beta\sigma) k \; d\theta \; d\tau. \tag{2.11}$$

The main issue associated with this approach is that, like the K-S model, it does not help in the understanding of how intracellular signalling processes act to produce observed population scale behaviour. Another criticism that has been levelled at this model is that it is difficult to solve analytically [104] and thus it does not necessarily lend itself to gaining insight into how certain behaviours emerge.

Other stochastic approaches have also been considered such as Othmer et al. [105] and Rivero et al. [106]. The Rivero et al. [106] (RTBL) model considered a cell population moving either left or right along an infinite one-dimensional line with a constant velocity v (other methods of motility are discussed by Codling et al. [102]). Considering turning probabilities of the form $r_1 = \lambda_1 \tau$ and $r_2 = \lambda_2 \tau$ for left- and right-moving cells, respectively, it can be shown that under certain conditions the problem simplifies to the biased telegraph equation along a one-dimensional line

$$\frac{\partial^2 p}{\partial t^2} + (\lambda_1 + \lambda_2)\frac{\partial p}{\partial t} + v(\lambda_2 - \lambda_1)\frac{\partial p}{\partial x} = v^2 \frac{\partial^2 p}{\partial x^2}, \tag{2.12}$$

where $p(x,t)$ denotes the total population density. This model sought to address the main issue of the Alt [103] model, namely the difficulty of finding an analytical solution. In addition to this, by considering temporal changes in receptor occupancy, this model represented a starting point down the road toward explaining how single cell traits affect population level behaviour.

There is also a fair amount of work within the literature that compares the models discussed above. Ford & Cummings [107] chose to compare three models, namely those due to Alt [103], Segel [97] and Rivero et al. [106]. In doing so they were able to find a number of conditions under which these various models were analogous to one another. Ford & Cummings [107] conducted a similar comparison using numerical simulations, concluding that while the Alt [103] model gives a detailed description of chemotactic aggregation, the RTBL model [106] gives an adequate continuum based description of chemotaxis.

Similar to the K-S models discussed in Section 2.3.1, the stochastic modelling approaches discussed here have been useful in certain areas, although they do not explain how intracellular signalling processes affect the emergence of population level phenomena.

2.4 Multi-Scale Modelling Approaches

Within Sections 2.2 and 2.3 we have summarised some of the key literature that has sought to understand both single cell and population level processes. Each of these approaches has yielded a number of interesting mathematical problems as well as insights that may not have been available via experimental approaches alone. However, with recent advancements in laboratory equipment and methods there is an ever increasing amount of experimental data available to mathematical modellers. Coupling this with vast increases in computing power, it has increasingly become possible to devise mathematical methods that bridge the gap between the different scales associated with bacterial chemotaxis.

To date there has been relatively little work focused on using single cell experimental data and mathematical models to inform how population level behaviour is controlled. Here we review some of the main approaches used to tackle this problem.

2.4.1 Agent-Based Models

Agent-based (also known as individual-based) modelling is one of the main approaches that has been utilised in order to understand the ways in which single cell behaviour affects that of a cell population. Such models are computational in nature and use a set of ‘rules’ allowing the effects of individual cell attributes/behaviours to be measured on the population scale.

One such example is that of Emonet et al. [108] which sought to examine how stochasticity in the chemotactic signalling network would impact upon population level behaviour. The main drawback of this model is that it failed to investigate how variation in other features such as receptor sensitivity, adaptation and motor response may impact upon population level behaviour. Additionally it failed to properly account for the relationship between the flagellar rotation bias and the CheY-P concentration.

Bray and colleagues have produced agent-based models (ABMs) that give a good level of agreement with experimental data. In contrast to the model used by Emonet et al. [108], these models aimed to capture a greater level of realism by incorporating a highly detailed ($\sim$90 ODE) model of the *E. coli* chemotaxis signalling cascade [109, 110]. They also incorporated a significant amount of biophysical data into their model so as to produce a good approximation of the swimming behaviour of individual cells.

The models summarised here have been able to provide good agreement with various experimental literature. It is clear therefore that models such as these could be of great benefit in allowing us to bridge the gap between single cell and population scale models. Each of these models, however, has their own drawbacks. In particular, these models either do not properly account for all processes or contain large ODE models making it difficult to understand which processes (eg. receptors, signalling pathway, motors) are affected by any given variation and how this will affect the output. It is clear therefore that further work is required to understand how the single cell and population scales are linked.

2.4.2 Equation-Free Models

In this section we briefly summarise another of the main methods that has sought to explain how single cell behaviour affects that of a cell population, namely equation-free methods. Models such as this consider behaviour on a coarse scale as well as incorporating a more detailed approach. It is worth noting here that, in spite of the name, such models do consist of a set of equations.

Equation-free methods have notably been used by Erban & Othmer [100, 91] and Setayeshgar et al. [111]. Setayeshgar et al. [111] showed that the large separation in excitation and adaptation time allows evolution of the cell population to be coarse grained. Erban & Othmer [100] incorporated a simplified microscopic model of the *E. coli* chemotaxis signalling pathway into a telegraph process (similar to equation (2.12)), subsequently showing that the chemotactic response vanished as the adaptation time tends to zero. Results were then generalised to higher dimensions.

Models such as those discussed here clearly go some way toward bridging the gap between single cell and population scale behaviour. These also address the important issue of being able to simulate such behaviour within a reasonable time [91]. However, the main drawback of such models is that they do not produce the same degree of detail as other approaches (such as agent-based modelling) since they consider a coarse grained representation of the population scale behaviour. Thus, they neglect to include the individual behaviour of each individual cell in favour of a continuum approximation of the overall behaviour.

2.5 Chapter Summary

In this chapter we presented a summary of some key experimental and theoretical literature associated with chemotaxis in *E. coli.* We began by outlining the biological processes that allow cells to exhibit chemotaxis. This was followed by brief reviews of mathematical approaches used to model chemotaxis both of individual cells and cell populations. We subsequently summarised some of the key work that has sought to link the two scales.

The review of single cell chemotaxis modelling focused on work related to ligand binding and adaptation; chemoreceptor sensitivity and gain; and the phosphorylation cascade. In doing so, it is clear that the issue of model complexity has been a significant limitation associated with many of the models discussed here.

Within the summary of population scale theoretical work we focused on two main approaches, namely Keller-Segel and stochastic models. Each of these approaches was shown to be useful in certain settings. However, they clearly possess one main drawback in that they do not consider how single cell features impact upon the overall behaviour of the population.

Multi-scale modelling approaches were then discussed since these seek to bridge the gap between single cell and population models. Here we focused on agent-based models and equation-free methods. Each of these was shown to have both good and bad points. In particular, agent-based models have been produced which display clear potential but either fail to account for certain processes or suffer from the same model complexity issue as many single cell models.

It is clear that many different approaches have been considered when modelling chemotaxis. Each of these has their own respective advantages and disadvantages. Within the more recent literature, however, a number of these drawbacks have started to be addressed. In particular, MWC modelling has helped to remove significant amounts of complexity from models whilst retaining (simplified) representations of all key biological processes. In addition to this, the emergence of agent-based modelling has allowed researchers to investigate how single cell phenomena affect the behaviour of populations.

Within the remainder of this thesis we seek to take advantage of these developments. In particular, the use of MWC models opens up avenues of investigation that were previously hampered by model complexity whilst agent-based modelling allows single cell investigations to be examined in terms of their impact on the population scale.

Chapter 3

Mathematical Model of *E. coli* Chemotaxis Signalling

Within this chapter we analyse a recent mathematical model of the *E. coli* chemotaxis signalling cascade [70]. Firstly, the model is presented alongside a discussion of how it allows for further investigation of intracellular phenomena where other models from the literature do not. A rigorous analysis of the model is then conducted. In particular, we compare a number of functions describing the methylation dependent free-energy (offset energy) of chemoreceptors against experimental data in order to choose the most suitable form for use throughout this thesis. A non-dimensional re-scaling of the model system is then considered. Using this non-dimensional model it is demonstrated that the system of equations possesses just one biologically feasible steady-state that is subsequently shown to be asymptotically stable. Finally, the eigenvalues of this system are examined in order to assess the stiffness of the model system, thus impacting on the choice of numerical methods used in this thesis.

3.1 *E. coli* Chemotaxis Signalling Pathway Model

In this section we summarise and discuss a mathematical model of the *E. coli* chemotaxis signalling pathway due to Clausznitzer et al. [70]. In particular, we begin by focusing on the motivation behind this choice of model before outlining the model and briefly mentioning the methods utilised by Clausznitzer et al. [70] in its derivation.

3.1.1 Motivation

Within the literature there has been a wide range of approaches considered for the modelling of chemotaxis signalling cascades in individual *E. coli* cells (see Chapter 2). Here we briefly summarise the key strengths and weaknesses of the main approaches considered by way of motivation for the use of the Clausznitzer et al. [70] model presented within this work.

Many mathematical models have focused on describing the processes associated with ligand binding and adaptation [39, 40, 41, 45]. A common approach here was to consider separate model variables for each possible receptor state (i.e. ligand bound/unbound and each possible methylation state). This approach was able to produce good agreement with some experimental data. However the resulting large numbers of non-linear ODEs do not lend themselves to insightful analysis.

In stark contrast to the above approach, it is also common within the literature to produce a minimal model that is capable of demonstrating certain behaviour. One such example is that of Tu et al. [63] which consists of just one ODE for receptor methylation level and a decoupled read-out variable representing CheY-P concentration. Such models have been used to great effect within the literature in order to elucidate specific features of chemotaxis signalling pathways. In spite of this, these models are not always suitable since they do not include a detailed enough representation of the whole signalling pathway.

Within this thesis we analyse the model due to Clausznitzer et al. [70]. This recent mathematical model takes a different approach to those previously considered by making use of the law of mass action and the Monod-Wyman-Changeux (MWC) model. The result is a model of five non-linear ODEs that has been shown to produce a good fit to a range of experimental data. In particular, this model includes a representation of all key signalling pathway processes, agrees with experimental data and is small enough as to be amenable to analysis. Since each reaction/process within this model relates directly to a physical process, it is also simple to utilise the vast experimental literature in order to produce a fully parameterised system (see Table 3.1). As such, this model has opened up many avenues of investigation that were not previously possible.

3.1.2 Model Details

Here we outline the Clausznitzer et al. [70] model of the *E. coli* chemotaxis signalling cascade that is studied within this thesis. Within this section we provide a summary of the model alongside a brief description of its derivation. In addition to this we mention the key assumptions underlying this model and state a number of simplifications considered in this work.

Within *E. coli* cells the spatial distribution of proteins within the chemotaxis signalling cascade is not generally considered to play a significant role. As a result of this it is possible to consider here a mathematical model that assumes no spatial heterogeneity exists for any of the signalling proteins. This lack of spatial heterogeneity is one of the key assumptions underlying the law of mass action (see Appendix A). As such Clausznitzer et al. [70] applied the law of mass action to reactions within the intracellular signalling cascade (see Figure 2.2), giving the non-linear ODEs

$$\begin{aligned}
\frac{d[A_p]}{dt} &= \Phi k_1([A_T]-[A_p]) - k_2[A_p]([Y_T]-[Y_p]) - k_3[A_p]([B_T]-[B_p]), && (3.1)\\
\frac{d[Y_p]}{dt} &= k_2[A_p]([Y_T]-[Y_p]) - k_A[Y_p]([Z_T]-[Y_pZ]) + k_D[Y_pZ], && (3.2)\\
\frac{d[Y_pZ]}{dt} &= k_A[Y_p]([Z_T]-[Y_pZ]) - (k_D+k_4)[Y_pZ], && (3.3)\\
\frac{d[B_p]}{dt} &= k_3[A_p]([B_T]-[B_p]) - k_5[B_p], && (3.4)
\end{aligned}$$

within which Φ denotes the receptor signalling team activity, k_i $(i = 1, 2, 3, 4, 5, A, D)$ indicate the kinetic rates of each reaction and [...] denote the concentrations of the appropriate proteins with subscripts T and p indicating the total and phosphorylated amounts, respectively. This original model of Clausznitzer et al. [70] considers a separate model variable describing binding of CheY-P with CheZ to form a transient complexed state ($[Y_pZ]$) and the subsequent dephosphorylation and unbinding. Since this is a transient process (i.e. it occurs on a very short timescale), here we assume that it occurs rapidly enough as to be modelled as a one step process. This allows it to be represented by a single term within equation (3.6). We also include the autodephosphorylation of protein CheY-P into this model (with a kinetic rate constant denoted k_6), giving set of non-linear ODEs

$$\begin{aligned}
\frac{d[A_p]}{dt} &= \Phi k_1([A_T]-[A_p]) - k_2[A_p]([Y_T]-[Y_p]) - k_3[A_p]([B_T]-[B_p]), && (3.5)\\
\frac{d[Y_p]}{dt} &= k_2[A_p]([Y_T]-[Y_p]) - k_4[Y_p][Z_T] - k_6[Y_p], && (3.6)\\
\frac{d[B_P]}{dt} &= k_3[A_p]([B_T]-[B_p]) - k_5[B_p]. && (3.7)
\end{aligned}$$

In addition to the above description of the phosphorylation cascade, Clausznitzer et al. [70] produced a description of receptor methylation kinetics. More specifically, by considering the time immediately after a step change in the attractant concentration (as in equation (2.7)), they were able to compare a number of possible models to experimental data. In doing so it was found that one particular representation produced a good model fit. This is given by

$$\frac{dm}{dt} = g_R[R_T](1-\Phi) - g_B[B_p]^2\Phi, \qquad (3.8)$$

where m is the average methylation level of a receptor within the signalling team and g_R, g_B denote the kinetic rates of receptor methylation by CheR and demethylation by CheB-P, respectively. Further to this Clausznitzer et al. [70] define receptor signalling team activity (Φ) in equations (3.5) and (3.8) as a Monod-Wyman-Changeux (MWC) description of receptor clustering [58]. This is given by

$$\Phi = \frac{1}{1 + e^{F}}, \tag{3.9}$$

in which F, the free-energy of a receptor signalling team, is defined by

$$F = N \left[h(m) + \nu_a \ln \left(\frac{1 + [L]/K_a^{off}}{1 + [L]/K_a^{on}} \right) + \nu_s \ln \left(\frac{1 + [L]/K_s^{off}}{1 + [L]/K_s^{on}} \right) \right], \tag{3.10}$$

where N is the number of chemoreceptors in the signalling team, $\nu_a : \nu_s$ denotes the ratio of Tar to Tsr receptors in the signalling team, $h(m)$ represents the offset energy (i.e. the contribution to F from the addition/removal of one methyl group) and the remainder of the expression is the free-energy contributions from Tar and Tsr receptors with dissociation constants for a single ligand denoted $K_{a/s}^{on/off}$ for active/inactive Tar (a) and Tsr (s) receptors and $[L]$ is the extracellular ligand concentration. Parameter values used in this model are listed in Table 3.1. This mathematical model has been shown to be a good fit to experimental data in spite of the low copy numbers associated with CheR and CheB, suggesting that the continuum limit holds. With this being the case we do not anticipate stochastic effects would alter any conclusions drawn within this work.

The chemoreceptor signalling team free-energy expression in equation (3.10) contains contributions from two different types of chemoreceptors when challenged with aspartate stimuli. It is known that both Tar and Tsr chemoreceptors may sense aspartate stimuli albeit with Tsr binding aspartate with a much lower affinity than Tar chemoreceptors. As such, it has been shown by Mello & Tu [33] that aspartate binding to Tsr chemoreceptors has little or no effect at small to intermediate aspartate concentrations. There is a noticeable effect from aspartate binding to Tsr on the overall sensitivity of a receptor signalling team in cases with a very high concentration of aspartate. With this being the case, for the majority of the subsequent work within this thesis we shall restrict our attention to cases in which there are small to intermediate aspartate concentrations with respect to the relevant dissociation constants. This allows us to simplify equation (3.10) such that it becomes

$$F = N \left[h(m) + \ln \left(\frac{1 + [L]/K_a^{off}}{1 + [L]/K_a^{on}} \right) \right]. \tag{3.11}$$

Within this expression we have neglected the contribution of Tsr receptors binding aspartate. It is also worth noting that since Clausznitzer et al. [70] state the ratio of Tar to Tsr (ν_a:ν_s) receptors is equal to 1:1.4, the symbol ν_a may be neglected since it is a multiple equal to one and thus will have no effect on the overall free-energy. The work of Clausznitzer et al. [70] compared the mathematical model laid out above to a range of experimental data. This demonstrated that the model represents a detailed enough description of the chemotaxis signalling pathway of *E. coli* cells as to adequately represent the biological processes involved. Whilst Clausznitzer et al. [70] clearly explored this model in the context of large amounts of experimental data, they provided very little in the way of mathematical analysis. As such, for the remainder of this chapter we shall present a rigorous analysis of this model forming a platform upon which the more detailed examination in subsequent chapters may be based.

Table 3.1: A base set of *E. coli* signalling pathway parameter values from the literature.

Symbol	Definition	Value	Source
$[A_T]$	Total concentration of CheA	7.9μM	[112]*
$[B_T]$	Total concentration of CheB	0.28μM	[112]*
$[R_T]$	Total concentration of CheR	0.16μM	[112]*
$[Y_T]$	Total concentration of CheY	9.7μM	[112]*
$[Z_T]$	Total concentration of CheZ	3.8μM	[112]*
k_1	CheA autophosphorylation	34s^{-1}	[113]
k_2	Phosphotransfer to CheY	$100\mu\text{M}^{-1}\text{s}^{-1}$	[114]
k_3	Phosphotransfer to CheB	$15\mu\text{M}^{-1}\text{s}^{-1}$	[114]
k_4	CheY-P dephosphorylation by CheZ	$1.6\mu\text{M}^{-1}\text{s}^{-1}$	[112]
k_5	Dephosphorylation of CheB-P	0.7s^{-1}	[115]
k_6	Dephosphorylation of CheY-P	0.085s^{-1}	[116]
g_R	Methylation by CheR	$0.0375\mu\text{M}^{-1}\text{s}^{-1}$	[70]
g_B	Demethylation by CheB-P	$3.14\mu\text{M}^{-2}\text{s}^{-1}$	[70]
N	Number of Tar receptors in a signalling team	18	[62]
K_a^{on}	Dissociation constant of an active Tar receptor	0.5mM	[60]
K_a^{off}	Dissociation constant of an inactive Tar receptor	0.02mM	[60]
K_s^{on}	Dissociation constant of an active Tsr receptor	100mM	[70]
K_s^{off}	Dissociation constant of an inactive Tsr receptor	10^6mM	[70]
ν_a	Tar component of receptor type ratio ($\nu_a : \nu_s$)	1	[70]
ν_s	Tsr component of receptor type ratio ($\nu_a : \nu_s$)	1.4	[70]

* All total protein concentrations are calculated from experimental values in [112] assuming a cellular volume of 1.4fl, as per [117].

3.2 Model Analysis

3.2.1 Choosing an Offset Energy Function

Before progressing to further analysis of this mathematical model it is necessary to define a representation of the "offset energy" function, denoted $h(m)$. This may also be thought of as being the methylation dependent part of the free-energy of a chemoreceptor signalling team. Within the literature there have been a number of forms taken for this function [63, 70, 118]. However, as shall be the case within this section, these offset energy functions have never been assessed against a collection of the available experimental data.

The offset energy functions of Tu et al. [63], Clausznitzer et al. [70] and Shimizu et al. [118]

all take the same basic form

$$h(m) = \alpha(m_0 - m). \tag{3.12}$$

In this expression α is a scaling value and m_0 represents an initial methylation level (in the absence of any ligand). Tu et al. [63] takes $\alpha = 2$ and $m_0 = 1$ to give

$$h(m) = 2(1 - m), \tag{3.13}$$

whilst Clausznitzer et al. [70] consider $\alpha = 0.5$ and $m_0 = 2$, i.e.

$$h(m) = 1 - \frac{m}{2}, \tag{3.14}$$

and Shimizu et al. [118] take $\alpha = 2$ and $m_0 = 0.5$, giving

$$h(m) = 2\left(\frac{1}{2} - m\right). \tag{3.15}$$

Each of these functional forms can be compared to experimental data found in the literature [62, 118, 119] as shown in Figure 3.1. In particular, we compare each $h(m)$ function to the range of methylation states possible for each receptor. It is worth noting that each receptor possesses four methylation sites, which leads to a limit of eight methyl groups per receptor dimer and an upper limit of 48 methyl groups for an assistance neighbourhood of 6 receptor dimers.

The experimental data of both Endres et al. [62] and Vaknin & Berg [119] use similar experimental techniques, whereby Tar receptors were genetically engineered to have either a glutamate (E) or a glutamine (Q) at each of their four methylation sites [62, 119]. In experiments such as these, a glutamine has similar functionality to a methylated glutamate [120]. Each data set considers slightly different combinations of modification states with the first set considering Tar receptors in QQQQ, QEQQ, QEQE and QEEE states [62]. The other data set gives results for Tar receptors in QQQQ, QEQQ, QEEE and EEEE states [62, 119]. In order to make a comparison between the representations of $h(m)$ and the experimental data for amidation modifications (see glossary for definition) it is necessary to convert them into the same form since one source gives data for individual receptors and the other gives data for receptor dimers. This means one data set, namely that of Vaknin & Berg [119], refers to QEEE as equivalent to a methylation level of $m = 1$ whereas Endres et al. [62] would consider this as $m = 2$. This conversion requires a simple scaling of the m values used for Figure 3.1.

It is clear upon inspection of Figure 3.1 that the two data sets give similar results and that a good fit to this data will produce a model with a good level of realism. Examination of the three

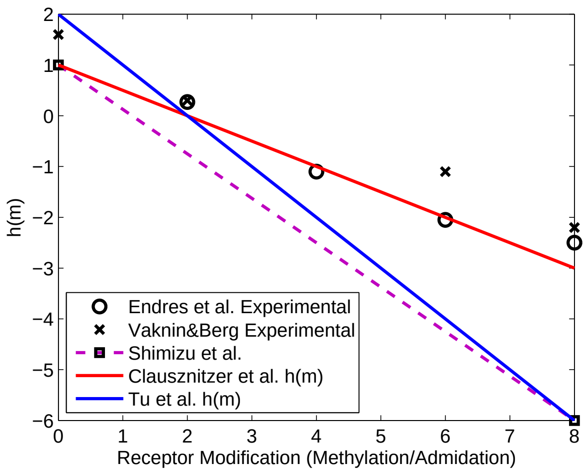


Figure 3.1: A plot showing experimental data and various model approximations (equations (3.13)-(3.15)). Experimental data for amidation modifications is from Endres et al. [62] (crosses) as well as Vaknin & Berg [119] (circles). Data for methylation modifications from Shimizu et al. [118] is also displayed (squares).

choices of $h(m)$ reveals that the representation considered by Clausznitzer et al. [70] produces a better fit to the two sets of amidation modification data. Since the mathematical model of Clausznitzer et al. [70] has been shown to give a good fit to FRET experimental data, we adopt the form of $h(m)$ defined by equation (3.14) for the remainder of this thesis.

3.2.2 Non-Dimensionalisation

As is common in the modelling of biological systems we consider a non-dimensionalisation (re-scaling) of the governing system of equations. In this case we choose to re-scale each of the key protein concentrations with respect to the total concentration within the cell, i.e. $[A_p] = a_p[A_T]$, $[B_p] = b_p[B_T]$ and $[Y_p] = y_p[Y_T]$. We then re-scale time with respect to the rate of spontaneous de-phosphorylation of the protein CheB-P, i.e. $t = \tau/k_5$. Substitution of these re-scaled variables into equations (3.5)-(3.8) yields the system of non-dimensional ODEs

$$\frac{dm}{d\tau} = \gamma_R(1-\Phi) - \gamma_B b_p^2\Phi \qquad = f_1(m, b_p), \tag{3.16}$$

$$\frac{da_p}{d\tau} = \Phi\bar{k}_1(1-a_p) - \bar{k}_2(1-y_p)a_p - \bar{k}_3(1-b_p)a_p \qquad = f_2(m, a_p, y_p, b_p), \tag{3.17}$$

$$\frac{dy_p}{d\tau} = \alpha_1\bar{k}_2(1-yp)a_p - \left(\bar{k}_4 + \bar{k}_6\right)y_p \qquad = f_3(a_p, y_p), \tag{3.18}$$

$$\frac{db_p}{d\tau} = \alpha_2\bar{k}_3(1-b_p)a_p - \bar{k}_5 b_p \qquad = f_4(a_p, b_p). \tag{3.19}$$

Within this set of non-dimensionalised equations parameters have been grouped into simplified versions such as $\bar{k}_2 = k_2[Y_T]/k_5$ (see Table 3.2 for all definitions and values). The main reason for doing this is that it often acts to reduce the number of parameters contained in the model, thus making it simpler to apply, among others, parameter fitting techniques. However, in this particular example, there exists a vast experimental literature within which all of the parameters in this model have been measured to a good degree of accuracy.

3.2.3 Model Steady-States

In order to find the steady-state (equilibrium) of the mathematical model given in Section 3.2.2 we seek the state(s) whereby the rate of change in time is equal to zero. Thus, we set the derivatives in equations (3.16)-(3.19) equal to zero, giving our model steady-state(s) as

$$
\begin{aligned}
0 &= \gamma_R(1-\Phi^*) - \gamma_B b_p^{*2}\Phi^*, && (3.20)\\
0 &= \Phi^*\bar{k}_1(1-a_p^*) - \bar{k}_2(1-y_p^*)a_p^* - \bar{k}_3(1-b_p^*)a_p^*, && (3.21)\\
0 &= \alpha_1\bar{k}_2(1-y_p^*)a_p^* - \left(\bar{k}_4+\bar{k}_6\right)y_p^*, && (3.22)\\
0 &= \alpha_2\bar{k}_3(1-b_p^*)a_p^* - \bar{k}_5 b_p^*, && (3.23)
\end{aligned}
$$

within which the superscript asterisk (*) indicates the steady-state value of the relevant model variable. We now examine the form of the solution to this model.

We begin by rearranging equation (3.22), such that

$$y_p^* = \frac{\alpha_1\bar{k}_2 a_p^*}{\alpha_1\bar{k}_2 a_p^* + \bar{k}_4 + \bar{k}_6} = \frac{k_2[A_T]a_p^*}{k_2[A_T]a_p^* + k_4[Z_T] + k_6}. \tag{3.24}$$

Similarly, from equation (3.23)

$$b_p^* = \frac{\alpha_2\bar{k}_3}{\alpha_2\bar{k}_3 + \bar{k}_5} = \frac{k_3[A_T]a_p^*}{k_3[A_T]a_p^* + k_5}, \tag{3.25}$$

Table 3.2: A base set of non-dimensional *E. coli* signalling pathway parameter values, calculated from values in Table 3.1.

Symbol	Value
$\bar{k}_1 = k_1/k_5$	48.571
$\bar{k}_2 = k_2[Y_T]/k_5$	1385.714
$\bar{k}_3 = k_3[B_T]/k_5$	6
$\bar{k}_4 = k_4[Z_T]/k_5$	8.686
$\bar{k}_5 = k_5/k_5$	1
$\bar{k}_6 = k_6/k_5$	0.121
$\alpha_1 = [A_T]/[Y_T]$	0.814
$\alpha_2 = [A_T]/[B_T]$	28.214
$\gamma_R = g_R[R_T]/k_5$	8.57×10^{-3}
$\gamma_B = g_B[B_T]^2/k_5$	0.352

and from equation (3.20)

$$\Phi^* = \frac{\gamma_R}{\gamma_R + \gamma_B b_p^{*2}} = \frac{1}{1 + \frac{\gamma_B}{\gamma_R} b_p^{*2}} = \frac{g_R[R_T]}{g_R[R_T] + g_B[B_T]^2 b_p^{*2}} = \frac{1}{1 + \frac{g_B[B_T]^2 b_p^{*2}}{g_R[R_T]}}. \tag{3.26}$$

Using this expression alongside equation (3.9) we see that

$$e^{F^*} = \frac{\gamma_B}{\gamma_R} b_p^{*2} = \frac{g_B[B_T]^2 b_p^{*2}}{g_R[R_T]}, \tag{3.27}$$

within which F is given by equation (3.11). Now, rearranging this equality we are able to obtain the steady-state expression for the average chemoreceptor methylation level

$$\begin{aligned} m^* &= 2\left[1 + \ln\left(\frac{1 + [L]/K_a^{off}}{1 + [L]/K_a^{on}}\right) - \frac{1}{N}\ln\left(\frac{\gamma_B b_p^{*2}}{\gamma_R}\right)\right] \\ &= 2\left[1 + \ln\left(\frac{1 + [L]/K_a^{off}}{1 + [L]/K_a^{on}}\right) - \frac{1}{N}\ln\left(\frac{g_B[B_T]^2 b_p^{*2}}{g_R[R_T]}\right)\right]. \end{aligned} \tag{3.28}$$

Each of the steady-state expressions obtained thus far have been done so by simple rearranging of expressions. These equations may also be utilised to obtain a steady-state expression for a_p (i.e. for equation (3.21)). In this case we substitute the steady-state equations (3.24)-(3.26) into equation (3.21) and multiply through by the denominators of each term. This yields a fifth order polynomial in a_p^*, the roots of which represent values for a_p^*. This takes the form

$$p(a_p^*) = 0 = Aa_p^{*5} + Ba_p^{*4} + Ca_p^{*3} + Da_p^{*2} + Ea_p^* + F, \tag{3.29}$$

where A, B, C, D, E and F are coefficients of the polynomial $p(a_p^*)$ that are defined as follows

$$A = -g_R[R_T][A_T]^4 k_1 k_2 k_3^3, \tag{3.30}$$

$$\begin{aligned} B = & -g_B[B_T]^3[A_T]^3 k_2 k_3^3 k_5 - 3g_R[R_T][A_T]^3 k_1 k_2 k_3^2 k_5 - g_R[R_T][A_T]^3[Z_T] k_1 k_3^3 k_4 \\ & - g_B[B_T]^2[A_T]^3[Y_T] k_2 k_3^3 k_6 - g_R[R_T][A_T]^3[B_T] k_2 k_3^3 k_5 - g_B[B_T]^2[A_T]^3[Y_T][Z_T] k_2 k_3^3 k_4 \\ & - g_R[R_T][A_T]^3[Y_T][Z_T] k_2 k_3^3 k_4 + g_R[R_T][A_T]^4 k_1 k_2 k_3^3 - g_R[R_T][A_T]^3[Y_T] k_2 k_3^3 k_6 \\ & - g_R[R_T][A_T]^3 k_1 k_3^3 k_6, \end{aligned} \tag{3.31}$$

$$\begin{aligned} C = & -g_B[B_T]^3[A_T]^2 k_3^3 k_5 k_6 + 3g_R[R_T][A_T]^3 k_1 k_2 k_3^3 k_5 - 3g_R[R_T][A_T]^2 k_1 k_3^2 k_5 k_6 \\ & - g_R[R_T][A_T]^2[B_T] k_3^3 k_5 k_6 - g_R[R_T][A_T]^2[B_T][Z_T] k_3^3 k_4 k_5 - 3g_R[R_T][A_T]^2 k_1 k_2 k_3 k_5^2 \end{aligned} \tag{3.32}$$

$$
\begin{aligned}
&- \quad 2g_R[R_T][A_T]^2[B_T]k_2k_3^2k_5^2 + g_R[R_T][A_T]^3[Z_T]k_1k_3^3k_4 - 3g_R[R_T][A_T]^2[Y_T]k_2k_3^2k_5k_6 \\
&- \quad 3g_R[R_T][A_T]^2[Y_T][Z_T]k_2k_3^2k_4k_5 - g_B[B_T]^3[A_T]^2[Z_T]k_3^3k_4k_5 + g_R[R_T][A_T]^3k_1k_3^3k_6 \\
&- \quad 3g_R[R_T][A_T]^2[Z_T]k_1k_3^2k_4k_5 - g_B[B_T]^2[A_T]^2[Y_T]k_2k_3^2k_5k_6 \\
&- \quad g_B[B_T][A_T]^2[Y_T][Z_T]k_2k_3^2k_4k_5,
\end{aligned}
$$

$$
\begin{aligned}
D \quad = \quad & -2g_R[R_T][A_T][B_T]k_3^2k_5^2k_6 - 3g_R[R_T][A_T]k_1k_3k_5^2k_6 \\
- \quad & 2g_R[R_T][A_T][B_T][Z_T]k_3^2k_4k_5^2 - g_R[R_T][A_T]k_1k_2k_5^3 + 3g_R[R_T][A_T]^2[Z_T]k_1k_3^2k_4k_5 \\
- \quad & g_R[R_T][A_T][B_T]k_2k_3k_5^3 + 3g_R[R_T][A_T]^2k_1k_3^2k_5k_6 - 3g_R[R_T][A_T][Y_T]k_2k_3k_5^2k_6 \\
- \quad & 3g_R[R_T][A_T][Z_T]k_1k_3k_4k_5^2 + 3g_R[R_T][A_T]^2k_1k_2k_3k_5^2 - 3g_R[R_T][A_T][Y_T][Z_T]k_2k_3k_4k_5^2,
\end{aligned}
\tag{3.33}
$$

$$
\begin{aligned}
E \quad = \quad & g_R[R_T][A_T]k_1k_2k_5^3 - g_R[R_T][B_T][Z_T]k_3k_4k_5^3 - g_R[R_T][B_T]k_3k_5^3k_6 \\
+ \quad & 3g_R[R_T][A_T]k_1k_3k_5^2k_6 - g_R[R_T]k_1k_5^3k_6 - g_R[R_T][Y_T][Z_T]k_2k_4k_5^3 \\
- \quad & g_R[R_T][Z_T]k_1k_4k_5^3 + 3g_R[R_T][A_T][Z_T]k_1k_3k_4k_5^2 - g_R[R_T][Y_T]k_2k_5^3k_6,
\end{aligned}
\tag{3.34}
$$

$$
F \quad = \quad g_R[R_T][Z_T]k_1k_4k_5^3 + g_R[R_T]k_1k_5^3k_6. \tag{3.35}
$$

Since there is no generally applicable analytical solution form for quintic equations such as this, we must consider an alternative approach to finding the a_p steady-state value. In the first instance, we substitute the parameter values from Table 3.1 into these polynomial coefficients, thus giving the numerical values displayed in Table 3.3

Table 3.3: A table of coefficient values for the CheA-P steady-state polynomial. In the case of $p(-a_p^*)$ we have absorbed the change of sign into the polynomial coefficients.

	A	B	C	D	E	F
$p(a_p^*)$	2.682×10^8	2.256×10^9	8.815×10^6	-5.861×10^4	-261.564	-0.431
$p(-a_p^*)$	-2.682×10^8	2.256×10^9	-8.815×10^6	-5.861×10^4	261.564	-0.431

It is now possible to utilise Descartes' rule of signs (see Appendix B) in order to examine the nature of the solutions (roots) to equation (3.21). Since the number of sign changes between consecutive polynomial coefficients in $p(a_p^*)$ is equal to one, we know for sure that there must be exactly one positive root. Examining the polynomial $p(-a_p^*)$, we find four sign changes meaning that there exists a maximum of four negative roots. However, there may also exist a number of complex roots. Since these roots must exist in complex conjugate pairs there are three possible combinations of roots

- 1 positive root, 4 negative roots and 0 complex roots,

- 1 positive root, 2 negative roots and 2 complex roots,
- 1 positive root, 0 negative roots and 4 complex roots.

Given that the steady-state value for CheA-P must be positive in order to be biologically feasible (since a negative or complex concentration does not make sense) we are able to see that there may only be one biologically feasible steady-state for CheA-P. Therefore, under the assumption that initial conditions are chosen within a biologically feasible range, we may have just one steady-state for the system. In fact we find that the steady-state expressions of CheA-P, CheB-P and CheY-P are independent of the extracellular ligand concentration or the average chemoreceptor methylation level. In contrast to this, the steady-state value for the average chemoreceptor methylation level (m) will vary depending upon the ambient ligand concentration as well as the CheB-P steady-state.

Using equations (3.24), (3.25) and (3.29) in addition to the parameter values listed in Table 3.1 we may obtain, using the MATLAB 'roots' function, the following non-dimensional steady-state values

$$a_p^* = 5.5769 \times 10^{-3}, \quad b_p^* = 4.8562 \times 10^{-1}, \quad y_p^* = 4.1679 \times 10^{-1},$$

given to four decimal places. Each of these values will remain the same for any ambient extracellular ligand concentration or average chemoreceptor methylation level. We may also calculate the steady-state value for the average chemoreceptor methylation level. Under the assumption that there is no ambient ligand concentration (i.e. $[L] = 0$) we obtain

$$m = 1.7478,$$

to four decimal places. However, this steady-state methylation level will rise/fall in order to re-set the signalling pathway to a pre-stimulus state when the extracellular ligand concentration varies. This proceeds in the manner described by equation (3.28) and displayed in Figure 3.2.

3.2.4 Model Stability Analysis

From Section 3.2.3 we now know that given the parameter set in Table 3.1 there may only be one biologically feasible steady-state. We now consider the stability of the steady-state determined in the previous section. In order to do this it is necessary to analyse the eigenvalues

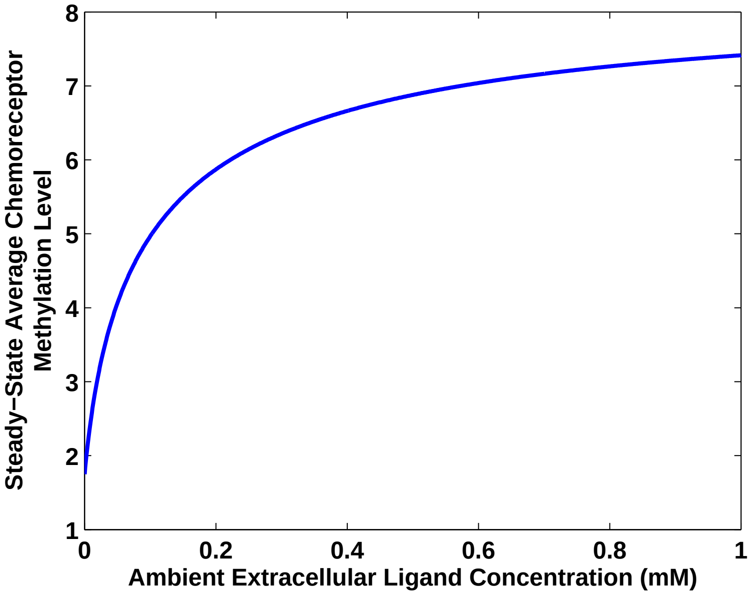


Figure 3.2: A plot showing how the steady-state value for the average chemoreceptor methylation level rises in relation to the ambient extracellular ligand concentration. This result is similar to those given by Hansen et al. [69] and Endres & Wingreen [68].

of the Jacobian matrix for this system which is defined as

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial f_1}{\partial m} & 0 & 0 & \frac{\partial f_1}{\partial b_p} \\ \frac{\partial f_2}{\partial m} & \frac{\partial f_2}{\partial a_p} & \frac{\partial f_2}{\partial y_p} & \frac{\partial f_2}{\partial b_p} \\ 0 & \frac{\partial f_3}{\partial a_p} & \frac{\partial f_3}{\partial y_p} & 0 \\ 0 & \frac{\partial f_4}{\partial a_p} & 0 & \frac{\partial f_4}{\partial b_p} \end{pmatrix}, \tag{3.36}$$

within which the partial derivatives are given by

$$\frac{\partial f_1}{\partial m} = \frac{-Ne^F\Phi^2}{2k_5}\left(g_R[R_T] + g_B[B_T]^2 b_p^2\right), \tag{3.37}$$

$$\frac{\partial f_1}{\partial b_p} = \frac{-2g_B[B_T]^2 b_p \Phi}{k_5}, \tag{3.38}$$

$$\frac{\partial f_2}{\partial m} = \frac{Nk_1(1-a_p)e^F\Phi}{2k_5}, \tag{3.39}$$

$$\frac{\partial f_2}{\partial a_p} = \frac{k_1\Phi}{k_5} - \frac{k_2[Y_T](1-y_p)}{k_5} - \frac{k_3[B_T](1-b_p)}{k_5}, \tag{3.40}$$

$$\frac{\partial f_2}{\partial y_p} = \frac{k_2[Y_T]a_p}{k_5}, \tag{3.41}$$

$$\frac{\partial f_2}{\partial b_p} = \frac{k_3[B_T]a_p}{k_5}, \tag{3.42}$$

$$\frac{\partial f_3}{\partial a_p} = \frac{k_2[A_T](1-y_p)}{k5}, \tag{3.43}$$

$$\frac{\partial f_3}{\partial y_p} = \frac{-k_2[A_T]a_p}{k_5} - \frac{k_4[Z_T]}{k_5} - \frac{k_6}{k_5}, \tag{3.44}$$

$$\frac{\partial f_4}{\partial a_p} = \frac{k_3[A_T](1-b_p)}{k_5}, \tag{3.45}$$

$$\frac{\partial f_4}{\partial b_p} = \frac{-k_3[A_T]a_p}{k_5} - 1. \tag{3.46}$$

We now investigate the eigenvalues of this system by substituting into the Jacobian each of the steady-state values before solving

$$\det|\boldsymbol{J} - \lambda \boldsymbol{I}| = 0, \tag{3.47}$$

within which $\boldsymbol{I}$ is the identity matrix with the same dimensions as the Jacobian and λ denotes an eigenvalue of the system. Utilising the parameter values given in Table 3.1 we are able to obtain the four eigenvalues of the system, which are

$$\lambda_1 = -822.1086, \quad \lambda_2 = -8.8146, \quad \lambda_3 = -1.7996, \quad \text{and} \quad \lambda_4 = -0.1942.$$

It is clear that each of these four eigenvalues of the system is such that $\lambda_i < 0$ and hence we are able to conclude that this system is asymptotically stable at the parameter values stated in Table 3.1.

3.2.5 Model Stiffness

It has clearly been shown that given the parameter set in Table 3.1 there exists four real, negative eigenvalues which denotes an asymptotically stable system. However, the relative magnitudes of these eigenvalues can tell us about the stiffness of the system. In this case we may obtain the stiffness coefficient of the system by calculating

$$s = \frac{|\mathrm{Re}(\overline{\lambda})|}{|\mathrm{Re}(\underline{\lambda})|}, \tag{3.48}$$

in which s denotes the stiffness coefficient, whilst $\overline{\lambda}$ and $\underline{\lambda}$ indicate the eigenvalues with the largest and smallest real part (in absolute value), respectively. In the current example the stiffness coefficient will be given by

$$s = \frac{|\mathrm{Re}(\lambda_1)|}{|\mathrm{Re}(\lambda_4)|} = 4233.31, \tag{3.49}$$

to two decimal places. This is clearly rather large and so many numerical methods may be subject to significant restrictions on the length of time step that may be used when numerical methods are applied to this problem. As such we shall utilise the inbuilt MATLAB ODE solver ode15s for numerical approximations of this system. This particular ODE solver is specifically designed for stiff systems of ODEs and allows the definition of error tolerances. The algorithm adjusts the length of each time step in order to maintain these tolerances as well as the numerical stability of the scheme used.

3.3 Summary & Discussion

Within this chapter we began by presenting a recent mathematical model of the *E. coli* intracellular signalling pathway from the literature. The remainder of this chapter then sought to produce a rigorous analysis of the model.

This analysis began by examining three different offset energy functions from the literature in the context of the available experimental data. Upon doing so it was shown that the function of Clausznitzer et al. [70] produced the best fit and will thus be used for all *E. coli* chemotaxis modelling work in this thesis.

A non-dimensional re-scaling of the model was then presented and utilised in order to demonstrate that there exists just one biologically feasible steady-state for this system. This equilibrium state was subsequently shown to be asymptotically stable, with eigenvalues demonstrating that this is a stiff system.

The analysis conducted within this chapter demonstrates a number of features of this mathematical model that had previously only been assumed within the literature. In addition to this, the features analysed here will prove useful within subsequent chapters.

Chapter 4

Fold-Change Detection in *Escherichia coli*

In this chapter we investigate a feature of the *E. coli* chemotaxis signalling cascade known as fold-change detection (FCD). We begin by giving a brief summary of some background literature by way of motivation for this study. Here we analyse a more detailed model of the *E. coli* chemotaxis signalling pathway than previously examined in the literature, namely that described in Chapter 3. In order to do so we utilise both theoretical and numerical techniques to investigate the ability of this model to match experimental observations. Parameter sensitivity analysis then shows that FCD is invariant to changes in signalling pathway parameters assuming they remain constant in time. Consideration is then given to the ability of this mathematical model to display FCD when the chemotactic response is governed by multiple receptor types. This chapter concludes with a discussion regarding the biological feasibility of the conditions imposed in order to obtain FCD.

4.1 Motivation

FCD is the phenomenon whereby a system will exhibit the same response amplitude, shape and recovery time in response to equal fold-changes in input signal relative to the background level [121] (as shown in Figure 4.1). This has, in recent years, been a popular research area and has now been shown to exist within a number of different systems.

Mesibov et al. [24] were first to hint at the existence of FCD in *E. coli* in the early 1970s when they noted that cells produced a similar response to fractional increases in ligand concentration. More recently Lazova et al. [122] were able to confirm using *in vivo* fluorescence resonance energy transfer (FRET) techniques that *E. coli* cells do indeed display FCD. In particular this work showed the existence of two distinct regimes within which FCD would be exhibited, namely

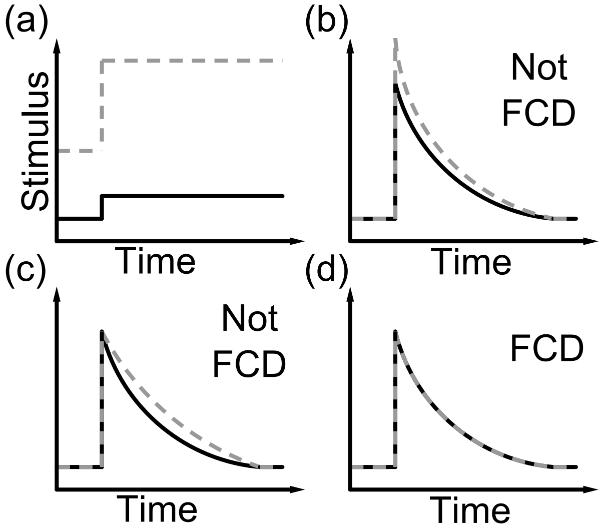


Figure 4.1: After [121], a graphical representation of the definition of FCD. The three responses (b)-(d) are possible given the input stimulus (change in ligand concentration) (a). Figure (d) shows two responses which are exactly the same in terms of response amplitude, shape and recovery time. Figure (b) shows responses with different amplitude and (c) a different response shape.

FCD1 ($[L]_0 = 0.018 - 0.23$mM) and FCD2 ($[L]_0 = 0.82 - 10.3$mM), where $[L]_0$ is an initial ligand (e.g. MeAsp) concentration.

Within the literature it has been proposed that adopting FCD may be beneficial to chemotactic bacteria. The main reason for this is that the process of a sensory search whereby cells move through their environment becomes invariant to the amplitude of the field itself [121]. This allows scalar multiples to be picked out from the extracellular environment, thus allowing cells to focus on the characteristic shape of the ligand field. It has also been suggested that a sensory search using FCD may be optimal in situations where the discovery of any nutrient source is of benefit [122]. In other words, the payoff for cells of reaching what may well be a weak nutrient source may be larger than the payoff if cells were to attempt to discriminate in favour of richer nutrient sources.

Previous theoretical work has shown that FCD will be evident in a simplified model of the *E. coli* chemotaxis signalling pathway [121] as proposed by Tu et al. [63]. This particular mathematical model considers a Monod-Wyman-Changeux (MWC) representation of receptor clustering and couples it with ODEs describing receptor methylation and CheY-P concentration. Within this work the simplified model was shown to exhibit FCD, however they did not show whether or not the model produced a good fit with experimental observations. In the remainder of this chapter we demonstrate the ability of a more complete model to exhibit FCD and subsequently show a good fit to experimental data.

4.2 Conditions for FCD

In this section we outline theoretical conditions required for a mathematical model of the form

$$\begin{aligned} \dot{x} &= f(x(m), \Phi, [L]), && (4.1) \\ \Phi &= g(x(m), \Phi, [L]), && (4.2) \end{aligned}$$

to exhibit FCD as recently shown by Shoval and colleagues [121, 123]. Here Φ and $[L]$ are as previously defined whilst $x(m)$ is an effective (methylation dependent) dissociation constant.

A set of four sufficient conditions are given for the existence of FCD [121, 123]. The first two of these are that the model must exhibit a stable steady-state and exact adaptation (the ability to return precisely to the pre-stimulus state). In addition to this a pair of homogeneity conditions are given relating to a model of the form displayed in equations (4.1) and (4.2), namely

$$f(px(m), \Phi, p[L]) = pf(x(m), \Phi, [L]), \tag{4.3}$$

$$g(px(m), \Phi, p[L]) = g(x(m), \Phi, [L]), \tag{4.4}$$

in which p is some constant such that $p > 0$. A proof of these two conditions as well as a generalisation allowing a function $\varphi(p, x)$ to be used in place of a constant p are given in the supplementary information of [121].

4.3 FCD in a Whole-Pathway Model

We now progress to examining whether the whole-pathway mathematical model laid out in Chapter 3 can exhibit FCD. This model is a more complete representation of the signalling cascade within an *E. coli* cell than previously considered in analysing FCD in the *E. coli* chemotaxis literature [63]. Within this work we initially consider N (receptor signalling team size) to be constant before discussing the implications for FCD when it depends upon the initial ligand concentration.

As is common in the modelling of chemotactic systems, for now we shall consider only the cell response to a single chemoattractant. The majority of the chemotaxis literature focuses on either α-methylaspartate (MeAsp) or serine. These are mostly chosen due to the fact that they are non-metabolisable (in the case of MeAsp) or are sensed by the most abundant chemoreceptors (in the case of serine). Within this work we consider the cell response to MeAsp as it is non-metabolisable, thus its concentration remains constant in time.

Within the theoretical work of Shoval and colleagues [121, 123], a key assumption made is that $K_a^{off} \ll [L] \ll K_a^{on}$, i.e. the extracellular ligand concentration must be much larger than the ligand dissociation constant of an inactive chemoreceptor whilst also being much smaller than the dissociation constant for an active chemoreceptor. Following the example of Shoval et al. [121] this allows us to make two simplifications to the logarithmic terms in equation (3.11), namely

$$\ln\left(1 + \frac{[L]}{K_a^{on}}\right) \approx \ln(1) = 0, \tag{4.5}$$

and

$$\ln\left(1 + \frac{[L]}{K_a^{off}}\right) \approx \ln\left(\frac{[L]}{K_a^{off}}\right), \tag{4.6}$$

which, alongside the fact that $\nu_a = 1$, reduces equation (3.11) to

$$F = N\left[\left(1 - \frac{m}{2}\right) + \ln\left(\frac{[L]}{K_a^{off}}\right)\right]. \tag{4.7}$$

Substituting this expression into equation (3.9), we obtain

$$\Phi = \frac{1}{1 + \left(\frac{[L]}{x(m)}\right)^N}, \tag{4.8}$$

in which $x(m) = K_a^{off} \exp\left((m-2)/2\right)$.

It is now possible to assess whether the model considered here will satisfy the four sufficient conditions for FCD [121, 123]. These conditions only apply to two expressions which may appear to be problematic given the size of the model considered here. The composition of the model, however, means that it is only two expressions (those for Φ and $x(m)$) that must be taken into account when attempting to show the existence of FCD.

The first two sufficient conditions are that the mathematical model must exhibit a stable steady-state and exact adaptation. That this model displays a stable steady-state was previously demonstrated in Sections 3.2.3 and 3.2.4. It has also been shown within the literature that the model in question here exhibits exact adaptation [70].

In order to examine the sufficient condition stated in equation (4.3) we must consider our expression for $x(m)$. This is the effective dissociation constant of the receptor to the chemoattractant [121]. Testing this condition requires us to examine the temporal derivative of $x(m)$

which, upon use of the chain rule, is given by

$$\dot{x} = \frac{\partial x}{\partial m}\frac{dm}{d\tau} = \frac{1}{2}x(m)\frac{dm}{d\tau} = f(x(m), \Phi, [L]). \tag{4.9}$$

The sufficient condition given by equation (4.3) may then be tested by multiplying $x(m)$ and $[L]$ by some constant $p > 0$ which gives

$$f(px(m), \Phi, p[L]) = \frac{1}{2}px(m)\frac{dm}{d\tau} = p\left(\frac{1}{2}x(m)\frac{dm}{d\tau}\right) = pf(x(m), \Phi, [L]), \tag{4.10}$$

which clearly satisfies equation (4.3).

Examination of equation (4.8) is necessitated by condition (4.4). Noticing that

$$\Phi = \frac{1}{1 + \left(\frac{[L]}{x(m)}\right)^N} = g(x(m), \Phi, [L]), \tag{4.11}$$

allows this condition to be tested in a similar manner to equation (4.3). Multiplication of the $x(m)$ and $[L]$ terms by the same constant $p > 0$ yields

$$g(px(m), \Phi, p[L]) = \frac{1}{1 + \left(\frac{p[L]}{px(m)}\right)^N} = \frac{1}{1 + \left(\frac{[L]}{x(m)}\right)^N} = g(x(m), \Phi, [L]). \tag{4.12}$$

As with the first homogeneity condition this gives clear evidence that the sufficient condition (4.4) holds.

It has now clearly been demonstrated that each of the four sufficient conditions for FCD are satisfied for this model. It is also worth noting here that, according to the conditions given by Lazova et al. [122], any linear choice of the offset energy function $h(m)$ should not affect the ability of the model to exhibit FCD. Having satisfied these conditions we would now expect each chemotaxis protein within the model to exhibit FCD since, as shown in Section 3.2.3, the ODEs for each chemotaxis protein are not ligand dependent.

4.3.1 Numerical Simulations

Here we use numerical simulations to provide examples of the model exhibiting FCD. The model equations ((3.16)-(3.19), (3.9) and (3.11)) were solved in MATLAB (MathWorks) using the ordinary differential equation solver ode15s and the parameter values in Tables 3.1 and 3.2. We allow our system to evolve to steady-state. The resultant values of m, a_p, y_p and b_p are used as our initial conditions in subsequent simulations. If we now assume that each chemoreceptor signalling team is such that $N \gg 1$, then in the range of ligand concentrations considered here

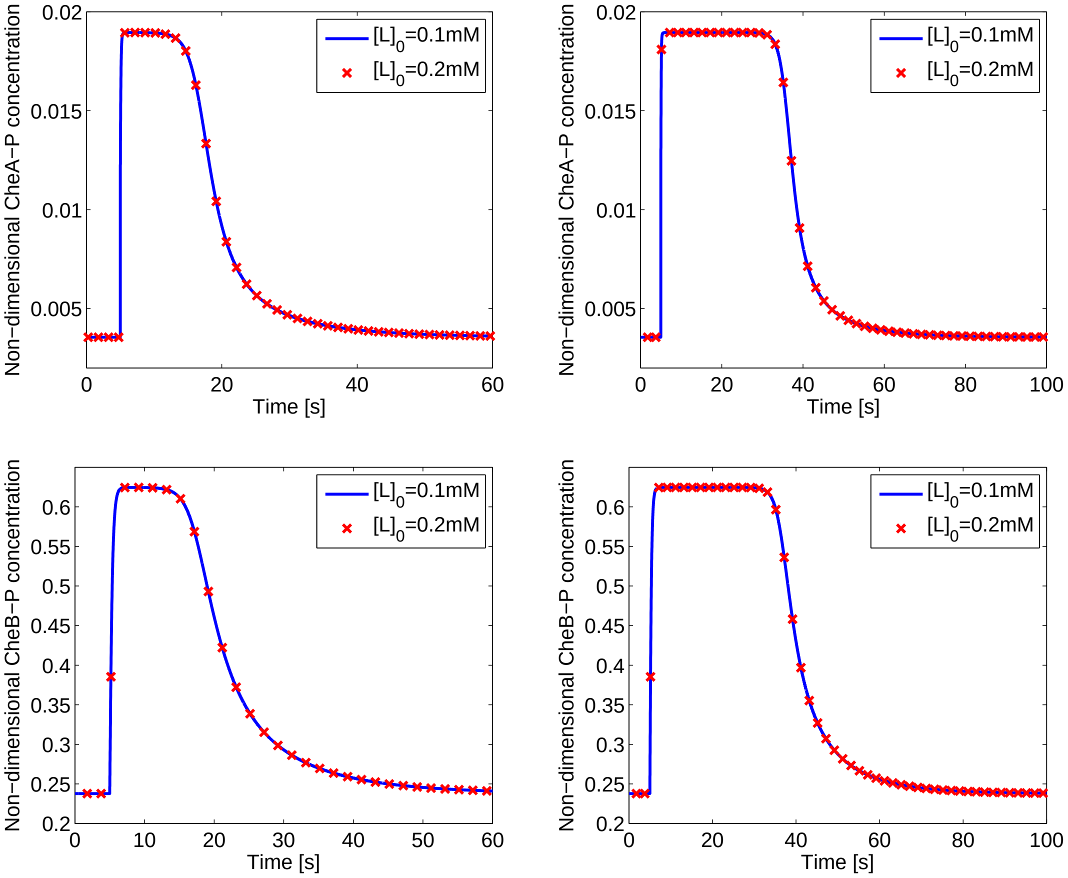


Figure 4.2: Plots displaying the behaviour of each chemotaxis protein in response to two-fold (left column) and five-fold (right column) changes in extracellular ligand concentration (applied at $t = 5$ seconds). Blue lines represent the response for an initial ligand concentration of 0.1mM and red stars for an initial ligand concentration of 0.2mM. Note that a relative error tolerance ('RelTol') of 10^{-8} and a maximum time step ('MaxStep') of 0.1 were used in the production of these simulations.

($K_a^{off} \ll [L] \ll K_a^{on}$) our steady-state average chemoreceptor methylation level may be given by

$$m^* \approx 2\left[1 - \ln\left(\frac{K_a^{off}}{[L]}\right)\right]. \tag{4.13}$$

This indicates that the average chemoreceptor methylation level is dominated by the ratio of K_a^{off} to $[L]$. It is clear that in cases considered here, in which we have $K_a^{off} \ll [L]$, this will yield a fairly large steady-state methylation level. It is often stated that each Tar receptor dimer possesses eight methylation sites [69]. At first glance this would seem to make our high methylation levels seem unrealistic. However, this is not the case since each CheR and CheB-P is able to bind to a 35 amino-acid tether attached to a Tar receptor [124]. Once bound to this tether, each CheR or CheB-P is able to methylate/demethylate approximately six nearby receptor dimers [68]. This group of receptor dimers is termed an assistance neighbourhood [125]. The effect of these assistance neighbourhoods is to increase the ladder of possible methylation states from 8 to ∼48 [68], meaning that high methylation levels, as defined by equation (4.13), are feasible.

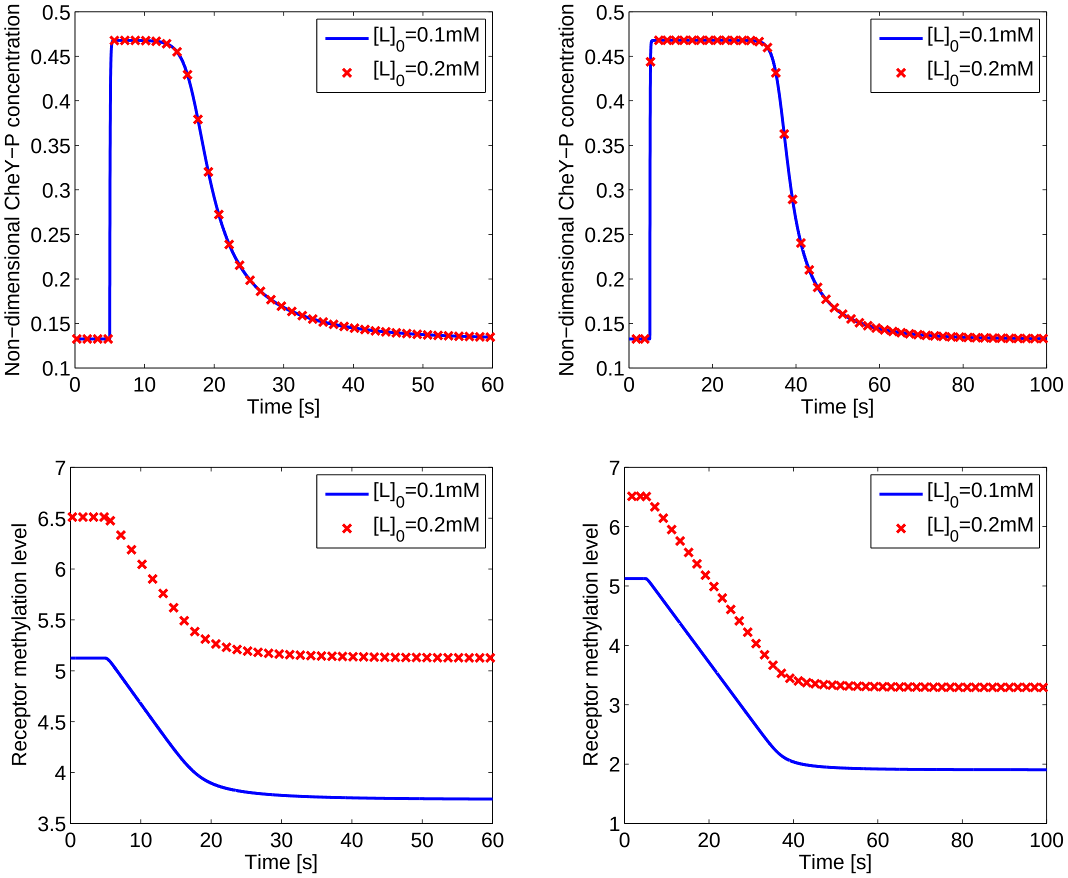


Figure 4.3: Plots displaying the behaviour of CheY-P and average receptor methylation level of a signalling team in response to two-fold (left column) and five-fold (right column) changes in extracellular ligand concentration. Blue lines represent the response for an initial ligand concentration of 0.1mM and red stars for an initial ligand concentration of 0.2mM.

A number of numerical simulations are used here to demonstrate the existence of FCD in the range of initial ligand concentrations 0.1mM$\leq$ $[L]_0 \leq$ 0.7mM. Figures 4.2 and 4.3 show the behaviour of each phosphorylated chemotaxis protein and the receptor methylation level in response to two- and five-fold changes in ligand concentration for initial ligand concentrations of 0.1mM and 0.2mM. From the results in Figures 4.2 and 4.3 FCD is evident for each chemotaxis protein; each shows the same shape, amplitude and recovery time in response to equal fold-changes in ligand concentration.

As mentioned earlier, a number of simplifying assumptions have been applied to the original mathematical model of Clausznitzer et al. [70]. In spite of these assumptions the model still gives good agreement with the full Clausznitzer et al. [70] model and thus also with their experimental FRET measurements. We note in particular the similarity between model and experiment in the time taken for the system to fully adapt to a change and the amplitude and profile of the response curves obtained.

4.3.2 Parameter Sensitivity Analysis

The results of Section 4.3 suggest that the property of FCD will be robust to variation in parameters in the signalling cascade. In order to test this result we undertook a local sensitivity analysis of the kinetic rate parameters and total protein concentrations. Each of the kinetic rate constant values and total protein concentrations in Table 3.1 were both increased and decreased by an order of magnitude such that $0.1x \leq x \leq 10x$, where x represents the initial parameter value. Numerical results obtained for each parameter variation showed that FCD still holds for all elements of the signalling cascade irrespective of which parameter value was varied, thus confirming the theoretical results of Section 4.3.

Non-Constant Receptor Signalling Team Size

Interestingly, there is experimental evidence that the size of a receptor signalling team varies with the extracellular ligand concentration [62]. With this being the case Clausznitzer et al. [70] state an expression for the size of a receptor signalling team, which is

$$N = a_0 + a_1[L]_0, \tag{4.14}$$

where a_0 and a_1 are fitting parameters (given values of $a_0 = 17.5$ and $a_1 = 3.35$, respectively) and $[L]_0$ represents the initial extracellular ligand concentration. Consideration of a model such as this would result in the use of a different value of N for each simulation. Essentially this would amount to changing the value of a parameter between consecutive numerical simulations of the model. Based on the theory presented in previous sections we would expect that this would not result in FCD. This is expected since the set of model parameters is essentially altered upon the application of each change in ligand concentration. This prediction is verified by the results shown in Figure 4.4. Given the dependence of the signalling team size N on the initial ligand concentration a small variation was observed between responses for the same fold-change. In order to quantify this effect we consider two pairs of numerical simulations. The first considered the application of a two-fold increase applied to two initial ligand concentrations falling in the FCD1 regime. The second pair of simulations was similar but for initial concentrations lying in the FCD2 regime. In each case the variation in observed behaviour, i.e. the extent to which FCD fails, is rather small ($< 5\%$ of the initial CheY-P response magnitude for two-fold changes in ligand concentration, for both the FCD1 and FCD2 regimes). We would expect that this variation is not likely to be detectable in an experimental setting.

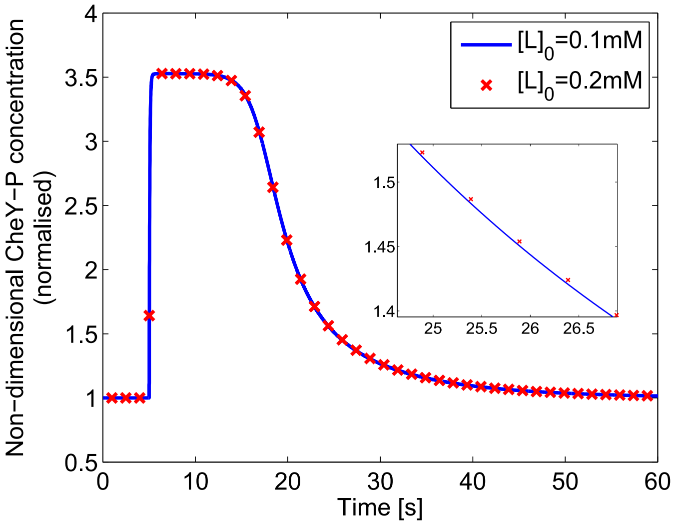


Figure 4.4: FCD does not hold exactly when considering a non-constant signalling team size. Simulations are displayed for initial ligand concentrations of 0.1mM and 0.2mM in response to a two-fold change in ligand concentration, normalised with respect to the initial steady-state concentration. The inset shows more clearly that FCD does not hold since the shape of the response is not the same in both simulations.

4.4 How Realistic are the Required Dissociation Constants?

In Section 4.3 it was assumed that $K_a^{off} \ll [L] \ll K_a^{on}$, which allowed the model to meet the conditions defined by equations (4.3) and (4.4).

In order to determine whether this condition holds for all cases *in vivo* we conducted a literature search and obtained dissociation constant values for Tar receptors, as detailed in Table 4.1. It is clear that these values vary significantly depending on the source. In particular we can see that dissociation constant values for an inactive Tar chemoreceptor (K_a^{off}) vary in the range 1.7-30μM whilst active Tar chemoreceptors (K_a^{on}) have been proposed to take dissociation constant values in the range 12-3000μM. We would expect that the dissociation constant values utilised by Tu and colleagues [63, 33, 126], Sourjik & Berg [127] and Endres, Wingreen and colleagues [70, 66, 69, 60, 68, 128, 65] yield a large enough separation between dissociation constants as to satisfy $K_a^{off} \ll [L] \ll K_a^{on}$, albeit with different sizes of FCD regimes. The final range of dissociation constant values were proposed by Bray and colleagues who suggest that the two dissociation constants should be very close together. This will clearly make it difficult for a model of the form of equations (3.16)-(3.19) to exhibit FCD from a model using these values.

The range of values used for the dissociation constants within the literature is clearly very large. However, in the majority of cases shown in Table 4.1 we would expect the condition $K_a^{off} \ll [L] \ll K_a^{on}$ to hold for some range of extracellular ligand concentrations ($[L]$).

Table 4.1: Dissociation constants for active and inactive Tar receptors, as reported in the literature.

Dissociation Constant		
Inactive Receptor (K_a^{off})	Active Receptor (K_a^{on})	Source
∼18μM	∼3000μM	[63, 33, 126]
20μM-25μM	500μM	[70, 66, 69, 60, 68, 128, 65]
30μM	1000μM	[127]
1.7μM	12μM	[129, 53]

4.4.1 Multiple Receptor Types

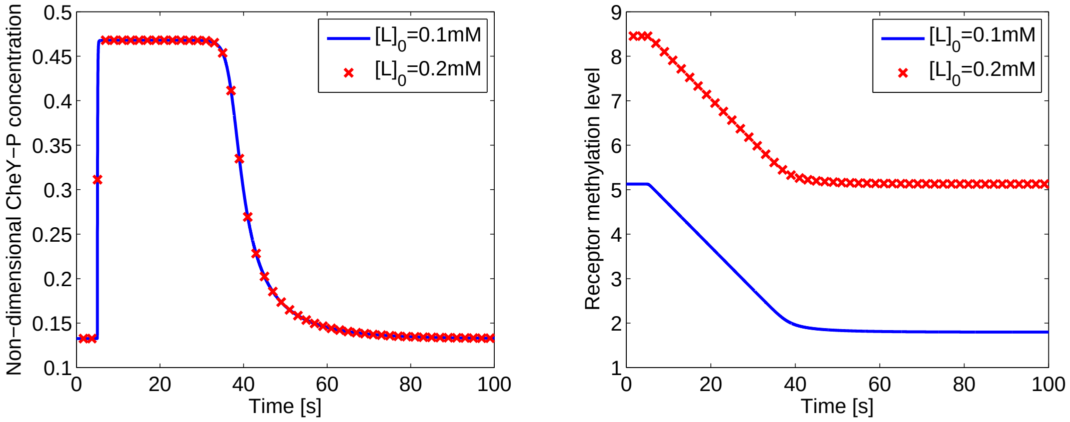
Within Section 4.3 we assumed that only Tar receptors are able to detect a chemoattractant gradient. However, there is evidence that *E. coli* cells will respond using both Tar and Tsr chemoreceptors. Tsr are predominantly serine sensing chemoreceptors, however they are also able to bind MeAsp (see Table 4.2), albeit with a much lower affinity [70]. As such we now consider whether it is feasible that a cell could exhibit FCD in the case where two or more receptor types can detect the same attractant gradient.

Theoretically it can be shown that FCD will hold for a model with two types of chemoreceptor as in equation (3.10). In this case, one must consider the sufficient conditions detailed in Section 4.2, however, the form of this model requires the use of a generalised function $\varphi(p, x)$ [121] of the form $\varphi = p^R$ (A. Hamadeh and E. Sontag, personal communication), where R represents the sum of all ν_i for each receptor type ($i = a, s$). The proof of this takes a very similar form to that in Section 4.3. More details are given in Appendix E. We undertook numerical simulations of our model using this assumption with the receptor free-energy described by equation (3.10) and found that all proteins in the signalling network demonstrated FCD as per the theoretical predictions. These numerical simulations are shown in Figure 4.5. Recent experimental work has shown FCD may be observed in cells expressing a wild-type receptor configuration [122]. This same work demonstrated the existence of two regimes of FCD behaviour, namely FCD1 ($[L]_0 = 0.018 - 0.23$mM) and FCD2 ($[L]_0 = 0.82 - 10.3$mM), where $[L]_0$ is an initial ligand (e.g. MeAsp) condition. In light of these results and those displayed in Figure 4.5 we consider the dissociation constant values displayed in Table 4.2 to see whether these two FCD regimes may be explained theoretically.

It is clear upon inspection of the dissociation constant values in Table 4.2 that it would be

Table 4.2: Dissociation constants of Tar and Tsr receptors for MeAsp, as reported in the literature.

Type	$K_{a/s}^{off}$	$K_{a/s}^{on}$	Source
Tar (a)	1.7×10^{-3}mM	0.5mM	[70, 129]
Tsr (s)	100mM	10^6mM	[70]



(a) CheY-P (b) Receptor methylation

Figure 4.5: Theoretically *E. coli* cells are able to display FCD in the presence of two receptor types. Plots displayed here are for a model containing both Tar and Tsr receptors. Responses are for a two-fold change in extracellular ligand concentration with initial concentrations of 0.1mM and 0.2mM. Panel (a) shows the response of CheY-P and (b) the receptor methylation level. Plots were created using K_s^{off} =0.1mM and $\nu_s = 1.4$.

difficult in the case where multiple receptors bind the same ligand to choose a ligand concentration that would ensure $K_i^{off} \ll [L] \ll K_i^{on}$ would hold for each receptor type. However, the conditions $[L] \ll K_i^{off} \ll K_i^{on}$ and $K_i^{off} \ll K_i^{on} \ll [L]$ may also result in similar receptor free-energy expressions to those given by equations (4.5) and (4.6) (see Table 4.3 for details). Hence, FCD may potentially be found even in cases where $K_i^{off} \ll [L] \ll K_i^{on}$ does not hold. Here this results in five possible FCD regimes, as seen in Figure 4.6. In particular, we may use the values in Table 4.2 to explain the two regimes of FCD behaviour. The FCD1 regime ($[L]_0$ =0.018-0.23mM) clearly satisfies $K_a^{off} \ll [L] \ll K_a^{on}$ for Tar receptors and $[L] \ll K_s^{off} \ll K_s^{on}$ for Tsr receptors, when binding MeAsp. Similarly, for the FCD2 regime ($[L]_0 = 0.82 - 10.3$mM) we satisfy $K_a^{off} \ll K_a^{on} \ll [L]$ for Tar receptors and $[L] \ll K_s^{off} \ll K_s^{on}$ for Tsr receptors. The relationship between these regimes of FCD behaviour and the sensitivity curve of a chemoreceptor signalling team is shown in Figure 4.7. Within this figure we can clearly see that the two experimental regimes of FCD behaviour correspond to maxima and minima of the sensitivity curve. This is perhaps unsurprising since at these points the sensitivity of the chemoreceptor signalling team is essentially constant for a range of ambient ligand concentrations.

Table 4.3: Effects of different ligand and dissociation constant conditions on the free energy expressions of active and inactive chemoreceptors.

Condition (K_i^{off})	Reduction to Active Receptor Free Energy	Reduction to Inactive Receptor Free Energy
$[L] \ll K_i^{off} \ll K_i^{on}$	$\ln(1+[L]/K_i^{on}) \approx \ln(1) = 0$	$\ln(1+[L]/K_i^{off}) \approx \ln(1) = 0$
$K_i^{off} \ll [L] \ll K_i^{on}$	$\ln(1+[L]/K_i^{on}) \approx \ln(1) = 0$	$\ln(1+[L]/K_i^{off}) \approx \ln([L]/K_i^{off})$
$K_i^{off} \ll K_i^{on} \ll [L]$	$\ln(1+[L]/K_i^{on}) \approx \ln([L]/K_i^{on})$	$\ln(1+[L]/K_i^{off}) \approx \ln([L]/K_i^{off})$

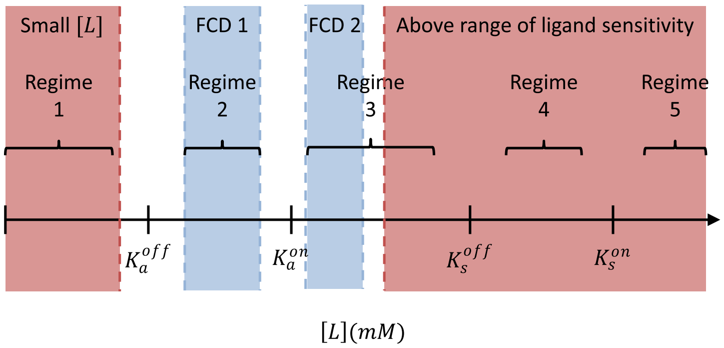
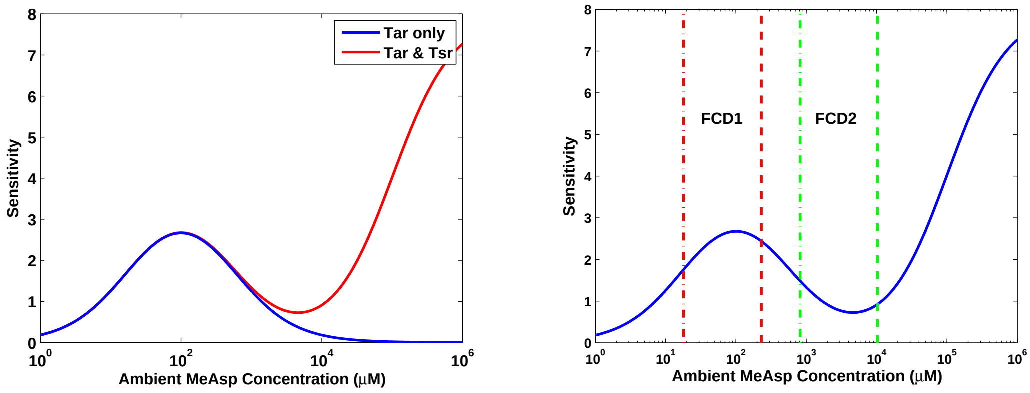
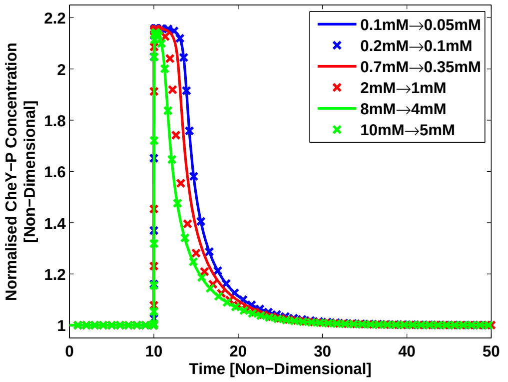


Figure 4.6: For a cell with two receptor types capable of sensing the same ligand we would predict five separate regimes of FCD behaviour (braces). Experimental work was able to demonstrate just two regimes of FCD behaviour, likely to be due to the cell's range of ligand sensitivity and small ligand concentrations not considered experimentally (red boxes). As such, we have two remaining regimes of FCD behaviour, namely FCD1 and FCD2 (blue boxes). Note that the regime FCD2 is truncated somewhat by the ligand sensitivity range of the cell.



(a) Tar Only vs. Tar & Tsr

(b) FCD and the Sensitivity Curve



(c) FCD Regime Demonstration

Figure 4.7: (a)After Mello & Tu [33], a plot showing the sensitivity of a chemoreceptor signalling team for different extracellular chemoattractant concentrations. The blue line shows sensitivity when only Tar receptors bind MeAsp whereas red shows the sensitivity when both Tar and Tsr receptors bind MeAsp. The effect of low affinity binding of MeAsp to Tsr receptors is demonstrated by the difference between the two curves. (b) Examining the sensitivity of *E. coli* receptor clusters within different ambient ligand concentrations shows that the two regimes of FCD behaviour predicted experimentally correspond to maxima and minima of the sensitivity curve. (c) Numerical simulations of the mathematical model using the full receptor free-energy expression, i.e. equation (3.10) demonstrate that FCD holds inside the two regimes of FCD behaviour (blue and green lines & crosses) whereas it fails outside of these (red line & crosses).

4.4.2 Multiple Different Ligand Types

In light of the results of Section 4.4.1 we now consider the effect of the cell detecting two different ligands (MeAsp and serine). This situation is of special interest since the affinity of Tsr receptors to MeAsp is so low, meaning MeAsp binding to Tsr receptors generally has very little effect. In this case we consider a situation whereby both receptor types are present in similar numbers and bind separate ligands. Parameter values have already been given for the binding of MeAsp to Tar receptors in Table 4.2. For the particular case considered here we must define a new chemoreceptor free-energy expression containing the effects of binding two different ligands. This is of the form

$$F = \left[\left(1 - \frac{m}{2}\right) + N_{Tar} \ln\left(\frac{[L_a]}{K_a^{off}}\right) + N_{Tsr} \ln\left(\frac{[L_s]}{K_{ser}^{off}}\right) \right], \tag{4.15}$$

within which $[L_{a,s}]$ denote the MeAsp and serine concentrations, respectively. The ligand (serine) dissociation constant of an inactive Tsr receptor is chosen such that $K_{ser}^{on} = 6 \times 10^{-3}$mM [64]. As per Kalinin et al. [64] we consider receptors to be present in different quantities, namely $N_{Tar} = 6$ and $N_{Tsr} = 12$, where N_{Tar} is the number of Tar receptors and N_{Tsr} is the number of Tsr receptors present in a signalling team.

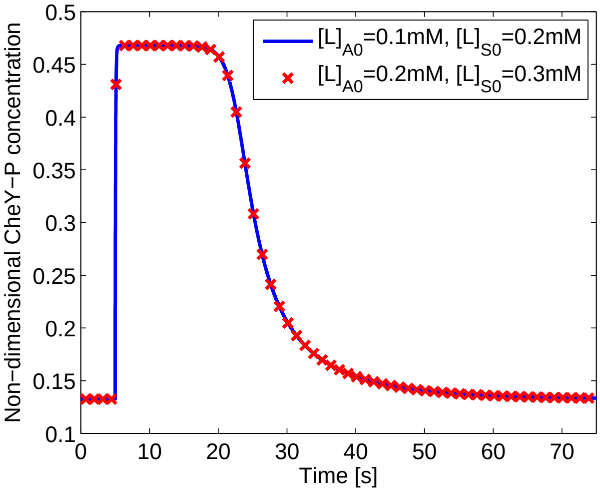


Figure 4.8: FCD may be displayed in situations where multiple receptor types act to bind separate ligands. Responses are for initial ligand concentrations of $[L]_{a0} = 0.1$mM with $[L]_{s0} = 0.2$mM and $[L]_{a0} = 0.2$mM with $[L]_{s0} = 0.3$mM. Ligand concentration changes were two-fold in MeAsp and three-fold in serine.

In order to test whether FCD can be displayed when cells sense two different ligand types we perform numerical simulations of the model utilising the chemoreceptor free-energy expression given by equation (4.15). Here we shall consider a different fold-change for each ligand in addition to the case where each ligand is subject to an equal fold-change. As with the previous multiple receptor case, theoretically we discover that FCD will hold when each receptor satisfies either

$[L] \ll K_i^{off} \ll K_i^{on}$, $K_i^{off} \ll [L] \ll K_i^{on}$ or $K_i^{off} \ll K_i^{on} \ll [L]$. Figure 4.8 displays numerical results verifying that FCD holds in this signalling team configuration. It should therefore be biologically feasible given experimentally determined values for the ligand dissociation constants of Tar and Tsr receptors to MeAsp and serine in both cases where receptors bind either the same or separate ligands.

4.5 Summary & Discussion

Within this chapter we have demonstrated, using both theoretical and numerical techniques, that the property of FCD is present within the model of the *E. coli* chemotaxis signalling pathway described in Chapter 3. In particular, it has been shown that FCD may hold in a number of different cases. These include situations whereby one receptor type binds one ligand type, two receptor types bind to one ligand type and finally, two receptor types separately bind two ligand types.

Using a parameter sensitivity analysis we were able to show that FCD is a robust property of the *E. coli* chemotaxis signalling network. However, considering a model within which receptor signalling team size is dependent upon the ambient ligand concentration causes FCD to fail. Importantly though we find only a small variation from FCD behaviour within the two experimentally determined FCD regimes.

Here we have demonstrated that the two experimentally determined regimes of FCD behaviour are reproduced by the mathematical model when we allow both Tar and Tsr chemoreceptors to bind the chemoattractant MeAsp. In particular, we may predict that for a cell with n different receptor types, able to bind the same ligand, there could exist $2n + 1$ separate regimes of FCD behaviour. However, examining results within the experimental literature we can justify the elimination of the three predicted additional regimes, leaving just those discovered by Lazova et al. [122].

The results in this chapter raise interesting questions as to the occurrence of FCD in more complicated cell signalling architectures. Results here tend to suggest that so long as the sensing components of the cascade (most commonly those areas of the cell where receptor-ligand binding is controlled) can exhibit FCD then the underlying signalling cascade should not affect the FCD response. This has interesting implications for other intracellular signalling cascades in prokaryotic and eukaryotic systems.

Chapter 5

Overshoot in *Escherichia coli*

In this chapter we seek to elucidate the mechanism(s) behind the phenomenon of overshoot within the chemotaxis signalling pathway of *E. coli* cells. We begin by summarising some key background literature in order to demonstrate the motivation behind this study. Here we hypothesise that overshoot may be modelled as damped oscillatory behaviour and seek conditions under which this hypothesis holds. A numerical routine is developed and utilised in order to search each possible two-parameter space for such conditions. Using this numerical routine we find that total protein concentrations and not kinetic rates must be responsible for the emergence of overshoot. A number of model reductions are subsequently considered, allowing an analytical condition to be found describing when overshoot will occur. Using this condition we are then able to identify a balance between receptor dynamics and CheB-P feedback as important in causing overshoot. Finally, results are compared to experimental data and discussed with reference to their biological feasibility.

5.1 Motivation

In response to a change in the extracellular ligand concentration, the phosphorylation level of proteins in the cell will undergo a rapid increase or decrease. After this, the phosphorylation level of proteins is usually assumed to gradually return its pre-stimulus value. However, where the phenomenon of overshoot is evident, cellular phosphorylation levels will transiently exceed their pre-stimulus levels in the opposing direction of the rapid initial change (see Figure 5.1).

The phenomenon of overshoot within signalling systems is not widely discussed in either the experimental or theoretical literature. However, some examples of each do exist.

Experimentally, overshoot was first observed in the late 1970s and 1980s. This work noted that if *E. coli* cells were challenged with an impulse stimulus (i.e. one lasting a very short period of time), they would exhibit overshoot [130, 36, 37, 46].

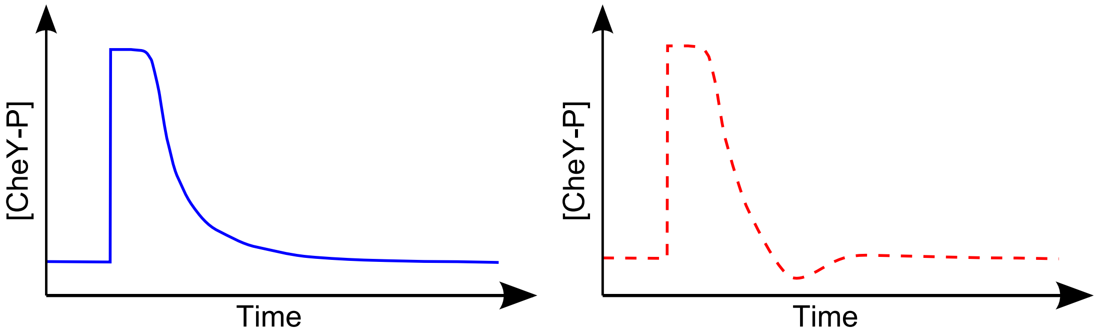


Figure 5.1: Diagrams showing two examples of the *E. coli* chemotactic response. (Left) In a cell that does not display overshoot, the initial rapid response is followed by a period of smooth transient behaviour in which the cell returns to pre-stimulus levels. (Right) When a cell does display overshoot there exists a transient period during which the cellular phosphorylation level exceeds its steady-state in the opposite direction from the initial rapid response.

Min et al. [131] have shown overshoot to exist in response to step-changes in the extracellular ligand concentration [131], i.e. where a stimulus persists indefinitely. This work utilised an experimental method known as 'optical tweezers' to demonstrate that the overshoot amplitude (the degree to which the cell exceeds the pre-stimulus value) is dependent upon the size of stimulus applied [131, 132].

Min et al. [131] also give alternative ideas as to why the phenomenon is not widely observed within experimental work. These are as follows.

- **Inadequate experimental duration:** Experimentally it is necessary to choose a time period over which to monitor features of interest. This time period may often be chosen to be as short as possible given the exact feature of interest so as to minimise both financial and time costs associated with the relevant experimental work. Thus in many cases experiments may not be of a sufficient duration to observe overshoot behaviour.

- **Insufficient temporal resolution:** Another necessary choice when designing experiments is that of the temporal resolution. In other words, it is necessary to choose how many data points must be captured within a given time frame so as to enable conclusions to be drawn from the resulting data. A number of limitations including equipment availability and setup times may limit the resolution chosen for a particular experiment. Since overshoot is a transient feature, the use of lower temporal resolutions may mean that there are not sufficient data points as to be able to observe overshoot.

- **Ligand dependency of overshoot amplitude:** The experimental work of Min et al. [131] shows that the amplitude of overshoot (i.e. the extent to which phosphorylation transiently exceeds steady-state) is dependent upon the size of ligand stimulus applied. It is likely therefore, that much of the experimental literature has considered stimuli that would yield either no or very small overshoot amplitudes that would be difficult to classify without prior knowledge of where it may occur.

Whilst there has been little experimental investigation into overshoot, there exist examples in the theoretical literature. A number of different mechanisms have been postulated as to the cause of overshoot.

One such example is that due to Hansen et al. [128] that considers a Monod-Wyman-Changeux (MWC) model of receptor clustering which has been adapted to allow transient interactions between small receptor teams within a larger fixed chemoreceptor array. Within this work the possibility of an overshoot in the methylation level of chemoreceptors was predicted. More specifically, when a cell responds to a stimulus composed of just one ligand it is suggested that non-ligand binding chemoreceptor types may become transiently methylated due to their coupling with ligand binding receptors. An example of this is that when challenged with an aspartate stimulus, Tar receptors would methylate as normal, however Tsr (serine sensing chemoreceptors) may also become methylated due to the coupling of the two receptor types.

The work of Lan et al. [133] predicted a similar mechanism using a local adaptation model. This work predicted that the activity of one chemoreceptor type may overshoot due to cross-talk in receptor methylation. This work gives an example in which the proposed mechanism yields overshoot in response to an aspartate stimulus. In this case, the average methylation level of Tar chemoreceptors will monotonically reach a new steady-state level whilst the average methylation level of Tsr chemoreceptors would change only transiently before returning to its previous steady-state. As a result of this, the average Tar chemoreceptor activity would relax monotonically to steady-state whereas that of Tsr chemoreceptors would overshoot before fully returning to it pre-stimulus steady-state.

Whilst it is clear that overshoot has been mentioned within the theoretical literature, it has mainly been referred to within a study into other features of the chemotactic signalling pathway. As such, the mechanisms associated with overshoot have not been clearly identified. Thus within this chapter we analyse the mathematical model from Chapter 3 and investigate the roles of pathway signalling dynamics, negative feedbacks and their associated timescales and how these combine to form a mechanism by which cells exhibit overshoot.

5.2 Hypothesis

Prior to analysing mathematical models, it is often useful to form hypotheses in advance which may act as a sense check of the results obtained. As such we begin by considering the network structure of the *E. coli* chemotaxis signalling pathway. Upon doing so we note the existence of a negative feedback loop in which the protein CheB-P acts to reduce the methylation level of the chemoreceptors. This particular feature of the network is of interest since negative feedback is known to have the potential for creating oscillatory behaviour within monotone dynamical systems (i.e. those in which proteins activating certain processes do not repress that process at a different concentration and vice versa) [134, 135, 136].

Motivated by the existence of this negative feedback loop, we postulate that overshoot can be described as damped oscillatory behaviour. As such, in the remainder of this chapter, we consider the model laid out in Chapter 3 and ask under what conditions it exhibits overshoot.

5.3 Methodology

In this section we consider methods that are needed in order to test the hypothesis stated in Section 5.2. The hypothesis that overshoot may be modelled as damped oscillatory behaviour requires an investigation into how variation in parameter values affect eigenvalues of the system steady-state. Thus we appeal to the theory of asymptotic stability analysis (see Appendix C).

Within Section 3.2.4, an asymptotic stability analysis for the model laid out in Chapter 3 showed that it displays (non-oscillatory) stable behaviour for the parameter values in Table 3.1. In this case, for an eigenvalue described as

$$\lambda = a + ib, \tag{5.1}$$

the magnitudes of the real (a) and imaginary (b) parts are such that $a < 0$ and $b = 0$, respectively. However, here we are interested in the emergence of damped oscillatory (overshoot) behaviour. As such we are seeking conditions leading to eigenvalues with $a < 0$ and $b \neq 0$, respectively.

In order to find such conditions we consider variation in each of the parameters in the system. In particular we wish to examine the effects of variation in the eight kinetic rate parameters (denoted k_i for $i = 1, 2, 3, 4, 5, 6$ and g_j where $j = B, R$) and the total intracellular protein concentrations (denoted $[X]_T$ where $X = A, B, R, Y, Z$). For the purposes of simple visualisation and interpretation of results we consider pair-wise variations of parameters, i.e. we vary two parameters at a time and examine for which combinations overshoot is exhibited. As such twenty eight separate pairs of kinetic rates must be examined in addition to ten distinct pairs of total protein concentrations.

It will clearly not be practical to study the asymptotic stability of the required 38 parameter pairs by hand for a four dimensional non-linear ODE model. As such a numerical routine was formulated that is capable of examining the asymptotic stability in the region about the steady-state of a system over some pre-defined range for any pair of parameters in the system. This stability analysis routine takes the following form.

Algorithm 1

```
Define a base set of parameters
Loop over values of selected kinetic rate/protein concentration
    Loop over values of selected kinetic rate/protein concentration
        Calculate steady-state values of each variable
        Construct the Jacobian matrix
        Solve for the eigenvalues
        Store required parts of the eigenvalues
    End loop
End loop
Plot results
```

The routine discretises a two parameter space into a grid of equally spaced values. Using a similar approach to that in Section 3.2.3, the routine calculates the system steady-state. A stability analysis about this steady-state is then undertaken, using the approach in Section 3.2.4, for each point in that discretised parameter space. Results obtained from this routine can then be used to see for which combinations of those two chosen parameters damped oscillatory behaviour is possible.

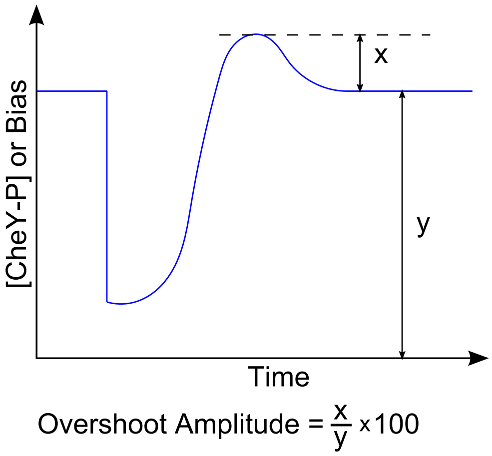


Figure 5.2: Overshoot amplitude is calculated as a percentage of either the steady-state CW flagellar rotation bias or the CheY-P concentration. Shown is an example of overshoot amplitude calculated for a step-increase in external ligand concentration. The magnitude of the overshoot is divided by the steady-state CheY-P concentration or CW rotational bias and then multiplied by 100, giving a value for the overshoot amplitude.

In addition to the numerical routine defined here we also wish to calculate overshoot amplitude. This is necessary so that we may compare the theoretical results obtained within this chapter to the experimental data of Min et al. [131]. Taking a lead from this experimental work we consider overshoot amplitude to be calculated as a percentage of the steady-state CW bias or CheY-P concentration, as shown in Figure 5.2. For a step-increase in the extracellular ligand concentration we take the peak amplitude of the overshoot i.e. the maximum CW bias or CheY-P concentration minus their steady-state value. This is then divided by the relevant steady-state and multiplied by 100, giving overshoot amplitude as a percentage of the steady-state value. Other definitions of overshoot amplitude could also have been considered, however, we chose this method so as to allow a fair comparison between our results and those of Min et al. [131].

5.4 Full Four-Dimensional Model Results

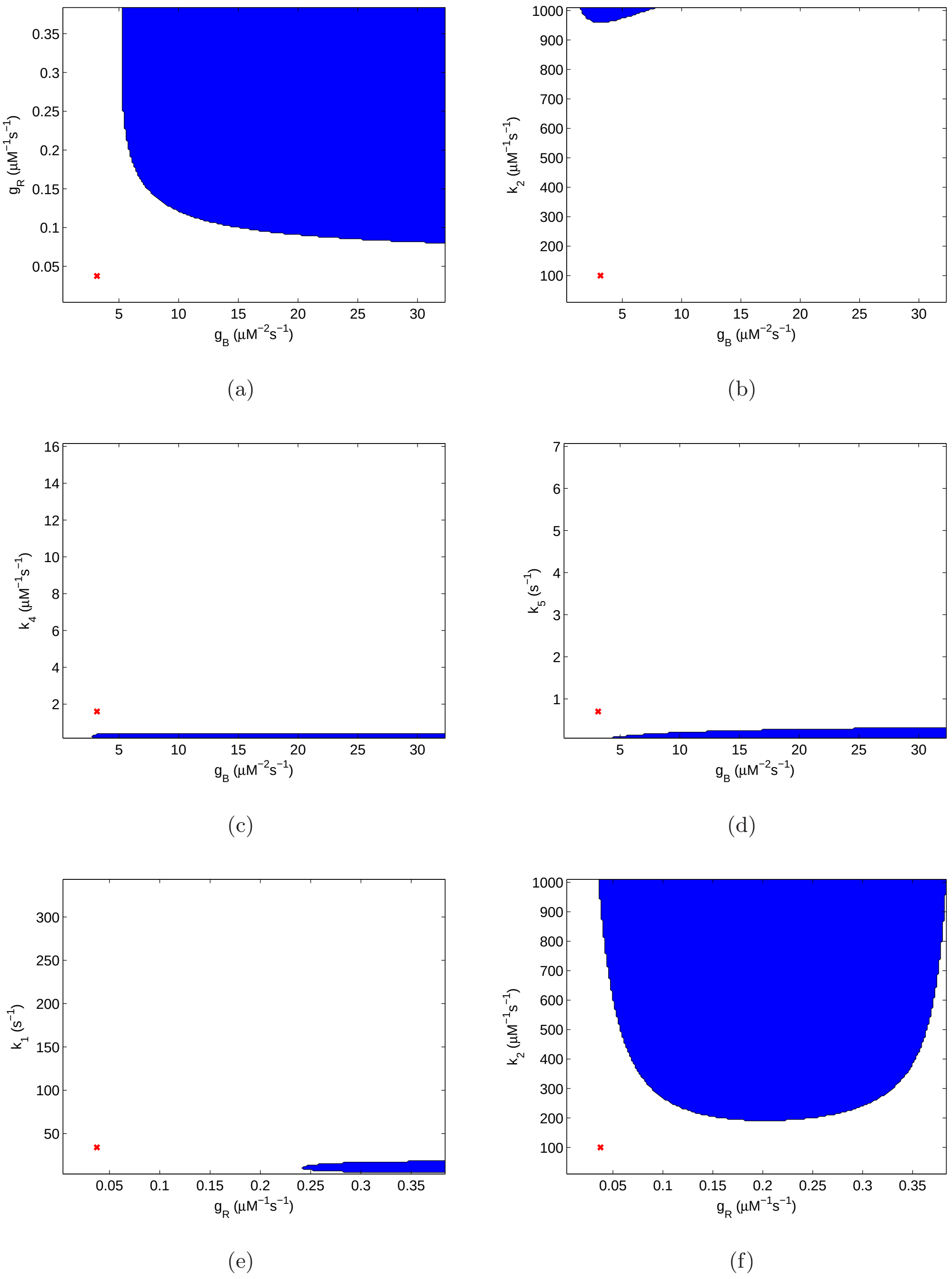
Within this section we utilise the numerical routine from Section 5.3 in order to study the effects of pair-wise variations in kinetic rate parameters and total protein concentrations.

Application of the numerical routine to the model from Chapter 3 yields the results displayed in Figures 5.3-5.5.

In support of the hypothesis that the observed negative feedback may be the cause of overshoot in the chemotactic response of *E. coli* cells, we find that damped oscillatory behaviour is observed for a number of pair-wise variations in kinetic rate parameters as seen in Figures 5.3-5.5. In particular we find that 19 out of 28 possible pairs of kinetic rates yielded damped oscillations, while the remaining 9 did not when each parameter was varied over both a ten-fold range upward and downward. However, each protein molecule within the cell is essentially identical and as such we would not expect the kinetic rates associated with protein-protein reactions to vary to the extent required here. As such, the ten-fold (upward and downward) ranges considered in Figures 5.3-5.5 do not represent biologically feasible ranges. Upon closer inspection of these results we notice that in order for any of the base parameter values (red crosses) to be moved into a blue region of damped oscillatory behaviour would require a larger variation than we would expect to be biologically feasible.

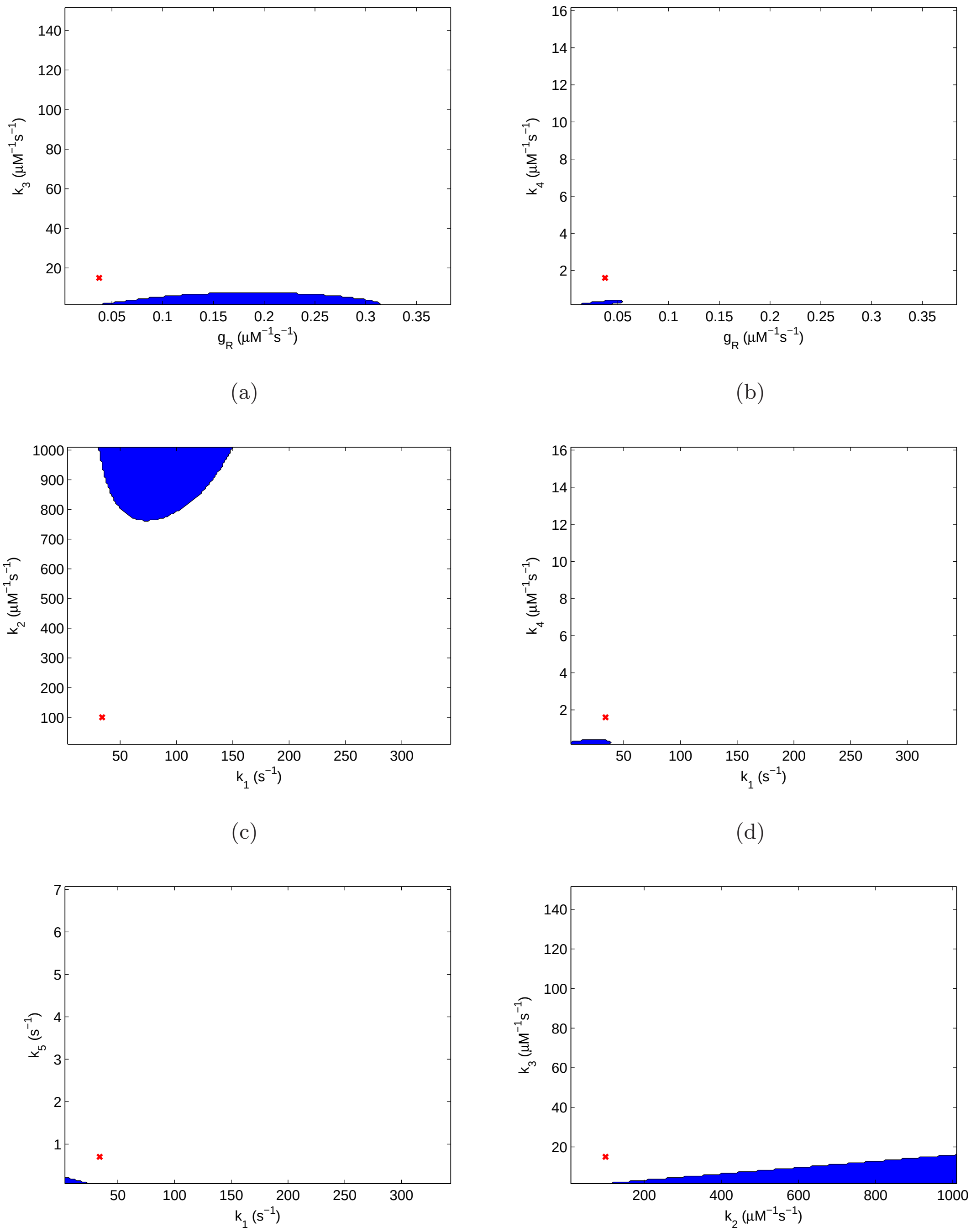
Since pair-wise variations in the kinetic rate parameters of this system do not appear to be able to explain the emergence of overshoot behaviour, we consider an alternative explanation. As such we consider the experimental literature, within which it is known that intracellular protein concentrations are subject to significant stochastic variation between cells [137]. In particular, the concentration of each signalling protein has been shown to vary by as much as ten-fold depending on strain and growth conditions, albeit with certain concentration ratios being more tightly maintained than others [112].

Based upon these experimental observations, pair-wise variation over a ten-fold range in each of the chemotaxis signalling proteins was considered. This yielded results demonstrating that overshoot may be found under eight out of ten separate pair-wise protein concentration variations, as seen in Figure 5.6. Interestingly, we may observe from these results that pair-wise variations involving the proteins CheB, CheR, CheY and CheZ allowed overshoot to occur more



(a) (b) (c) (d) (e) (f)

Figure 5.3: Plots showing areas of parameter space in which oscillatory behaviour may be found by varying two kinetic rate parameters (blue). These are regions in which at least two eigenvalues of the system have non-zero imaginary part. A red cross is used to show the location of our base parameter set.

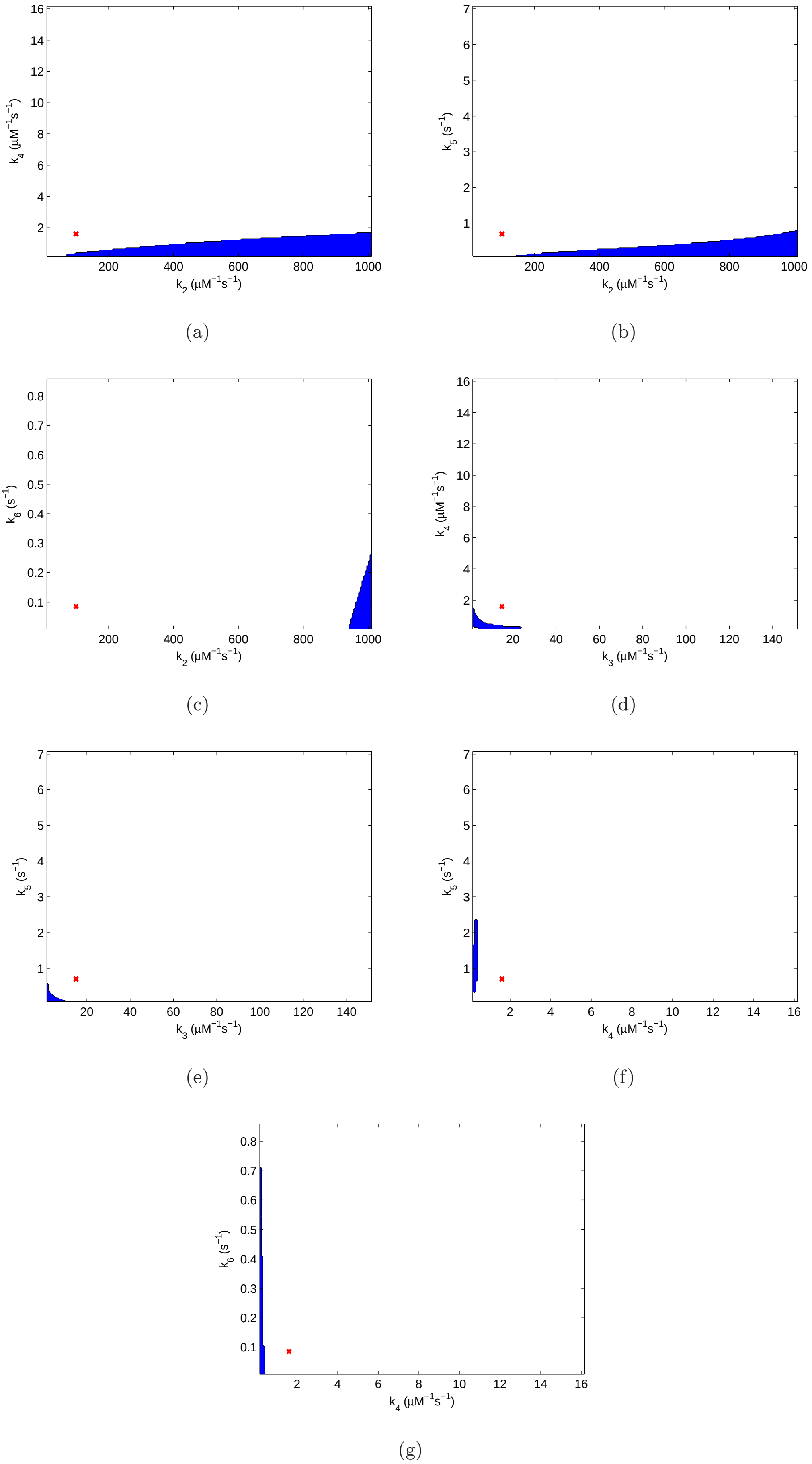


(a) (b)

(c) (d)

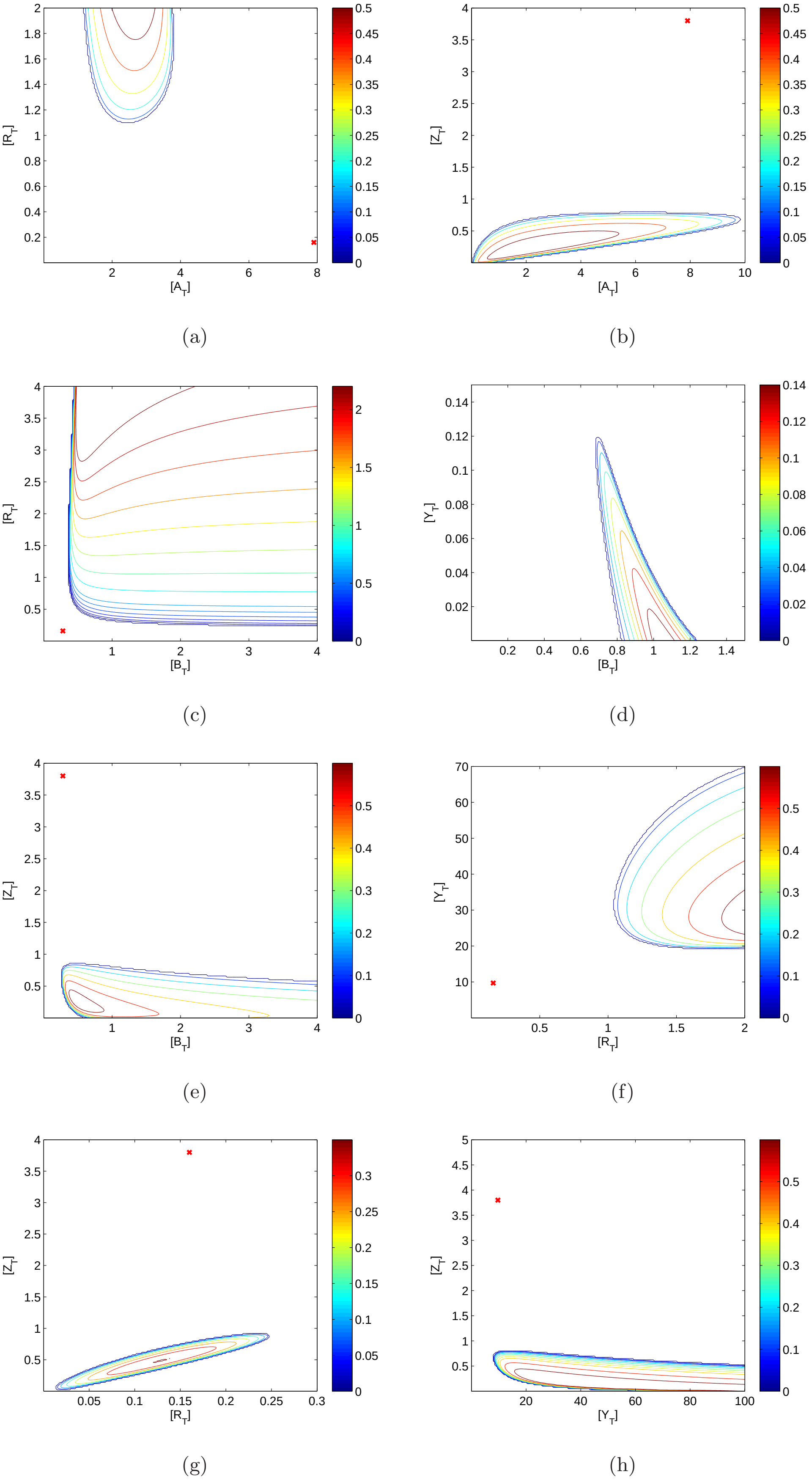
(e) (f)

Figure 5.4: Plots showing areas of parameter space in which oscillatory behaviour may be found by varying two kinetic rate parameters (blue). These are regions in which at least two eigenvalues of the system have non-zero imaginary part. A red cross is used to show the location of our base parameter set.



(a) (b) (c) (d) (e) (f) (g)

Figure 5.5: Plots showing areas of parameter space in which oscillatory behaviour may be found by varying two kinetic rate parameters (blue). These are regions in which at least two eigenvalues of the system have non-zero imaginary part. A red cross is used to show the location of our base parameter set.



(a)

(b)

(c)

(d)

(e)

(f)

(g)

(h)

Figure 5.6: Areas of parameter space in which oscillatory behaviour may be found by varying the total concentrations of two chemotaxis proteins. These are regions in which at least two eigenvalues of the system have a non-zero imaginary part. The colours of the contour lines represent the magnitudes of the imaginary parts of the eigenvalues obtained from the fourth-order system. Note that red crosses indicate the location of our base parameter set (outside of the plotted range in (d)). All concentrations are expressed in μM.

readily than those pairs that included variation of the CheA concentration. Of particular interest amongst the results obtained here is that where the total concentrations of CheB and CheY were varied (see Figure 5.6(d)). In this particular case we note that in order to achieve overshoot, the total concentrations of CheB and CheY must be increased and decreased, respectively from the experimentally determined population average (i.e. from the values in Table 3.1). This indicates that the ratio in total concentrations of CheB and CheY is important for observing overshoot. Since phosphoryl groups are transferred from CheA-P to both CheB and CheY, the ratio of these protein concentrations will clearly affect how many phosphoryl groups are passed to each protein which, in turn, will affect the timescale and strength of CheB-P feedback onto the receptor state.

The observation that certain pairs of proteins in the signalling cascade are more able to produce overshoot than others led us to consider the ways in which protein concentrations are managed within the cell. Proteins in the chemotaxis signalling pathway of *E. coli* cells are known to be encoded by two separate operons (see glossary for definition), namely *mocha* (CheA, CheW) and *meche* (CheB, CheR, CheY, CheZ) [138]. This operon structure is a key mechanism by which cells are able to maintain suitable protein levels and ratios. Where proteins are encoded in such a manner, we would expect the ratios of protein concentrations from the same operon to be approximately fixed. Variation would be expected to exist between proteins of different operons. Based upon these observations we therefore choose to group proteins as per the operon in which they are encoded and allow variation over a ten-fold range. This entails varying the concentration of CheA against concentrations of CheB, CheR, CheY and CheZ which are varied in such a manner as to maintain constant ratios between them, the result of which is shown in Figure 5.7.

We may see from the results in Figure 5.7 that a ∼3-fold increase in all protein concentrations was sufficient to produce overshoot behaviour. This falls well within the biologically realistic ten-fold variation. While individual total protein concentrations have been shown to vary by as much as ten-fold, the ratios between protein concentrations have not been shown to vary by much more than ∼30% [112]. Protein co-expression only limits stochastic fluctuations in the ratios of proteins encoded within the same operon. Hence we would expect the maximal variation to occur between proteins encoded by separate operons. Thus a biologically feasible range within Figure 5.7 would be such that $0.7\alpha \leq \beta \leq 1.3\alpha$ where α is a fold-change in the proteins of one operon and β a fold-change in the proteins of the other operon. It is also noteworthy here that beyond a ∼4-fold increase in the concentration of CheA, the vast majority of change in magnitude of the imaginary parts of the eigenvalues appears to be caused by variation in *meche* operon proteins. This is supported by the earlier observation that pairs of proteins involving

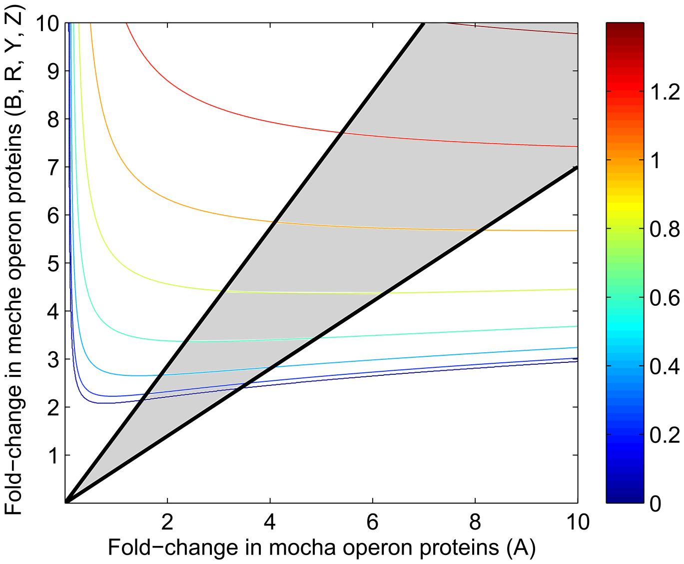


Figure 5.7: Operon-wise variation in protein concentration can produce overshoot behaviour. A plot showing the magnitude of imaginary parts of the eigenvalues obtained when taking operon wise fold-changes in chemotaxis protein concentrations. The biologically feasible region is shaded in grey.

CheB, CheR, CheY and CheZ were more readily able to produce overshoot than those where CheA was included.

5.5 Model Reduction Analysis

In this section we consider a number of reductions to the original model. This was done in order to further investigate the roles of network structure, negative feedback and dynamic timescales of the *E. coli* chemotaxis signalling cascade. In particular, we restrict our attention here to the study of variation in total intracellular protein concentration. Using these reduced mathematical models it should become more feasible to obtain analytical results that will lead to a greater understanding of the mechanism(s) causing overshoot to occur.

When handling high order dynamical systems, a common approach is to use model reduction techniques in order to produce lower order systems. Such reduced models are more amenable to mathematical analysis due to the fact they include less equations and most likely fewer parameters. However, there is a trade off between the convenience associated with these simpler models and their ability to retain all key features of the full system (see Flach & Schnell [139]).

5.5.1 Model 1 - Third-Order System: QSSA Applied to CheY-P

Examination of the non-dimensional parameter set reveals that the rate of phosphotransfer onto CheY is very fast in comparison to all other rates in the system (see Table 3.2). As such we assume that this rate is fast enough that the protein CheY-P may be assumed to equilibrate

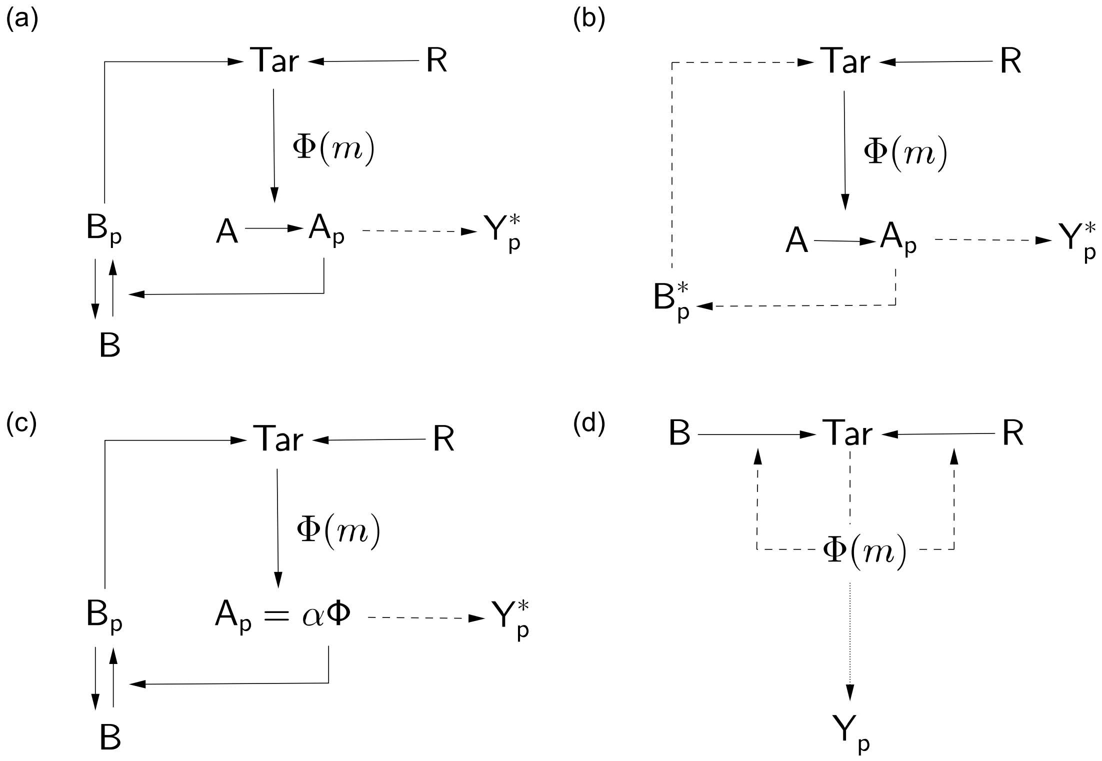


Figure 5.8: Schematic representations of the four model reductions considered. (a) Reduction to third-order by applying the quasi-steady-state approximation to CheY-P. (b) Reduction to second-order via application of the quasi-steady-state approximation to both CheY-P and CheB-P. (c) Reduction to second-order by assuming CheA-P may be represented by a multiple scaling of receptor signalling team activity and representing CheY-P as a decouplable read-out variable. (d) A first-order model due to Tu et al. [63]. Here solid lines indicate interactions while dashed lines indicate quasi-steady-state/read-out variables. The dotted line in (d) represents the decoupled expression for CheY-P.

instantly (see Fig. 5.8(a)). This results in the application of the quasi-steady-state approximation (QSSA) to equation (3.18) which yields a third-order dynamical system of the form

$$\frac{dm}{d\tau} = \gamma_R(1-\Phi) - \gamma_B b_p^2 \Phi \qquad = r_1(m, b_p), \tag{5.2}$$

$$\frac{da_p}{d\tau} = \Phi \bar{k}_1(1-a_p) - \bar{k}_2(1-y_p^*)a_p - \bar{k}_3(1-b_p)a_p \qquad = r_2(m, a_p, b_p), \tag{5.3}$$

$$\frac{db_p}{d\tau} = \alpha_2 \bar{k}_3(1-b_p)a_p - b_p \qquad = r_3(a_p, b_p), \tag{5.4}$$

within which Φ is defined by equations (3.9) and (3.11) whilst y_p^* represents the non-dimensional steady-state for CheY-P, as given by

$$y_p^* = \frac{k_2[A_T]a_p}{k_2[A_T]a_p + k_4[Z_T] + k_6}. \tag{5.5}$$

Comparing results obtained from this reduced model and the full system shows that there are only small changes in the steady-state values obtained, as can be seen in Figure 5.9.

In order to compare the stability characteristics of this reduced system and the full fourth-

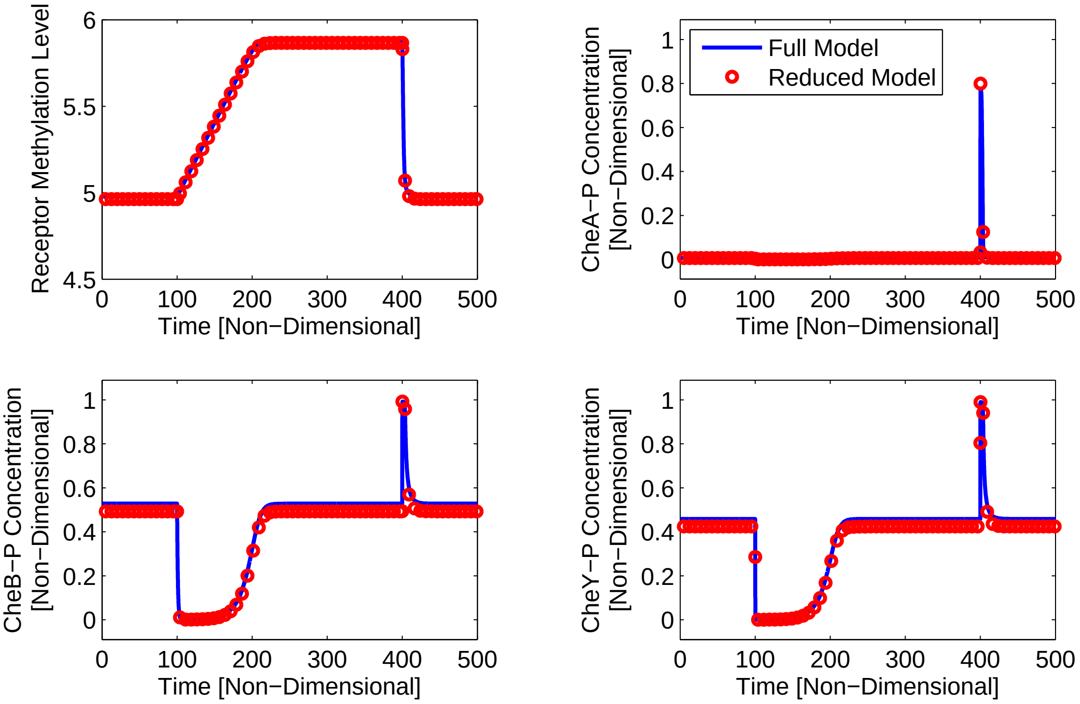


Figure 5.9: A plot displaying a behavioural comparison between the full fourth-order system and the reduced third-order system of Section 5.5.1. Here a blue line represents the behaviour of the full-system whereas red circles display the behaviour of the reduced model system. In this case, in order to obtain the reduced model representation of CheY-P, equation (5.5) has been solved at each time point.

order system we utilise our numerical routine from Section 5.3. As with the analysis of the full fourth-order system we must consider the eigenvalues of the model's Jacobian matrix. This is given by

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial r_1}{\partial m} & \frac{\partial r_1}{\partial a_p} & \frac{\partial r_1}{\partial b_p} \\ \frac{\partial r_2}{\partial m} & \frac{\partial r_2}{\partial a_p} & \frac{\partial r_2}{\partial b_p} \\ \frac{\partial r_3}{\partial m} & \frac{\partial r_3}{\partial a_p} & \frac{\partial r_3}{\partial b_p} \end{pmatrix}, \tag{5.6}$$

within which

$$\frac{\partial r_1}{\partial m} = \frac{-g_R[R_T]Ne^F\Phi^2 - g_B[B_T]^2 b_p^2 Ne^F\Phi^2}{2k_5} = \rho_1, \tag{5.7}$$

$$\frac{\partial r_1}{\partial b_p} = \frac{-2g_B[B_T]^2 b_p\Phi}{k_5} = \rho_2, \tag{5.8}$$

$$\frac{\partial r_2}{\partial m} = \frac{k_1(1-a_p)Ne^F\Phi^2}{2k_5} = \rho_3, \tag{5.9}$$

$$\frac{\partial r_2}{\partial a_p} = \frac{-1}{k_5}\left(k_1\Phi + k_2[Y_T](1-y_p^*) + k_2[Y_T]\left(y_p^{*2} - y_p^*\right) + k_3[B_T](1-b_p)\right) = \rho_4, \tag{5.10}$$

$$\frac{\partial r_2}{\partial b_p} = \frac{k_3[B_T]a_p}{k_5} = \rho_5, \tag{5.11}$$

$$\frac{\partial r_3}{\partial a_p} = \frac{k_3[A_T](1-b_p)}{k_5} \qquad = \rho_6, \tag{5.12}$$

$$\frac{\partial r_3}{\partial b_p} = \frac{-k_3[A_T]a_p}{k_5} - 1 \qquad = \rho_7, \tag{5.13}$$

are the partial derivatives of equations (5.2)-(5.4) with respect to each of the three system variables. In order to obtain the eigenvalues of the system it is necessary to find the characteristic polynomial of this Jacobian matrix. In this case the eigenvalues are obtained by calculating the roots of the characteristic polynomial given by

$$p(\lambda) = \det|\boldsymbol{J} - \lambda\boldsymbol{I}| = \begin{vmatrix} \rho_1 - \lambda & 0 & \rho_2 \\ \rho_3 & \rho_4 - \lambda & \rho_5 \\ 0 & \rho_6 & \rho_7 - \lambda \end{vmatrix}. \tag{5.14}$$

Upon examining the results obtained using our numerical routine, we may observe that only small differences exist between the stability characteristics of this reduced system and the full fourth-order system. Since there is very little difference in the steady-state and stability characteristics of this reduced system when compared to the full fourth-order system, we may conclude that analysis of this system may be of use in examining the cause of overshoot in the chemotactic response of *E. coli* cells. As such, this model is considered in greater detail in Section 5.6.

5.5.2 Model 2 - Second-Order System: QSSA Applied to CheB-P & CheY-P

Analysis of dynamical systems becomes significantly simpler as their dimensionality is reduced. As such, in spite of the apparent validity of the third-order reduction in Section 5.5.1 we investigate whether further reductions may be suitable.

Upon examining the parameter values in Table 3.1 we see that the rate of phosphorylation of CheB-P is quite large in relation to a number of other rates. With this being the case we now investigate the applicability of the quasi-steady-state approximation to both CheY-P and CheB-P (as in Fig. 5.8(b)). Doing so results in a second-order dynamical system of the form

$$\frac{dm}{d\tau} = \gamma_R(1-\Phi) - \gamma_B b_p^{*2}\Phi \qquad = r_4(m, a_p), \tag{5.15}$$

$$\frac{da_p}{d\tau} = \Phi\bar{k}_1(1-a_p) - \bar{k}_2(1-y_p^*)a_p - \bar{k}_3(1-b_p^*)a_p \qquad = r_5(m, a_p), \tag{5.16}$$

where Φ is given in equations (3.9) and (3.11) whilst y_p^* and b_p^* represent the non-dimensional steady-states of CheY-P and CheB-P, respectively. The non-dimensional steady-state of CheB-P is given by

$$b_p^* = \frac{k_3[A_T]a_p}{k_3[A_T]a_p + k_5}, \tag{5.17}$$

while y_p^* is given by equation (5.5).

In order to investigate the stability characteristics of this system it is necessary to investigate the eigenvalues of the Jacobian matrix. In this particular case the Jacobian matrix is of the form

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial r_4}{\partial m} & \frac{\partial r_4}{\partial a_p} \\ \frac{\partial r_5}{\partial m} & \frac{\partial r_5}{\partial a_p} \end{pmatrix}, \tag{5.18}$$

where

$$\frac{\partial r_4}{\partial m} = -\frac{N g_R [R_T] e^F \Phi^2}{2k_5} - \frac{N [A]_T^2 [B]_T^2 g_B k_3^2 a_p^2 e^F \Phi^2}{2k_5(k_3[A]_T a_p + k_5)} = \rho_8, \tag{5.19}$$

$$\frac{\partial r_4}{\partial a_p} = \frac{2 g_B k_3^3 [A]_T^3 [B]_T^2 a_p^2 \Phi}{k_5(k_3[A]_T a_p + k_5)^3} - \frac{2 g_B k_3^2 [A]_T^2 [B]_T^2 a_p \Phi}{k_5(k_3[A]_T a_p + k_5)^2} = \rho_9, \tag{5.20}$$

$$\frac{\partial r_5}{\partial m} = \frac{N k_1 (1 - a_p) e^F \Phi^2}{2k_5} = \rho_{10}, \tag{5.21}$$

$$\begin{aligned} \frac{\partial r_5}{\partial a_p} = & -\frac{k_1 \Phi}{k_5} - \frac{k_2 [Y]_T (1 - y_p^*)}{k_5} - \frac{k_2 [Y]_T (y_p^{*2} - y_p^*)}{k_5} - \frac{k_3 [B]_T (1 - b_p^*)}{k_5} \\ & - \frac{k_3 [B]_T (b_p^{*2} - b_p^*)}{k_5} = \rho_{11}, \end{aligned} \tag{5.22}$$

are the partial derivatives with respect to each of the system variables. Thus, eigenvalues of this system are obtained by solving

$$\lambda^2 - \lambda\left(\rho_8 + \rho_{11}\right) + \rho_8 \rho_{11} - \rho_9 \rho_{10} = 0. \tag{5.23}$$

Analysis of eigenvalues obtained from equation (5.23), found using the numerical routine laid out in Section 5.3, shows that the region of parameter space in which damped oscillations are found is significantly altered. In particular, we require a larger fold-change in all total protein concentrations in order for Model 2 to exhibit overshoot. We must therefore conclude that this particular reduction is not valid. In fact it would appear that the cause of this is that we have removed the timescale upon which the feedback of CheB-P acts on the chemoreceptors. This supports the assertion of the main text that the CheB-P feedback timescale is important for overshoot to be observed.

5.5.3 Model 3 - Second-Order System: CheA-P is a Multiple of Receptor Activity

Due to the failure of the second-order model reduction considered in Section 5.5.2 we now investigate an alternative method of reducing this model to a second-order dynamical system. In this particular case we consider the concentration of CheA-P to be a simple multiplicative scaling of the receptor signalling team activity, i.e. $\Lambda\Phi \approx [A_p]$, in which Λ is calculated at steady-state from a numerical simulation of the full system using parameters from Table 3.1 (see Fig. 5.8(c)) and Φ is described by equations (3.9) and (3.11). In addition to this we consider CheY-P to be a decouplable read-out variable as in the model of Tu et al. [63]. This results in a second-order dynamical system of the form

$$\frac{dm}{d\tau} = \gamma_R(1-\Phi) - \gamma_B b_p^2 \Phi \qquad = r_6(m, b_p), \tag{5.24}$$

$$\frac{db_p}{d\tau} = \alpha_2 \bar{k}_3 \Lambda\Phi(1-b_p) - b_p \qquad = r_7(m, b_p), \tag{5.25}$$

in which Λ is the multiplicative scaling of Φ such that $\Lambda\Phi \approx a_p$.

In this case the assumption that CheY-P concentration is a (decoupled) output variable means that all phosphoryl groups produced by CheA must transfer to CheB. Clearly this means that the sharing of phosphoryl groups between CheB and CheY has been removed from the system. We can also see that the timescale upon which CheA autophosphorylates has been removed.

Once again, in order to analyse the stability characteristics of this system we must investigate the eigenvalues of the Jacobian matrix. For this particular reduced model the Jacobian matrix is of the form

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial r_6}{\partial m} & \frac{\partial r_6}{\partial b_p} \\ \frac{\partial r_7}{\partial m} & \frac{\partial r_7}{\partial b_p} \end{pmatrix}, \tag{5.26}$$

within which the partial derivatives with respect to each of the system variables are given by

$$\frac{\partial r_6}{\partial m} = -\frac{N g_R [R_T] e^F \Phi^2}{2k_5} - \frac{N g_B [B]_T^2 b_p^2 e^F \Phi^2}{2k_5} \qquad = \rho_{12}, \tag{5.27}$$

$$\frac{\partial r_6}{\partial b_p} = -\frac{2 g_B [B]_T^2 b_p \Phi}{k_5} \qquad = \rho_{13}, \tag{5.28}$$

$$\frac{\partial r_7}{\partial m} = \frac{N \Lambda k_3 [A]_T (1-b_p) e^F \Phi^2}{2k_5} \qquad = \rho_{14}, \tag{5.29}$$

$$\frac{\partial r_7}{\partial b_p} = -1 - \frac{k_3 \Lambda [A]_T \Phi}{k_5} \qquad = \rho_{15}. \tag{5.30}$$

Here eigenvalues are obtained upon solving

$$\lambda^2 - \lambda\left(\rho_{12} + \rho_{15}\right) + \rho_{12}\rho_{15} - \rho_{13}\rho_{14} = 0, \tag{5.31}$$

in which ρ_{12-15} are given by equations (5.27)-(5.30). Analysing these eigenvalues shows that a large fold-change (∼9.5 fold increase) in all protein concentrations is required for this model to display overshoot behaviour, as seen in Figure 5.10. Clearly this represents a significant change from the original fourth-order dynamical system. Thus we conclude that Model 3 does not represent a valid reduction for the system. This supports the notion that the timescales of various reactions in the system as well as the sharing of phosphoryl groups between CheB and CheY are important features in the mechanism causing damped oscillatory behaviour to be observed.

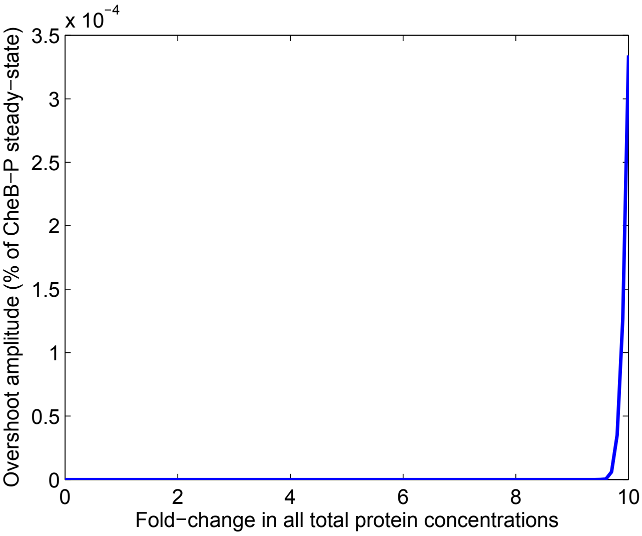


Figure 5.10: A large fold-change in all total protein concentrations is required to obtain overshoot from a two ordinary differential equation model (equations (5.24) and (5.25)) in which the CheA-P concentration is taken to be a multiplicative scaling of the receptor signalling team activity. The line indicates the overshoot amplitude obtained within the protein CheB-P. Here CheB-P is used in place of CheY-P since this has been eliminated from the model.

5.5.4 Model 4 - First-Order System: Tu et al. Model

In addition to the model reductions considered in Sections 5.5.1-5.5.3 we also analyse an example contained within the theoretical literature. The particular model considered is that due to Tu et al. [63]. This model consists of an ODE for the methylation kinetics of the chemoreceptors that is defined by

$$r_8 = \frac{dm}{dt} = k^R_{cat}[R_T]\frac{1-\Phi}{1-\Phi+K^R_M} - k^B_{cat}[B_T]\frac{\Phi}{\Phi+K^B_M}, \tag{5.32}$$

in which Φ is defined as per equations (3.9) and (3.11) whilst $k_{cat}^{R/B}$ and $K_M^{R/B}$ are catalytic rates and Michaelis-Menten constants of CheR and CheB, respectively. Here the concentration of CheY-P is described by

$$\frac{d[Y_p]}{dt} = k_a \Phi - \frac{[Y_p]}{\tau_z}, \tag{5.33}$$

in which k_a is the rate of phosphotransfer from CheA-P onto CheY and τ_z is the dephosphorylation time of the protein CheY-P. This model (see Figure 5.8(d)) is based upon a number of assumptions including CheB acting only on active receptors, CheR only acting upon inactive receptors and that CheY-P decouples from equation (5.32). In addition to this we note that this model utilises only the total concentration of proteins CheR and CheB, implying that the phosphorylated quantity is not important in determining the receptor state. In terms of our analysis we need only investigate equation (5.32) since equation (5.33) decouples from equation (5.32). Due to the failure of the second-order systems considered in Sections 5.5.2 and 5.5.3, as well as the lack of negative feedback in this model, we do not expect it to yield the damped oscillatory behaviour examined in this chapter.

Since equation (5.33) decouples from this system we need only consider the stability of equation (5.32). This is done by considering

$$\lambda = \frac{\partial r_8}{\partial m} = \frac{N k_{cat}^B [B_T] e^F \Phi^3}{2(K_M^B + \Phi)^2} - \frac{N k_{cat}^B [B_T] e^F \Phi^2}{2(K_M^B + \Phi)} - \frac{N k_{cat}^R [R_T] e^F \Phi^2}{2(K_M^R - \Phi + 1)} + \frac{N k_{cat}^R [R_T] e^F (1 - \Phi)\Phi^2}{2(K_M^R - \Phi + 1)^2}, \tag{5.34}$$

which is akin to a 1×1 real valued Jacobian matrix for the system. In this case the model will yield just one eigenvalue. Since we only have one eigenvalue, it would be impossible to obtain a complex conjugate pair of eigenvalues and so oscillatory behaviour cannot occur in this model. In order to confirm this we considered a ten-fold increase and decrease in all relevant total protein concentrations (varied by operon groupings) and as expected, found no oscillatory behaviour. Whilst this model has been used to good effect in studies of other phenomena, it is clear that it fails to capture the potential of the chemotaxis signalling pathway to display damped oscillatory behaviour. The failure of this model to capture the stability characteristics of the full fourth-order system led us to the conclusion that this particular model is not suitable for investigating the causes of overshoot. However, the fact that this model does not include protein phosphorylation would suggest that feedback of CheB-P onto the receptor state is critical in producing overshoot.

5.6 Understanding Key Principles of the Overshoot Response

Using model reduction analysis we have found some support for the idea that the dynamic timescales of intracellular processes are important for the ability of a cell to display overshoot. In order to further examine this, we seek to deduce within this section an analytic expression that is capable of reproducing the region of damped oscillatory behaviour displayed in Figure 5.7.

As discussed in Section 5.5, the use of lower-order dynamical systems may significantly simplify analytical work. Based on the analysis of various reduced model forms, here we shall make use of the third-order reduction considered in Section 5.5.1. Since we seek to explain the emergence of overshoot we must consider here the eigenvalues of the Jacobian matrix for this system, as given by the roots of the characteristic polynomial. In order to do this we consider the different analytical solution forms that may be obtained from a cubic polynomial of the form

$$p(\lambda) = \lambda^3 + A\lambda^2 + B\lambda + C = 0, \tag{5.35}$$

in which λ is an eigenvalue of the system and

$$\begin{aligned} A &= -\left(\rho_1 + \rho_4 + \rho_7\right), && (5.36)\\ B &= \rho_1\rho_4 + \rho_1\rho_7 + \rho_4\rho_7 + \rho_6\rho_6, && (5.37)\\ C &= -\left(\rho_2\rho_3\rho_6 + \rho_1\rho_4\rho_7 + \rho_1\rho_5\rho_6\right), && (5.38) \end{aligned}$$

are the polynomial coefficients (found as per equation (5.14)) where ρ_{1-7} are defined by equations (5.7)-(5.13). These analytical solution forms are listed by Murray [140] and require consideration of

$$A = 3a, \quad B = 3b, \quad \alpha = a^2 - b \quad \text{and} \quad \beta = 2a^3 - 3ab + C. \tag{5.39}$$

The relative sizes of these A, B, α and β terms determine the appropriate solution form for the polynomial $p(\lambda)$.

In order to understand overshoot we are seeking a solution form capable of yielding at least one pair of eigenvalues with non-zero imaginary part. As such we restrict our attention to those combinations of A, B, α and β which yield such eigenvalues. Using the parameter values in Table 3.1 reveals that just one condition is capable of producing damped oscillatory behaviour from the third-order dynamical system. As such we find that $\beta > 2\alpha^{3/2}$ must hold in order for us to obtain damped oscillatory behaviour. This condition may then be applied to our third-order dynamical system, giving the condition

$$\begin{aligned}
&\frac{2}{27}\frac{k_2^3[Y_T]^3(1-y_p^*)^3}{k_5^3}-\frac{1}{2}\frac{(g_R[R_T]+g_B[B_T]^2b_p^{*2})Ne^{F^*}\Phi^{*2}k_2[Y_T](1-y_p^*)(-k_3[A_T]a_p^*-k_5)}{k_5^3}\\
-\;&k_2[Y_T](1-y_p^*)\left[\frac{1}{6}\frac{(g_R[R_T]+g_B[B_T]^2b_p^{*2})Ne^{F^*}\Phi^{*2}k_2[Y_T](1-y_p^*)}{k_5^3}\right.\\
-\;&\left.\frac{1}{3}\frac{k_2[Y_T](1-y_p^*)(-k_3[A_T]a_p^*-k_5)}{k_5^3}\right]+\frac{k_1(1-a_p^*)Ne^{F^*}\Phi^{*3}k_3[A_T](1-b_p^*)g_B[B_T]^2b_p^*}{k_5^3}\\
-\;&2\left[\frac{1}{3}\frac{k_2[Y_T](1-y_p^*)(-k_3[A_T]a_p^*-k_5)}{k_5^2}-\frac{1}{6}\frac{(g_R[R_T]+g_B[B_T]^2b_p^{*2})Ne^{F^*}\Phi^{*2}k_2[Y_T](1-y_p^*)}{k_5^2}\right.\\
+\;&\left.\frac{1}{9}\frac{k_2^2[Y_T]^2(1-y_p^*)^2}{k_5^2}\right]^{3/2}>0. \qquad (5.40)
\end{aligned}$$

It is then possible to examine the magnitudes of each term within this expression by substituting parameter values from Table 3.1. Upon doing so we find that all terms lie in the range $O(10^4)$ to $O(10^7)$. We then choose to neglect the terms of the lowest order (i.e. those $\leq O(10^4)$) which yields the condition

$$\mathrm{Re}\left(\frac{2}{27}k_2^3[Y_T]^3(1-y_p^*)^3\left[1-(1-x)^{3/2}\right]\right)>5\times10^5, \tag{5.41}$$

in which

$$x=\frac{\frac{3N}{2}e^{F^*}\Phi^{*2}\left(g_R[R_T]+g_B[B_T]b_p^{*2}\right)+k_3[A_T]a_p^*+k_5}{k_2[Y_T](1-y_p^*)}, \tag{5.42}$$

and 5×10^5 is an approximation of the magnitude of the largest remaining term of less than $O(10^7)$. Within this expression N is the number of chemoreceptors in a signalling team which has activity Φ, chemoreceptor free-energy F, $[..._T]$ denotes the total concentration of the relevant protein and a_p^* and b_p^* are the (non-dimensional) steady-state concentrations of proteins CheA-P and CheB-P.

In order to investigate the ability of this analytical condition to reproduce the numerically obtained region of damped oscillatory behaviour we compare the results obtained from equation (5.41) to those obtained upon application of the numerical routine from Section 5.3 to the full fourth-order model (as in Section 5.4) and Model 1 (see Section 5.5.1). Results of this comparison are shown in Figure 5.11.

Clearly, upon comparing the prediction of this condition to the numerical results obtained from the full-fourth order and the third-order systems, we observe good agreement between these and equation (5.41) apart from a small region close to zero. This is likely due to the terms neglected in forming equation (5.41). However, this small region does not affect the

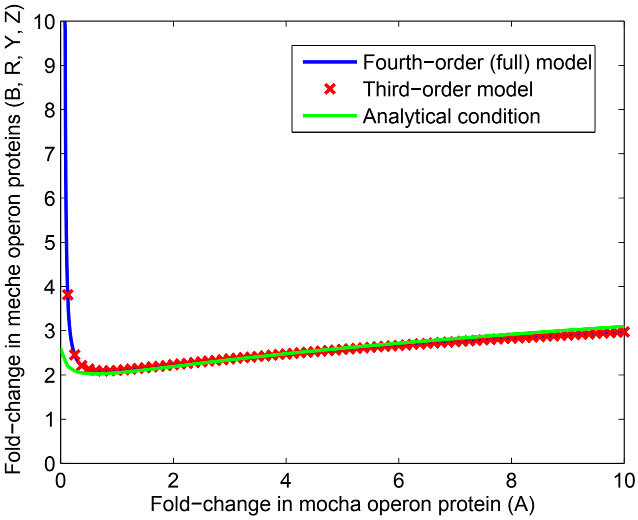


Figure 5.11: A comparison of numerical and analytical approximations to the region in which overshoot behaviour is found. The area above each of these lines signifies the region in which the model exhibits damped oscillatory behaviour. The blue line indicates the region of oscillatory behaviour found from the full fourth-order dynamical system. Red crosses show the region in which oscillatory behaviour is found in the third-order case in which the quasi-steady-state approximation has been applied to the concentration of CheY-P. The green line shows the region predicted by the analytical condition given by equation (5.41).

results considered here since overshoot is only observed at higher fold-changes in total protein concentrations. This suggests that key processes involved in causing overshoot are captured by expression (5.41). Despite its relative simplicity we are able to see that this expression includes processes associated with receptor dynamics, the feedback timescale of CheB-P and phosphotransfer to CheY (see Table 3.1 for parameter definitions), suggesting they each play a role in overshoot.

We may therefore be confident that the most important features responsible for causing overshoot are contained within equation (5.41). As such we now seek to simplify this further to investigate whether we may narrow down the potential cause(s) of overshoot. The term $(2/27)k_2^3[Y_T]^3(1-y_p^*)^3$ is neglected since it is an approximately exponential multiplier, when subjected to simultaneous equal fold-changes in all total protein concentrations. To leading order, an asymptotic expansion of remaining terms (assuming $x \ll 1$) gives

$$1-(1-x)^{3/2} \approx \frac{3}{2}x + ..., \tag{5.43}$$

where ... indicates the addition of lower order terms. This leaves $3x/2$ (where x is defined by equation (5.42)) which is further simplified upon neglecting $(2/3)k_2[Y_T](1-y_p^*)$ since it is approximately linear for equal fold-changes in all total protein concentrations. We therefore restrict our attention to the numerator of equation (5.41), as given by

$$\psi = \underbrace{\frac{3N}{2} e^{F^*} \Phi^{*2} \left(g_R[R_T] + g_B[B_T]^2 b_p^{*2} \right)}_{\text{Receptor dynamics}} + \underbrace{k_3[A_T]a_p^* + k_5}_{\text{CheB-P feedback}} . \tag{5.44}$$

Figure 5.12 displays values obtained for ψ (equation (5.44)) in addition to the overshoot amplitude which is calculated as per Figure 5.2. In particular the overshoot amplitude is calculated using numerical simulations of the reduced third-order system in Section 5.5.1 obtained for equal fold-changes in the total concentrations of all protein concentrations. Interestingly, it may be observed that the minimum of the curve obtained from equation (5.44) corresponds to the initial fold-increase in all total protein concentrations at which a non-zero overshoot amplitude is observed.

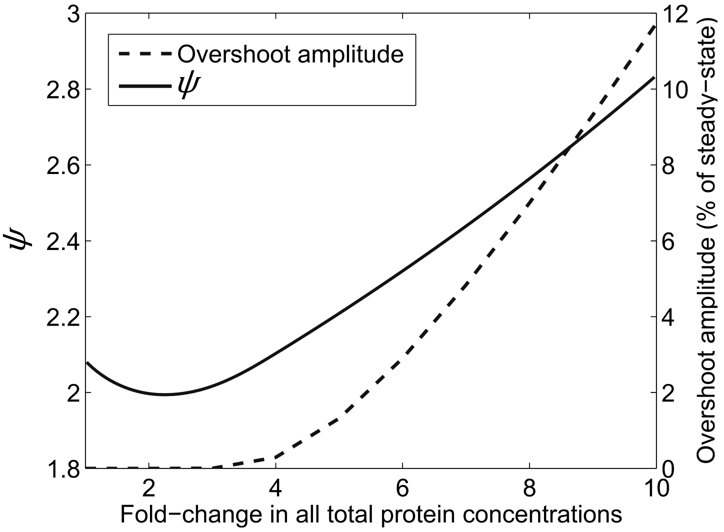


Figure 5.12: Receptor dynamics and CheB-P feedback timescale are critical in the occurrence of overshoot behaviour. The solid line shows values obtained from equation (5.44) and dashed lines show overshoot amplitudes calculated from numerical simulations, as per Figure 5.2. These are obtained under equal fold-changes in the total concentrations of all chemotaxis signalling proteins and are expressed as a percentage of the steady-state CheY-P concentration. The location of the minimum of the solid line corresponds to the fold-change required in order to obtain a non-zero overshoot amplitude.

We earlier noted that the first underbrace of equation (5.44) represents a contribution from the receptor state, specifically methylation and demethylation rates are important, whilst the second shows the importance of CheB-P feedback onto the receptor state, as shown by the presence of k_3, k_5, $[A_T]$ and a_p^*. In particular, $[A_T]a_p^*$ represents the number of phosphoryl groups available for transfer from CheA-P onto CheB at steady-state while k_3 shows how quickly the phosphoryl groups may be transferred around the system, causing demethylation of chemoreceptors. As such we conclude that the balance between chemoreceptor dynamics and CheB-P feedback is important in causing overshoot.

5.7 Comparison with Experimental Data

Recent experimental work sought to quantify the overshoot observed in the chemotactic response of *E. coli* cells [131]. In particular, this work found that overshoot may be observed in response to step-changes (in both directions) of the extracellular chemoattractant concentration. The overshoot observed was shown to have an amplitude that is dependent upon the magnitude of the ligand stimulus. For low and high extremes of ligand stimulus the overshoot amplitude observed was negligible whereas a peak of as much as ∼20% of the pre-stimulus clockwise (CW) flagellar rotation bias was observed for intermediate stimuli, here defined as those in the range 5-50μM.

Within this section we shall focus on exploring the overshoot observed in response to step-up ligand stimuli since there exists a greater number of experimental data points with which we may compare our results. The experimental data for step-up stimuli also span a wider range of ligand concentrations which is useful in terms of comparing our theoretical results with experimental data, however the model utilised here is capable of producing overshoot in response to both step-up and step-down ligand stimuli.

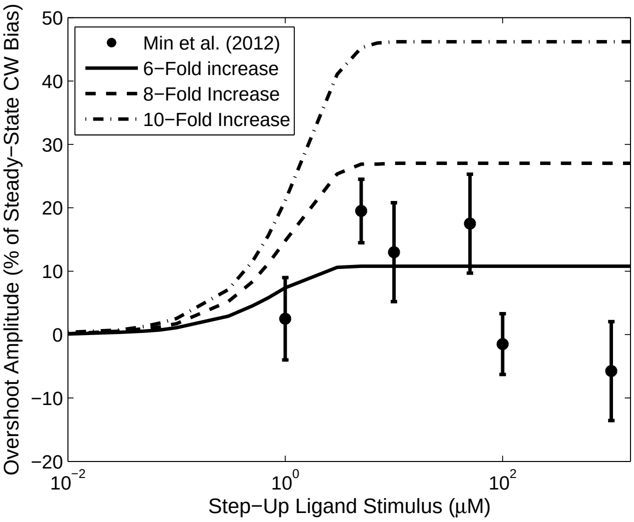


Figure 5.13: Overshoot amplitudes are dependent on the size of a ligand stimulus. Circles with error bars display the results of Min et al. [131] for step increases in ligand concentration. These results are expressed as a percentage of the steady-state tumble bias of cells. Black lines indicate overshoot amplitudes obtained from numerical simulations under equal fold-changes in the total concentration of all proteins with solid, dashed and dash-dot lines indicating six-, eight- and ten-fold increases in all total protein concentrations, respectively, compared with those given in Table 3.1.

Utilising numerical simulations of the full fourth-order dynamical system we find overshoot amplitudes, for ten-fold variation in all total protein concentrations, of up to ∼50% of the pre-stimulus CW flagellar rotation bias (calculated as per Morton-Firth & Bray [141]) in response to a step-increase ligand stimulus of 50μM, as seen in Figure 5.13. However, we do not believe such high fold-changes in protein concentration are often likely to occur. We would therefore not expect to achieve overshoot amplitudes in excess of the ∼25% attainable with an eight-fold variation in all protein concentrations. Within a population of cells the average overshoot amplitude is likely to be biased by those cells with higher protein concentrations (see Figure 5.12 and 5.13).

Results given within the experimental literature were calculated using a certain number of run and tumble events [131]. Here we have calculated the peak value of overshoot amplitude as a percentage of the steady-state CW flagellar rotation bias obtained from numerical simulations of the model system. This may lead to some differences in results since those of Min et al. [131] may not necessarily capture the peak value for the overshoot amplitude if the numbers of run and tumble events do not coincide precisely with the peak overshoot amplitude. It is notable here that adaptation times are significantly altered by variation in total protein concentrations (see Figure 5.14) and as such there is no guarantee that counting a fixed number of run and tumble events from the time of stimulus application will align with the point of maximal overshoot amplitude.

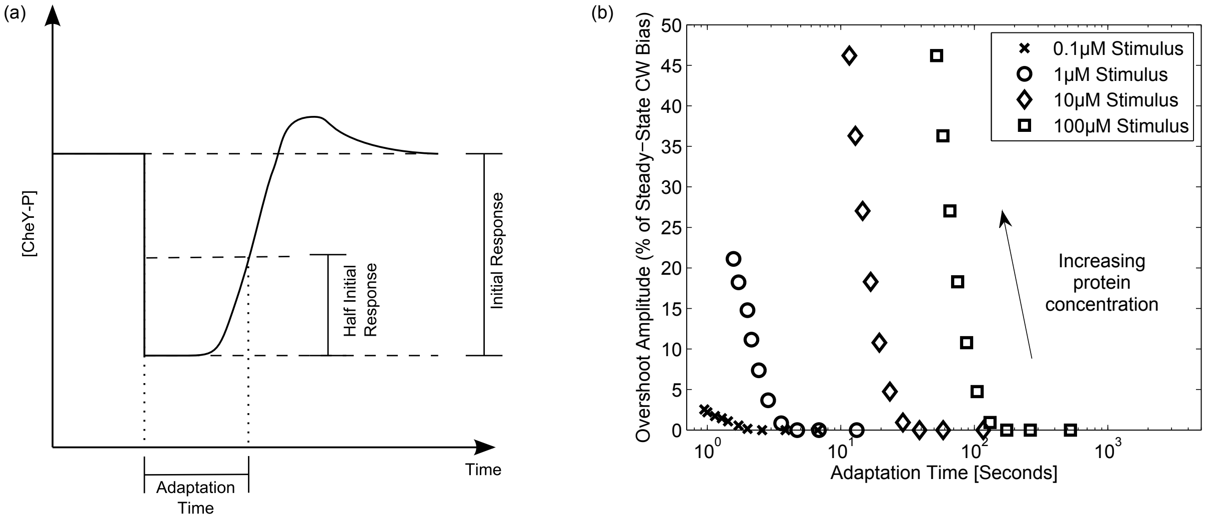


Figure 5.14: (a) Here the adaptation time is chosen to be the time necessary for a cell to recover from half of the initial response. (b) Shown here are the overshoot amplitudes and associated adaptation times for different size of step-up ligand stimuli, namely 0.1μM (represented by crosses), 1μM (circles), 10μM (diamonds) and 100μM (squares). The different data points for each stimulus refer to cells with different (1-10) fold-increases in the concentration of all proteins. Shorter adaptation times are associated with larger fold-increases in all protein concentrations. Clearly, cells with shorter adaptation times display larger overshoot amplitudes, however there is also a dependence on the size of ligand stimulus applied, as noted by Min et al. [131].

It can be seen, upon examination of the results in Figure 5.13, that we do not capture the negligible overshoot amplitude obtained in response to high extremes in chemoattractant stimuli. One potential explanation of this is that during experimental studies chemoreceptor clusters have been observed to break apart during responses to large, saturating stimuli and reappear once the cell has enough time to adapt in both *E. coli* and *Bacillus subtilis* [142, 143]. In fact, this observation likely explains the difference between the experimentally obtained results for step-up and step-down stimuli. In the case of a step-up stimulus it is possible for chemoreceptors to become saturated and break apart resulting in negligible overshoot amplitudes. However, it is not possible to saturate a chemoreceptor cluster via a step-down stimulus thus the overshoot amplitude will not tail off for large stimuli. To our knowledge this phenomenon has not been modelled within the literature and could therefore prove to be an interesting area for future work.

5.8 Summary & Discussion

This chapter has demonstrated that within a mathematical model of the *E. coli* chemotaxis signalling cascade both kinetic rate parameters and total protein concentrations can affect the cellular response. Here it was discovered that pair-wise variations of kinetic rate parameters were theoretically able to produce overshoot, however the magnitude of variation required was too large to be biologically feasible. As such, this work focused on studying how network structure and total protein levels affect receptor state, dynamic time scales and sharing of phosphoryl groups in order to produce overshoot. In particular, examining the effects of operon-wise variation in signalling protein concentrations, we found a balance between receptor state and the timescale of CheB-P receptor demethylation to be a key feature responsible for overshoot within the *E. coli* chemotactic response.

Within the literature the phenomenon of overshoot behaviour has most commonly been discussed as a likely result of interaction between different chemoreceptor types. In particular, past theoretical work has proposed overshoot to be caused by cross-talk in receptor methylation levels [128, 133]. Also discussed theoretically is that a significant increase in the catalytic rate of the protein CheR can result in increased overshoot amplitudes in response to impulse stimuli [144]. However, we would not expect a kinetic rate (such as the CheR catalytic rate) to vary by a large degree and thus this may not be the only mechanism responsible for overshoot behaviour.

Numerous theoretical studies of *E. coli* chemotaxis signalling have failed to observe overshoot [70, 63, 30, 129]. There are likely to be numerous different reasons for this. Firstly, it is common in mathematical modelling of such systems to utilise experimentally determined average protein concentrations. In terms of the average behaviour of a cell population this would appear to be a reasonable approach, however it fails to account for the effects of the significant stochastic variation in protein concentrations observed over a cell population [145]; effects which have been shown here to be important determinants of the transient cell response. Secondly, simplified mathematical models give a number of benefits, particularly in terms of the ease with which analytical results may be obtained. However, such simplifications not only alter the network structure but the ability of the system to exhibit previously observed transient behaviour as demonstrated in Section 5.5. In particular the ability of the model to fit with biological observations may be greatly altered or lost altogether. One such example is the model due to Tu et al. [63] which has been used with some success in a number of studies such as that by Kalinin et al. [64]. However, it has been shown here that this particular model cannot exhibit damped oscillatory behaviour.

This chapter gives a number of theoretical results that may be investigated in an experimental setting. In particular, the results presented in Figures 5.6, 5.7, 5.12 and 5.13 are testable by appealing to certain under/overexpression mutant cells. In the case of Figure 5.6 it is possible to create cells under/overexpressing certain proteins. To investigate Figure 5.7 the *meche* and *mocha* operons may be under/overexpressed. Finally, Figures 5.12 and 5.13 would require cells under/overexpressing all of the chemotaxis signalling proteins. In each case, once under/overexpression mutant cells are available, analysis of flagellar rotation time-courses during which cells are challenged with a step-change in ligand concentration should reveal the magnitude of overshoot exhibited by cells (if any).

Many biological systems exhibit a large degree of individual variability across their populations. This is usually put down to genetic differences, environment and history. However, even cells identical under these criteria display behavioural variability [146]. This is likely to be caused by the low copy numbers of components including DNA and key regulatory molecules, leading to stochastic effects [147]. A significant stochastic cell-to-cell variation in concentrations of the *E. coli* chemotaxis signalling proteins has regularly been referred to within the literature [112, 145, 148, 149, 150, 151]. When faced with such cell-to-cell variation, it has been suggested that a reliable signal processing system will maintain key features [152]. A number of features, such as precise adaptation, have been shown to be robust in *E. coli* chemotaxis signalling [44, 76]. There are, however numerous consequences associated with the stochastic variation discussed within the literature [153]. Most commonly studied are the effects of variation in the concentrations of proteins CheB and CheR; those directly involved in adaptation. In studying these effects it was noted that they have a significant impact on the adaptation times of cells [151], a feature noticeable within results obtained here (see Figure 5.14). It has also been shown that varying the total signalling protein concentrations can result in different steady-state phosphorylation levels [149]. This was observed within our work, which suggests that sharing of phosphoryl groups between CheB and CheY may be important in the occurrence of phenomena including overshoot.

Also, demonstrated here is that protein concentration is an important factor in affecting the temporal response of an intracellular signalling cascade. Based upon the results obtained within this work we believe that there are three key ingredients for determining variation in a network. These are driven by the network structure, kinetic rates and the concentrations of its constitutive elements. We would expect these principles to be relevant in explaining similar phenomena within other biological systems. The first two of these aspects are well founded results within the analysis of signalling cascades [154]. The third allows us to determine how the transient cell response will vary given different intracellular conditions, namely protein concentration in this example. Results obtained here indicate that the simplification of using population average values may mask the inherent effects of cell-to-cell variability. It would therefore seem sensible, in addition to studying population behaviour using average values, to consider the potential effects of cell-to-cell variation when considering transient behaviour in cellular systems.

In terms of a wider picture, cell-to-cell variability could be vital for population survival, especially those cells subjected to a wide range of environmental conditions [150]. It could therefore be the case that variation in protein concentrations may be beneficial to the survival of a population. This will be investigated further in Chapter 6.

Chapter 6

An Agent-Based Model of *E. coli* Chemotaxis

Within Chapters 4 and 5 the mechanisms behind the phenomena of FCD and overshoot in the chemotactic responses of single *E. coli* cells were investigated. The implications of these two features have been studied in the context of a single cell chemotactic response. It is not necessarily clear what impact such single cell phenomena will have upon the overall chemotactic behaviour of an *E. coli* population. In order to do this, we present within this chapter the formulation and analysis of an agent-based model (ABM) of *E. coli* chemotaxis (see Fig. 6.1). Here we choose to focus primarily on the mathematical details of the model and subsequent results rather than the exact numerical implementation.

6.1 Agent-Based Model Formulation

Within this section we describe the formulation of an ABM describing the chemotactic behaviour of *E. coli* cells. In order to do so, we begin by giving a general overview of the ABM algorithm before providing a more in depth description of each model component.

6.1.1 ABM Algorithm

The ABM discussed within this chapter contains a description of many different aspects of the *E. coli* chemotactic response. Combining these aspects allows us to formulate an algorithm that may be used to simulate the chemotactic behaviour of cell populations. As such, we provide here a basic outline of the algorithm used for the ABM simulations conducted in this chapter.

This algorithm is composed of five main stages that proceed in a cyclical manner over a time-scale appropriate for the given application. These stages are as follows.

1. Calculate ligand concentration (at the cell location);

2. Update intracellular signalling pathway;

3. Calculate flagellar rotational bias;

4. Simulate cell movement - straight swim (run) or turn and swim (tumble);

5. Return to step 1.

For a graphical summary of this algorithm, see Figure 6.1. The details of ABM components 1-4 as well as their respective modelling assumptions and simplifications are examined in-depth within subsequent sections. We note that component 5 in the algorithm summary is not discussed further since it simply refers to the saving of results and the beginning of a new time step (i.e. a return to component 1).

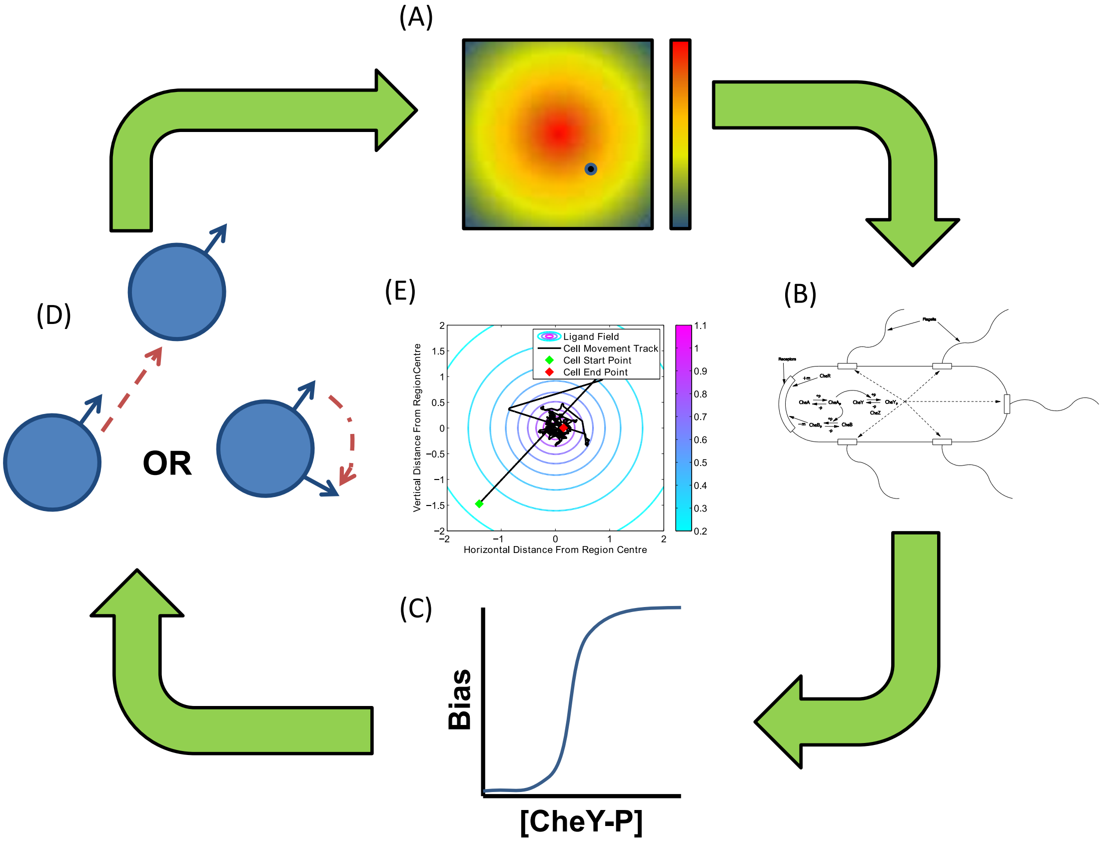


Figure 6.1: A cartoon showing the workings of the agent-based model. (A) An initial location is chosen for a simulated *E. coli* cell. (B) The appropriate ligand concentration feeds into an ODE model of the intracellular signalling pathway and progresses one time step forward. (C) The rotational bias of the simulated cell's flagella is calculated and a (uniformly distributed) random number generator used to choose a "run" or a "tumble". (D) A new location is defined if the cell "runs" or a new direction of travel and a new location are chosen if the cell "tumbles". The new ligand concentration is calculated and the process repeated for the desired number of time steps, producing results as in (E).

6.1.2 Calculating the Ligand Concentration

In order to accurately represent any particular situation of interest it is necessary to carefully choose a spatial domain and ligand profile. This helps to make sure that ABM simulations will provide a good approximation of real-world experimentation.

Within this work we choose to conduct all ABM simulations within a two-dimensional square spatial domain. This choice is common within ABM studies of chemotaxis, such as those of Zonia & Bray [109] and Emonet et al. [108], as it allows for simple interpretation of results. The size of this domain was chosen to be arbitrarily large, and is described by

$$(x, y) = \{x, y \in \mathbb{R} : [-2, 2]\}, \tag{6.1}$$

where x and y are the horizontal and vertical Cartesian coordinates, respectively.

It is also necessary to consider a ligand profile that is defined across this spatial domain. A number of different ligand profiles have been studied within both the experimental and theoretical literature. The most common examples are those with exponential, linear and zero gradient profiles. Within this chapter we mainly focus on exponential ligand gradients of the form

$$[L] = l_0 + \exp\left(-\sqrt{x^2 + y^2}\right), \tag{6.2}$$

within which $[L]$ denotes the ligand concentration (in this case MeAsp), x and y are the horizontal and vertical Cartesian coordinates of the domain, respectively and l_0 is a minimum ligand concentration (assigned a value of $l_0 = 0.1$mM unless otherwise stated).

We consider here a number of simplifying assumptions that allow for either easier computation or a more intuitive understanding of results. One such example is that for the majority of this chapter we consider the response of cells to just one attractant. Choosing the attractant MeAsp here allows for simpler computation since it is known to not be metabolised by cells. As such we do not need to consider how the number of cells in different areas will degrade the ligand concentration. We also choose to consider a stationary ligand profile (i.e. one in which the ligand profile does not evolve in time). Without these two simplifications it would be necessary to solve a reaction-diffusion equation describing ligand evolution in time, however, here we may simply calculate concentrations directly from equation (6.2). Whilst the consideration of the cellular response to multiple attractants is more biologically realistic, here we mostly restrict our attention to the more simple one attractant case. This allows us to obtain a more intuitive understanding from results obtained here.

6.1.3 Updating the Signalling Pathway

Within the ABM described here we consider the behaviour of an *E. coli* cell population within the ligand profile and spatial domain discussed in Section 6.1.2. In doing so we describe the response of each individual cell with the model in Chapter 3. The key assumptions and simplifications of this model were discussed in Section 3.1.2.

Within the ABM formulated here, the intracellular signalling cascade ODE model is updated using an inbuilt MATLAB stiff ODE solver (ode15s). In contrast to previous chapters, here we restrict this solver such that it progresses forward by the length of one ABM time-step. This allows a new ligand concentration to be calculated and input into the ODE model depending upon the location of each cell.

Using this ODE solver we are able to obtain a complete representation of the internal state of each simulated cell for every model time-step. As such, we are able to observe the response of CheA-P, CheB-P, CheY-P and the receptor methylation level for each cell over the entire period of an ABM simulation.

6.1.4 Calculating the Cell Response

The internal signalling cascade information calculated in Section 6.1.3 is then used to calculate the response of each individual cell. In particular, it is known that the CheY-P concentration acts to regulate the rotational behaviour (bias) of the flagellar motors in *E. coli* cells. Within the literature there exist some examples whereby this CheY-P and flagellar rotational bias relationship is modelled. Here we shall consider two of the most commonly used models of this relationship, namely those of Cluzel et al. [29] and Morton-Firth & Bray [155], and explain how these functions are used within the context of the ABM.

Within the literature it was Cluzel et al. [29] that were first able to experimentally quantify this relationship. In doing so it was found that there exists a sigmoidal relationship between CheY-P concentration and CW (clockwise or tumble) bias. This was subsequently shown to be fitted by a Hill function of the form

$$\text{Bias}_C = \frac{K^n}{K^n + \left(\frac{[Y_p]}{[Y_T]}\right)^n}, \tag{6.3}$$

with $[Y_p]$ the CheY-P concentration calculated in Section 6.1.3, $[Y_T]$ is the total (phosphorylated and non-phosphorylated) CheY concentration, n denotes a Hill coefficient ($n \approx 10.3 \pm 1.1$) and K is a dissociation constant ($K = 3.1\mu\text{M}$).

Morton-Firth & Bray [155] considered a similar, sigmoidal function to describe this relationship. This is given by

$$\text{Bias}_M = \frac{1}{1+\frac{3}{7}\left(\frac{[Y_p]}{[Y_p]_0}\right)^{5.5}}, \tag{6.4}$$

where $[Y_p]$ is the CheY-P concentration calculated in Section 6.1.3 and $[Y_p]_0$ is the concentration in absence of any stimulus.

Examination of the results in Figure 6.2 clearly shows that equation (6.4) yields a shallower curve than that of equation (6.3). Whilst the two expressions clearly demonstrate fairly similar relationships, in the context of this work, equation (6.4) appears to be more suitable. This is due to its incorporation of the ability of cells to shift the bias curve such that the steady-state CheY-P concentration remains in the sensitive (steeply sloping) part. Thus we utilise equation (6.4) throughout the remainder of this chapter.

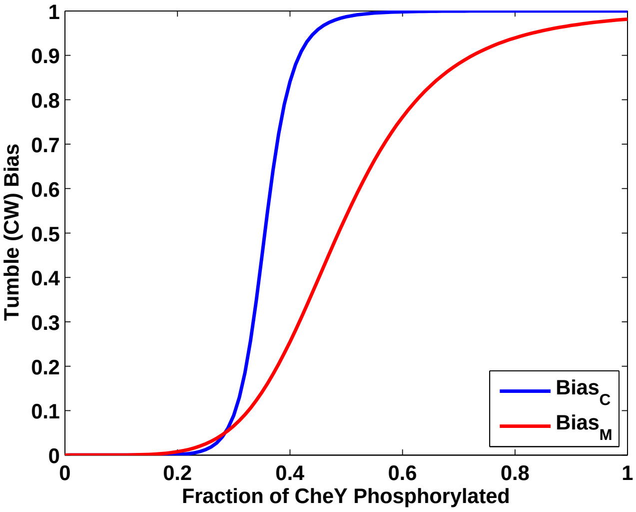


Figure 6.2: A plot demonstrating the differences between the CW (tumble) bias functions of Cluzel et al. [29] (blue) and Morton-Firth & Bray [155] (red). It is worth noting that the Morton-Firth & Bray bias curve shown here is calculated using a value of $[Y_p]_0 = 0.4168$ corresponding to the CheY-P steady-state in equation (3.24) when utilising the base parameter set of Table 3.1. Within this work the expression due to Morton-Firth & Bray [155] is utilised since it includes the ability for the curve to shift for differing CheY-P steady-state values.

6.1.5 Simulating Cell Swimming

Within this section we outline how swimming behaviour of individual cells is considered in the ABM. In order to accurately represent this swimming behaviour it is necessary to represent the stochastic nature of flagellar motor switching and the subsequent run and tumble swimming pattern.

The ability of *E. coli* cells to produce the observed run and tumble swimming pattern stems from the flagella and the motors controlling their rotation. Explicitly modelling this process would require significant computational cost. Thus, instead we consider a simplified approach that still represents this process to a good degree.

Here we consider the flagellar rotational bias expression from Section 6.1.4 (i.e. equation (6.4)). This tumble bias denotes the probability that a cell will produce a tumble for any given CheY-P concentration. We therefore utilise a uniformly distributed random number generator to choose a number $0 \leq r \leq 1$ for each simulated cell and assign swimming behaviour according to

$$\text{Flagella Direction} = \begin{cases} \text{CW (Tumble)}, & \text{if } \text{Bias}_M > r \\ \\ \text{CCW (Run)}, & \text{otherwise} \end{cases}$$

in which the Bias_M value has been calculated according to equation (6.4). Using this simple approach we represent the stochastic nature of flagellar motor switching without the need for more complex stochastic equations.

In addition to assigning the type of swimming (run or tumble) behaviour for individual cells, we also consider the resultant movement within the spatial domain from Section 6.1.2. During the run phase cells are known to swim in (approximately) a straight line. Mathematically we define this by

$$\frac{dx}{dt} = c \cdot \sin(\theta_n), \tag{6.5}$$

$$\frac{dy}{dt} = c \cdot \cos(\theta_n), \tag{6.6}$$

where c is the swimming speed during a run, θ_n is the angular direction of travel and x, y are the horizontal and vertical location of the simulated cell in the domain of interest (see Figure 6.3). In order for this description to hold we have assumed that each cell travels with a fixed speed (i.e. there is no acceleration during a run). This particular physical regime implies that inertial forces are negligible (i.e. low Reynolds number), which is a valid assumption here due to the small size and relatively low swimming speeds of *E. coli* cells (see Purcell [156] for a more detailed explanation).

During a tumble we also include a turn component, i.e. a change in θ_n. This is achieved by considering

$$\theta_n = \theta_o + \theta_r, \tag{6.7}$$

within which n and o are subscripts denoting the new and old values, respectively whilst the subscript r indicates a turning angle.

In the case of a run the cell is not re-oriented (i.e. we choose $\theta_r = 0$) whereas for a tumble θ_r is chosen according to a uniformly distributed random turning angle distribution restricted to between ±18 and 98 degrees as per experimental findings summarised in Table 6.1. Since the

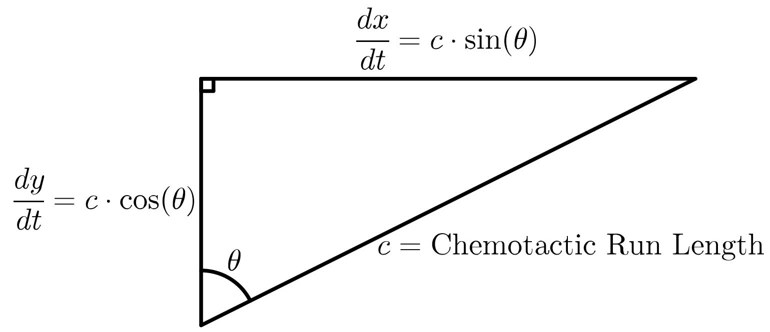


Figure 6.3: A schematic demonstrating the calculation of a chemotactic run. Here the length of a chemotactic run is denoted c whilst the direction of travel for a cell is given by θ. Using basic trigonometric results this run length may be split into a horizontal (dx/dt) and a vertical (dy/dt) component.

duration of a tumble is significantly shorter than that of a run we define a tumble event here as a change in direction from equation (6.7) combined with the movement defined by equations (6.5) and (6.6).

Consideration must also be given here to the behaviour of cells at the boundary of the spatial domain. Specifically, we require rules governing the behaviour of cells when they pass outside of the spatial domain from Section 6.1.2. Within the literature there are two main examples considered. These are as follows.

- **Periodic:** Cells swimming outside of the spatial domain are assumed to re-appear on the opposite side. In the case of the domain in Section 6.1.2, a cell leaving the domain at $(x, y) = (1, 2)$ will re-enter the domain at $(x, y) = (1, -2)$.
- **Wall:** Here cells swimming outside of the domain are returned to the boundary as if they swim into a solid wall. For example if, at the end of a time-step, a simulated cell is positioned at $(x, y) = (2.05, 1)$ then it will be returned to the boundary at $(x, y) = (2, 1)$. This is demonstrated in Figure 6.4.

In the remainder of this chapter we consider the wall boundary. This is intended to replicate the behaviour of cells in a bounded region such as a petri dish where they will swim into the side wall.

In biologically realistic situations, cells would clearly not be able to occupy the same space, i.e. the presence of one cell excludes others from occupying that same space. In the context of this ABM we would be required to implement an exclusion radius for each simulated cell. However, here we consider a large enough spatial domain relative to the size of an *E. coli* cell that we assume these effects can be neglected.

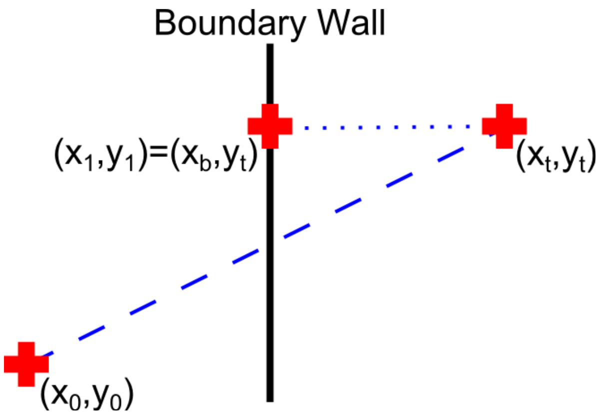


Figure 6.4: A schematic demonstrating the implementation of the wall boundary condition. A cell (red cross) is initially located at position (x_0, y_0) within the spatial domain. This cell then undergoes a simulated chemotactic run (dashed blue line) which results in the cell crossing the boundary of the domain (black line) to a position at (x_t, y_t), where t indicates that these are temporary coordinates. According to the wall boundary condition this cell is then returned (dotted blue line) to the boundary of the domain (i.e. the position $(x_1, y_1) = (x_b, y_t)$, where b indicates a location upon the boundary).

Table 6.1: A set of parameter values describing the swimming behaviour of *E. coli* cells.

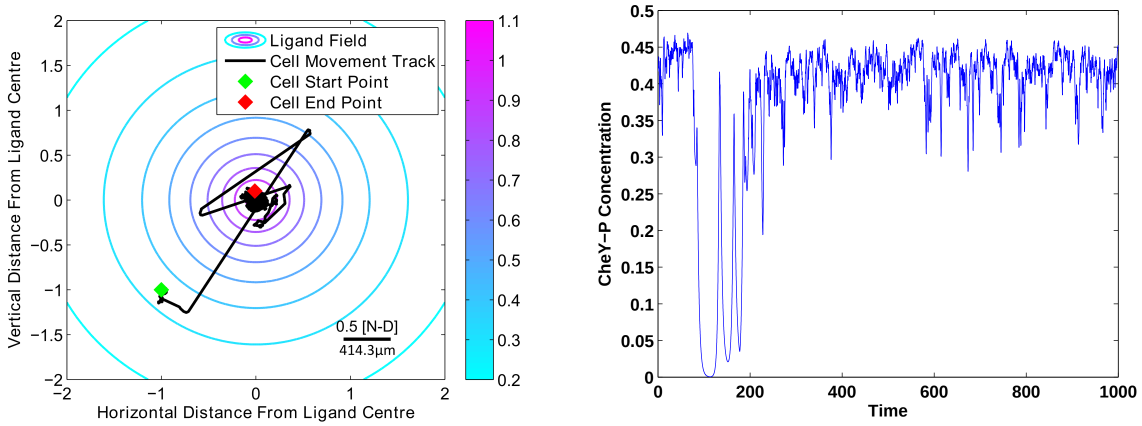
Symbol	Definition	Value	Source
c	Swimming speed during a run	29±6μm/s	[157]
θ_r	Angle turned during a tumble	58±40°	[158]

6.2 Model Validation

In order to assess whether the ABM formulated here is suitable for further analysis of chemotactic behaviour in cell populations it is necessary to undergo a process of model validation (or to be more precise, model non-invalidation). This will require the comparison of results obtained using this model against results and observations within both the previous experimental and theoretical literature. It is worth noting, however, that there is a lack of experimental literature looking at chemotactic swimming behaviour of single *E. coli* cells and as such direct comparisons are difficult to make.

The first stage in addressing the suitability of this ABM for the study of *E. coli* cell populations is to compare the results obtained to the qualitative behaviour described by Berg & Brown [11]. This provides an initial sense check of the model, ensuring it produces appropriate chemotaxis-like behaviour. In order to do this a number of simulations were conducted for a single *E. coli* cell placed in a ligand gradient of the form in equation (6.2). An example of the simulated cell behaviour is displayed in Figure 6.5.

Upon examination of the results in Figure 6.5(a) it can be seen that the simulated cell was able to direct its movement toward the region with the highest ligand concentration. Also, the results of Berg & Brown [11] observed, using a three-dimensional tracking microscope, that chemotactic cells are able to modulate their swimming behaviour such that they tumble less



(a) Cell Trace

(b) CheY-P

Figure 6.5: An example of a simulated *E. coli* cell within an exponential ligand field. (a) A plot showing a trace of the swimming pattern displayed by the cell. (b) The CheY-P concentration time course leading to the behaviour in panel (a). Each of the variables displayed in these plots are expressed in non-dimensional form. A scale bar is included in (a) to indicate the relationship between non-dimensional units (denoted N-D) and dimensional values. Note that the results shown here represent a typical example selected from a pool of 50 simulated cells.

when travelling up a ligand gradient. In addition to this Macnab & Koshland [159] showed that *S. typhimurium* cells tumble more frequently when travelling down a ligand gradient, a feature that is now commonly accepted to exist in *E. coli* also. Both of these features of the chemotactic response can clearly be seen in Figure 6.5(a) where the straight runs are much longer when travelling in a beneficial direction and significantly shortened when travelling down a ligand gradient. Here, we have shown a phenomenological agreement with the observations described by Berg & Brown [11]. However, a direct comparison between the results is not possible since Berg & Brown [11] observed *E. coli* cells swimming in either a spatially homogeneous environment (i.e. cells will not display chemotaxis) or a time-varying gradient (not yet considered in this ABM).

Figure 6.5(b) displays the CheY-P concentration time course associated with the swimming behaviour in Figure 6.5(a). In particular, a number of prolonged drops in the concentration of protein CheY-P can be observed in the early part of this time course. These are associated with the extended runs visible in Figure 6.5(a). Due to experimental limitations, results displaying both the swimming behaviour of an individual cell and the relevant intracellular signalling dynamics are not available for comparison. However, this result can be explained by examining equation (6.4) in which a lower CheY-P concentration leads to a greater CCW bias and thus the cell tumbles less frequently. Similar behaviour is observed for each cell across 50 runs of this simulation.

The ABM formulated here clearly displays phenomenological agreement with the descriptions of Berg & Brown [11] and Macnab & Koshland [159]. In addition to these results, we also

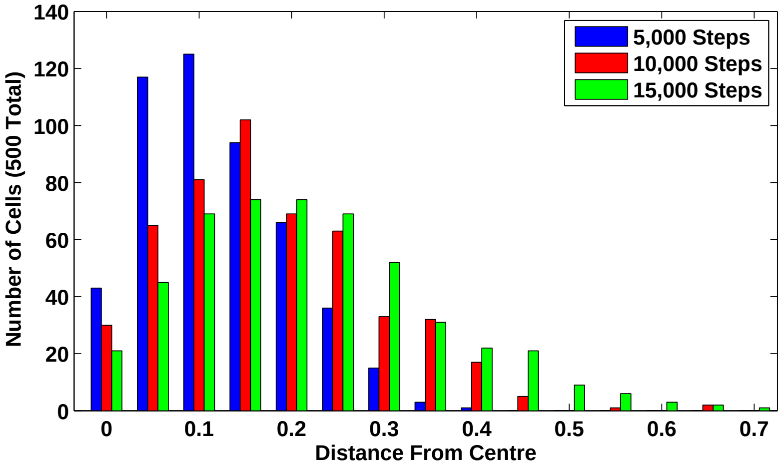


Figure 6.6: A bar chart showing the development of the spatial distribution of an *E. coli* cell population in a domain with a spatially homogeneous ligand concentration, when all simulated cells begin at the centre of the domain, i.e. $(x, y) = (0, 0)$. Here the number of cells are grouped as per their distance from the centre of the domain and bars are plotted accordingly. In particular, results are shown for the same population of cells after 5,000 (blue bars), 10,000 (red bars) and 15,000 (green bars) non-dimensional time steps (approximately equal to 1.2 minutes, 2.4 minutes and 3.6 minutes, respectively).

compare the ABM with previous theoretical work in the area. In order to do this we consider a population of cells placed within a domain containing a spatially homogeneous ligand field (i.e. the ligand concentration is constant, at $[L] = 0$mM, across the entire domain). This allows the visualisation of the development of the population distribution. In particular, we shall consider a population consisting of 500 simulated *E. coli* cells (10 model runs each simulating a 50 cell population) that is initially present at the centre of the domain, i.e. $(x, y) = (0, 0)$. These cells are then allowed to swim freely within the domain, however there is no spatial heterogeneity in the ligand field that can bias the movement of these cells. Results obtained from the ABM in this configuration are displayed in Figure 6.6.

Analysing the results in Figure 6.6 we can see that the population moves outward from the centre of the domain at an approximately constant velocity. These results may now be compared to those in the theoretical literature. In particular, numerous methods of motility have been discussed by Codling et al. [102] alongside descriptions of the appropriate governing equations. It is stated within this work that the biased telegraph equation (such as that in equation (2.12)) may be used to describe the type of chemotactic behaviour discussed within this thesis. In support of the validity of the ABM considered here it can be seen that the results in Figure 6.6 are consistent with the solution of the telegraph equation except for a stochastic component that is most likely associated with the limited population size considered within this example.

It is clear that the ABM considered here compares well with experimental observations in addition to cell population modelling work. In addition to this a comparison may also be made

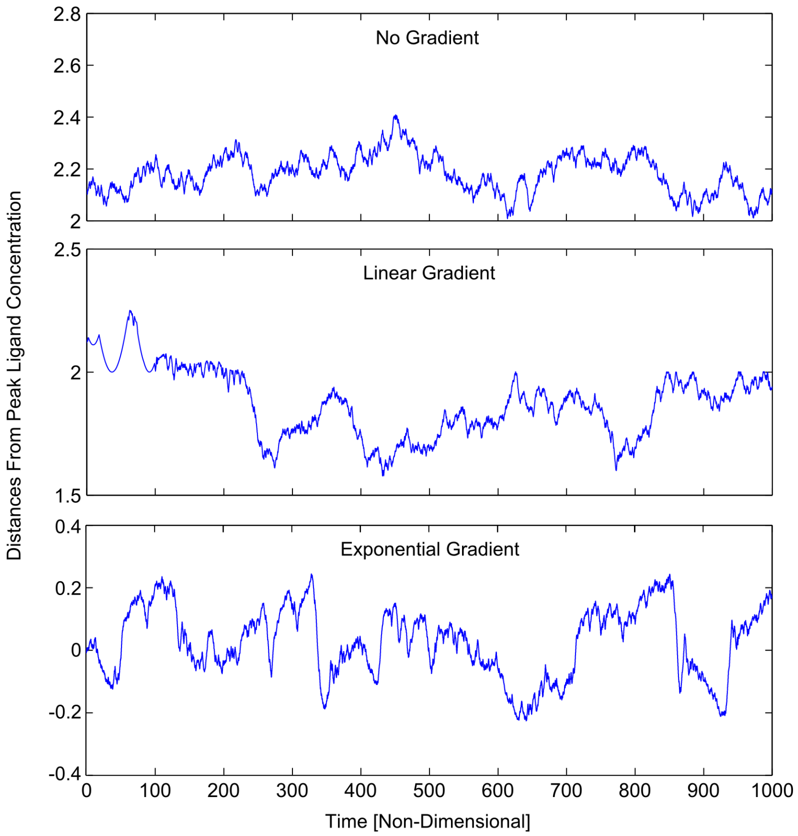


Figure 6.7: Plots showing the typical simulated behaviour of chemotactic *E. coli* cells in environments containing no spatial ligand heterogeneity (top), a linear MeAsp concentration field (middle) and an exponential MeAsp concentration field (bottom). These results display qualitatively similar behaviour to simulations conducted in the work of Zonia & Bray [109] although an exact comparison is not possible. Note that results shown here represent a typical example selected from a pool of 50 simulated cells in each different ligand gradient.

between results obtained from this ABM and those of another ABM that utilises a similar formulation, namely that due to Zonia & Bray [109]. This work considered simulations within a number of different ligand fields of single cell *E. coli* chemotactic behaviour. In order to compare the results obtained from the two models simulations of single *E. coli* cell chemotaxis are conducted in situations whereby there is no spatial ligand heterogeneity, a linear ligand field and an exponential ligand field are considered. The results obtained from these simulations are shown in Figure 6.7. We note, however, that the work of Zonia & Bray [109] does not give exact mathematical definitions of the ligand gradients considered, thus making a direct comparison of results impossible.

In spite of the difficulties in making a direct comparison, it appears that the results in Figure 6.7 show a qualitative agreement with the closest examples from the work of Zonia & Bray [109]. Within the ABM considered in this thesis we take a four ODE model of the *E. coli* chemotaxis signalling cascade whereas Zonia & Bray [109] take a model consisting of ∼90 ODEs. Whilst the ∼90 ODE model contains details of all known reactions in the *E. coli* chemotaxis signalling cascade, the use of such a large ODE model can make it difficult to gain an intuitive understanding of how changes within the signalling network will affect the cell response.

Within this section we have shown that the ABM formulated here cannot be invalidated based upon comparisons with either experimental observations or theoretical results obtained from population level differential equations and an alternative ABM. None of these tests have displayed results suggesting that the ABM considered here may be producing false results. As such we shall utilise this model within the remaining sections of this chapter and consider any results obtained to be a reasonable approximation to the true behaviour. As such we are now able to consider how the intracellular signalling phenomena discussed in Chapters 4 and 5 affect the behaviour of an overall cell population and attempt to understand what benefits (if any) they have in terms of the development and survival of a cell population.

6.3 Chemotaxis in Different Ligand Gradients

Within Chapter 5 a number of features relating to the phenomenon of overshoot in the chemotactic response of *E. coli* cells were discussed. In particular it was noted that populations of cells display a significant amount of non-genetic variability. In the context of the overshoot response it was proposed that cell-to-cell differences in the total intracellular chemotaxis signalling protein concentrations could be responsible for the emergence of overshoot behaviour. Specifically, by utilising variations in the total protein concentrations we were able to find conditions under which overshoot should occur. This section explores, using the ABM formulated in Section 6.1, how individual changes in the cell response (e.g. overshoot) via total protein concentration, affect the collective behaviour of a cell population.

In order to study the effects of variation in total protein concentrations associated with overshoot, we investigate how it affects cells' ability to respond in ligand gradients of varying steepness. In order to do this we shall consider a range of different multiples of the total signalling protein concentrations of the form

$$[X]_T = \beta [X]_{T0}, \tag{6.8}$$

where $[X]_T$ ($X = A, B, R, Y, Z$) represents the total protein concentration used, the additional subscript 0 indicates the population average concentration totals as listed in Table 3.1 and β ($= 1/4, 1/2, 1, 2, 4, 6, 8, 10$) denotes the multiple considered for the total concentration of all chemotaxis signalling proteins. It is clear upon inspection of the results in Figure 6.8 that total protein concentration variations, such as those in equation (6.8), can produce significantly different behaviour. In particular, from Figure 6.8 it is clear that cells with larger total protein concentrations display

- lower fractions of phosphorylated proteins at steady-state,
- shorter adaptation times,
- smaller initial response amplitudes (in terms of phosphorylated fraction),

than those with smaller total protein concentrations. Such variation in total protein concentrations is then applied in the signalling pathway ODE model from Chapter 3 which is used to simulate the behaviour of *E. coli* cells in the ABM.

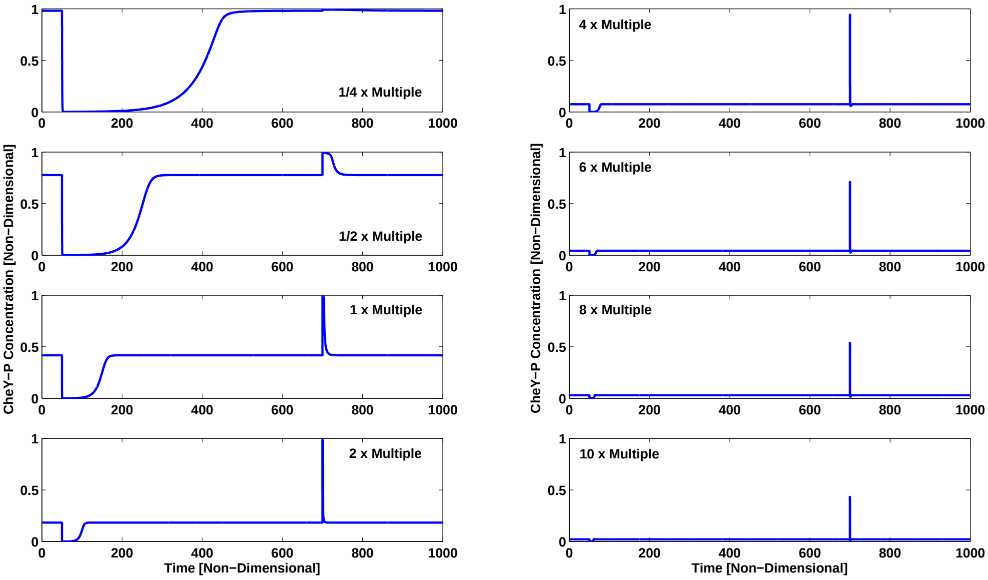
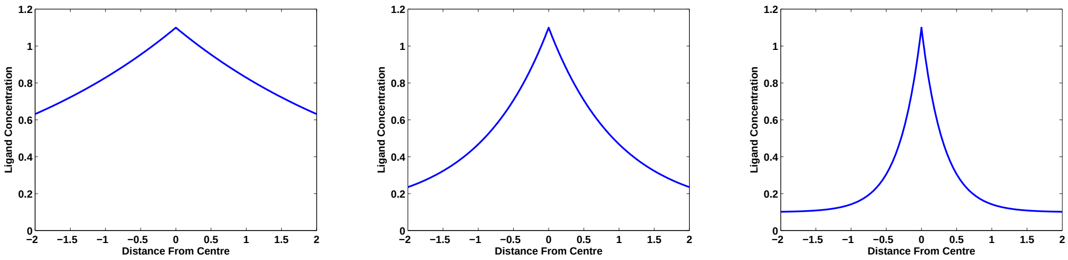


Figure 6.8: A collection of figures showing the intracellular signalling cascade response to step-changes in the extracellular ligand concentration. At non-dimensional time $\tau = 50$ the MeAsp concentration is increased from $[L] = 0.1$mM to $[L] = 0.2$mM and this is subsequently reversed at $\tau = 700$. The different panels here correspond to cells with different total signalling protein concentrations. Shown here are the responses of cells in which each total protein concentration has been multiplied by some constant factor ($\alpha = 1/4, 1/2, 1, 2, 4, 6, 8, 10$).

It is now possible to simulate the behaviour of chemotactic *E. coli* cells under different intracellular conditions. We wish to understand how these intracellular variations affect the ability of a cell population to respond to different extracellular chemical environments. As such we consider different exponential ligand gradients of the form

$$[L] = l_0 + \exp\left(-\sqrt{\frac{x^2 + y^2}{d}}\right), \tag{6.9}$$

within which $[L]$ denotes the ligand concentration, l_0 represents a minimum ligand concentration for the domain, x and y are the horizontal and vertical spatial coordinates, respectively and d is used in order to vary the steepness of the ligand gradient. Specifically, within this section we shall consider three different exponential ligand fields which are achieved by considering $d = 10$



(a) Shallow (b) Intermediate (c) Steep

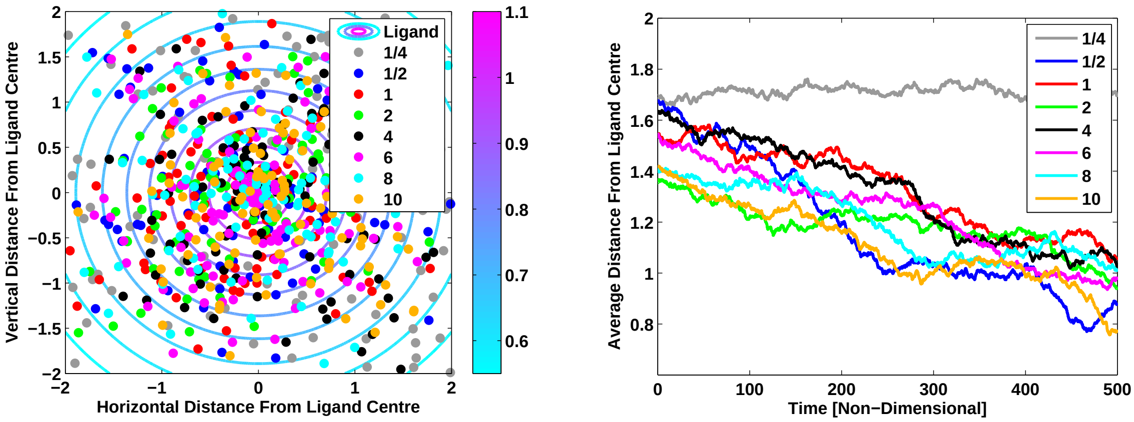
Figure 6.9: Cross-sectional plots taken across the centre of the three ligand gradients used for agent-based simulations of *E. coli* chemotactic behaviour. Each gradient is of the form $[L] = l_0 + \exp(-\sqrt{((x^2 + y^2)/d)})$ and takes a different value of d in order to vary the steepness. Here $d = 10$ produces a shallow gradient (a), $d = 1$ yields an intermediate gradient (b) and $d = 0.1$ gives a steep gradient (c).

for a shallow gradient, $d = 1$ for an intermediate gradient and $d = 0.1$ for a steep gradient, as shown in Figure 6.9.

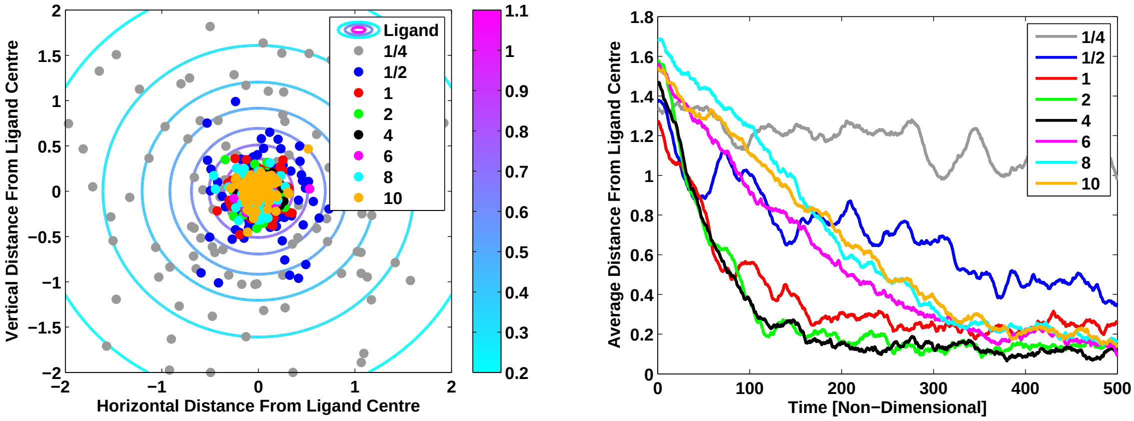
In order to understand how the variation in intracellular signalling protein concentrations affects the ability of the cell to respond in different ligand gradients a number of ABM simulations were conducted. The ABM was used to simulate the behaviour of 100 individual cells within each of the three ligand gradients as well as for each of the eight sets of intracellular protein concentrations considered (i.e. a total of 24 ABM runs). Within each of these simulations all 100 cells were assigned random values for both their initial locations within the domain and the initial direction of travel. The behaviour of each of the 100 individual cells within each population is simulated for 50,000 model time steps equating to approximately 12 minutes (long enough for most simulations to reach an approximate equilibrium). Results obtained from these simulations are displayed in Figure 6.10.

Upon examination of the results displayed in Figure 6.10 a number of interesting features may be observed. Firstly, it is clear from the average distance to the peak ligand concentration that the performance of each cell population is not equal in all gradients. An example of this is that the population of cells with an eight-fold increase in all signalling protein concentrations (i.e. $\beta = 8$) is the best performing population when placed into the steep gradient yet only the fifth best in the shallow gradient, as seen in Figure 6.11. however this result can vary dependent upon the gradients considered.

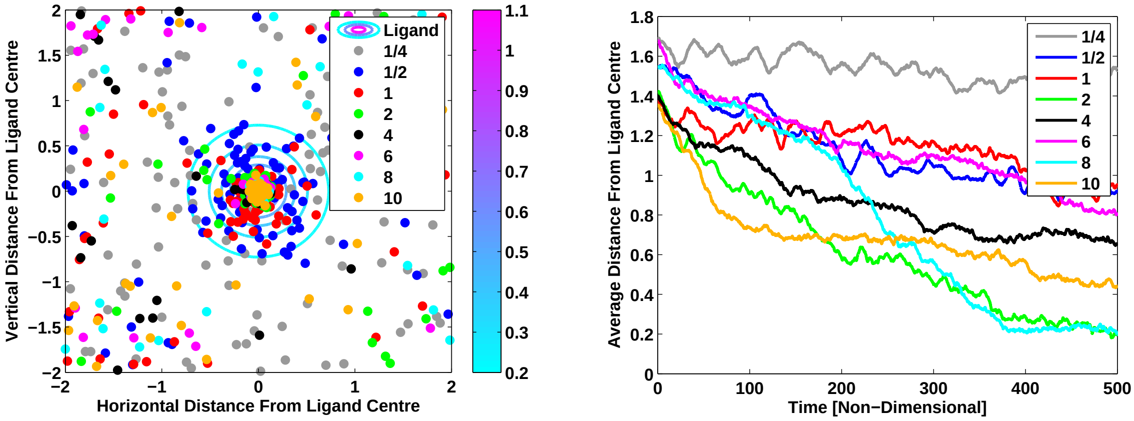
In order to understand why cell populations with different protein concentrations may be better at responding to certain gradients than others, it is useful to consider the signalling cascade response of cells in each population, as seen in Figure 6.8. From these results we may observe a significantly shorter adaptation time for those cells with larger β factors. In particular, within these cells there will exist larger concentrations of CheB and CheR which are



(a) Shallow Extracellular Ligand Gradient



(b) Intermediate Extracellular Ligand Gradient



(c) Steep Extracellular Ligand Gradient

Figure 6.10: Plots showing (left column) the final locations of simulated *E. coli* cells in different ligand gradients and (right column) the development in time of the average distance to the peak ligand concentration for each cell population. Results are shown for (a) a shallow gradient, (b) an intermediate gradient and (c) a steep gradient. (Left) Each dot refers to the final location of a single simulated *E. coli* cell. (Right) Each line shows the average behaviour of a cell population in time. The colour of each dot/line denotes the multiple of all total signalling protein concentrations used in the simulation of that cell population. Considered here are multiples of 1/4 (grey), 1/2 (blue), 1 (red), 2 (green) 4 (black), 6 (pink), 8 (cyan) and 10 (gold).

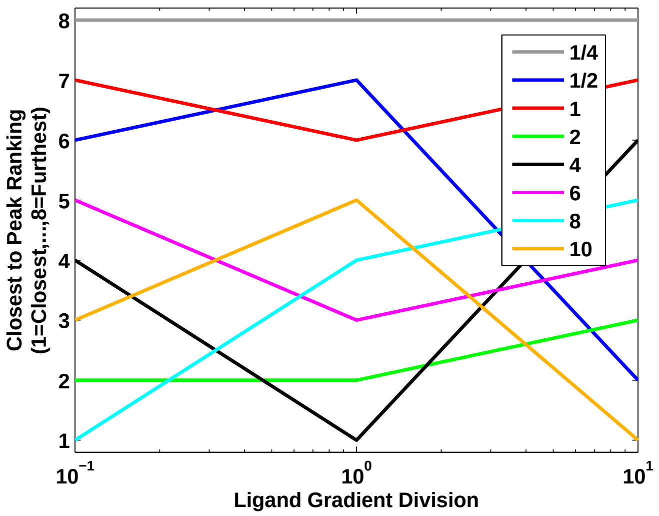


Figure 6.11: A plot comparing the relative abilities of different cell populations to accumulate about the peak concentration within different exponential ligand fields. In particular we consider here a steep (left, 10^{-1}), intermediate (centre, 10^{0}) and shallow (right, 10^{1}) ligand gradient, where the x-axis values correspond to d in equation (6.9). Coloured lines show the rankings of each cell population with 8 representing the worst performing population and 1 the best. The colours of lines indicate the multiples (shown in the figure legend) of all total protein concentrations used in order to create different cell populations. Rankings displayed in this figure simply relate to the average distance of each population from the peak ligand concentration at the end of the ABM simulation.

responsible for controlling adaptation. As such, with more of these proteins available methylation and demethylation occur more rapidly and so the cell can reset itself more quickly following a stimulus. This is usually considered to be a desirable property in chemotactic systems as it allows the cell to prepare for sensing a new stimulus. So, why is it then that cells with larger protein concentrations do not always respond more effectively than those with lower concentrations? To answer this we again look to the responses of individual cells in Figure 6.8. Within these results different adaptation times may be observed. Whilst a shorter adaptation time may be beneficial to chemotactic cells it will also result in a shorter period where the CheY-P concentration is out of equilibrium. Thus a shorter adaptation time will alter the ability of cells to produce long chemotactic runs. However, an inability to produce long extended runs is not always detrimental. For example, within a steep ligand gradient as considered here ($d = 0.1$) shorter runs may be beneficial as they allow the cell to tumble more quickly as it starts to travel down the ligand gradient, an occurrence commonly observed in steeper gradients due to large concentration changes in short spaces.

If a cell with a shorter adaptation time really is more effective at remaining on a beneficial course, why is this not always helpful? To answer this we consider results obtained here for the shallow ligand gradient. In this case, cells begin to travel down the ligand gradient less often due to the fact there is less rapid spatial variation in the concentration gradient. Thus it becomes beneficial for the cell to be able to produce longer runs in order to make the most of subtle changes in the ligand concentration. Examining the single cell responses in Figure 6.8 shows that the ability to produce long runs is associated with those cells displaying longer adaptation times, i.e. those with smaller total protein concentrations. We therefore expect slower adapting cells to perform more efficiently within shallower ligand gradients as seen in Figure 6.10(a).

Within the theoretical literature Vladimirov et al [160] studied a similar problem using an MWC-type model and the RapidCell agent-based modelling program. In particular, they note that experimentally and theoretically it can be observed that cells with too little CheB and CheR tend toward running and fail to display tumbles, thus failing to display effective chemotaxis behaviour. This likely corresponds to the $\beta = 1/4$ population in Figure 6.10 which do not appear to display any significant chemotactic accumulation about the peak ligand concentrations of any gradient considered. In addition to this, it is stated that cells with too much CheB and CheR lose sensitivity to the gradient. Vladimirov et al. [160] also note that it is optimal for cells to display faster rates of adaptation when placed in steeper ligand gradients. To some extent this corresponds to the behaviour observed in Figure 6.11 within which we see that a number of populations with large protein concentrations perform fairly well. However, the results obtained here are slightly skewed by the initial random distribution of cells. This is due to the shallow region seen on the outside of the steep gradient in Figure 6.9(c). In particular, in order for fast adapting cells to respond efficiently within steep gradients they must first find their way into the steep region of the gradient around the centre. Once again, this skews results since the faster adapting cells are less effective in shallow gradients and as they such may struggle to find their way into the steep portion.

6.4 Chemotactic Response to Multiple Ligands

Previous sections of this chapter have considered the ways in which features of the intracellular signalling pathway affect the chemotactic response of *E. coli* cells. In each of these cases it has been assumed that the only chemoattractant present in the environment was MeAsp. This simplifying assumption has clear benefits within a theoretical framework, however more biologically representative is the study of cells when two chemoattractants are present. As such, within this section we shall investigate the chemotactic response of simulated *E. coli* cells in a domain containing both MeAsp and serine - substances for which *E. coli* cells possess the most abundant chemoreceptors.

In order to do this we take inspiration from the original model of Clausznitzer et al. [70] that describes a response for both Tar and Tsr receptor types to MeAsp. It has previously been shown that for small to intermediate MeAsp concentrations, the effect of MeAsp binding to Tsr receptors is negligible (see Figure 4.7(a)) and as such this effect has previously been ignored. However, by considering a slight adjustment to equation (3.10) it is possible to incorporate MeAsp and serine binding to Tar and Tsr receptors, respectively. This results in a free-energy expression of the form

$$F = N\left[1 - \frac{m}{2} + \nu_a \ln\left(\frac{1 + [L_a]/K_a^{off}}{1 + [L_a]/K_a^{on}}\right) + \nu_s \ln\left(\frac{1 + [L_s]/K_{ser}^{off}}{1 + [L_s]/K_{ser}^{on}}\right)\right], \tag{6.10}$$

within which $[L_a]$ and $[L_s]$ represent the concentrations of MeAsp and serine, respectively while $K_{ser}^{on} = 1mM$ and $K_{ser}^{off} = 0.0025mM$ denote the ligand dissociation constants of Tsr chemoreceptors to serine in their active and inactive states, respectively [69]. In addition to this ν_a and ν_s denote the abundance ratio of the two chemoreceptor types, defined as $\nu_a : \nu_s = 1 : 1.4$ [70]. All other model parameters retain their previous values (see Table 3.1) whilst the remainder of the model is taken to be that outlined in Chapter 3.

In order to ensure that the alteration considered in equation (6.10) produces a mathematical model that is capable of allowing chemotactic responses to both MeAsp and serine gradients we utilise two numerical simulations. Firstly we conduct a simulation within a domain in which there exists an exponential gradient of MeAsp in addition to a constant concentration of serine over the entire domain. We then consider the reverse of this case in which there is an exponential gradient of serine alongside a constant concentration of MeAsp. This gave the results displayed in Figure 6.12 which simulate the behaviour of ten *E. coli* cells that are allocated a random initial location within the domain.

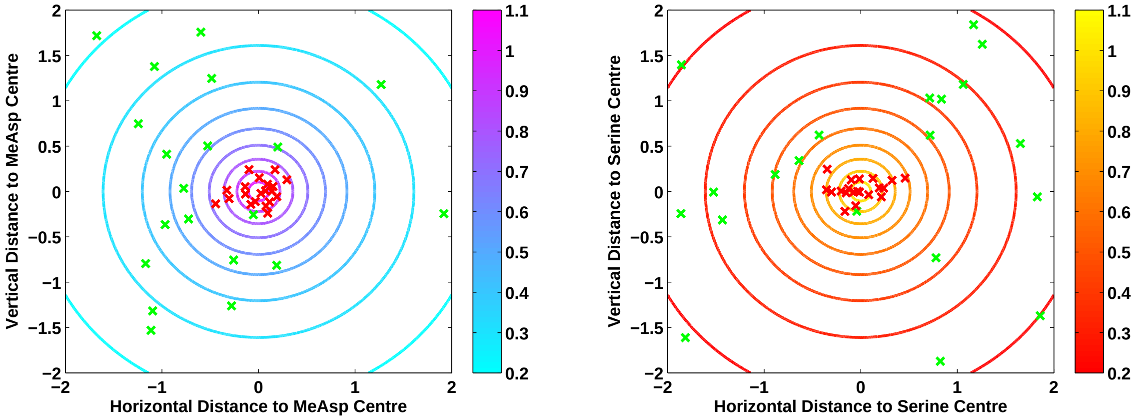
It is clear upon inspection of the results in Figure 6.12 that simulated *E. coli* cells are able to accumulate about the peak of both a MeAsp and a serine gradient. This would tend to suggest that equation (6.10) is able to adequately produce a response for multiple, simultaneously occurring ligands, in this case MeAsp and serine.

It is now possible to examine the effects upon the chemotactic behaviour of a cell population caused by the presence of two distinct ligand gradients. In order to do this we consider two separate exponential ligand gradients, one for each chemoattractant. These are given by the expressions

$$[L_a] = \omega\left(l_{a0} + \exp\left(-\sqrt{(x + x_a)^2 + y^2}\right)\right), \qquad \text{(MeAsp)} \tag{6.11}$$

$$[L_s] = \upsilon\left(l_{s0} + \exp\left(-\sqrt{(x + x_s)^2 + y^2}\right)\right), \qquad \text{(Serine)} \tag{6.12}$$

where $[L_{a/s}]$ denote the concentration of a chemoattractant, l_0 indicates a minimum chemoat-



(a) Exponential MeAsp, Constant Serine

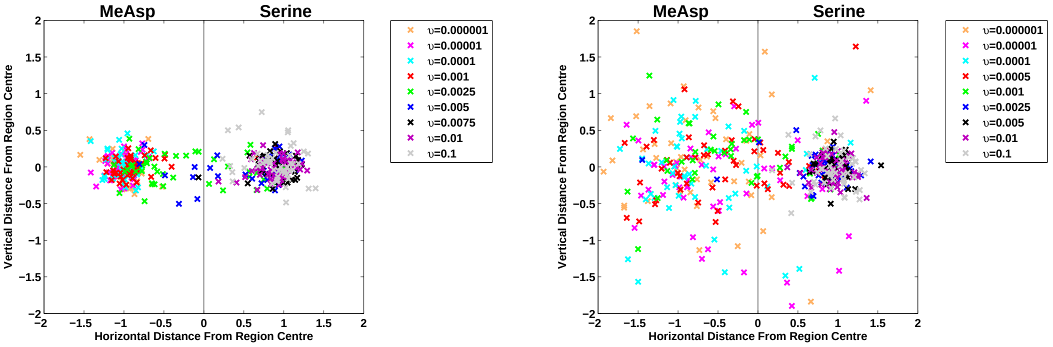
(b) Constant MeAsp, Exponential Serine

Figure 6.12: Plots demonstrating the ability of simulated *E. coli* cells to accumulate in response to both MeAsp and serine concentration gradients. Shown here are cases in which both MeAsp and serine are present within the same domain. In one example (left) we consider a constant concentration of serine across the whole domain, thus demonstrating the ability of simulated cells to respond to a MeAsp concentration gradient. In the other example (right) we consider a constant concentration of MeAsp across the entire domain, thus showing that cells are able to respond to a serine concentration gradient. Within each plot green and red crosses indicate the starting and final locations of each simulated cell, respectively whilst contour lines and colour bars show the concentration of the non-constant ligand across the domain.

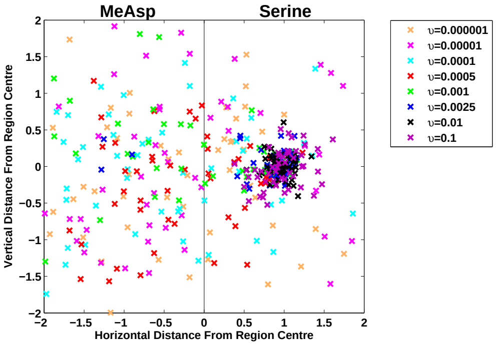
tractant concentration, x and y denote horizontal and vertical coordinates, respectively. Two scaling parameters, namely ω and υ, are used in order to alter the relative concentrations within the MeAsp and serine gradients, respectively. Within these equations the subscripts a and s indicate a reference to aspartate or serine, respectively whilst x_a and x_s are used in order to move the peaks of the ligand gradients apart. In order to do this we choose $x_a = 1$ and $x_s = -1$, resulting in an exponential MeAsp gradient centred about $(x, y) = (-1, 0)$ and an exponential serine gradient centred about $(x, y) = (1, 0)$. Finally, it is possible to utilise the parameters ω and υ as multiplicative scaling factors in order to create ligand gradients with greater peak concentrations.

Using the two scaling parameters it is possible to assess which chemoattractant cells will accumulate about under a variety of different conditions. In particular we consider three different scalings for the MeAsp gradient, namely $\omega = 1$, $\omega = 5$ and $\omega = 10$. For each of these three MeAsp gradient scalings we consider a range of scalings using various values of υ. For each pair of ω and υ values we conduct a simulation of a 50 cell population with a random initial distribution of cells within the domain. Results obtained from these simulations are displayed in Figure 6.13.

It can clearly be seen from the results in Figure 6.13 that there are a number of conditions on ω and υ which result in vastly different numbers of cells being attracted to each gradient. Using these agent-based model simulations we may count the number cells accumulating toward



(a) $\omega = 1$ (b) $\omega = 5$



(c) $\omega = 10$

Figure 6.13: Plots showing the final positions of simulated cells after 50,000 model time-steps (∼12 minutes). In each case the chemotactic response is to a gradient of MeAsp centred in the left half of the domain and a serine gradient centred in the right half of the domain. Coloured crosses show the final position of each simulated *E. coli* cell with the colour indicating the multiple of the serine gradient (i.e. the value of υ within equation (6.12)). The three separate panels relate to the multiple applied to the MeAsp gradient (i.e. the ω value in equation (6.11)). In particular, considered here are values of $\omega = 1$ (top left), $\omega = 5$ (top right) and $\omega = 10$ (bottom).

each attractant. For simplicity we consider a cell to be attracted to MeAsp if the final location is such that $x < 0$ or for $x > 0$ we say the cell was attracted to the serine gradient. In order to more clearly show how the relationship between ω and υ affects the accumulation of cells about the two different ligands are summarised in Figure 6.14.

In order to understand these results it is necessary to consider them together rather than each in isolation. Firstly, it is clear from Figure 6.14 that there is a critical υ value below which some cells will begin to be attracted to the MeAsp gradient. It also appears that for larger values of ω this critical value decreases. At first this may appear counter intuitive - why would a greater MeAsp concentration be overcome by a smaller concentration of serine? In order to answer this question we look back to Figure 4.7(a) and consider this in the context of the peak ligand concentrations utilised within these simulations. In particular for values of $\omega = 1$, $\omega = 5$ and $\omega = 10$ we have peak MeAsp concentrations of 1.1mM, 5.5mM and 11mM, respectively. Examining these in the context of the Mello & Tu [33] sensitivity curve (see Figure 4.7(a)) it can be seen that for the three examples considered here, an $\omega = 1$ scaling produces the greatest

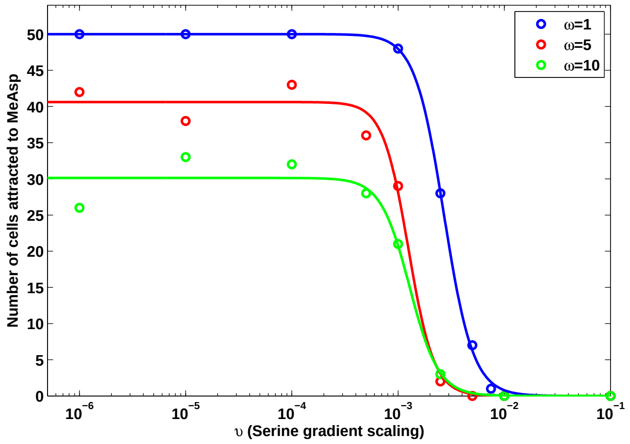


Figure 6.14: A plot summarising the accumulation of simulated *E. coli* cells towards gradients of two separate ligands with different peak concentrations. Results displayed here represent a summary of those in Figure 6.13 with cells considered to accumulate to MeAsp if they end with $x < 1$ and to serine where they finish with $x > 1$. Circles represent the data points drawn from Figure 6.13 with the colour indicating the MeAsp gradient scaling factor where $\omega = 1$ (blue), $\omega = 5$ (red) and $\omega = 10$ (green). Since the agent based model used to accumulate this data contains a number of stochastic components, lines are used to display the general trend of the data. In particular a Hill function is fitted to each set of data using a simple least-squares fit in order to obtain the relevant parameter values.

sensitivity and $\omega = 10$ (due to saturation of receptors) produces the least. As such, there is clear theoretical support for the idea that increasing the peak concentration of one ligand will not necessarily require increasing amounts of another ligand in order to overcome cells being attracted to it. In fact, upon examining the results of Figure 6.13 it seems clearer that the receptor sensitivity is playing a role in determining the ability of cells to accumulate about the MeAsp peak concentration. In Figure 6.13(a), the example with the greatest sensitivity to MeAsp, it is clear that there is strong accumulation toward the peak concentration located at $(x, y) = (-1, 0)$. This differs from panel (b) and (c) where the accumulation is clearly less strong, as evidenced by the reduced clustering about $(x, y) = (-1, 0)$. It is in fact possible to observe weak clustering in panel (b), corresponding to $\omega = 5$ whereas panel (c) ($\omega = 10$) displays virtually no clustering toward the MeAsp peak concentration at all which strongly suggests that the link between the ability to accumulate toward a certain ligand concentration and the receptor sensitivity is causing the emergence of the behaviour observed within these simulations.

Upon further consideration of the results displayed here it is clear that the sensitivity of chemoreceptors to the chemoattractant MeAsp does not quite tell the whole story. It is clear that the sensitivity of chemoreceptors to MeAsp is responsible for the ability of cells to accumulate about a peak MeAsp concentration. This however will not directly affect the ability of cells to accumulate in response to a serine chemoattractant gradient apart from the fact that the two chemoreceptor types share a common intracellular signalling pathway in order to produce a single response. In order to consider the ability of cells to accumulate about serine we look to Figure 6.14. It is clear from these results that the υ value at which cells begin to accumulate toward MeAsp is fairly similar in each of the three examples. This would suggest that a ligand sensitivity curve similar to that for MeAsp is acting to control the sensitivity of the response to serine. In particular, for values of $\upsilon > 10^{-1}$ it is clear that there must be a high level of sensitivity to the serine gradient since all 50 cells in each example are attracted toward the serine peak. For $\upsilon < 10^{-4}$ we would expect a low sensitivity toward the peak aspartate concentration since this is the region in which the fewest cells are attracted to the serine gradient. We should therefore expect that in the range $10^{-4} < \upsilon < 10^{-1}$ we should observe a decreasing sensitivity to the serine peak concentration as the value of υ is decreased.

6.5 Summary & Discussion

Within this chapter we have formulated an ABM of *E. coli* chemotactic cell populations. Upon comparison to both theoretical and experimental literature sources, this model could not be invalidated. This suggests that at the very least it is capable of producing an adequate approximation of the chemotactic behaviour of an *E. coli* cell population.

Within Chapter 5 it was demonstrated that the total concentrations of signalling proteins within cells are important in determining their chemotactic response. In particular, these protein concentrations were observed to lead to different steady-state phosphorylated amounts, significantly different adaptation times as well as the emergence of the experimentally observed overshoot response. In this chapter, at the population scale, it was shown that variation in intracellular protein concentrations can lead to significant differences in the ability of the cells to accumulate about the peak concentration of a ligand gradient. More specifically cells in which protein concentrations produced shorter adaptation times performed more effectively in steep ligand gradients whereas those with longer adaptation times were more effective in shallow ones. This is due to the fact that faster adapting cells are better able to deal with sharp changes in ligand concentration, thus ensuring they maintain a beneficial swimming direction whereas cells with longer adaptation times can produce longer runs which are more beneficial in shallower ligand gradients.

Whilst the majority of work within this thesis focuses on the ability of cells to form a chemotactic response to one chemoattractant (namely MeAsp), Section 6.4 went a step further and examined how cells respond to two different chemoattractants. It was shown here that the response of a cell population would be determined by the sensitivity of the chemoreceptors to the precise chemoattractant concentrations present within the environment. Cells will accumulate toward a ligand concentration for which they are most sensitive rather than the largest absolute concentration. In the case of two competing gradients it is necessary to compare the sensitivity of cells to each in order to assess which gradient will be preferred, with some intermediate regime in which some cells will be attracted to the peak of each gradient.

The results discussed within this chapter demonstrate some of the potential uses of ABM in the study of bacterial chemotaxis. In fact, they would tend to suggest that approaches such as that demonstrated here could even help in the study of as yet understudied systems. ABMs of systems that are not fully understood experimentally could provide an initial round of model rejection in which models that do not produce the experimentally observed behaviour may be identified more rapidly than may be the case in more conventional single cell studies.

Chapter 7

R. sphaeroides Chemotaxis Signalling

In this chapter we investigate the chemotaxis signalling pathways of *R. sphaeroides* cells. We begin with a summary of the experimental and theoretical literature that represents the current understanding of chemotaxis in *R. sphaeroides* cells. A recent mathematical model of this system is then presented alongside a discussion of the main assumptions and simplifications used in its formulation [161]. We then analyse this model in a manner similar to that in Chapter 3. In particular, we consider a non-dimensional re-scaling. This is then used to demonstrate that there exists just one biologically feasible equilibrium state. Stability analysis then shows this equilibrium state to be asymptotically stable and the resulting eigenvalues are examined in order to assess the stiffness of the model system. Finally, we identify the need for new ideas by showing that this model fails to capture behaviour observed in recent experimental work [162].

7.1 Background

7.1.1 Biology Underlying Chemotaxis in *R. sphaeroides*

Here we summarise the key biological processes allowing *R. sphaeroides* cells to exhibit bacterial chemotaxis. Within the experimental literature a number of features of this system have been studied in depth. However, some other areas are less well understood. We do not therefore possess a full understanding of the system from start to finish. As such, within this section we begin by explaining what is known both about the chemotactic swimming behaviour of *R. sphaeroides* cells and how this is controlled by the cells' intracellular signalling pathways. In particular, here we focus on explaining well understood features of the system and note areas where further work is required.

It was discussed in Section 2.1 that *E. coli* cells use bidirectional flagellar motors to produce a run and tumble swimming pattern. In contrast to this, *R. sphaeroides* possess only unidirectional flagellar motors [5]. As such, they are unable to display tumble-type behaviour and instead simply stop, allowing rotational diffusion to re-orient the cell instead [163]. This run and stop swimming pattern is controlled by the intracellular signalling pathways of the cell, however the exact mechanism(s) by which this occurs are not clear.

Whilst it is unclear how the swimming behaviour of *R. sphaeroides* cells is controlled, many intracellular features are better understood. One such feature is the degree of complexity displayed by *R. sphaeroides* cells. This complexity manifests itself in two main ways. The first of these is the existence of three operons that encode multiple homologues of the chemotaxis signalling proteins (four CheW, four CheA, six CheY, two CheB and three CheR) [164]. It is the case however, that proteins encoded in just two of these operons are essential for chemotaxis [164]. As such, within the remainder of this section we focus only on these essential chemotaxis proteins.

Another source of this increased complexity in *R. sphaeroides* is the existence of two separate clusters of MCP-like proteins (chemoreceptors) [165]. These have been shown to localise either to the poles of cells or in the cytoplasm [165]. These distinct receptor clusters are linked to two separate chemosensory pathways [166]. In particular, the two pathways will each begin at a different receptor cluster (see Figure 7.1).

The first of these chemosensory pathways is that beginning at the polar receptor cluster. Similar to those found in *E. coli*, receptors span the cytoplasmic membrane of the cell with the external part able to bind molecules of certain attractants/repellents. A signal is then passed into the cell. Here the intracellular domain of the chemoreceptors bind to the protein CheA$_2$ via the linker proteins CheW$_2$ and CheW$_3$ [167]. Similar to the HPK (histidine protein kinase) in *E. coli*, CheA$_2$ autophosphorylates (forming CheA$_2$-P) at a rate dependent upon the signal received.

Once phosphorylated, CheA$_2$-P may pass phosphoryl groups onto the response regulating proteins CheY$_3$, CheY$_4$, CheY$_6$, CheB$_1$ and CheB$_2$ each of which has been shown to be essential for chemotaxis [164]. Similar to *E. coli*, it is thought that an adaptation module exists to return the cell to its pre-stimulus state. Within this mechanism CheR$_2$ constantly methylates receptors to increase their activity whilst CheB$_1$-P demethylates receptors, lowering their activity [168].

The second signalling pathway of *R. sphaeroides* cells begins at a cytoplasmic cluster consisting of soluble chemoreceptors. This receptor cluster is thought to sense the metabolic state of the cell, although the mechanism(s) associated with this are unclear [5]. In order to pass this signal into the cell, the soluble chemoreceptors associate with proteins CheA$_3$ and CheA$_4$ via a linker CheW$_4$ [169]. The proteins CheA$_3$ and CheA$_4$ each lack some of the domains associated with traditional autophosphorylation. As such, together they form an atypical HPK [170]. Here, neither individual protein may autophosphorylate themselves, rather CheA$_4$ has been shown to

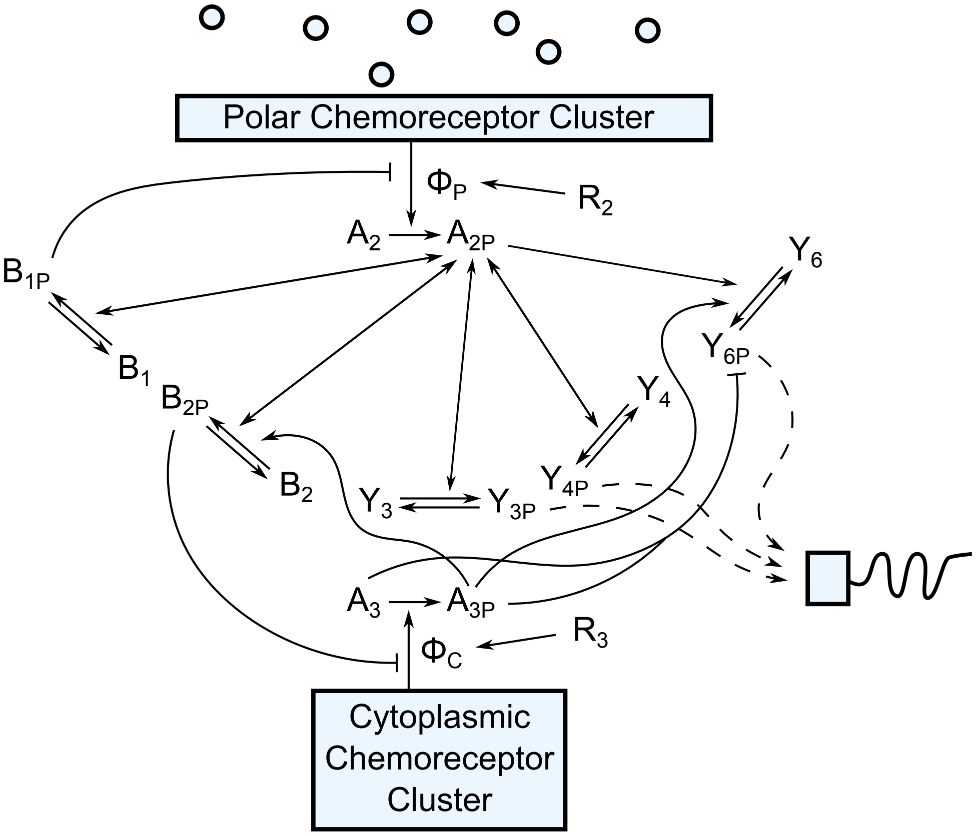


Figure 7.1: A schematic of the *R. sphaeroides* chemotaxis signalling pathway model due to Kojadinovic et al. [161]. This system consists of two chemoreceptor clusters, one located at the cell pole and the other in the cytoplasm of the cell. CheA$_2$ autophosphorylation occurs at the polar cluster and subsequently passes phosphoryl groups onto CheB$_1$, CheB$_2$, CheY$_3$, CheY$_4$ and CheY$_6$. Similarly, CheA$_3$ phosphorylation by CheA$_4$ (modelled as CheA$_3$ autophosphorylation) occurs at the cytoplasmic cluster and the phosphoryl groups are passed onto either CheB$_2$ or CheY$_6$. It is believed that CheY$_3$-P, CheY$_4$-P and CheY$_6$-P control the rotational behaviour of the flagellar motor. Proteins CheR$_2$ and CheR$_3$ along with CheB$_1$ and CheB$_2$ are likely to control the process of adaptation and thus the re-setting of the cell to its pre-stimulus state, however the mechanism by which this occurs is, as yet, unknown.

phosphorylate CheA$_3$ (forming CheA$_3$-P). Subsequent to this, phosphoryl groups may be passed on to a subset of the response regulating proteins, namely CheY$_6$ and CheB$_2$ [170]. It is thought that the proteins CheR$_2$ and CheB$_2$-P form an adaptation mechanism. Whilst this mechanism is not known, it has previously been assumed to function in a similar manner to that in *E. coli*.

It is clear that each receptor cluster can produce responses to different signals via the phosphorylation of certain response regulating proteins. However, since the cell possesses just one flagellum, the two separate chemosensory pathway signals must be integrated into a single response. In order to do this, the proteins CheY$_3$, CheY$_4$ and CheY$_6$ may all interact with the flagellar motor regulating protein FliM [171]. These protein interactions are thought to control the run and stop swimming behaviour of *R. sphaeroides* cells, however the mechanism by which this occurs is unknown.

7.1.2 Mathematical Modelling

In addition to the experimental work discussed in Section 7.1.1, there are also a number of examples within the literature whereby mathematical modelling has sought to further understanding of chemotaxis in *R. sphaeroides* cells. Each of the examples discussed here utilise ODE models to investigate various aspects of the *R. sphaeroides* chemotaxis signalling network. As such, here we focus on the main findings of each paper and provide just a brief summary of the key differences between the theoretical approaches considered.

Early examples of theoretical work on this system sought to elucidate the network of reactions responsible for chemotaxis in *R. sphaeroides.* In particular, the works of Roberts et al. [172] and Hamadeh et al. [173] each constructed a number of mathematical models containing different combinations of possible network connections (i.e. phosphorylation and methylation reactions). In each case the models were compared against experimental data in order to assess which was the most effective model of *R. sphaeroides* chemotaxis. However, whilst these models produce a reasonable fit to some experimental data, they have subsequently been shown to contain reactions that cannot occur biologically.

Tindall et al. [174] sought to utilise mathematical modelling to elucidate the roles of various components of the *R. sphaeroides* chemotaxis signalling pathways. This work was able to demonstrate that CheA$_3$ functions as a bi-functional kinase-phosphatase protein. Biologically, this means that CheA$_3$-P phosphorylates CheY$_6$ and CheB$_2$ (kinase action), however both CheA$_3$ and CheA$_3$-P act to dephosphorylate CheY$_6$-P (phosphatase action). This work also showed that a number of reverse phosphorylation reactions occur, i.e. phosphorylated response regulator proteins (CheB and CheY homologues) may transfer their phosphoryl groups on to non-phosphorylated CheA proteins. In particular, a reverse phosphorylation reaction was found to exist for each phosphotransfer reaction except for one - CheY$_6$-P does not transfer phosphoryl groups onto CheA$_2$.

The work of Amin et al. [175] modelled only reactions occurring at the cytoplasmic cluster of *R. sphaeroides* cells. As such, their model included proteins CheA$_3$, CheA$_4$ and CheY$_6$ in their phosphorylated and non-phosphorylated states as well as the transient complexes formed during each reaction. Using this model it was shown that the atypical HPK present in *R. sphaeroides* cells is capable of producing both ultrasensitivity and bistability under certain parameter regimes.

Martin et al. [168] investigated the mechanisms behind adaptation in *R. sphaeroides* cells by studying reactions involving CheR and CheB proteins. Whilst no mechanisms were proposed in this work, it was concluded that CheR and CheB homologues are likely to be responsible for adaptation. They also went on to state that in spite of utilising homologues of the same proteins, the mechanism for adaptation in *R. sphaeroides* does not appear to be the same as that of *B. subtilis* or *E. coli*.

Recently, the work of Hamadeh et al. [176] and Kojadinovic et al. [161] demonstrated the existence of fold-change detection (FCD) within chemotactic *R. sphaeroides* cells. The theoretical work of Kojadinovic et al. [161] is of particular interest here since it extends the work of Tindall et al. [174] and couples this with an MWC-based adaptation mechanism similar to those considered for *E. coli* chemotaxis modelling for both the polar and cytoplasmic receptor clusters. This model therefore represents the most complete description of the chemotaxis signalling pathways of *R. sphaeroides* within the literature.

7.2 Motivation

Since the work summarised in Section 7.1, new experimental data has become available that has sought to provide new insight into the chemotaxis signalling pathways of *R. sphaeroides* cells [162]. In particular, this work examined the effects on chemotactic swimming behaviour of various different cellular mutations (alterations from the wild-type workings of the cell). These mutations are summarised in Table 7.1 whilst their impacts upon the steady-state fraction of time each cell type spends in a stopped state is shown in Figure 7.2.

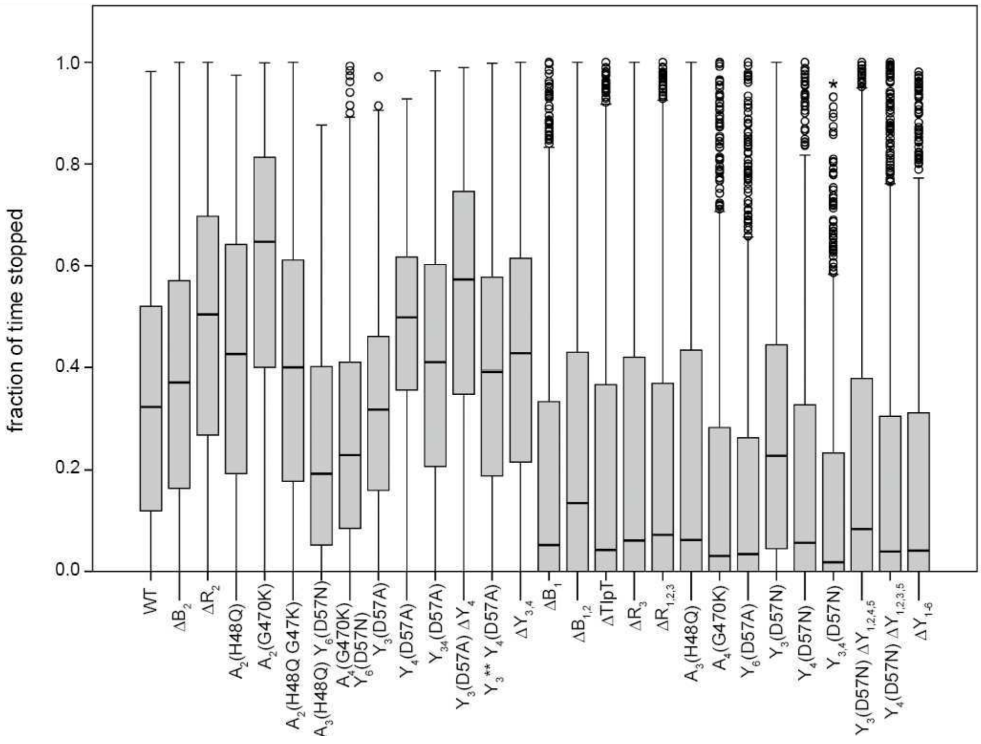


Figure 7.2: From the work of de Beyer [162], the fractions of time each chemotaxis mutant spends stopped. Box plot edges indicate first and third quartile, box centre indicates median, whiskers indicate minimum and maximum and circles indicate outliers.

It is clear from Section 7.1 that a number of mathematical models have been used to study certain features of *R. sphaeroides* chemotaxis signalling pathways. Each of these examples has relied on a number of assumptions or hypotheses, the effects of which have not been been considered in depth.

Table 7.1: Descriptions of each *R. sphaeroides* mutant utilised within the experimental work.

	Mutation	Description
#1	WT	Wild-type strain
#2	ΔCheB$_2$	CheB$_2$ deletion
#3	ΔCheR$_2$	CheR$_2$ deletion
#4	CheA$_2$(H48Q)	CheA$_2$ cannot be phosphorylated
#5	CheA$_2$(G470K)	CheA$_2$ is unable to autophosphorylate
#6	CheA$_2$(H48Q,G470K)	CheA$_2$ has no kinase activity
#7	CheA$_3$(H48Q)CheY$_6$(D57N)	CheA$_3$ cannot be phosphorylated & CheY$_6$ full phosphorylation mimic
#8	CheA$_4$(G470K)CheY$_6$(D57N)	CheA$_4$ cannot autophosphorylate & CheY$_6$ full phosphorylation mimic
#9	CheY$_3$(D57A)	CheY$_3$ inactive and non-phosphorylatable
#10	CheY$_4$(D57A)	CheY$_4$ inactive and non-phosphorylatable
#11	CheY$_{3,4}$(D57A)	CheY$_3$ inactive and non-phosphorylatable & CheY$_4$ inactive and non-phosphorylatable
#12	CheY$_3$(D57A)ΔCheY$_4$	CheY$_3$ inactive and non-phosphorylatable & CheY$_4$ deletion
#13	CheY$_3$**CheY$_4$(D57A)	CheY$_3$ inactivated & CheY$_4$ inactive and non-phosphorylatable
#14	ΔCheY$_{3,4}$	CheY$_3$ and CheY$_4$ deletions
#15	ΔCheB$_1$	CheB$_1$ deletion
#16	ΔCheB$_{1,2}$	CheB$_1$ and CheB$_2$ deletions
#17	ΔTlpT	Deletion of TlpT (chemoreceptor)
#18	ΔCheR$_3$	CheR$_3$ deletion
#19	ΔCheR$_{1,2,3}$	CheR$_1$, CheR$_2$ and CheR$_3$ deletion
#20	CheA$_3$(H48Q)	CheA$_3$ cannot be phosphorylated
#21	CheA$_4$(G470K)	CheA$_4$ cannot autophosphorylate
#22	CheY$_6$(D57A)	CheY$_6$ inactive and non-phosphorylatable
#23	CheY$_3$(D57N)	CheY$_3$ full phosphorylation mimic
#24	CheY$_4$(D57N)	CheY$_4$ full phosphorylation mimic
#25	CheY$_{3,4}$(D57N)	CheY$_3$ full phosphorylation mimic & CheY$_4$ full phosphorylation mimic
#26	CheY$_3$(D57N)ΔCheY$_{1,2,4,5}$	CheY$_3$ full phosphorylation mimic & Deletion of CheY$_{1,2,4,5}$
#27	CheY$_4$(D57N)ΔCheY$_{1,2,3,5}$	CheY$_4$ full phosphorylation mimic & Deletion of CheY$_{1,2,3,5}$
#28	ΔCheY$_{1-6}$	Deletion of CheY$_{1-6}$
#29	CheY$_6$(D57N)	CheY$_6$ full phosphorylation mimic
#30	ΔCheY$_4$	CheY$_4$ deletion
#31	ΔCheY$_3$	CheY$_3$ deletion
#32	ΔCheY$_6$	CheY$_6$ deletion

In light of this new experimental data, we are presented with an opportunity to further analyse existing mathematical models. As such, we consider here a number of altered models representing mutant cell types. Using these models we are able to demonstrate that the most complete model to date does not represent an adequate description of chemotaxis signalling in *R. sphaeroides* and identify areas for further work.

7.3 *R. sphaeroides* Chemotaxis Signalling Model

Within Section 7.1 we summarised the key experimental and theoretical literature that represents the current understanding of *R. sphaeroides* signalling pathways. One model of particular interest is that due to Kojadinovic et al. [161]. This represents the most complete description of *R. sphaeroides* chemotaxis signalling to date and as such is subjected to further analysis within the remainder of this chapter.

This model takes an ODE model of known signalling pathway reactions (formulated using the law of mass action, see Appendix A) and couples this with MWC descriptions of both the polar and cytoplasmic receptor clusters. In addition to this, the adaptation mechanism of *E. coli* cells is assumed to function at each of the polar and cytoplasmic receptor clusters. Due to the methods used in formulating this model it is subject to the assumptions/limitations detailed in Appendix A. Further details of the Kojadinovic et al. [161] model are summarised in the remainder of this section.

The polar chemoreceptor cluster is modelled borrowing the MWC approach from *E. coli* that is as follows. The chemoreceptor activity (Φ_P) is given by

$$\Phi_P = \frac{1}{1 + e^{F_P}}, \tag{7.1}$$

within which F_P denotes the free-energy of the chemoreceptor signalling team. This takes the form

$$F_P = N_P \left[1 - \frac{m_P}{2} + \ln \left(\frac{1 + [L]/K_P^{off}}{1 + [L]/K_P^{on}} \right) \right], \tag{7.2}$$

where N_P denotes the number of receptors forming the polar cluster signalling team, $[L]$ is the extracellular ligand concentration, $K_P^{on/off}$ are the ligand dissociation constants of polar chemoreceptors in their active and inactive states, respectively and m_P denotes the methylation level of the polar chemoreceptors. The behaviour in time of this receptor methylation level is described by the ODE

$$\frac{dm_P}{dt} = g_{R2}(1 - \Phi_P) - g_{B1}[B_{1P}]^2 \Phi_P, \tag{7.3}$$

within which g_{R2} and g_{B1} represent the kinetic rate parameters of receptor methylation and demethylation by the proteins CheR$_2$ and CheB$_1$-P, respectively.

In a manner similar to that of the polar chemoreceptor cluster, the cytoplasmic receptor cluster activity (Φ_C) is also borrowed from *E. coli* and is described by the expression

$$\Phi_C = \frac{1}{1+e^{F_C}}, \tag{7.4}$$

in which F_C represents the free-energy of a cytoplasmic cluster chemoreceptor signalling team. This free-energy is represented by

$$F_C = N_C \left[1 - \frac{m_C}{2} + \ln\left(\frac{1+[L]/K_C^{off}}{1+[L]/K_C^{on}}\right)\right], \tag{7.5}$$

where N_C indicates how many receptors constitute a signalling team within the cytoplasmic cluster, $[L]$ is the ligand concentration, $K_C^{on/off}$ are the ligand dissociation constants of cytoplasmic cluster chemoreceptors in their active and inactive states, respectively and m_C is the methylation level of the cytoplasmic chemoreceptors. The dynamic behaviour of the chemoreceptor methylation level is given by

$$\frac{dm_C}{dt} = g_{R3}(1-\Phi_C) - g_{B2}[B_{2P}]^2\Phi_C, \tag{7.6}$$

within which g_{R3} and g_{B2} represent the kinetic rate parameters of receptor methylation and demethylation by the proteins CheR$_3$ and CheB$_2$-P, respectively.

In addition to the MWC representations of the polar and cytoplasmic chemoreceptor clusters, it is also necessary to consider a set of ODEs for the behaviour in time of the chemotaxis proteins CheA$_2$, CheA$_3$, CheY$_3$, CheY$_4$, CheY$_6$, CheB$_1$ and CheB$_2$. Based upon the reactions outlined in Figure 7.1, the law of mass action (see Appendix A) was utilised and conservation laws identified in order to produce a pair of ODEs describing the behaviour of proteins CheA$_2$-P and CheA$_3$-P. These take the form

$$\begin{aligned}
\frac{d[A_{2p}]}{dt} &= k_1\Phi_P\left([A_2]_T-[A_{2p}]\right) - k_3[A_{2p}]\left([Y_3]_T-[Y_{3p}]\right) + k_{-3}\left([A_2]_T-[A_{2p}]\right)[Y_{3p}] \qquad (7.7)\\
&- k_4[A_{2p}]\left([Y_4]_T-[Y_{4p}]\right) + k_{-4}\left([A_2]_T-[A_{2p}]\right)[Y_{4p}] - k_5[A_{2p}]\left([Y_6]_T-[Y_{6p}]\right)\\
&- k_6[A_{2p}]\left([B_1]_T-[B_{1p}]\right) + k_{-6}\left([A_2]_T-[A_{2p}]\right)[B_{1p}] - k_7[A_{2p}]\left([B_2]_T-[B_{2p}]\right)\\
&+ k_{-7}\left([A_2]_T-[A_{2p}]\right)[B_{2p}],\\
\frac{d[A_{3p}]}{dt} &= k_2\Phi_C\left([A_3]_T-[A_{3p}]\right) - k_8[A_{3p}]\left([Y_6]_T-[Y_{6p}]\right) + k_{-8}\left([A_3]_T-[A_{3p}]\right)[Y_{6p}] \qquad (7.8)\\
&- k_9[A_{3p}]\left([B_2]_T-[B_{2p}]\right) + k_{-9}\left([A_3]_T-[A_{3p}]\right)[B_{2p}].
\end{aligned}$$

In addition to the ODEs describing the dynamics of the CheA proteins, the law of mass action

Table 7.2: A base set of *R. sphaeroides* parameter values from experimental and theoretical literature sources. Here the superscript a denotes a value taken directly from experimental data, b shows that a parameter was obtained by fitting a mathematical model to experimental data and c indicates that a value was inferred from *E. coli* experimental values.

Symbol	Description	Value	Source
k_1	Rate of CheA$_2$ autophosphorylation	0.12s^{-1}	$[177]^a$
k_2	Rate of CheA$_3$ phosphorylation by CheA$_4$	0.98s^{-1}	$[170]^a$
k_3	Phosphotransfer CheA$_2$-P to CheY$_3$	$6.60\times10^3(\text{Ms})^{-1}$	$[174]^b$
k_{-3}	Reverse phosphotransfer CheY$_3$-P to CheA$_2$	$1.17\times10^4(\text{Ms})^{-1}$	$[174]^b$
k_4	Phosphotransfer CheA$_2$-P to CheY$_4$	$8.85\times10^5(\text{Ms})^{-1}$	$[174]^b$
k_{-4}	Reverse phosphotransfer CheY$_4$-P to CheA$_2$	$2.32\times10^5(\text{Ms})^{-1}$	$[174]^b$
k_5	Phosphotransfer CheA$_2$-P to CheY$_6$	$1.54\times10^3(\text{Ms})^{-1}$	$[174]^b$
k_6	Phosphotransfer CheA$_2$-P to CheB$_1$	$1.78\times10^6(\text{Ms})^{-1}$	$[174]^b$
k_{-6}	Reverse phosphotransfer CheB$_1$-P to CheA$_2$	$2.85\times10^6(\text{Ms})^{-1}$	$[174]^b$
k_7	Phosphotransfer CheA$_2$-P to CheB$_2$	$3.07\times10^3(\text{Ms})^{-1}$	$[174]^b$
k_{-7}	Reverse phosphotransfer CheB$_2$-P to CheA$_2$	$1.53\times10^3(\text{Ms})^{-1}$	$[174]^b$
k_8	Phosphotransfer CheA$_3$-P to CheY$_6$	$7.75\times10^5(\text{Ms})^{-1}$	$[174]^b$
k_{-8}	Reverse phosphotransfer CheY$_6$-P to CheA$_3$	$2.83\times10^3(\text{Ms})^{-1}$	$[174]^b$
k_9	Phosphotransfer CheA$_3$-P to CheB$_2$	$6.15\times10^4(\text{Ms})^{-1}$	$[174]^b$
k_{-9}	Reverse phosphotransfer CheB$_2$-P to CheA$_3$	$3.10\times10^3(\text{Ms})^{-1}$	$[174]^b$
k_{10}	Autodephosphorylation of CheY$_3$-P	$5.39\times10^{-2}\text{s}^{-1}$	$[177]^a$
k_{11}	Autodephosphorylation of CheY$_4$-P	$3.36\times10^{-2}\text{s}^{-1}$	$[177]^a$
k_{12}	Autodephosphorylation of CheY$_6$-P	$1.60\times10^{-1}\text{s}^{-1}$	$[177]^a$
k_{13}	Autodephosphorylation of CheB$_1$-P	$4.70\times10^{-3}\text{s}^{-1}$	$[177]^a$
k_{14}	Autodephosphorylation of CheB$_2$-P	$1.01\times10^{-2}\text{s}^{-1}$	$[177]^a$
k_{15}	Dephosphorylation of CheY$_6$-P by CheA$_3$	$5.20\times10^3(\text{Ms})^{-1}$	$[178]^a$
k_{16}	Dephosphorylation of CheY$_6$-P by CheA$_3$-P	$5.20\times10^3(\text{Ms})^{-1}$	$[178]^a$
g_{R2}	Rate of polar receptor methylation by CheR$_2$	$6.00\times10^{-3}\text{s}^{-1}$	$[161]^b$
g_{R3}	Rate of cytoplasmic methylation by CheR$_3$	$6.00\times10^{-3}\text{s}^{-1}$	$[161]^b$
g_{B1}	Rate of polar receptor methylation by CheB$_1$	$1.57\times10^8(\text{M}^2\text{s})^{-1}$	$[161]^b$
g_{B2}	Rate of cytoplasmic demethylation by CheB$_2$	$1.57\times10^9(\text{M}^2\text{s})^{-1}$	$[161]^b$
$[A_2]_T$	Total concentration of CheA$_2$	$89.93\mu\text{M}$	$[179]^a$
$[A_3]_T$	Total concentration of CheA$_3$	$89.93\mu\text{M}$	$[179]^a$
$[B_1]_T$	Total concentration of CheB$_1$	$81.20\mu\text{M}$	$[179]^a$
$[B_2]_T$	Total concentration of CheB$_2$	$20.78\mu\text{M}$	$[179]^a$
$[R_2]_T$	Total concentration of CheR$_2$	$0.16\mu\text{M}$	$[112]^c$
$[R_3]_T$	Total concentration of CheR$_3$	$0.16\mu\text{M}$	$[112]^c$
$[Y_3]_T$	Total concentration of CheY$_3$	$3.46\mu\text{M}$	$[179]^a$
$[Y_4]_T$	Total concentration of CheY$_4$	$13.84\mu\text{M}$	$[179]^a$
$[Y_6]_T$	Total concentration of CheY$_6$	$225\mu\text{M}$	$[179]^a$
N_P	Receptors in a polar signalling team	18	$[62]^c$
N_C	Receptors in a cytoplasmic signalling team	18	$[62]^c$
K_P^{on}	Dissociation constant: active polar receptor	$5\times10^{-1}\text{mM}$	$[70]^c$
K_P^{off}	Dissociation constant: inactive polar receptor	$2\times10^{-2}\text{mM}$	$[70]^c$
K_C^{on}	Dissociation constant: active cytoplasmic receptor	$5\times10^{-1}\text{mM}$	$[161]^c$
K_C^{off}	Dissociation constant: inactive cytoplasmic receptor	$2\times10^{-2}\text{mM}$	$[161]^c$

and conservation laws may also be applied to the reactions involving CheY proteins, yielding the following ODEs for the behaviour of CheY$_3$-P, CheY$_4$-P and CheY$_6$-P. These are of the form

$$\begin{aligned}
\frac{d[Y_{3p}]}{dt} &= k_3[A_{2p}]\left([Y_3]_T - [Y_{3p}]\right) - k_{-3}[Y_{3p}]\left([A_2]_T - [A_{2p}]\right) - k_{10}[Y_{3p}], \qquad (7.9)\\
\frac{d[Y_{4p}]}{dt} &= k_4[A_{2p}]\left([Y_4]_T - [Y_{4p}]\right) - k_{-4}[Y_{4p}]\left([A_2]_T - [A_{2p}]\right) - k_{11}[Y_{4p}], \qquad (7.10)\\
\frac{d[Y_{6p}]}{dt} &= k_5[A_{2p}]\left([Y_6]_T - [Y_{6p}]\right) + k_8[A_{3p}]\left([Y_6]_T - [Y_{6p}]\right) - k_{-8}\left([A_3]_T - [A_{3p}]\right)[Y_{6p}] \qquad (7.11)\\
&- k_{12}[Y_{6p}] - k_{15}\left([A_3]_T - [A_{3p}]\right)[Y_{6p}] - k_{16}[A_{3p}][Y_{6p}].
\end{aligned}$$

It is worth noting that in equation (7.11) the terms $-k_{15}([A_{3T}] - [A_{3p}])[Y_{6p}]$ and $-k_{16}[A_{3p}][Y_{6p}]$ could be written as $-k_{15}[A_{3T}][Y_{6p}]$ since the rate constants k_{15} and k_{16} are equal. However, in this case the separate terms are retained to explicitly include the effects of each reaction. Finally, the law of mass action and conservation laws may again be utilised in order to derive a pair of ODEs describing the dynamic behaviour of the proteins CheB$_1$-P and CheB$_2$-P which take the form

$$\begin{aligned}
\frac{d[B_{1p}]}{dt} &= k_6[A_{2p}]\left([B_1]_T - [B_{1p}]\right) - k_{-6}[B_{1p}]\left([A_2]_T - [A_{2p}]\right) - k_{13}[B_{1p}], \qquad (7.12)\\
\frac{d[B_{2p}]}{dt} &= k_7[A_{2p}]\left([B_2]_T - [B_{2p}]\right) - k_{-7}\left([A_2]_T - [A_{2p}]\right)[B_{2p}] \qquad (7.13)\\
&+ k_9[A_{3p}]\left([B_2]_T - [B_{2p}]\right) - k_{-9}\left([A_3]_T - [A_{3p}]\right)[B_{2p}] - k_{14}[B_{2p}].
\end{aligned}$$

Within equations (7.1)-(7.13) protein concentrations are denoted [...] with subscripts P and T indicating the phosphorylated and total concentrations of the relevant protein, respectively. The kinetic rate parameters of each reaction are (unless otherwise stated) denoted k_i ($i = 1, 2, 3, -3, 4, -4, 5, 6, -6, 7, -7, 8, -8, 9, -9, 10, 11, 12, 13, 14, 15, 16$) and full descriptions of each parameter are listed in Table 7.2.

In order to complete this model, a set of initial conditions are also considered. These take the form

$$m_P = m_{P0}, \quad m_C = m_{C0}, \quad [A_{2p}] = [A_{2p}]_0, \quad [A_{3p}] = [A_{3p}]_0,$$

$$[Y_{3p}] = [Y_{3p}]_0, \quad [Y_{4p}] = [Y_{4p}]_0, \quad [Y_{6p}] = [Y_{6p}]_0, \quad [B_{1p}] = [B_{1p}]_0, \quad \text{and} \quad [B_{2p}] = [B_{2p}]_0.$$

Values for these initial conditions may take any value within the biologically feasible range. For each protein this requires, for example, $0 \leq [A_{2p}]_0 \leq [A_2]_T$, while each methylation level must be chosen such that, for example, $0 \leq m_{P0} \leq m_{max}$, where m_{max} denotes the greatest possible methylation level of a chemoreceptor.

7.4 Model Analysis

7.4.1 Non-Dimensionalisation

Here we consider a non-dimensionalisation (re-scaling) of the mathematical model presented in equations (7.1)-(7.13). In particular we make the choice to re-scale each of the chemotaxis protein concentrations with respect to the relevant total concentration, eg. $[A_{2p}] = a_{2p}[A_2]_T$. We also consider a re-scaling of time with respect to the rate of CheA$_2$-P autophosphorylation, i.e. $t = \tau/k_1$, where τ represents non-dimensional time. Substitution of these re-scaled variables into equations (7.1)-(7.13) results in the system of non-dimensional ODEs

$$\frac{dm_P}{d\tau} = \gamma_{R2}(1-\Phi_P) - \gamma_{B1}b_{1p}^2\Phi_P = g_1, \tag{7.14}$$

$$\frac{dm_C}{d\tau} = \gamma_{R3}(1-\Phi_C) - \gamma_{B2}b_{2p}^2\Phi_C = g_2, \tag{7.15}$$

$$\begin{aligned}\frac{da_{2p}}{d\tau} &= \bar{k}_1\Phi_P\left(1-a_{2p}\right) - \bar{k}_3a_{2p}\left(1-y_{3p}\right) + \bar{k}_{-3}\left(1-a_{2p}\right)y_{3p}\\ &- \bar{k}_4a_{2p}\left(1-y_{4p}\right) + \bar{k}_{-4}\left(1-a_{2p}\right)y_{4p} - \bar{k}_5a_{2p}\left(1-y_{6p}\right)\\ &- \bar{k}_6a_{2p}\left(1-b_{1p}\right) + \bar{k}_{-6}\left(1-a_{2p}\right)b_{1p} - \bar{k}_7a_{2p}\left(1-b_{2p}\right)\\ &+ \bar{k}_{-7}\left(1-a_{2p}\right)b_{2p} = g_3,\end{aligned} \tag{7.16}$$

$$\begin{aligned}\frac{da_{3p}}{d\tau} &= \bar{k}_2\Phi_C\left(1-a_{3p}\right) - \bar{k}_8a_{3p}\left(1-y_{6p}\right) + \bar{k}_{-8}\left(1-a_{3p}\right)y_{6p}\\ &- \bar{k}_9a_{3p}\left(1-b_{2p}\right) + \bar{k}_{-9}\left(1-a_{3p}\right)b_{2p} = g_4,\end{aligned} \tag{7.17}$$

$$\frac{dy_{3p}}{d\tau} = \alpha_1\bar{k}_3a_{2p}\left(1-y_{3p}\right) - \alpha_1\bar{k}_{-3}y_{3p}\left(1-a_{2p}\right) - \bar{k}_{10}y_{3p} = g_5, \tag{7.18}$$

$$\frac{dy_{4p}}{d\tau} = \alpha_2\bar{k}_4a_{2p}\left(1-y_{4p}\right) - \alpha_2\bar{k}_{-4}y_{4p}\left(1-a_{2p}\right) - \bar{k}_{11}y_{4p} = g_6, \tag{7.19}$$

$$\begin{aligned}\frac{dy_{6p}}{d\tau} &= \alpha_3\bar{k}_5a_{2p}\left(1-y_{6p}\right) + \alpha_6\bar{k}_8a_{3p}\left(1-y_{6p}\right) - \alpha_6\bar{k}_{-8}\left(1-a_{3p}\right)y_{6p}\\ &- \bar{k}_{12}y_{6p} - \bar{k}_{15}\left(1-a_{3p}\right)y_{6p} - \bar{k}_{16}a_{3p}y_{6p} = g_7,\end{aligned} \tag{7.20}$$

$$\frac{db_{1p}}{d\tau} = \alpha_4\bar{k}_6a_{2p}\left(1-b_{1p}\right) - \alpha_4\bar{k}_{-6}b_{1p}\left(1-a_{2p}\right) - \bar{k}_{13}b_{1p} = g_8, \tag{7.21}$$

$$\begin{aligned}\frac{db_{2p}}{d\tau} &= \alpha_5\bar{k}_7a_{2p}\left(1-b_{2p}\right) - \alpha_5\bar{k}_{-7}\left(1-a_{2p}\right)b_{2p} + \alpha_7\bar{k}_9a_{3p}\left(1-b_{2p}\right)\\ &- \alpha_7\bar{k}_{-9}\left(1-a_{3p}\right)b_{2p} - \bar{k}_{14}b_{2p} = g_9,\end{aligned} \tag{7.22}$$

within which lower case a_i $(i = 2,3)$, y_i $(i = 3,4,6)$ and b_i $(i = 1,2)$ denote the (non-dimensional) fractions of the relevant chemotaxis proteins that are phosphorylated, $\bar{k}_i$ $(i = 1,2,3,-3,4,-4,5,6,-6,7,-7,8,-8,\ 9,-9,10,11,12,13,14,15,16)$ are non-dimensional parameters (as defined in Table 7.3), α_i $(i = 1,2,3,4,5,\ 6,7)$ represent total protein concentration ratios and all other symbols retain their previous definitions.

Table 7.3: A base set of non-dimensional *R. sphaeroides* parameter values calculated using values in Table 7.2.

Parameter	Value
$\bar{k}_1 = k_1/k_1$	1
$\bar{k}_2 = k_2/k_1$	8.167
$\bar{k}_3 = k_3[Y_3]_T/k_1$	0.190
$\bar{k}_{-3} = k_{-3}[Y_3]_T/k_1$	0.337
$\bar{k}_4 = k_4[Y_4]_T/k_1$	102.070
$\bar{k}_{-4} = k_{-4}[Y_4]_T/k_1$	26.757
$\bar{k}_5 = k_5[Y_6]_T/k_1$	2.888
$\bar{k}_6 = k_6[B_1]_T/k_1$	1204.467
$\bar{k}_{-6} = k_{-6}[B_1]_T/k_1$	1928.500
$\bar{k}_7 = k_7[B_2]_T/k_1$	0.532
$\bar{k}_{-7} = k_{-7}[B_2]_T/k_1$	0.265
$\bar{k}_8 = k_8[Y_6]_T/k_1$	1453.125
$\bar{k}_{-8} = k_{-8}[Y_6]_T/k_1$	5.306
$\bar{k}_9 = k_9[B_2]_T/k_1$	10.650
$\bar{k}_{-9} = k_{-9}[B_2]_T/k_1$	0.537
$\bar{k}_{10} = k_{10}/k_1$	44916.667
$\bar{k}_{11} = k_{11}/k_1$	0.280
$\bar{k}_{12} = k_{12}/k_1$	0.133
$\bar{k}_{13} = k_{13}/k_1$	0.014
$\bar{k}_{14} = k_{14}/k_1$	0.084
$\bar{k}_{15} = k_{15}[A_3]_T/k_1$	3.897
$\bar{k}_{16} = k_{15}[A_3]_T/k_1$	3.897
$\gamma_{R2} = g_{R2}/k_1$	0.050
$\gamma_{R3} = g_{R3}/k_1$	0.050
$\gamma_{B1} = g_{B1}[B_1]_T^2/k_1$	8.626
$\gamma_{B2} = g_{B2}[B_2]_T^2/k_1$	5.649
$\alpha_1 = [A_2]_T/[Y_3]_T$	25.991
$\alpha_2 = [A_2]_T/[Y_4]_T$	6.498
$\alpha_3 = [A_2]_T/[Y_6]_T$	0.400
$\alpha_4 = [A_2]_T/[B_1]_T$	1.108
$\alpha_5 = [A_2]_T/[B_2]_T$	4.328
$\alpha_6 = [A_3]_T/[Y_6]_T$	0.400
$\alpha_7 = [A_3]_T/[B_2]_T$	4.328

In considering a non-dimensionalisation such as this, the initial conditions must also be re-scaled. As such, the initial conditions become

$$a_{2p} = a_{2p0}, \quad a_{3p} = a_{3p0}, \quad b_{1p} = b_{1p0}, \quad b_{2p} = b_{2p0},$$

$$y_{3p} = y_{3p0}, \quad y_{4p} = y_{4p0} \quad \text{and} \quad y_{6p} = y_{6p0}.$$

Those conditions for variables m_P and m_C remain unaffected. As in Section 7.3, these initial conditions must be chosen to satisfy biologically realistic conditions, for example, $0 \leq a_{2p0} \leq 1$.

7.4.2 Model Steady-States

Similar to the work in Section 3.2.3, here we seek to find the steady-state(s) of the mathematical model outlined in Section 7.4.1. As such, we seek states whereby the temporal rate of change is zero. Thus, we set the derivatives in equations (7.15)-(7.22) to zero, giving our model steady-state(s) as

$$\begin{aligned}
0 &= \gamma_{R2}(1-\Phi_P^*) - \gamma_{B1} b_{1p}^{*2} \Phi_P^*, &(7.23)\\
0 &= \gamma_{R3}(1-\Phi_C^*) - \gamma_{B2} b_{2p}^{*2} \Phi_C^*, &(7.24)\\
0 &= \bar{k}_1 \Phi_P^* \left(1-a_{2p}^*\right) - \bar{k}_3 a_{2p}^* \left(1-y_{3p}^*\right) + \bar{k}_{-3}\left(1-a_{2p}^*\right) y_{3p}^* &(7.25)\\
&- \bar{k}_4 a_{2p}^* \left(1-y_{4p}^*\right) + \bar{k}_{-4}\left(1-a_{2p}^*\right) y_{4p}^* - \bar{k}_5 a_{2p}^* \left(1-y_{6p}^*\right)\\
&- \bar{k}_6 a_{2p}^* \left(1-b_{1p}^*\right) + \bar{k}_{-6}\left(1-a_{2p}^*\right) b_{1p}^* - \bar{k}_7 a_{2p}^* \left(1-b_{2p}^*\right)\\
&+ \bar{k}_{-7}\left(1-a_{2p}^*\right) b_{2p}^*,\\
0 &= \bar{k}_2 \Phi_C^* \left(1-a_{3p}^*\right) - \bar{k}_8 a_{3p}^* \left(1-y_{6p}^*\right) + \bar{k}_{-8}\left(1-a_{3p}^*\right) y_{6p}^* &(7.26)\\
&- \bar{k}_9 a_{3p}^* \left(1-b_{2p}^*\right) + \bar{k}_{-9}\left(1-a_{3p}^*\right) b_{2p}^*,\\
0 &= \alpha_1 \bar{k}_3 a_{2p}^* \left(1-y_{3p}^*\right) - \alpha_1 \bar{k}_{-3} y_{3p}^* \left(1-a_{2p}^*\right) - \bar{k}_{10} y_{3p}^*, &(7.27)\\
0 &= \alpha_2 \bar{k}_4 a_{2p}^* \left(1-y_{4p}^*\right) - \alpha_2 \bar{k}_{-4} y_{4p}^* \left(1-a_{2p}^*\right) - \bar{k}_{11} y_{4p}^*, &(7.28)\\
0 &= \alpha_3 \bar{k}_5 a_{2p}^* \left(1-y_{6p}^*\right) + \alpha_6 \bar{k}_8 a_{3p}^* \left(1-y_{6p}^*\right) - \alpha_6 \bar{k}_{-8}\left(1-a_{3p}^*\right) y_{6p}^* &(7.29)\\
&- \bar{k}_{12} y_{6p}^* - \bar{k}_{15}\left(1-a_{3p}^*\right) y_{6p}^* - \bar{k}_{16} a_{3p}^* y_{6p}^*,\\
0 &= \alpha_4 \bar{k}_6 a_{2p}^* \left(1-b_{1p}^*\right) - \alpha_4 \bar{k}_{-6} b_{1p}^* \left(1-a_{2p}^*\right) - \bar{k}_{13} b_{1p}^*, &(7.30)\\
0 &= \alpha_5 \bar{k}_7 a_{2p}^* \left(1-b_{2p}^*\right) - \alpha_5 \bar{k}_{-7}\left(1-a_{2p}^*\right) b_{2p}^* + \alpha_7 \bar{k}_9 a_{3p}^* \left(1-b_{2p}^*\right) &(7.31)\\
&- \alpha_7 \bar{k}_{-9}\left(1-a_{3p}^*\right) b_{2p}^* - \bar{k}_{14} b_{2p}^*.
\end{aligned}$$

We begin here by rearranging equation (7.27) such that y_{3p}^* is expressed in terms of other parameters and variables. This gives the steady-state of CheY$_3$-P as

$$y_{3p}^* = \frac{k_3[A_2]_T a_{2p}^*}{k_3[A_2]_T a_{2p}^* + k_{-3}[A_2]_T(1-a_{2p}^*) + k_{10}}. \tag{7.32}$$

By application of a very similar process to equation (7.28) we are able to find that the steady-state expression for CheY$_4$-P is of the form

$$y^*_{4p} = \frac{k_4[A_2]_T a^*_{2p}}{k_4[A_2]_T a^*_{2p} + k_{-4}[A_2]_T(1 - a^*_{2p}) + k_{11}}. \tag{7.33}$$

Equation (7.29) may then be investigated from which we find that the steady-state expression for CheY$_6$-P takes the form

$$y^*_{6p} = \frac{k_5[A_2]_T a^*_{2p} + k_8[A_3]_T a^*_{3p}}{k_5[A_2]_T a^*_{2p} + k_8[A_3]_T a^*_{3p} + k_{-8}[A_3]_T(1 - a^*_{3p}) + k_{12} + k_{15}[A_3]_T(1 - a^*_{3p}) + k_{16}[A_3]_T a^*_{3p}}. \tag{7.34}$$

The steady-state expression for protein CheB$_1$-P may be obtained upon rearranging equation (7.30) which yields

$$b^*_{1p} = \frac{k_6[A_2]_T a^*_{2p}}{k_6[A_2]_T a^*_{2p} + k_{-6}[A_2]_T(1 - a^*_{2p}) + k_{13}}. \tag{7.35}$$

We may also apply this same method to equation (7.31) which gives the steady-state expression for CheB$_2$-P as

$$b^*_{2p} = \frac{k_7[A_2]_T a^*_{2p} + k_9[A_3]_T a^*_{3p}}{k_7[A_2]_T a^*_{2p} + k_9[A_3]_T a^*_{3p} + k_{-7}[A_2]_T(1 - a^*_{2p}) + k_{-9}[A_3]_T(1 - a^*_{3p}) + k_{14}}. \tag{7.36}$$

It is also possible to obtain steady-state expressions for the chemoreceptor signalling team activities for both of the polar and cytoplasmic receptor clusters. These are found upon rearranging equations (7.23) and (7.24) and are given by

$$\Phi^*_P = \frac{g_{R2}}{g_{R2} + g_{B1}[B_1]^2_T b^{*2}_{1p}} = \frac{1}{1 + \left(\frac{g_{B1}[B_1]^2_T}{g_{R2}} b^{*2}_{1p}\right)}, \tag{7.37}$$

and

$$\Phi^*_C = \frac{g_{R3}}{g_{R3} + g_{B2}[B_2]^2_T b^{*2}_{2p}} = \frac{1}{1 + \left(\frac{g_{B2}[B_2]^2_T}{g_{R3}} b^{*2}_{2p}\right)}. \tag{7.38}$$

Upon comparing these two expressions with equations (7.1) and (7.4) it can be seen that

$$e^{F_P} = \frac{g_{B1}[B_1]^2_T}{g_{R2}} b^{*2}_{1p}, \tag{7.39}$$

and

$$e^{F_C} = \frac{g_{B2}[B_2]^2_T}{g_{R3}} b^{*2}_{2p}, \tag{7.40}$$

in which F_P is given by equation (7.2) and F_C by equation (7.5). It is now possible to re-arrange these two expressions in order to obtain the steady-state expressions for the average chemoreceptor methylation level at the polar and cytoplasmic clusters. As such, these are given by

$$m_P^* = 2\left[1 + \ln\left(\frac{1+[L]/K_P^{off}}{1+[L]/K_P^{on}}\right) - \frac{1}{N_P}\ln\left(\frac{g_{B1}[B_1]_T^2}{g_{R2}}b_{1p}^{*2}\right)\right], \tag{7.41}$$

and

$$m_C^* = 2\left[1 + \ln\left(\frac{1+[L]/K_C^{off}}{1+[L]/K_C^{on}}\right) - \frac{1}{N_C}\ln\left(\frac{g_{B2}[B_2]_T^2}{g_{R3}}b_{2p}^{*2}\right)\right]. \tag{7.42}$$

It is still clearly necessary to obtain the steady-state expressions for CheA$_2$-P and CheA$_3$-P since these must be substituted into equations (7.32)-(7.42) in order to obtain specific steady-state values. However, upon substituting these a_{2p} and a_{3p} dependent steady-state expressions into equations (7.25) and (7.26) and rearranging, we obtain coupled polynomial expressions of the form

$$p(a_{2p}^*, a_{3p}^*) = c_1 a_{2p}^{*8} + c_2 a_{2p}^{*7} + c_3 a_{2p}^{*6} + c_4 a_{2p}^{*5} + c_5 a_{2p}^{*4} + c_6 a_{2p}^{*3} + c_7 a_{2p}^{*2} + c_8 a_{2p}^{*} + c_9 = 0, \tag{7.43}$$

and

$$q(a_{2p}^*, a_{3p}^*) = c_{10} a_{3p}^{*5} + c_{11} a_{3p}^{*4} + c_{12} a_{3p}^{*3} + c_{13} a_{3p}^{*2} + c_{14} a_{3p}^{*} + c_{15} = 0, \tag{7.44}$$

in which the polynomial coefficients c_{1-9} depend upon the value of a_{3p}^* and coefficients c_{10-15} are dependent upon the a_{2p}^* value.

We begin our analysis of the above polynomial equations by examining equation (7.44), the roots (solutions) of which represent the possible steady-state values for a_{3p}^*. We may investigate the nature of solutions to this polynomial equation utilising Descartes' rule of signs (see Appendix B) and the parameter values listed in Table 7.2. Upon substitution of these parameters into equation (7.44) we find that

$$c_{10} = -59.1490, \tag{7.45}$$

$$c_{11} = -977.0201 + 18.7423 a_{2p}^*, \tag{7.46}$$

$$c_{12} = -79.3906 - 116.3834 a_{2p}^* + 2.1571 a_{2p}^{*2}, \tag{7.47}$$

$$c_{13} = -1.8894\times10^{-1} - 7.4581 a_{2p}^* - 4.3907 a_{2p}^{*2} + 5.6289\times10^{-2} a_{2p}^{*3}, \tag{7.48}$$

$$\begin{aligned} c_{14} &= 4.3715\times10^{-3} + 5.8141\times10^{-3} a_{2p}^* - 1.7374\times10^{-1} a_{2p}^{*2} - 4.7975\times10^{-2} a_{2p}^{*3} \\ &+ 4.8953\times10^{-5} a_{2p}^{*4}, \end{aligned} \tag{7.49}$$

$$\begin{aligned} c_{15} &= 3.3667 \times 10^{-4} + 4.4652 \times 10^{-4} a_{2p}^{*} + 2.2002 \times 10^{-4} a_{2p}^{*2} + 3.0993 \times 10^{-4} a_{2p}^{*3} \quad (7.50) \\ &+ 7.9690 \times 10^{-5} a_{2p}^{*4}, \end{aligned}$$

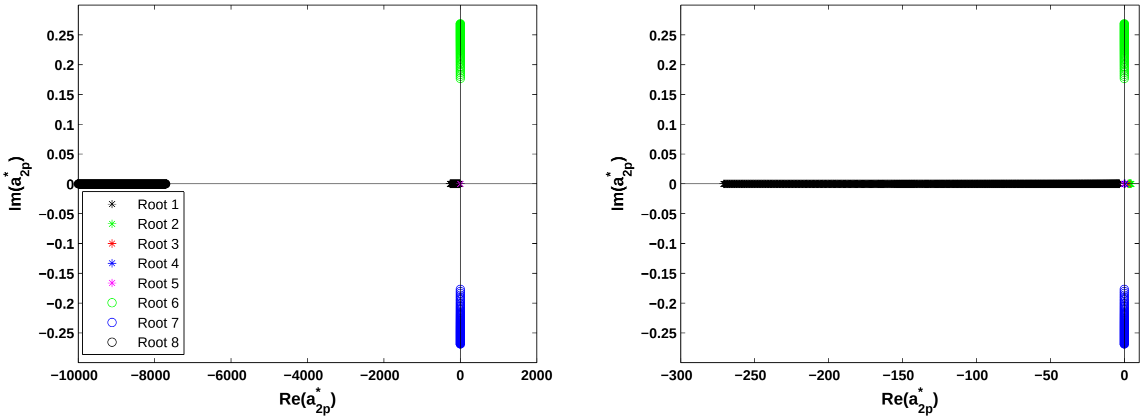
where each term is expressed to four decimal places. Examining these values shows that in the range $0 \leq a_{2p}^{*} \leq 1$ coefficients c_{10}-c_{13} are negative whilst c_{15} is positive. Coefficient c_{14} may be either positive or negative depending upon the magnitude of a_{2p}^{*}. In spite of the fact that the sign of c_{14} may change it is clear that there exists just one sign change between consecutive polynomial coefficients. Hence there will be just one positive steady-state value for this equation. Furthermore we may examine the coefficients of the polynomial $q(a_{2p}^{*}, -a_{3p}^{*})$, within which the coefficients c_{10}, c_{12} and c_{14} will take the opposite sign to the previous case. As such we find that there are four changes in sign between consecutive polynomial coefficients. It will therefore be the case that this polynomial will have either:

- **Case i:** one positive real, four negative real and zero complex roots;
- **Case ii:** one positive real, two negative real and two complex roots; or
- **Case iii:** one positive real, zero negative real and four complex roots.

However, in each of the above three cases these four roots are not biologically feasible and as such they may be neglected.

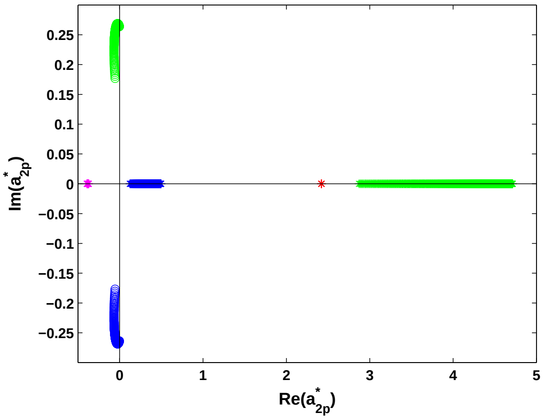
Analysis of the coefficients in the polynomial equation (7.43) using Descartes' rule of signs is inconclusive, i.e. the signs of the coefficients are not all clear. As such we must seek an alternative method to investigate the steady-state(s) of CheA$_2$-P. Since we have already shown that there is only one biologically feasible steady-state for CheA$_3$-P, we may investigate the roots of equation (7.43) using parameter values listed in Table 7.2 and varying the value of a_{3p}^{*} over the whole biologically feasible range $0 \leq a_{3p}^{*} \leq 1$. This investigation was carried out using the 'roots' function in MATLAB which calculates all possible roots of a polynomial equation given values for each of the coefficients. This analysis yields the results displayed in Figure 7.3.

Upon inspection of the results in Figure 7.3 it is clear that there are three positive real roots, two complex roots and three negative real roots to equation (7.43). Interestingly, whilst the values of these roots vary, their nature does not change as a_{3p}^{*} varies. Within this work we are only interested in those roots which lie in the biologically feasible range $0 \leq a_{2p}^{*} \leq 1$. As such we may neglect each of the negative real and complex roots (i.e. root 1 and roots 5-8 in Figure 7.3). Inspecting the positive real roots we are able to see (in Figure 7.3(c)) that there is just one root lying within the biologically feasible range. Hence we may neglect roots 2 and 3 in Figure 7.3 and conclude that the steady-state value for a_{3p}^{*} will be that given by root 4.



(a) All Roots

(b) Partial Zoom



(c) Full Zoom

Figure 7.3: Plots showing all possible roots that may be obtained from equation (7.43) when taking parameter values as in Table 7.2 and varying a^*_{3p} over the biologically feasible range $0 \leq a^*_{3p} \leq 1$. Plot (a) displays all eight roots of the polynomial equation. Plot (b) focuses in on the region of biological feasibility. Plot (c) further focuses on this region of interest. Different coloured symbols indicate the eight different roots of this polynomial, only one of which lies in the biologically feasible range $0 \leq a^*_{2p} \leq 1$.

7.4.3 Asymptotic Stability Analysis

Section 7.4.2 has clearly demonstrated that given the parameter set in Table 7.2 there is only one biologically feasible steady-state. It is now useful to consider the asymptotic stability of this particular steady-state. Proceeding as in Section 3.2.4 we obtain the Jacobian matrix, as given by

$$\boldsymbol{J} = \begin{pmatrix} \frac{\partial g_1}{\partial m_P} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial g_1}{\partial b_{1p}} & 0 \\ 0 & \frac{\partial g_2}{\partial m_C} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\partial g_2}{\partial b_{2p}} \\ \frac{\partial g_3}{\partial m_P} & 0 & \frac{\partial g_3}{\partial a_{2p}} & 0 & \frac{\partial g_3}{\partial y_{3p}} & \frac{\partial g_3}{\partial y_{4p}} & \frac{\partial g_3}{\partial y_{6p}} & \frac{\partial g_3}{\partial b_{1p}} & \frac{\partial g_3}{\partial b_{2p}} \\ 0 & \frac{\partial g_4}{\partial m_C} & 0 & \frac{\partial g_4}{\partial a_{3p}} & 0 & 0 & \frac{\partial g_4}{\partial y_{6p}} & 0 & \frac{\partial g_4}{\partial b_{2p}} \\ 0 & 0 & \frac{\partial g_5}{\partial a_{2p}} & 0 & \frac{\partial g_5}{\partial y_{3p}} & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial g_6}{\partial a_{2p}} & 0 & 0 & \frac{\partial g_6}{\partial y_{4p}} & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial g_7}{\partial a_{2p}} & \frac{\partial g_7}{\partial a_{3p}} & 0 & 0 & \frac{\partial g_7}{\partial y_{6p}} & \frac{\partial g_7}{\partial b_{1p}} & \frac{\partial g_7}{\partial b_{2p}} \\ 0 & 0 & \frac{\partial g_8}{\partial a_{2p}} & 0 & 0 & 0 & 0 & \frac{\partial g_8}{\partial b_{1p}} & 0 \\ 0 & 0 & \frac{\partial g_9}{\partial a_{2p}} & \frac{\partial g_9}{\partial a_{3p}} & 0 & 0 & 0 & 0 & \frac{\partial g_9}{\partial b_{2p}} \end{pmatrix}. \tag{7.51}$$

It is clear upon inspection that this Jacobian matrix is much more complex than that obtained when studying the *E. coli* model system (equation (3.36)) and as such there are significantly more partial derivatives to be evaluated. In this particular case, these partial derivatives are of the form

$$\frac{\partial g_1}{\partial m_P} = -\frac{N_P e^{F_P} \Phi_P^2}{2} \left(\gamma_{R2} + \gamma_{B1} b_{1p}^2\right), \tag{7.52}$$

$$\frac{\partial g_1}{\partial b_{1p}} = -2\gamma_{B1} b_{1p} \Phi_P, \tag{7.53}$$

$$\frac{\partial g_2}{\partial m_C} = -\frac{N_C e^{F_C} \Phi_C^2}{2} \left(\gamma_{R3} + \gamma_{B2} b_{2p}^2\right), \tag{7.54}$$

$$\frac{\partial g_2}{\partial b_{2p}} = -2\gamma_{B2} b_{2p} \Phi_C, \tag{7.55}$$

$$\frac{\partial g_3}{\partial m_P} = \frac{1}{2}(1 - a_{2p}) N_P e^{F_P} \Phi_P^2, \tag{7.56}$$

$$\begin{aligned} \frac{\partial g_3}{\partial a_{2p}} &= -\Phi_P - \bar{k}_3(1 - y_{3p}) - \bar{k}_{-3} y_{3p} - \bar{k}_4(1 - y_{4p}) - \bar{k}_{-4} y_{4p} \\ &- \bar{k}_5(1 - y_{6p}) - \bar{k}_6(1 - b_{1p}) - \bar{k}_{-6} b_{1p} - \bar{k}_7(1 - b_{2p}) - \bar{k}_{-7} b_{2p}, \end{aligned} \tag{7.57}$$

$$\frac{\partial g_3}{\partial y_{3p}} = \bar{k}_3 a_{2p} + \bar{k}_{-3}(1 - a_{2p}), \tag{7.58}$$

$$\frac{\partial g_3}{\partial y_{4p}} = \bar{k}_4 a_{2p} + \bar{k}_{-4}(1 - a_{2p}), \tag{7.59}$$

$$\frac{\partial g_3}{\partial y_{6p}} = \bar{k}_5 a_{2p}, \tag{7.60}$$

$$\frac{\partial g_3}{\partial b_{1p}} = \bar{k}_6 a_{2p} + \bar{k}_{-6}(1 - a_{2p}), \tag{7.61}$$

$$\frac{\partial g_3}{\partial b_{2p}} = \bar{k}_7 a_{2p} + \bar{k}_{-7}(1 - a_{2p}), \tag{7.62}$$

$$\frac{\partial g_4}{\partial m_C} = \bar{k}_2 \frac{(1 - a_{3p}) N_C e^{F_C} \Phi_C^2}{2}, \tag{7.63}$$

$$\frac{\partial g_4}{\partial a_{3p}} = -\bar{k}_2\Phi_C - \bar{k}_8(1-y_{6p}) - \bar{k}_{-8}y_{6p} - \bar{k}_9(1-b_{2p}) - \bar{k}_{-9}b_{2p}, \tag{7.64}$$

$$\frac{\partial g_4}{\partial y_{6p}} = \bar{k}_8a_{3p} + \bar{k}_{-8}(1-a_{3p}), \tag{7.65}$$

$$\frac{\partial g_4}{\partial b_{2p}} = \bar{k}_9a_{3p} + \bar{k}_{-9}(1-a_{3p}), \tag{7.66}$$

$$\frac{\partial g_5}{\partial a_{2p}} = \alpha_1\bar{k}_3(1-y_{3p}) + \alpha_1\bar{k}_{-3}y_{3p}, \tag{7.67}$$

$$\frac{\partial g_5}{\partial y_{3p}} = -\alpha_1\bar{k}_3a_{2p} + \alpha_1\bar{k}_{-3}(1-a_{2p}) - \bar{k}_{10}, \tag{7.68}$$

$$\frac{\partial g_6}{\partial a_{2p}} = \alpha_2\bar{k}_4(1-y_{4p}) + \alpha_2\bar{k}_{-4}y_{4p}, \tag{7.69}$$

$$\frac{\partial g_6}{\partial y_{4p}} = -\alpha_2\bar{k}_4a_{2p} + \alpha_2\bar{k}_{-4}(1-a_{2p}) - \bar{k}_{11}, \tag{7.70}$$

$$\frac{\partial g_7}{\partial a_{2p}} = \alpha_3\bar{k}_5(1-y_{6p}), \tag{7.71}$$

$$\frac{\partial g_7}{\partial a_{3p}} = \alpha_6\bar{k}_8(1-y_{6p}) + \alpha_6\bar{k}_{-8}(1-y_{6p}) + \bar{k}_{15}y_{6p} - \bar{k}_{16}y_{6p}, \tag{7.72}$$

$$\begin{aligned}\frac{\partial g_7}{\partial y_{6p}} &= -\alpha_3\bar{k}_5a_{2p} - \alpha_6\bar{k}_8a_{3p} - \alpha_6\bar{k}_{-8}(1-a_{3p}) - \bar{k}_{12} - \bar{k}_{15}(1-a_{3p}) \\ &- \bar{k}_{16}a_{3p},\end{aligned} \tag{7.73}$$

$$\frac{\partial g_8}{\partial a_{2p}} = \alpha_4\bar{k}_6(1-b_{1p}) + \alpha_4\bar{k}_{-6}b_{1p}, \tag{7.74}$$

$$\frac{\partial g_8}{\partial b_{1p}} = -\alpha_4\bar{k}_6a_{2p} - \alpha_4\bar{k}_{-6}(1-a_{2p}) - \bar{k}_{13}, \tag{7.75}$$

$$\frac{\partial g_9}{\partial a_{2p}} = \alpha_5\bar{k}_7(1-b_{2p}) + \alpha_5\bar{k}_{-7}b_{2p}, \tag{7.76}$$

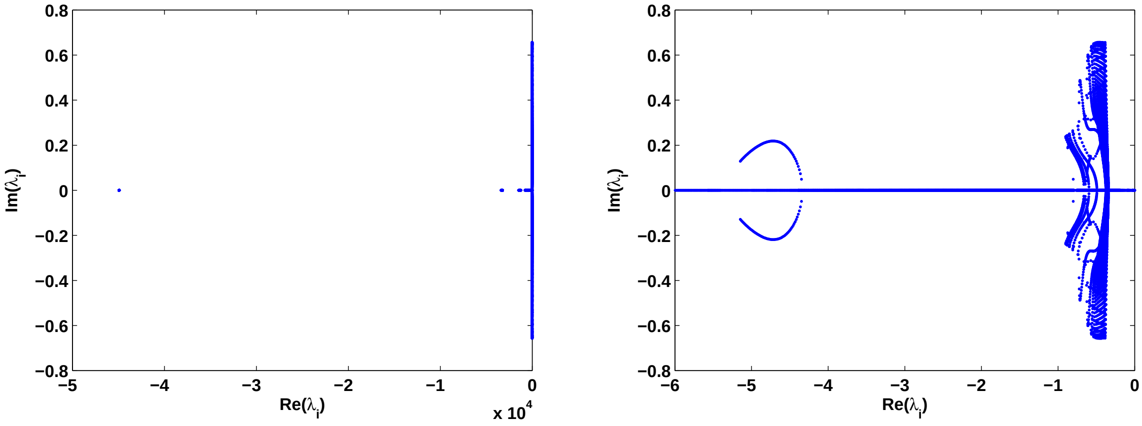
$$\frac{\partial g_9}{\partial a_{3p}} = \alpha_7\bar{k}_9(1-b_{2p}) + \alpha_7\bar{k}_{-9}b_{2p}, \tag{7.77}$$

$$\frac{\partial g_9}{\partial b_{2p}} = -\alpha_5\bar{k}_7a_{2p} - \alpha_5\bar{k}_{-7}(1-a_{2p}) - \alpha_7\bar{k}_9a_{3p} - \alpha_7\bar{k}_{-9}(1-a_{3p}) - \bar{k}_{14}. \tag{7.78}$$

Each of the partial derivatives in the Jacobian matrix that are not included in the above list are equal to zero. In order to investigate the asymptotic stability of the system equilibrium found in Section 7.4.2 we must examine the eigenvalues of the Jacobian matrix evaluated at this point. This is done by solving

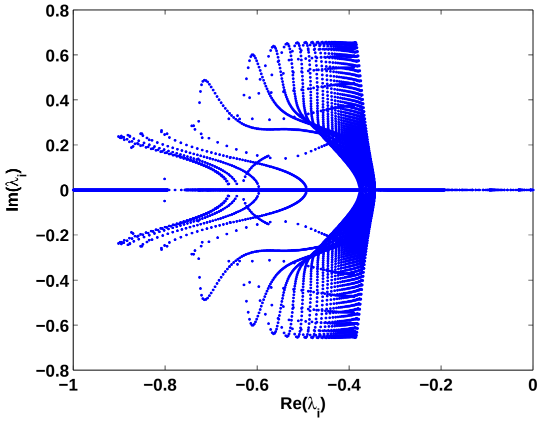
$$\det|\boldsymbol{J} - \lambda\boldsymbol{I}| = 0, \tag{7.79}$$

within which $\boldsymbol{I}$ denotes the identity matrix with the same dimensions as the Jacobian matrix. Examining the system equilibrium stated in Section 7.4.2 it is clear that the steady-state values of CheA$_2$-P and CheA$_3$-P depend upon one another and also that all other protein concentration steady-states also depend upon the CheA$_2$-P and CheA$_3$-P steady-states. We therefore utilise the parameter values listed in Table 7.2 and examine the eigenvalues obtained from evaluating equation (7.79) for values over the ranges $0 \leq a_{2p}^* \leq 1$ and $0 \leq a_{3p}^* \leq 1$. Upon doing so we obtain a set of nine eigenvalues for each combination of a_{2p}^* and a_{3p}^* as shown in Figure 7.4.



(a) All Values

(b) Partial Zoom



(c) Full Zoom

Figure 7.4: Plots showing the real and imaginary parts of eigenvalues (λ_i) of the model in equations (7.14)-(7.22), evaluated at steady-state values in the ranges $0 \leq a_{2p}^* \leq 1$ and $0 \leq a_{3p}^* \leq 1$. Plot (a) shows the location of all eigenvalues whilst (b) focuses about the origin and (c) focuses yet further. All eigenvalues have $\text{Re}(\lambda_i) < 0$ and so the system is stable.

It is clear to see from the results in Figure 7.4 that each of the eigenvalues is such that $\text{Re}(\lambda_i) < 0$ and as such we may conclude that given the parameter set in Table 7.2 this system is asymptotically stable for all steady-states of the system with a_{2p}^* and a_{3p}^* in the ranges $0 \leq a_{2p}^* \leq 1$ and $0 \leq a_{3p}^* \leq 1$.

It is also clear that pairs of complex eigenvalues are common within this system. Examination of results displayed in Figure 7.3 clearly shows that in the range $0 \leq a_{3p}^* \leq 1$, it is not possible to obtain all values of a_{2p}^* in the range $0 \leq a_{2p}^* \leq 1$. Thus it is likely that if eigenvalues obtained from values within the ranges $0 \leq a_{2p}^* \leq 1$ and $0 \leq a_{3p}^* \leq 1$ that cannot feasibly be obtained are removed then it is likely that we will always have at least one pair of complex eigenvalues and as such damped oscillatory behaviour is likely a fixed feature of this system under realistic conditions.

7.4.4 Model Stiffness

As discussed in Section 3.2.5, the stiffness coefficient of a dynamical system can be determined from the magnitudes of the extreme eigenvalues of the system. In this case, the stiffness coefficient may be calculated as

$$s = \frac{\left|\mathrm{Re}(\bar{\lambda})\right|}{\left|\mathrm{Re}(\underline{\lambda})\right|} = \frac{\left|\mathrm{Re}(\lambda_1)\right|}{\left|\mathrm{Re}(\lambda_7)\right|}, \tag{7.80}$$

within which $\bar{\lambda} = \lambda_1$ and $\underline{\lambda} = \lambda_7$ represent the largest and smallest magnitude eigenvalues of the system, respectively. In this particular case we may utilise the asymptotic stability analysis of Section 7.4.3 in order to obtain eigenvalues of this system. This gives a stiffness coefficient of $s = 31,482.06$ which is clearly large. In actual fact, the use of an ODE solver for stiff systems (such as the previously mentioned ode15s solver in MATLAB) is definitely required for the numerical simulation of this system since this stiffness coefficient is $\sim$7.5 times larger than that of the *E. coli* system that was obtained in Section 3.2.5 which was also deemed to be a stiff system.

7.5 Deletion Mutants

Within this section we seek to test the validity of the Kojadinovic et al. [161] model by examining its ability to reproduce experimental results. In particular, we wish to study whether this model is sufficient to explain the behaviour of *R. sphaeroides* mutant strains studied within the recent experimental work of de Beyer [162]. Within this experimental work a range of different mutant strains were studied (see Table 7.1 for definitions).

The results of de Beyer [162] demonstrate that a number of mutations result in significantly different swimming behaviour at steady-state (i.e. the time spent in runs and stops differs). In order to assess the ability of the Kojadinovic et al. [161] model to accurately represent this data we consider a number of alterations. Each of these altered models represents a single mutant cell type. We can then obtain steady-state concentrations of CheY$_3$-P, CheY$_4$-P and CheY$_6$-P, thus allowing us to investigate whether these concentrations could produce the differences in swimming behaviour observed experimentally.

Within this chapter we produce mutant models for a subset of the strains considered in the work of de Beyer [162]. In order to capture the effects of these mutations within our altered

Table 7.4: Alterations to kinetic rates required for the creation of each *R. sphaeroides* chemotaxis mutant model used in this chapter.

Strain	Parameters set to zero	Other
#1	No alteration	-
#2	$k_7,\ k_{-7},\ k_9,\ k_{-9},\ k_{14},\ g_{B2}$	$b_{2p} = 0$
#3	g_{R2}	-
#6	$k_{-3},\ k_{-4},\ k_{-6},\ k_{-7},\ \Phi_P^*$	-
#9	$k_3,\ k_{-3},\ k_{10}$	$y_{3p} = 0$
#10	$k_4,\ k_{-4},\ k_{11}$	$y_{4p} = 0$
#14	$k_3,\ k_{-3},\ k_4,\ k_{-4},\ k_{10},\ k_{11}$	-
#15	$k_6,\ k_{-6},\ k_{13},\ g_{B1}$	$b_{1p} = 0$
#18	g_{R3}	-
#20	$k_{-8},\ k_{-9},\ \Phi_C^*$	-

models there are a number of options available. For example, there are two main methods that could be used to represent the deletion of protein CheY$_3$ (i.e. a ΔCheY$_3$ mutant cell). Firstly, we may set the relevant total protein concentration ($[Y_3]_T$) parameter to zero. Alternatively, by setting each kinetic rate associated with the deleted protein to zero we may remove the effects of that protein from the system. Here we consider only changes to kinetic rate parameters since setting total protein concentrations to zero can cause divisions by zero as a result of the non-dimensionalisation.

It is clear upon examination of Table 7.1 that not all mutations relate to the deletion of a protein. However, in most cases they may continue to be modelled as such. Take for example the CheY$_{3,4}$(D57A) mutant strain in which CheY$_3$ and CheY$_4$ cannot be phosphorylated. It is clear that these proteins will not directly influence chemotactic behaviour. If we also assume that there are no indirect effects (eg. via intracellular crowding, etc.) then the CheY$_{3,4}$(D57A) mutant can clearly be modelled in the same manner as a ΔCheY$_{3,4}$ deletion mutant (see Figure 7.2). Other mutant cell types may be considered in the same manner, for example, CheA$_2$(H48Q, G470K) and CheA$_3$(H48Q) mutants are considered here to be equivalent to ΔCheA$_2$ (CheA$_2$ deletion) and ΔCheA$_3$ (CheA$_3$ deletion) mutants, respectively.

A complete summary of all kinetic rate parameter alterations used to represent mutant strains in this work are listed in Table 7.4. It is now possible to obtain the steady-state values for these 'mutant models' following the procedure laid out in Section 7.4.2. It is worth noting however that in the case of deletion mutants, the denominators of the steady-state expressions will be altered since a number of the rate parameters will now equal zero. Upon carrying out this investigation we obtain the steady-state values shown in Table 7.5.

Examination of the experimental data shows that the phosphorylation levels of certain proteins can affect the swimming behaviour of *R. sphaeroides* cells. In particular, CheY$_6$(D57A) and CheY$_6$(D57N) mutants will have zero and full phosphorylation, respectively whilst a wild-

Table 7.5: Steady-state values for CheY$_3$-P, CheY$_4$-P and CheY$_6$-P in a number of *R. sphaeroides* chemotaxis protein deletion mutant models. Shown in this table are non-dimensional steady-state values as well as the equivalent dimensional values.

Strain	y^*_{3p}	$[Y^*_{3p}]$ (μM)	y^*_{4p}	$[Y^*_{4p}]$ (μM)	y^*_{6p}	$[Y^*_{6p}]$ (μM)
Wild-Type	0.0785	0.27	0.3783	5.22	0.2452	55.17
ΔCheA$_2$	0	0	0	0	0.3039	68.38
ΔCheA$_3$	0.07601	0.27	0.3703	5.11	0.1038	23.36
ΔCheB$_1$	0.1605	0.56	0.5798	8.00	0.1853	41.69
ΔCheB$_2$	0.1069	0.37	0.4620	6.38	0.6263	140.92
ΔCheR$_2$	0.004629	0.016	0.03196	0.44	0.3440	77.40
ΔCheR$_3$	0.0702	0.25	0.3504	4.84	0.0297	6.68
ΔCheY$_3$	0	0	0.3787	5.23	0.2451	55.15
ΔCheY$_4$	0.07997	0.28	0	0	0.2436	54.81
ΔCheY$_{3,4}$	0	0	0	0	0.2434	54.77

type cell has an intermediate phosphorylation level. Experimental results for these three cell types shows that greater CheY$_6$-P concentrations relate to a greater fraction of time spent in a stopped state by a cell's flagellar motor. As such, here we compare the model prediction of CheY$_6$-P concentration with the experimentally observed fraction of time spent stopped. In doing so, we would expect a sufficiently accurate model to display an increase in the fraction of time spent in a stopped state as the CheY$_6$-P concentration increases.

Upon investigating the results displayed in Table 7.5 (and visualised in Figure 7.5) it can be seen that this model does not adequately explain the experimental data since it does not predict suitable CheY$_6$-P concentrations for each mutant cell type. In particular, the results of Figure 7.5 show that most mutant cell types do not vary too far from the expected relationship. However, the ΔCheB$_2$ (CheB$_2$ deletion) mutant displays a predicted CheY$_6$-P concentration that is clearly far too high. As such, we must seek to understand why the behaviour observed in Figure 7.5 occurs. Firstly, it may be possible that CheY$_6$-P concentration is not exclusively responsible for setting the rotational behaviour of the flagellar motor. Alternatively, the mathematical model analysed here may not contain an adequate description of signalling pathway mechanisms.

In addition to the results displayed in Table 7.5 and Figure 7.5 a number of alternative functions were considered to control the flagellar rotation. Specifically we considered the concentrations of CheY$_3$-P, CheY$_4$-P, the sum CheY$_3$-P + CheY$_4$-P + CheY$_6$-P, the combination

$$\text{Bias} = [Y_{6p}]([Y_{3p}] + [Y_{4p}]), \tag{7.81}$$

and the rotational frequency function described by Hamadeh et al. [173] which is given by

$$f = -\frac{1}{0.125 + \phi\left([Y_{3p}], [Y_{4p}], [Y_{6p}]\right)^4}, \tag{7.82}$$

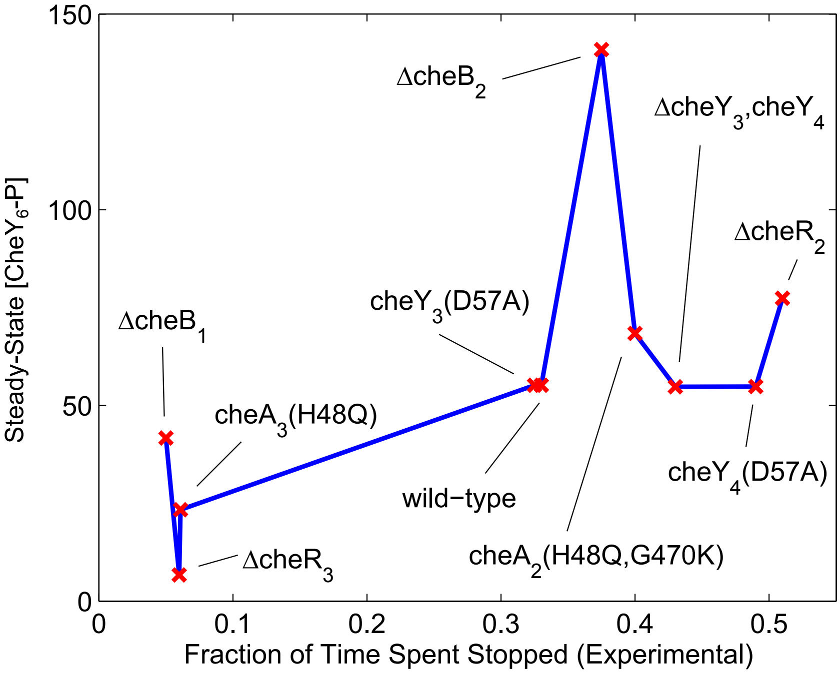


Figure 7.5: The model of Kojadinovic et al. [161] is not sufficient to represent the experimental data of de Beyer [162]. The horizontal axis displays the fraction of time that the flagellar motor spends in a stopped state for a number of chemotaxis mutant cells (labelled in the figure). The vertical axis displays the steady-state CheY$_6$-P concentration predicted by the mathematical model presented by Kojadinovic et al. [161].

within which

$$\phi\left([Y_{3p}],[Y_{4p}],[Y_{6p}]\right) = q[Y_{6p}]\frac{[Y_{3p}]+[Y_{4p}]}{0.1+[Y_{3p}]+[Y_{4p}]}. \tag{7.83}$$

Each of these different cases displayed a similar failure to represent the behaviour of the *R. sphaeroides* chemotaxis mutants. This is likely to suggest that the model considered here does not include an adequate description of some key feature(s) in the signalling pathway.

Analysis of Figure 7.5 clearly shows that the most severe failure of the model to display the expected behaviour occurs when the proteins involved with adaptation (i.e. the CheBs and CheRs) are deleted. This suggests that the mechanisms used here are not a suitable description of *R. sphaeroides* adaptation. CheB$_2$ and CheR$_3$ display the greatest deviation from the expected behaviour and are associated with adaptation in the cytoplasmic cluster. With this being the case, in Chapter 8 mathematical modelling approaches shall be used in order to analyse features of the chemotaxis signalling pathways that may be used in order to create improved models of *R. sphaeroides* chemotaxis signalling in future studies. In particular, this work focuses on investigating simplified models of *R. sphaeroides* chemotaxis signalling to draw conclusions relating to the role of signalling proteins as well as the relationship between the signalling pathway and rotational behaviour of flagellar motors.

7.6 Summary & Discussion

This chapter began by analysing a recent model of intracellular signalling in *R. sphaeroides* cells from the literature. In particular, we considered a non-dimensional re-scaling of this model for which it was found that there is just one biologically feasible steady-state. Following the example of Chapter 3 it was shown that this model equilibrium state is asymptotically stable and the associated eigenvalues demonstrate that this is a stiff model system.

The same mathematical model was then considered in the context of recent experimental data obtained for a number of mutant strains. Utilising a number of kinetic rate alterations it was possible to formulate models of mutant strains. Upon obtaining the steady-states of each mutant model it was possible to show that the recent Kojadinovic et al. [161] model does not represent an adequate description of the chemotaxis signalling pathways in *R. sphaeroides* cells.

Upon closer inspection it can in fact be seen that the main failings of this model are associated with CheB and CheR proteins, i.e. those associated with adaptation. We therefore concluded that the adaptation mechanism borrowed from *E. coli* models is not suitable for use here and thus further work is required in order to identify the true adaptation mechanisms of this species. In particular, it was noted that whilst the adaptation mechanisms associated with both receptor clusters appear inadequate, it was the deletion of cytoplasmic cluster adaptation proteins that produced the much more serious failing. Thus within the next chapter we seek to provide a better understanding of intracellular signalling processes occurring at the cytoplasmic cluster of *R. sphaeroides* cells. In addition to this a number of outstanding questions shall be addressed, providing further information that will be useful in both future theoretical and experimental studies.

Chapter 8

Understanding *R. sphaeroides* using Simplified Modelling Approaches

Within this chapter we further investigate mechanisms associated with chemotaxis in *R. sphaeroides*. In particular, we formulate and analyse two simplified mathematical models in order to elucidate the workings of various system components. Firstly, we formulate a simple model of the cytoplasmic cluster in order to understand the roles of processes acting there. We then consider a simplified Kojadinovic et al. [161] model, in which adaptation is removed. This is first used to verify that the correct intracellular reactions are considered regardless of whether or not adaptation mechanisms are included. In addition to this, the creation of various models based on mutations considered within the experimental data allows us to further our understanding of the cell's flagellar response. This chapter concludes with the proposal of a new *R. sphaeroides* chemotaxis model formulated according to results obtained both here and in Chapter 7.

8.1 Motivation

In Section 7.5 it was demonstrated that the main failing of the Kojadinovic et al. [161] was its failure to represent the behaviour of ΔCheB$_2$ (CheB$_2$ deletion) mutant cells. Thus, within this chapter we seek to formulate a new mathematical model that is able to provide better agreement with experimental data [162] (see Figure 7.2).

This experimental data allows us to examine a number of key features. In particular, within this chapter we seek to understand each process acting at the cytoplasmic cluster; verify the set of phosphorylation/dephosphorylation reactions considered; and investigate the link between protein phosphorylation levels and flagellar rotational behaviour. Each of these investigations is carried out in subsequent sections of this chapter, followed by the proposal of a new mathematical

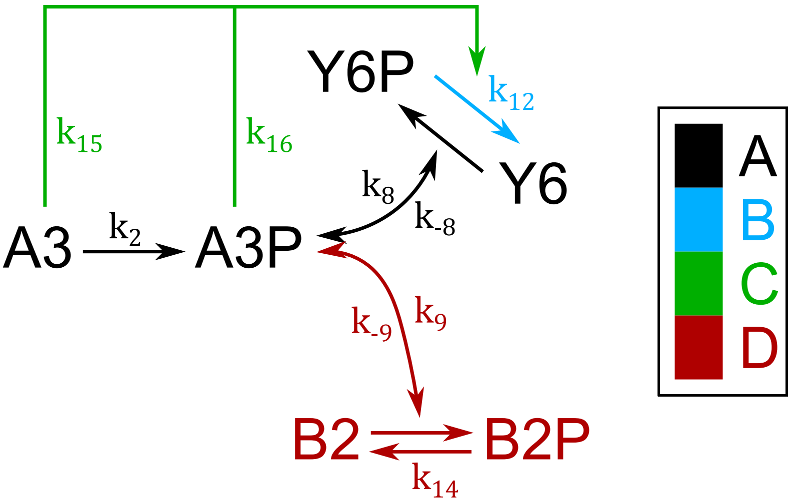


Figure 8.1: A schematic diagram of the simplified *R. sphaeroides* cytoplasmic cluster model. Here k_i $(i = 2, 8, -8, 9, -9, 12, 13, 15, 16)$ denote the kinetic rate constants of each reaction while arrows indicate the direction of phosphoryl groups passing around the system.

model that better represents the experimental data.

8.2 Simplified Mathematical Models

Within this chapter we outline the formulation of two simplified mathematical models of chemotaxis in *R. sphaeroides*. In particular, Section 8.2.1 considers a reduced model of cytoplasmic cluster processes while Section 8.2.2 lays out a simplified Kojadinovic et al. [161] model from which adaptation has been removed. Analysis of these models and results obtained are then discussed within later sections of this chapter.

8.2.1 Simplified Model 1: The Cytoplasmic Cluster

Within Section 7.5 it was demonstrated that the main failing of the Kojadinovic et al. [161] model was its failure to accurately capture the behaviour of ΔCheB$_2$ (CheB$_2$ deletion) mutant cells. It was thus suggested that the description of cytoplasmic cluster adaptation was not adequate. As such, here we consider a reduced model that contains only reactions occurring at the cytoplasmic cluster. This is then used to answer a number of outstanding questions on the workings of the cytoplasmic cluster.

In order to formulate this reduced model we consider only the set of reactions occurring at the cytoplasmic receptor cluster (shown in Figure 8.1), namely

- CheA$_3$ phosphorylation by CheA$_4$ (simplified here to CheA$_3$ autophosphorylation);
- phosphotransfer from CheA$_3$-P onto CheY$_6$ and CheB$_2$;
- reverse phosphotransfer from CheY$_6$-P and CheB$_2$-P onto CheA$_3$;

- autodephosphorylation of CheY$_6$-P and CheB$_2$-P; and
- CheA$_3$ and CheA$_3$-P phosphatase action on CheY$_6$-P.

Applying the law of mass action (see Appendix A) in addition to conservation laws of the form $[X]_T = [X] + [X_p]$ for each protein, we obtain three non-linear ODEs of the form

$$\begin{aligned}
\frac{d[A_{3p}]}{dt} &= k_2\left([A_{3T}] - [A_{3p}]\right) - k_8[A_{3p}]\left([Y_{6T}] - [Y_{6p}]\right) + k_{-8}\left([A_{3T}] - [A_{3p}]\right)[Y_{6p}] \qquad (8.1)\\
&- k_9[A_{3p}]\left([B_{2T}] - [B_{2p}]\right) + k_{-9}\left([A_{3T}] - [A_{3p}]\right)[B_{2p}],\\
\frac{d[Y_{6p}]}{dt} &= k_8[A_{3p}]\left([Y_{6T}] - [Y_{6p}]\right) - k_{-8}\left([A_{3T}] - [A_{3p}]\right)[Y_{6p}] - k_{12}[Y_{6p}] \qquad (8.2)\\
&- k_{15}\left([A_{3T}] - [A_{3p}]\right)[Y_{6p}] - k_{16}[A_{3p}][Y_{6p}],\\
\frac{d[B_{2p}]}{dt} &= k_9[A_{3p}]\left([B_{2T}] - [B_{2p}]\right) - k_{-9}\left([A_{3T}] - [A_{3p}]\right)[B_{2p}] - k_{14}[B_{2p}], \qquad (8.3)
\end{aligned}$$

within which $[X]$ (with $X = A_3, B_2, Y_6$) represents the concentration of the relevant protein with subscripts T and p denoting the total and phosphorylated concentrations, respectively and k_i ($i = 2, 8, -8, 9, -9, 12, 14, 15, 16$) indicating the kinetic rates of each process (see Table 7.2 for values). Here we have chosen not to assign new subscripts to each kinetic rate within this model so as to allow direct comparison with the model in Chapter 7.

Non-Dimensionalisation

Equations (8.1)-(8.3) are non-dimensionalised such that $[X_p] = [X]_T x_p$, where x_p is the associated non-dimensional protein concentration, $X \in [A_3, Y_6, B_2]$ and $x \in [a_3, y_6, b_2]$. We also re-scale time with respect to the timescale of CheA$_3$ autophosphorylation such that $t = \tau/k_2$. This yields

$$\begin{aligned}
\frac{da_{3p}}{d\tau} &= \underbrace{(1 - a_{3p})}_{\text{A}} - \underbrace{\tilde{k}_8 a_{3p}(1 - y_{6p})}_{\text{A}} + \underbrace{\tilde{k}_{-8}(1 - a_{3p})\,y_{6p}}_{\text{A}} \qquad (8.4)\\
&- \underbrace{\tilde{k}_9 a_{3p}(1 - b_{2p})}_{\text{D}} + \underbrace{\tilde{k}_{-9}(1 - a_{3p})\,b_{2p}}_{\text{D}},\\
\frac{dy_{6p}}{d\tau} &= \underbrace{\alpha_6\tilde{k}_8 a_{3p}(1 - y_{6p})}_{\text{A}} - \underbrace{\alpha_6\tilde{k}_{-8}(1 - a_{3p})\,y_{6p}}_{\text{A}} - \underbrace{\tilde{k}_{12}y_{6p}}_{\text{B}} \qquad (8.5)\\
&- \underbrace{\tilde{k}_{15}(1 - a_{3p})\,y_{6p}}_{\text{C}} - \underbrace{\tilde{k}_{16}a_{3p}y_{6p}}_{\text{C}},\\
\frac{db_{2p}}{d\tau} &= \underbrace{\alpha_7\tilde{k}_9 a_{3p}(1 - b_{2p})}_{\text{D}} - \underbrace{\alpha_7\tilde{k}_{-9}(1 - a_{3p})\,b_{2p}}_{\text{D}} - \underbrace{\tilde{k}_{14}b_{2p}}_{\text{D}}. \qquad (8.6)
\end{aligned}$$

Here, $\tilde{k}_i$ $(i = 8, -8, 9, -9, 12, 14, 15, 16)$ are non-dimensional kinetic rate parameters and α_j $(j = 6, 7)$ are protein concentration ratios. Values for each non-dimensional parameter are listed in Table 8.1.

Table 8.1: A base set of non-dimensional *R. sphaeroides* parameter values calculated using values in Table 7.2.

Parameter	Value
$\tilde{k}_2 = k_2/k_2$	1
$\tilde{k}_8 = k_8[Y_6]_T/k_2$	177.934
$\tilde{k}_{-8} = k_{-8}[Y_6]_T/k_2$	0.650
$\tilde{k}_9 = k_9[B_2]_T/k_2$	1.304
$\tilde{k}_{-9} = k_{-9}[B_2]_T/k_2$	0.0657
$\tilde{k}_{12} = k_{12}/k_2$	0.0163
$\tilde{k}_{14} = k_{14}/k_2$	0.0103
$\tilde{k}_{15} = k_{15}[A_3]_T/k_2$	0.477
$\tilde{k}_{16} = k_{15}[A_3]_T/k_2$	0.477
$\alpha_6 = [A_3]_T/[Y_6]_T$	0.400
$\alpha_7 = [A_3]_T/[B_2]_T$	4.328

To assist in understanding processes acting at the cytoplasmic cluster, terms associated with specific mechanisms have been split into four groups denoted A, B, C and D, as shown in equations (8.4)-(8.6). This model is analysed further in Section 8.3.1 of this chapter.

8.2.2 Simplified Model 2: No Adaptation at Either Cluster

Section 7.5 clearly showed that models of mutant cells in which CheB or CheR homologues (i.e. those associated with adaptation) were removed produced the greatest deviation from expected behaviour in terms of the CheY$_6$-P concentrations they predict. Within Section 8.2.1 a reduced model of the *R. sphaeroides* cytoplasmic cluster was outlined. This allows us to study the effects of cytoplasmic cluster processes. However, it is known that *R. sphaeroides* cells display two chemosensory pathways. Thus, in order to fully understand this system it is necessary to consider both of these pathways. As such, here we consider a simplified version of the Kojadinovic et al. [161] model, in which adaptation has been removed.

In order to formulate such a model we assume that each receptor cluster produces a constant level of kinase activity. This is achieved by considering the substitution of polar (Φ_P) and cytoplasmic (Φ_C) cluster kinase activities for their respective steady-state values Φ_P^* and Φ_C^*. In order to determine values for Φ_P^* and Φ_C it is necessary to utilise the steady-states given in equations (7.37) and (7.38), which yield values as displayed in Table 8.2. Consideration of such a simplification allows for the removal of all reactions relating to methylation and demethylation of chemoreceptors (i.e. $dm_P/dt = 0 = dm_C/dt$) as well as the MWC description of the receptor signalling teams. This yields the non-linear ODEs

$$\begin{aligned}
\frac{d[A_{2p}]}{dt} &= k_1\Phi_P^*\left([A_2]_T-[A_{2p}]\right)-k_3[A_{2p}]\left([Y_3]_T-[Y_{3p}]\right)+k_{-3}\left([A_2]_T-[A_{2p}]\right)[Y_{3p}] \quad (8.7)\\
&- k_4[A_{2p}]\left([Y_4]_T-[Y_{4p}]\right)+k_{-4}\left([A_2]_T-[A_{2p}]\right)[Y_{4p}]-k_5[A_{2p}]\left([Y_6]_T-[Y_{6p}]\right)\\
&- k_6[A_{2p}]\left([B_1]_T-[B_{1p}]\right)+k_{-6}\left([A_2]_T-[A_{2p}]\right)[B_{1p}]-k_7[A_{2p}]\left([B_2]_T-[B_{2p}]\right)\\
&+ k_{-7}\left([A_2]_T-[A_{2p}]\right)[B_{2p}],\\
\frac{d[A_{3p}]}{dt} &= k_2\Phi_C^*\left([A_3]_T-[A_{3p}]\right)-k_8[A_{3p}]\left([Y_6]_T-[Y_{6p}]\right)+k_{-8}\left([A_3]_T-[A_{3p}]\right)[Y_{6p}] \quad (8.8)\\
&- k_9[A_{3p}]\left([B_2]_T-[B_{2p}]\right)+k_{-9}\left([A_3]_T-[A_{3p}]\right)[B_{2p}],\\
\frac{d[Y_{3p}]}{dt} &= k_3[A_{2p}]\left([Y_3]_T-[Y_{3p}]\right)-k_{-3}[Y_{3p}]\left([A_2]_T-[A_{2p}]\right)-k_{10}[Y_{3p}], \quad (8.9)\\
\frac{d[Y_{4p}]}{dt} &= k_4[A_{2p}]\left([Y_4]_T-[Y_{4p}]\right)-k_{-4}[Y_{4p}]\left([A_2]_T-[A_{2p}]\right)-k_{11}[Y_{4p}], \quad (8.10)\\
\frac{d[Y_{6p}]}{dt} &= k_5[A_{2p}]\left([Y_6]_T-[Y_{6p}]\right)+k_8[A_{3p}]\left([Y_6]_T-[Y_{6p}]\right)-k_{-8}\left([A_3]_T-[A_{3p}]\right)[Y_{6p}] \quad (8.11)\\
&- k_{12}[Y_{6p}]-k_{15}\left([A_3]_T-[A_{3p}]\right)[Y_{6p}]-k_{16}[A_{3p}][Y_{6p}],\\
\frac{d[B_{1p}]}{dt} &= k_6[A_{2p}]\left([B_1]_T-[B_{1p}]\right)-k_{-6}[B_{1p}]\left([A_2]_T-[A_{2p}]\right)-k_{13}[B_{1p}], \quad (8.12)\\
\frac{d[B_{2p}]}{dt} &= k_7[A_{2p}]\left([B_2]_T-[B_{2p}]\right)-k_{-7}\left([A_2]_T-[A_{2p}]\right)[B_{2p}] \quad (8.13)\\
&+ k_9[A_{3p}]\left([B_2]_T-[B_{2p}]\right)-k_{-9}\left([A_3]_T-[A_{3p}]\right)[B_{2p}]-k_{14}[B_{2p}],
\end{aligned}$$

within which all parameters retain their earlier values and definitions (see Tables 7.2 and 7.3).

Table 8.2: Steady-state receptor activities obtained from the Kojadinovic et al. model.

Activity	Symbol	Value
Polar	Φ_P^*	0.413
Cytoplasmic	Φ_C^*	0.354

Model Analysis

Non-Dimensionalisation As with all models discussed throughout this thesis, analysis begins by considering the application of a non-dimensional re-scaling. Here we consider a similar non-dimensionalisation to that presented in Section 7.4.1.

Steady-States In Section 8.1 it was discussed that the experimental data of de Beyer [162] relates to the steady-state behaviour of *R. sphaeroides* mutant cell types. As such, here it is important to obtain steady-state expressions for each component of this model. The specific simplification of the Kojadinovic et al. [161] model considered here leaves the steady-states of CheY and CheB proteins unaffected, thus equations (7.32)-(7.36) remain valid. However, the CheA$_2$-P and CheA$_3$-P steady-state polynomials now reduce to

$$0 = A_1 a_{2p}^{*6} + A_2 a_{2p}^{*5} + A_3 a_{2p}^{*4} + A_4 a_{2p}^{*3} + A_5 a_{2p}^{*2} + A_6 a_{2p}^{*1} + A_7, \tag{8.14}$$

$$0 = B_1 a_{3p}^{*3} + B_2 a_{3p}^{*2} + B_3 a_{3p}^{*1} + B_4, \tag{8.15}$$

within which A_{1-7} and B_{1-4} denote the polynomial coefficients composed of the parameters of the original ODE system. These polynomial equations are very similar to those in equations (7.43) and (7.44). The only differences here come from the fact that the kinase activities here are considered to be constant and thus there is one less denominator to clear in each case. As such the result will be lower order steady-state polynomials than those in previous examples.

It has already been stated that the kinase activities utilised here are chosen to be equal to the steady-state values within the Kojadinovic et al. [161] model. As such, the steady-state values obtained from equations (8.14) and (8.15) will yield the same biologically feasible steady-state values found in Section 7.4.2. Thus it may be concluded without further work here that Simplified Model 2 will possess just one biologically feasible steady-state.

Stability Analysis It has been shown here that this simplified model will yield the same biologically feasible steady-state as the full Kojadinovic et al. [161] model. However, in spite of this it is not so clear how the asymptotic stability characteristics of the system analysed here will be affected.

In order to determine how a simplification to a model without adaptation affects the behaviour of the system an asymptotic stability analysis very similar to that in Section 7.4.3 was conducted. This yielded the eigenvalues

$$\lambda_1 = -44924.90604, \quad \lambda_2 = -3327.57927, \quad \lambda_3 = -1026.67477, \quad \lambda_4 = -284.56248,$$

$$\lambda_5 = -4.36955, \quad \lambda_6 = -3.92332, \quad \lambda_7 = -1.42700,$$

for the non-adapting model (Simplified Model 2) which are clearly all negative and real valued. From these eigenvalues it may be concluded that the non-adapting model system will be asymptotically stable. This is interesting since Section 7.4.3 suggested that in the adapting version of this model, damped oscillatory behaviour was likely to be a fixed property. It is therefore the case that this system loses some characteristics of the dynamic response as a result of removing adaptation. This is not surprising given the removal of negative feedbacks from the system.

8.3 Model Analysis and Results

Here we seek to utilise the simplified mathematical models presented in Section 8.2 in order to elucidate a number of features of the *R. sphaeroides* chemotaxis signalling pathways. Firstly, we utilise Simplified Model 1 (Section 8.2.1) to investigate the roles of processes acting at the cytoplasmic cluster. Simplified Model 2 (Section 8.2.2) is then used to validate the phosphorylation/dephosphorylation reactions within the signalling pathways and to help elucidate the link between CheY phosphorylation levels and flagellar rotational behaviour.

8.3.1 Elucidating the Roles of Cytoplasmic Cluster Processes

In this section we seek to utilise Simplified Model 1 (equations (8.4)-(8.6)) in order to help elucidate the roles of cytoplasmic cluster processes. In particular, we wish to answer the following questions.

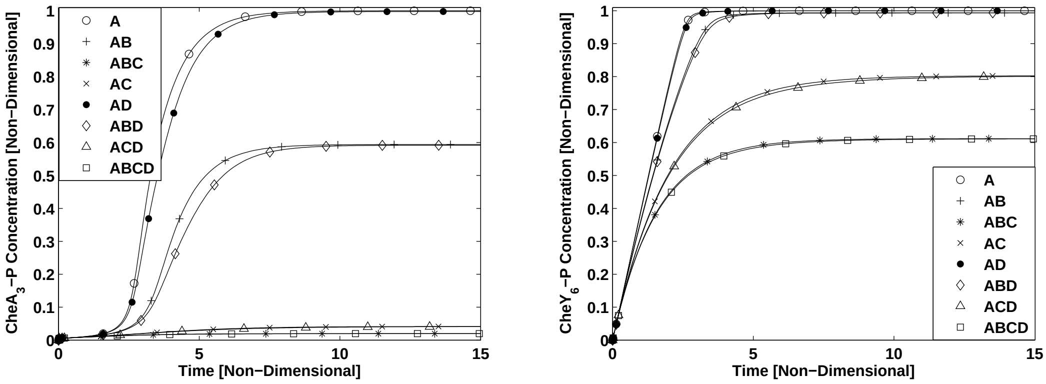
1. Does CheB$_2$ act as a phosphate sink?
2. How do CheA$_3$ and CheA$_3$-P phosphatase activities affect CheY$_6$-P levels?
3. Is adaptation necessary at the cytoplasmic cluster?

Here the role of each set of mechanisms (A-D) is examined independently and in detail to see how they affect CheY$_6$-P levels and thus the cell response. In particular, these terms are grouped as follows.

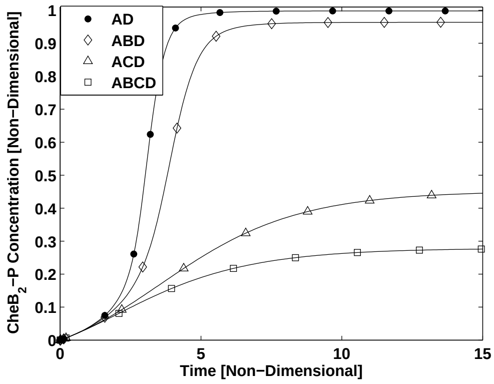
- **Group A:** Autophosphorylation of CheA$_3$ (k_2), phosphotransfer from CheA$_3$-P onto CheY$_6$ (k_8) and its related reverse phosphotransfer (k_{-8}). These terms are essential for the cell to function, thus they are included in each model.
- **Group B:** Autodephosphorylation of CheY$_6$-P (k_{12}).
- **Group C:** Phosphatase action of CheA$_3$ (k_{15}) and CheA$_3$-P (k_{16}) on CheY$_6$-P.
- **Group D:** Phosphotransfer from CheA$_3$-P onto CheB$_2$ (k_9) as well as the associated reverse phosphotransfer (k_{-9}) and autodephosphorylation (k_{14}). This also brings into the model the implied sharing of phosphoryl groups between CheY$_6$ and CheB$_2$.

As such, each model containing terms from groups A, AB, AC, AD, ABC, ABD, ACD and ABCD are numerically simulated. In order to create each of these models we set all kinetic rates in neglected groups to zero. Each of these simulations is conducted using the initial conditions

$$a_{3p} = 0, \quad b_{2p} = 0 \quad \text{and} \quad y_{6p} = 0, \tag{8.16}$$



(a) CheA$_3$-P (b) CheY$_6$-P



(c) CheB$_2$-P

Figure 8.2: Results obtained from the ODE model outlined in equations (8.4)-(8.6). Symbols indicate the combinations of terms considered in each case. These are A (circles), AB (pluses), ABC (asterisks), AC (crosses), AD (dots), ABD (diamonds), ACD (triangles) and ABCD (squares). See Section 8.2.1 for details of the meanings of A, B, C and D.

and run (using solver ode15s) until each system reaches its respective steady-state. The effect of the various combinations of terms of CheA$_3$-P, CheY$_6$-P and CheB$_2$-P levels are shown in Figure 8.2.

For CheA$_3$-P (see Figure 8.2(a)) combinations of terms A and AD result in essentially saturated CheA$_3$-P. The inclusion of CheY$_6$-P autodephosphorylation into these models (i.e. AB and ABD) results in a lower steady-state CheA$_3$-P concentration. This is due to the autodephosphorylation of CheY$_6$-P creating more CheY$_6$, thus allowing more phosphoryl groups to transfer away from CheA$_3$-P. Whilst this concentration is lowered, it would still appear that the CheA$_3$-P concentration is higher than expected for a HPK (histidine protein kinase), based on *E. coli* which has a very low a_p^* value. Combinations of terms AC, ACD, ABC and ABCD all yielded fairly similar results in terms of their steady-state CheA$_3$-P concentration. This lower concentration is due to the inclusion of more CheY$_6$-P dephosphorylation (via the phosphatase action of CheA$_3$ and CheA$_3$-P) and phosphotransfer onto CheB$_2$, each of which results in more phosphoryl groups passing away from CheA$_3$-P. These concentrations are more in line with our expectations based upon the HPKs of other systems such as *E. coli*.

Results for CheY$_6$-P (Figure 8.2(b)) show that the inclusion of autodephosphorylation has a much more subtle effect than was the case for CheA$_3$-P. This is due to autodephosphorylation creating more non-phosphorylated CheY$_6$. The result of this is that phosphotransfer from CheA$_3$-P increases, thus reducing the CheA$_3$-P concentration. However, the increase in this phosphotransfer from CheA$_3$-P onto CheY$_6$ partially compensates for the effects of the autodephosphorylation. In particular, here we see that each of the combinations A, AB, AD and ABD lead to what is essentially saturation of CheY$_6$-P. This is likely to be an undesirable trait for the system since CheY$_6$-P is essential for chemotaxis in *R. sphaeroides* (with changes in this level likely to be essential). The combinations of terms AC and ACD (i.e. those that include the experimentally observed CheA$_3$ and CheA$_3$-P phosphatase action on CheY$_6$-P) give steady-state CheY$_6$-P concentrations of approximately 75% of the total amount. The addition of CheY$_6$-P autodephosphorylation into these two systems (i.e. combinations ABC and $ABCD$) appears to further lower the steady-state concentration to around 60% of the total amount. Either of the 75% or 60% phosphorylation levels could feasibly allow significant enough changes to allow for chemotaxis. However, all of the reactions have been demonstrated using *in vitro* experimental techniques and are likely to be present for some purpose and as such it is likely that the 60% CheY$_6$-P phosphorylation level is more desirable for chemotaxis in *R. sphaeroides*.

Also shown in Figure 8.2(c) are results obtained for protein CheB$_2$-P. Here we see that combinations AB and ABD lead to essentially saturated CheB$_2$-P. This saturation would appear undesirable since it would likely mean CheB$_2$ loses the ability to participate in any cellular response. The other combinations (ACD and ABCD) produce more reasonable levels of phosphorylation, with the inclusion of CheY$_6$-P autodephosphorylation simply lowering the steady-state concentration by 40% of the total amount. The exact role of protein CheB$_2$ is not known, however it has been suggested that it is involved in adaptation at the cytoplasmic cluster. Another possibility is that it may act as a phosphate sink. This essentially means that CheB$_2$ may act to limit the CheY$_6$-P concentration via the sharing of phosphoryl groups transferred from CheA$_3$-P. Considering the results of Figure 8.2(b), this phosphate sink hypothesis may be ruled out since the inclusion of CheB$_2$ (i.e. group D) into the model has very little effect on the CheY$_6$-P concentration. Thus, it is likely that CheB$_2$ is involved in a cytoplasmic cluster adaptation mechanism which is not yet understood.

Within bacterial signalling pathways it is the dynamic behaviour of protein phosphorylation levels which allows cells to respond to stimuli. As such, we also consider the dynamic behaviour of each model system when subjected to a persistent change in the rate of CheA$_3$ autophosphorylation. This mimics the initial response displayed by the system to a change in the extracellular

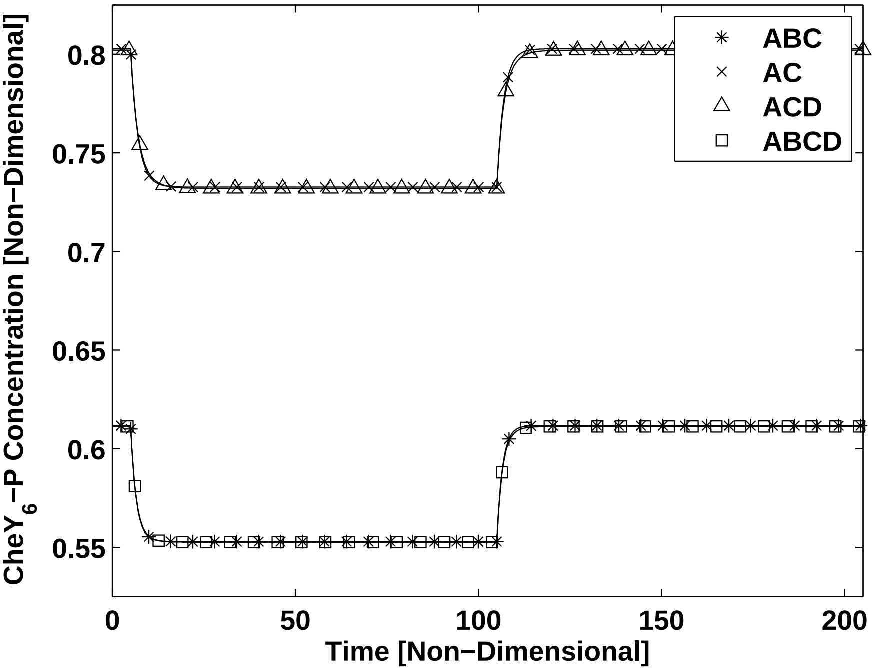


Figure 8.3: Dynamic results obtained from the ODE model outlined in equations (8.4)-(8.6). Results are shown under changes in the rate of CheA$_3$ autophosphorylation, scaled by L. In particular, L drops from $L = 1$ to $L = 0.95$ at time $\tau = 5$ and subsequently returns to $L = 1$ at $\tau = 105$. Symbols indicate the combinations of terms considered in each case. These are ABC (asterisks), AC (crosses), ACD (triangles) and ABCD (squares). See Section 8.2.1 for details of the meanings of A, B, C and D.

ligand concentration. Such effects are modelled here via an autophosphorylation term of the form $L(1 - a_{3p})$ rather than $(1 - a_{3p})$. Here, L is a scaling of the autophosphorylation rate used to represent a dynamic change within the system. To begin with we consider $L = 1$ and then reduce the autophosphorylation rate by choosing $L = 0.95$. We then consider the reverse change from $L = 0.95$ back to $L = 1$. Results obtained using this method are shown for protein CheY$_6$-P in Figure 8.3 for combinations AC, ABC, ACD and ABCD since these are most likely to yield experimentally realistic results, i.e. they do not produce saturation of phosphorylated proteins.

The results in Figure 8.3 indicate that the system comprising terms AC and ACD yield very similar results when the system is subjected to alterations in the rate of CheA$_3$ autophosphorylation (L). Groupings ABC and ABCD also show very similar results to one another, and indeed to AC and ACD, however they both have a different basal level of phosphorylation. Note that the inclusion of terms marked D in equations (8.4)-(8.6) leads to a slightly slower response to the change in L. The fact that the dynamic behaviour of each of these four systems is very similar suggests that these proteins do not play a significant role in setting the timescales associated with the initial, transient chemotactic response. Thus we conclude that CheB$_2$ is associated with the dynamic response of the cell via a possible adaptation mechanism.

Considering steady-state and dynamic results together, we may propose a number of roles for the chemotaxis signalling proteins studied here.

- **Group A:** These terms are required to ensure functioning of the signalling pathway.
- **Group B:** Autodephosphorylation acts to reduce the phosphorylated CheY$_6$ concentration slightly. The small degree to which this occurs is likely to suggest that this process represents an additional level of robustness, allowing the cell to produce a limited response even when the phosphatase action fails.
- **Group C:** Phosphatase action of CheA$_3$ and CheA$_3$-P on CheY$_6$-P is important in controlling the basal CheY$_6$-P concentration. It therefore prevents the basal CheY$_6$-P concentration from becoming too close to saturation as to hinder normal function.
- **Group D:** It is clear that CheB$_2$ acts to create a sharing of phosphoryl groups between itself and CheY$_6$. However, the effect on the CheY$_6$-P concentration is very small due to the very large CheY$_6$ concentration in *R. sphaeroides* cells.

How then may CheB$_2$ be involved in adaptation? In Chapter 7 it was shown that adaptation via an *E. coli* MWC mechanism leads to unrealistic CheY$_6$ phosphorylation levels in protein deletion mutant strains. In particular, the deletion of CheB$_2$ led to large increases in the phosphorylation level of CheY$_6$. Considering a model of the form given in equations (8.4)-(8.6) appears to limit this issue. We could perhaps then suggest that the adaptation mechanism at the cytoplasmic cluster of *R. sphaeroides* cells is much weaker or slower than in the case of *E. coli*, the mechanism of which was previously assumed to hold for *R. sphaeroides*. This would likely lead to a much smaller difference in the steady-state CheY$_6$-P concentration for a CheB$_2$ deletion mutant cell since the effect of CheB$_2$ on the system would be much more subtle than in the *E. coli* mechanism.

It is worth noting that this very simple system lacking in an adaptation mechanism appears to solve the problem associated with the ΔCheB$_2$ mutant cell type, since the CheY$_6$-P concentration remains almost unaffected. However, the experimental data of de Beyer [162] describes the behaviour of the ΔCheB$_2$ deletion mutant as being inhibited, with much longer adaptation times. A system with no adaptation mechanism such as that considered here cannot produce a response of this type under a CheB$_2$ deletion and thus another mechanism must be sought in future work. It is perhaps possible that the connection between the polar and cytoplasmic clusters via CheB$_2$, may act to produce an appropriate response however the exact mechanism for this is as yet unknown.

The results presented within this section clearly demonstrate the roles of each cytoplasmic cluster process. Whilst useful in its own right, this model considers only the cytoplasmic cluster. Thus, in order to improve our understanding of other cell features it is necessary to consider the two chemosensory pathways together in the same model.

8.3.2 Validation of the Signalling Pathways

In Section 8.3.1 a simplified mathematical model of processes occurring at the cytoplasmic cluster in *R. sphaeroides* cells was investigated. Studying the steady-state and dynamic properties of this simplified ODE model with various processes systematically removed allowed us to assign likely roles to each of the relevant signalling proteins acting at the cytoplasmic cluster. Within this section we utilise Simplified Model 2 (Section 8.2.2) in order to investigate a key outstanding question relating to the polar cluster. This is as follows.

- A reverse phosphotransfer from CheB$_2$-P onto CheA$_2$ has been observed in *in vitro* experimental work. However, to date, this has not been shown to occur *in vivo*. Would we expect to see this reaction occurring?

Within the previous literature, all phosphorylation reactions included within the Kojadinovic et al. [161] model have been shown to occur *in vitro*. In fact, all apart from the reverse phosphotransfer from CheB$_2$-P onto CheA$_2$ have also been demonstrated to exist *in vivo*. An investigation into the existence of this reaction allows us to create a full, validated set of phosphorylation/dephosphorylation reactions occurring within *R. sphaeroides* cells. This is explored in-depth within the next section.

Is Reverse Phosphotransfer from CheB2-P onto CheA2 Expected *in vitro*?

As discussed in Section 8.3.2, all phosphorylation/dephosphorylation reactions in the signalling pathways of *R. sphaeroides* cells have been validated except for one. This is the reverse phosphotransfer from CheB$_2$-P onto CheA$_2$ which has been shown to be possible *in vitro* but not demonstrated *in vivo*. As such, here we utilise mathematical modelling to ascertain whether or not this reaction will be expected to occur *in vivo*. This will then result in a full set of validated phosphorylation/dephosphorylation reactions for the two chemosensory pathways of *R. sphaeroides* cells.

In order to answer this question, the wild-type system described in Section 8.2.2 is compared with a system in which the reverse phosphotransfer in question has been eliminated (i.e. $k_{-7} = 0$). Analysis of each of these two models shows the existence of just one biologically feasible steady-state, each of which is asymptotically stable.

Using a steady-state analysis of the two systems discussed here, the effects of a reverse phosphotransfer from CheB$_2$-P onto CheA$_2$ can be investigated. In particular, the differences between the steady-states of each signalling protein can be studied which should allow a prediction to be made as to whether this reverse phosphotransfer will be present *in vivo*.

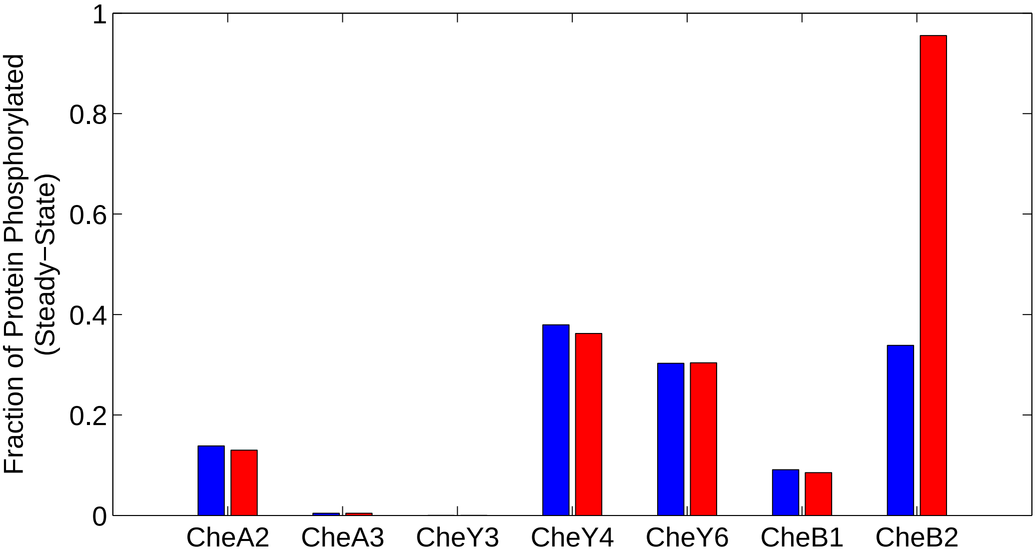


Figure 8.4: A bar chart showing the differences at steady-state between systems with (blue bars) and without (red bars) a reverse phosphotransfer from CheB$_2$-P onto CheA$_2$.

Inspecting the results of Figure 8.4, it can be seen that most of the steady-state phosphorylated protein concentrations remain almost unchanged. However the exception is CheB$_2$-P. Perhaps unsurprisingly this protein displays the largest change as a result of the removal of this reverse phosphotransfer. In fact, a wild-type cell with the reverse phosphotransfer acting will have around one third of its total CheB$_2$ phosphorylated at steady-state whereas in those where this reverse phosphotransfer has been eliminated, around 95% of all CheB$_2$ will be phosphorylated. Clearly this is rather a significant change in the CheB$_2$-P steady-state.

If a cell were to maintain a steady-state CheB$_2$-P concentration whereby ∼95% is phosphorylated (as in the $k_{-7} = 0$ case) then the ability of CheB$_2$-P to perform its intracellular function would be severely impaired in cases where this concentration is required to increase. In contrast to this, maintaining a steady-state in which ∼33% of CheB$_2$ is phosphorylated allows room for this protein to function in both cases whereby an increase or a decrease is required.

Upon considering the effect of removing the reverse phosphotransfer from CheB$_2$-P, it is clear that some process must be acting to maintain a reasonable CheB$_2$-P concentration *in vivo*. The results shown in Figure 8.4 demonstrate the ability of this reverse phosphotransfer to limit the CheB$_2$-P concentration. Using this result as well as the fact that the reverse phosphotransfer from CheB$_2$-P onto CheA$_2$ has been shown to occur *in vitro*, it can reasonably be expected that this process is acting to limit the $CheB_2$-P concentration.

8.3.3 Understanding the Motor Response

Within this section we seek to elucidate mechanisms underlying the control of flagellar rotation in *R. sphaeroides* cells. In particular, we consider the experimental data of de Beyer [162] that investigated the behaviour of a number of mutant cell types (discussed in Section 8.1). Here we produce variations of Simplified Model 2 (Section 8.2.2) in order to assess how intracellular protein phosphorylation levels may act to control the rotational behaviour of the cell's flagellar motor. Specifically, we consider two key outstanding questions:

1. The experimental data of de Beyer [162] showed that wild-type and $\Delta cheY_{3,4}$ ($CheY_3$ and $CheY_4$ deletion) mutant cells produce different steady-state swimming behaviour (see Figure 7.2). Why is this the case if $CheY_6$-P dominantly controls flagellar rotation as previously assumed in the literature?
2. Using model predictions of protein phosphorylation levels for each mutant cell type, can we understand how these are linked to the control of flagellar rotation?

The difference in steady-state swimming behaviour observed between wild-type and $\Delta cheY_{3,4}$ mutant cells suggests that $CheY_6$-P does not dominantly control flagellar rotation as previously assumed. An investigation into the mechanism behind this difference (Question 1) allows us to obtain a better understanding of how protein phosphorylation levels control flagellar rotational behaviour. This is the subject of further investigation in this section.

By producing a number of altered versions of Simplified Model 2 (Section 8.2.2), it is possible to obtain a predicted phosphorylation level for all proteins under each mutation considered. Using these predictions alongside the experimental data of de Beyer [162] (Figure 7.2) we may investigate how protein phosphorylation levels affect flagellar rotation.

Why do $\Delta cheY_{3,4}$ Cells Differ from Wild-Type?

As discussed above, wild-type and $\Delta cheY_{3,4}$ strains of *R. sphaeroides* display different steady-state swimming behaviour. It has previously been assumed that $CheY_6$-P dominantly controls the rotational behaviour of the cell's flagellum. How then do these cell types produce different behaviour? In order to investigate why this is the case, here we conduct steady-state analyses of two mathematical models (based on Simplified Model 2) one representing each cell type. Results of this are shown in Table 8.3.

Upon inspection of the steady-state $CheY_3$-P, $CheY_4$-P and $CheY_6$-P concentrations of wild-type and $\Delta CheY_{3,4}$ cells it can be seen that the $\Delta CheY_{3,4}$ mutation does not significantly alter

Table 8.3: Steady-state CheY concentrations for wild-type and ΔCheY$_{3,4}$ cell types. All concentrations listed in this table are expressed in μM and are shown to four decimal places.

Cell Type	[CheY$_3$-P]	[CheY$_4$-P]	[CheY$_6$-P]
W-T	5.27087×10^{-5}	5.2514	68.9646
ΔCheY$_{3,4}$	0	0	68.1697

the steady-state concentration of CheY$_6$-P. In fact, this particular mutation only leads to an increase of $\sim 0.8\mu M$ (of the total $225\mu M$) in the steady-state CheY$_6$-P concentration. We would anticipate that this small increase would not be sufficient to allow for the experimentally observed behavioural differences if CheY$_6$-P dominantly controls flagellar rotational behaviour. Thus, we must consider what is the cause of this behavioural difference if it is not a difference in CheY$_6$-P concentration?

One possible cause of the difference between W-T and ΔCheY$_{3,4}$ cells is that the flagellar motor of *R. sphaeroides* cells will change its action in the face of this mutation. Using the results obtained here it is not possible to rule this out since the relationship between intracellular protein concentrations and flagellar motor rotational behaviour is as yet unknown. Whilst we cannot rule out a change in motor action, we would perhaps see this as less likely as a cause of behavioural changes than differences in protein concentrations other than CheY$_6$-P. Since there is a difference of $\sim 5.25\mu M$ in the combined CheY$_3$-P and CheY$_4$-P concentration between these two cell types (see Table 8.3) we would anticipate that this is the more likely cause of behavioural differences.

The results in Table 8.3 could potentially be seen as evidence of CheY$_3$ and CheY$_4$ acting to re-start flagellar rotation. Assuming that the ΔCheY$_{3,4}$ mutation does not have any knock-on effects on other processes in the system then the only difference between the two cell types discussed here will be the absence of the proteins CheY$_3$ and CheY$_4$. Since removing these proteins does not have a significant effect on the steady-state concentrations of other proteins in the system (see Tables 8.3 and 8.5) we would expect that the difference in the steady-state swimming behaviour between these two cell types will be due to CheY$_3$-P and CheY$_4$-P. This difference could be caused by one of two effects. Firstly, CheY$_3$ and CheY$_4$ may compete with CheY$_6$ for the ability to bind to the motor driving protein. We would perhaps not expect this to be such a significant effect due to the relatively small concentrations of CheY$_3$ and CheY$_4$ in comparison to CheY$_6$ (see Table 7.2). The other, and perhaps more likely option, is that CheY$_3$ and CheY$_4$ have more of an active role in setting the flagellar rotational behaviour of *R. sphaeroides* cells. We would suggest that due to the fact that a ΔCheY$_{3,4}$ mutant cell spends a greater fraction of time spent stopped than a wild-type cell, CheY$_3$-P and CheY$_4$-P could play a role in re-starting the rotation of the flagellar motor.

How is Flagellar Rotation Controlled?

In this section we consider a number of alterations to Simplified Model 2 (Section 8.2.2), each of which represents one of the mutant cell types described in Table 7.1. In order to investigate how the flagellar rotational behaviour is controlled we conduct steady-state analyses of each mutant model. This, alongside the experimental data of de Beyer [162] (Figure 7.2) then allows us to investigate how protein phosphorylation levels act to control the flagellar rotational behaviour.

Table 8.4: Kinetic rate alterations required to create each *R. sphaeroides* mutant model.

Strain	Parameters set to zero	Other
#1	No alteration	-
#2	—	-
#3	—	-
#4	k_{-3} k_{-4} k_{-6} k_{-7}	-
#5	Φ_P^*	-
#6	k_{-3} k_{-4} k_{-6} k_{-7} Φ_P^*	-
#7	k_5 k_8 k_{-8} k_9 k_{-9} k_{12} k_{15} k_{16} Φ_C^*	$y_{6p} = 1$
#8	k_5 k_8 k_{-8} k_{12} k_{15} k_{16} Φ_C^*	$y_{6p} = 1$
#9	k_3 k_{-3} k_{10}	$y_{3p} = 0$
#10	k_4 k_{-4} k_{11}	$y_{4p} = 0$
#11	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	$y_{3p} = 0 = y_{4p}$
#12	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	$y_{3p} = 0$
#13	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	$y_{3p} = 0 = y_{4p}$
#14	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	-
#15	—	-
#16	—	-
#17	—	-
#18	—	-
#19	—	-
#20	k_{-8} k_{-9} Φ_C^*	-
#21	Φ_C^*	-
#22	k_5 k_8 k_{-8} k_{12} k_{15} k_{16}	$y_{6p} = 0$
#23	k_3 k_{-3} k_{10}	$y_{3p} = 1$
#24	k_4 k_{-4} k_{11}	$y_{4p} = 1$
#25	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	$y_{3p} = 1 = y_{4p}$
#26	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	$y_{3p} = 1$
#27	k_3 k_{-3} k_4 k_{-4} k_{10} k_{11}	$y_{4p} = 1$
#28	k_3 k_{-3} k_4 k_{-4} k_5 k_8 k_{-8} k_{10} k_{11} k_{12} k_{15} k_{16}	-
#29	k_5 k_8 k_{-8} k_{12} k_{15} k_{16}	$y_{6p} = 1$
#30	k_4 k_{-4} k_{11}	-
#31	k_3 k_{-3} k_{10}	-
#32	k_5 k_8 k_{-8} k_{12} k_{15} k_{16}	-

Similar to the work in Section 7.5, here we implement a number of changes to model parameters in order to represent each mutant cell type. These are detailed in Table 8.4. Using the method of obtaining the steady-state values laid out in section 7.4.2 it is possible to obtain steady-state chemotaxis protein concentrations for mutant cells examined within this work, see Table 8.5. In doing so each model formulated within this work (one for each mutant cell type considered) is found to display just one biologically feasible steady-state, i.e. all non-dimensional steady-state values lie in the range $0 \leq x_p^* \leq 1$. Further to this, analysis of the eigenvalues of each system (as discussed in Section 7.4.3), undertaken using the 'eig' function in MATLAB shows that each of these biologically feasible steady-states are asymptotically stable, as indicated by the existence of only negative eigenvalues (see Table 8.6). It is worth noting here that the number of eigenvalues relates to the number of variables within the system. Thus, in cases whereby a protein concentration is not able to change or has been deleted, less variables would exist. As such, for each constant concentration or deletion we would lose one eigenvalue.

In order for the results detailed in Table 8.5 to be of use it is necessary to define an output variable of interest. Within the earlier work on *E. coli* this output variable was considered to be either the concentration of the flagellar motor regulating protein (CheY-P) or the directional rotation bias of the cells' flagellar motors. Here it is not so simple to define a variable of interest since the relationship between intracellular proteins and the flagellar motor response is not known. Examining the experimental data of de Beyer [162] as well as Section 8.3.3, it appears likely that CheY_6-P is responsible for causing flagellar rotation to stop and the proteins CheY_3-P and CheY_4-P may be responsible for re-starting the rotation of the flagellar motor. As such, it is supposed here that the response of the flagellar motor in *R. sphaeroides* is likely controlled by the ratio of CheY_6-P concentration to the sum of concentrations of CheY_3-P and CheY_4-P, i.e.

$$\text{Ratio} = \frac{[\text{CheY}_3\text{-P}] + [\text{CheY}_4\text{-P}]}{[\text{CheY}_6\text{-P}]}. \tag{8.17}$$

Upon consideration of this ratio as our output variable we obtain the values in Table 8.7.

It is worth noting that within Table 8.7 some of the ratios are listed as 'Div By 0'. This indicates that the steady-state CheY_6-P concentration was zero within that model and thus in calculating the output ratio we would be required to divide by zero. It is unclear exactly how these cases should be handled and thus we arbitrarily assign a ratio value of one in each of these cases. Also, mutant strains #29 - #32 are created for the interest of modelling, but do not have a directly comparable experimental result and thus they are not considered any further here. Extracting the experimentally determined fraction of time stopped from Figure 7.2, we are able to create a comparison with model predicted ratios in Table 8.7.

Table 8.5: Steady-states of each *R. sphaeroides* mutant model.

	Steady-State Concentrations [Non-Dimensional]						
Description	CheA2-P	CheA3-P	CheY3-P	CheY4-P	CheY6-P	CheB1-P	CheB2-P
WT	0.1384	0.004322	1.523E-005	0.3794	0.3030	0.09115	0.3385
ΔCheB$_2$	-	-	-	-	-	-	-
ΔCheR$_2$	-	-	-	-	-	-	-
CheA$_2$(H48Q)	0	0.003961	0	0	0.2725	0	0.8865
CheA$_2$(G470K)	0.004584	0.003946	5.050E-07	0.01724	0.2722	0.002868	0.1418
CheA$_2$(H48Q,G470K)	0	0.003961	0	0	0.2725	0	0.8865
CheA$_3$(H48Q)CheY$_6$(D57N)	0.5912	0	6.509E-005	0.8460	1	0.4746	0.7437
CheA$_4$(G470K)CheY$_6$(D57N)	0.4256	4.500E-03	4.686E-005	0.7381	1	0.3164	0.6351
CheY$_3$(D57A)	0.1480	0.004349	0	0.3980	0.3051	0.09784	0.3509
CheY$_4$(D57A)	0.1441	0.004338	1.586E-005	0	0.3042	0.09514	0.3459
CheY$_{3,4}$(D57A)	0.1544	0.004367	0	0	0.3065	0.1024	0.3592
CheY$_3$(D57A)ΔCheY$_4$	0.1544	0.004367	0	0	0.3065	0.1024	0.3592
CheY$_3$**CheY$_4$(D57A)	0.1544	0.004367	0	0	0.3065	0.1024	0.3592
ΔCheY$_{3,4}$	0.1544	0.004367	0	0	0.3065	0.1024	0.3592
ΔCheB$_1$	-	-	-	-	-	-	-
ΔCheB$_{1,2}$	-	-	-	-	-	-	-
ΔTlpT	-	-	-	-	-	-	-
ΔCheR$_3$	-	-	-	-	-	-	-
ΔCheR$_{1,2,3}$	-	-	-	-	-	-	-
CheA$_3$(H48Q)	0.1159	0	1.276E-005	0.3330	0.03213	0.07569	0.2083
CheA$_4$(G470K)	0.1066	0.0001445	0.00001174	0.3124	0.03255	0.06934	0.1938
CheY$_6$(D57A)	0.7104	0.6982	7.821E-005	0.9030	0	0.6050	0.9901
CheY$_3$(D57N)	0.1480	0.004349	1	0.3980	0.3051	0.09784	0.3509
CheY$_4$(D57N)	0.1441	0.004338	1.586E-005	1	0.3042	0.09514	0.3459
CheY$_{3,4}$(D57N)	0.1544	0.004367	1	1	0.3065	0.1024	0.3592
CheY$_3$(D57N)ΔCheY$_{1,2,4,5}$	0.1544	0.004367	1	0	0.3065	0.1024	0.3592
CheY$_4$(D57N)ΔCheY$_{1,2,3,5}$	0.1544	0.004367	0	1	0.3065	0.1024	0.3592
ΔCheY$_{1-6}$	0.9765	0.9151	0	0	0	0.9629	0.9993
CheY$_6$(D57N)	0.7104	0.6982	7.821E-005	0.9030	1	0.6050	0.9901
ΔCheY$_4$	0.1441	0.004338	1.586E-005	0	0.3042	0.09514	0.3459
ΔCheY$_3$	0.1480	0.004349	0	0.3980	0.3051	0.09784	0.3509
ΔCheY$_6$	0.7104	0.6982	7.821E-005	0.9030	0	0.6050	0.9901

Table 8.6: Eigenvalues associated with the steady-state of each *R. sphaeroides* chemotaxis mutant model.

	Eigenvalues (to 5 d.p.)							
Description	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7	Stability
WT	-44924.90604	-3327.57927	-1026.67477	-284.56248	-4.36955	-3.92332	-1.42700	Stable
ΔCheB$_2$	-	-	-	-	-	-	-	-
ΔCheR$_2$	-	-	-	-	-	-	-	-
CheA$_2$(H48Q)	-44916.66667	-1308.88797	-1065.06796	-4.03404	-2.56477	-0.28000	-0.01417	Stable
CheA$_2$(G470K)	-44925.41736	-3378.97011	-1073.06437	-238.31540	-4.23978	-3.64706	-1.26113	Stable
CheA$_2$(H48Q,G470K)	-44916.66667	-1308.88797	-1065.06796	-4.03404	-2.56477	-0.28000	-0.01417	Stable
CheA$_3$(H48Q)CheY$_6$(D57N)	-44923.17534	-3232.05643	-480.06038	-2.72955	-2.10003	-0.29587	-	Stable
CheA$_4$(G470K)CheY$_6$(D57N)	-44923.80820	-3252.12794	-405.30908	-7.52107	-1.19343	-0.28136	-	Stable
CheY$_3$(D57A)	-3324.24919	-1023.50468	-288.23981	-4.37887	-3.93330	-1.33131	-	Stable
CheY$_4$(D57A)	-44924.88415	-3294.83388	-1024.77909	-4.37323	-3.99557	-1.61196	-	Stable
CheY$_{3,4}$(D57A)	-3291.71890	-1021.37678	-4.38325	-4.00038	-1.49361	-	-	Stable
CheY$_3$(D57A)ΔCheY$_4$	-3291.71890	-1021.37678	-4.38325	-4.00038	-1.49361	-	-	Stable
CheY$_3$**CheY$_4$(D57A)	-3291.71890	-1021.37678	-4.38325	-4.00038	-1.49361	-	-	Stable
ΔCheY$_{3,4}$	-3291.71890	-1021.37678	-4.38325	-4.00038	-1.49361	-	-	Stable
ΔCheB$_1$	-	-	-	-	-	-	-	-
ΔCheB$_{1,2}$	-	-	-	-	-	-	-	-
ΔTlpT	-	-	-	-	-	-	-	-
ΔCheR$_3$	-	-	-	-	-	-	-	-
ΔCheR$_{1,2,3}$	-	-	-	-	-	-	-	-
CheA$_3$(H48Q)	-44924.99185	-3335.72200	-1414.87227	-276.14834	-4.25441	-2.27629	-1.03334	Stable
CheA$_4$(G470K)	-44925.02750	-3339.06216	-1416.89595	-272.69133	-4.23927	-3.84851	-1.77068	Stable
CheY$_6$(D57A)	-44922.71980	-3230.11771	-534.92032	-35.54345	-0.47060	-0.33474	-	Stable
CheY$_3$(D57N)	-3324.24919	-1023.50468	-288.23981	-4.37887	-3.93330	-1.33131	-	Stable
CheY$_4$(D57N)	-44924.88415	-3294.83388	-1024.77910	-4.37323	-3.99557	-1.61196	-	Stable
CheY$_{3,4}$(D57N)	-3291.71890	-1021.37678	-1.49361	-4.38325	-4.00038	-	-	Stable
CheY$_3$(D57N)ΔCheY$_{1,2,4,5}$	-3291.71890	-1021.37678	-4.38325	-4.00038	-1.49361	-	-	Stable
CheY$_4$(D57N)ΔCheY$_{1,2,3,5}$	-3291.71890	-1021.37678	-4.38325	-4.00038	-1.49361	-	-	Stable
ΔCheY$_{1-6}$	-3254.83384	-45.25421	-0.40278	-0.26009	-	-	-	Stable
CheY$_6$(D57N)	-44922.71980	-3230.11771	-534.92032	-35.54345	-0.47060	-0.33474	-	Stable
ΔCheY$_4$	-44924.88415	-3294.83388	-1024.77909	-4.37323	-3.99557	-1.61196	-	Stable
ΔCheY$_3$	-3324.24919	-1023.50468	-288.23981	-4.37887	-3.93330	-1.33131	-	Stable
ΔCheY$_6$	-44922.71980	-3230.11771	-534.92032	-35.54345	-0.47060	-0.33474	-	Stable

Table 8.7: Values obtained from equation (8.17) for each *R. sphaeroides* mutant model considered.

Mutation	Ratio	Mutation	Ratio
WT	0.077034	ΔTlpT	-
ΔCheB$_2$	-	ΔCheR$_3$	-
ΔCheR$_2$	-	ΔCheR$_{1,2,3}$	-
CheA$_2$(H48Q)	0.203684	CheA$_3$(H48Q)	0.637575
CheA$_2$(G470K)	0.003896	CheA$_4$(G470K)	0.590262
CheA$_2$(H48Q,G470K)	0.000000	CheY$_6$(D57A)	Div By 0
CheA$_3$(H48Q)CheY$_6$(D57N)	0.052041	CheY$_3$(D57N)	0.130650
CheA$_4$(G470K)CheY$_6$(D57N)	0.045404	CheY$_4$(D57N)	0.202180
CheY$_3$(D57A)	0.080246	CheY$_{3,4}$(D57N)	0.250853
CheY$_4$(D57A)	8.02E-07	CheY$_3$(D57N)ΔCheY$_{1,2,4,5}$	0.050171
CheY$_{3,4}$(D57A)	0.000000	CheY$_4$(D57N)ΔCheY$_{1,2,3,5}$	0.200683
CheY$_3$(D57A)ΔCheY$_4$	0.000000	ΔCheY$_{1-6}$	Div By 0
CheY$_3$**CheY$_4$(D57A)	0.000000	CheY$_6$(D57N)	0.055543
ΔCheY$_{3,4}$	0.000000	ΔCheY$_4$	8.02E-07
ΔCheB$_1$	-	ΔCheY$_3$	0.080246
ΔCheB$_{1,2}$	-	ΔCheY$_6$	Div By 0

The ratio referred to in this table is calculated as Ratio=[CheY$_3$-P]+[CheY$_4$-P]/[CheY$_6$-P]. Those displaying 'Div By 0' in the Ratio column produce a steady-state CheY$_6$-P concentration of $0\mu M$.

It can be seen from the results of Figure 8.5 that there is a general correlation in that those cells with larger $[Y_{3p}]+[Y_{4p}]$:$[Y_{6p}]$ ratios correspond to those displaying a smaller fraction of time stopped within the experimental data. In addition to this, cells with smaller $[Y_{3p}]+[Y_{4p}]$:$[Y_{6p}]$ ratios appear to correspond to those for which experimental data shows a larger fraction of time spent stopped.

Within a deterministic modelling study of this type it would perhaps be expected that the work would produce results that do not display a fluctuation from the predicted relationship. Whilst this would be true in many cases, the simplistic nature of the models used within this work do not necessarily allow for an exact relation to be obtained, rather a somewhat naïve approximation of the true result is given. The reason for this is that a number of processes have been neglected/simplified here that would be present within the experimental data. In particular, it has already been discussed that adaptation is not included within this model. This could impact upon the results here since the distribution of phosphoryl groups around the system will affect the strength of demethylation feedback upon chemoreceptor state. It would be expected that a variation in the strength of this feedback would impact upon the steady-state chemoreceptor activity level, thus affecting the autophosphoylation rate and subsequently the steady-state phosphorylation levels of chemotaxis proteins.

In this work, the chemoreceptor activities at the polar and cytoplasmic clusters have been assigned a constant value determined from the steady-state values in the model of Kojadinovic et al. [161]. Whilst this would most likely be a good approximation in most of the mutant

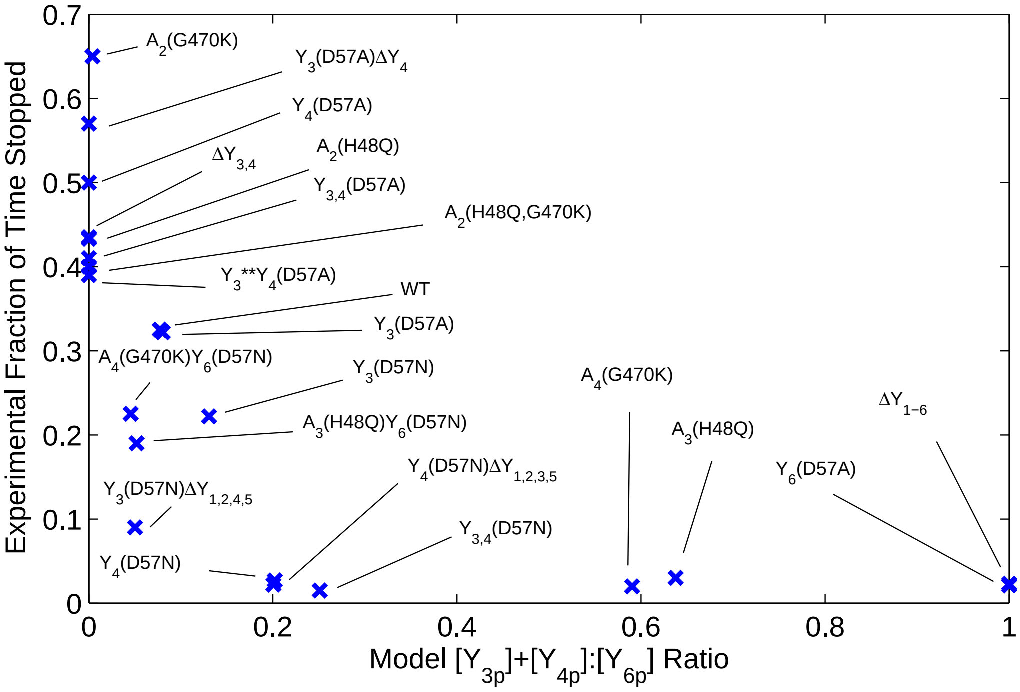


Figure 8.5: A comparison of the model predicted $[Y_{3p}] + [Y_{4p}]:[Y_{6p}]$ ratio with the experimentally measured fraction of time stopped for chemotaxis mutant cells. Blue crosses indicate data points for model predicted $[Y_{3p}] + [Y_{4p}]:[Y_{6p}]$ ratio and experimentally determined fraction of time stopped (for *R. sphaeroides* flagella). Experimental fraction of time stopped data points refer to the median of the experimental data, as displayed in Fig. 7.2.

cell types considered here, some examples may deviate further from these values causing the observed fluctuations in results. It is worth noting here that whilst fluctuations about the true relationship within our results may be caused by the mechanism used here, a number of mutant cell types that would directly impact upon methylation/demethylation have been removed, i.e. we have neglected those mutants involving alterations to the proteins CheR_2, CheR_3, CheB_1 and CheB_2. It is therefore anticipated that this will reduce the impact of variation in chemoreceptor activity, thus improving the validity of results in this chapter.

8.4 Proposed New Model

Within Section 7.5 it was shown that the Kojadinovic et al. [161] model did not provide an adequate description of chemotaxis in *R. sphaeroides* cells. As such, Section 8.3 of this chapter investigated various features associated with the chemosensory pathways of *R. sphaeroides* cells in order to help identify ways in which this model could be improved.

Using the results of these previous sections, in Section 8.4.1 we identify some alterations which could improve the performance of the Kojadinovic et al. [161] model. This results in the statement of a new mathematical model of chemotaxis in *R. sphaeroides*, some results of which are then presented in Section 8.4.2.

8.4.1 Mathematical Modelling

Within the results of Section 7.5 it was shown that the deletion of CheB and CheR homologues produced some deviation from expected steady-state CheY$_6$-P concentrations. By far the largest deviation resulted from the deletion of CheB$_2$ which is assumed to be responsible for demethylation of cytoplasmic cluster chemoreceptors within the model of Kojadinovic et al. [161]. By comparison, deviations from expected CheY$_6$-P steady-state concentrations as a result of CheB$_1$ and CheR$_2$ deletions were small. Thus, here we consider adaptation at the polar cluster to follow the *E. coli* mechanism, as in the Kojadinovic et al. [161] model. However, a new mechanism must be formulated for the cytoplasmic cluster.

Another area in which this model may be improved relates to the chemosensory behaviour of the cytoplasmic receptor cluster. Within the literature it is often stated that the cytoplasmic cluster is likely responsible for metabolic sensing [5]. However, within the Kojadinovic et al. [161] model the cytoplasmic cluster is assumed to detect the same external ligand signal as the polar cluster. As such it could potentially be seen as an improvement were a new model to sense an external signal at the polar cluster and a metabolic signal at the cytoplasmic cluster.

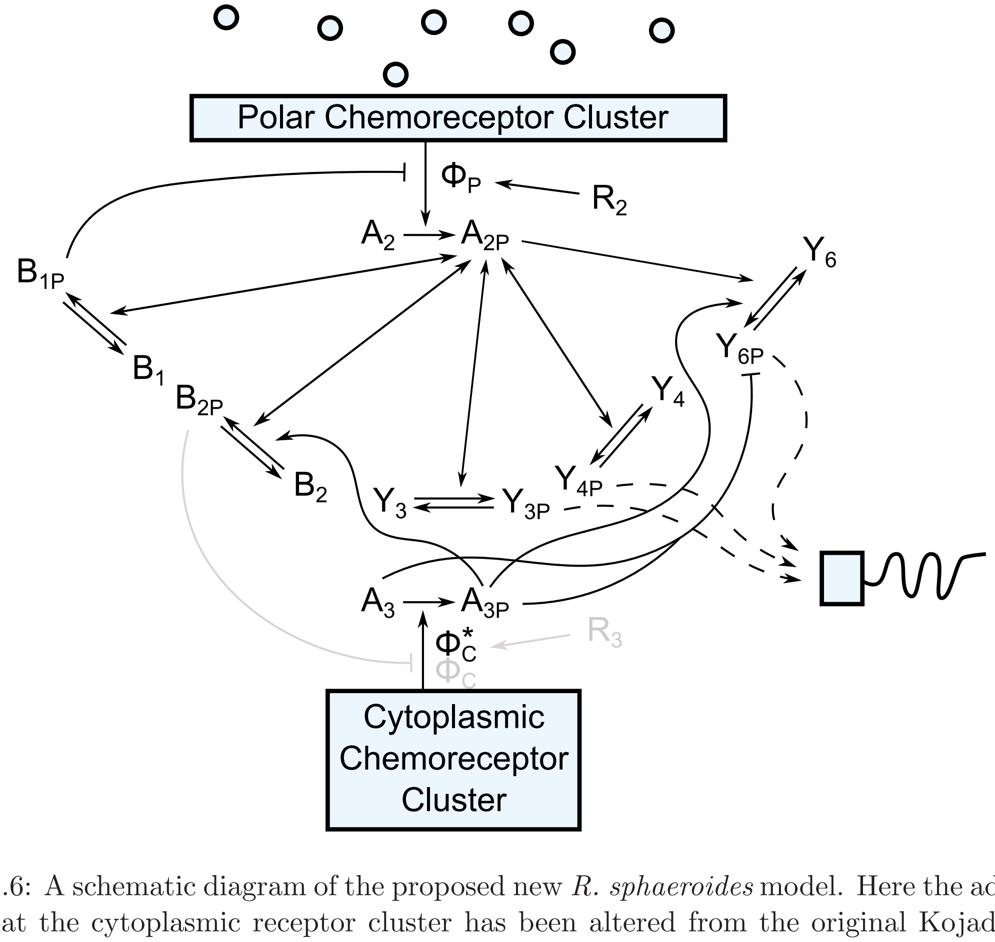


Figure 8.6: A schematic diagram of the proposed new *R. sphaeroides* model. Here the adaptation process at the cytoplasmic receptor cluster has been altered from the original Kojadinovic et al. model [161] to consider the detection of a metabolic signal. This metabolic signal acts by increasing or decreasing the steady-state receptor activity Φ^*_C. Note that greyed out elements in this diagram refer to those parts which have been altered from the original model.

Based on these suggestions we must now propose a mechanism for the cytoplasmic cluster that relieves the issue of CheY$_6$-P saturation when CheB$_2$ is deleted from the cell and also allows for the sensing of a metabolic signal (see Figure 8.6). In order to do this we look to the simplified cytoplasmic cluster model considered in Section 8.2.1. This mechanism clearly solved the issue associated with the deletion of CheB$_2$ and is easily adjusted to consider metabolic signals. As such, here we utilise the law of mass action (see Appendix A), conservation laws and a non-dimensionalisation similar to those in Section 7.4. This yields the set of non-dimensional ODEs

$$
\begin{aligned}
\frac{dm_P}{d\tau} &= \gamma_{R2}(1-\Phi_P) - \gamma_{B1} b_{1p}^2 \Phi_P, \qquad \text{(8.18)}\\
\frac{da_{2p}}{d\tau} &= \Phi_P(1-a_{2p}) - \bar{k}_3 a_{2p}(1-y_{3p}) + \bar{k}_{-3}(1-a_{2p})y_{3p} - \bar{k}_4 a_{2p}(1-y_{4p}) \qquad \text{(8.19)}\\
&+ \bar{k}_{-4}(1-a_{2p})y_{4p} - \bar{k}_5 a_{2p}(1-y_{6p}) - \bar{k}_6 a_{2p}(1-b_{1p}) + \bar{k}_{-6}(1-a_{2p})b_{1p}\\
&- \bar{k}_7 a_{2p}(1-b_{2p}) + \bar{k}_{-7}(1-a_{2p})b_{2p},\\
\frac{a_{3p}}{d\tau} &= \bar{k}_2 \Phi_C^*(1-a_{3p}) - \bar{k}_8 a_{3p}(1-y_{6p}) + \bar{k}_{-8}(1-a_{3p})y_{6p} - \bar{k}_9 a_{3p}(1-b_{2p}) \qquad \text{(8.20)}\\
&+ \bar{k}_{-9}(1-a_{3p})b_{2p},\\
\frac{dy_{3p}}{d\tau} &= \alpha_1 \bar{k}_3 a_{2p}(1-y_{3p}) - \alpha_1 \bar{k}_{-3}(1-a_{2p})y_{3p} - \bar{k}_{10} y_{3p}, \qquad \text{(8.21)}\\
\frac{dy_{4p}}{d\tau} &= \alpha_2 \bar{k}_4 a_{2p}(1-y_{4p}) - \alpha_2 \bar{k}_{-4}(1-a_{2p})y_{4p} - \bar{k}_{11}, \qquad \text{(8.22)}\\
\frac{dy_{6p}}{d\tau} &= \alpha_3 \bar{k}_5 a_{2p}(1-y_{6p}) + \alpha_6 \bar{k}_8 a_{3p}(1-y_{6p}) - \alpha_6 \bar{k}_{-8}(1-a_{3p})y_{6p} - \bar{k}_{12} y_{6p} \qquad \text{(8.23)}\\
&- \bar{k}_{15}(1-a_{3p})y_{6p} - \bar{k}_{16} a_{3p} y_{6p},\\
\frac{db_{1p}}{d\tau} &= \alpha_4 \bar{k}_6 a_{2p}(1-b_{1p}) - \alpha_4 \bar{k}_{-6}(1-a_{2p})b_{1p} - \bar{k}_{13} b_{1p}, \qquad \text{(8.24)}\\
\frac{db_{2p}}{d\tau} &= \alpha_5 \bar{k}_7 a_{2p}(1-b_{2p}) - \alpha_5 \bar{k}_{-7}(1-a_{2p})b_{2p} + \alpha_7 \bar{k}_9 a_{3p}(1-b_{2p}) \qquad \text{(8.25)}\\
&- \alpha_7 \bar{k}_{-9}(1-a_{3p})b_{2p} - \bar{k}_{14} b_{2p}, \qquad \text{(8.26)}
\end{aligned}
$$

in which all parameters retain their previous definitions (see Tables 7.3 and 8.2), Φ_P is the receptor signalling team activity at the polar cluster (as in equations (7.1) and (7.2)) and Φ_C^* is the steady-state receptor signalling team activity of chemoreceptors at the cytoplasmic cluster. We also consider initial conditions of the form

$$
m_P = m_{P0}, \quad a_{2p} = a_{2p0}, \quad a_{3p} = a_{3p0}, y_{3p} = y_{3p0}, \quad y_{4p} = y_{4p0},
$$

$$y_{6p} = y_{6p0}, \quad b_{1p} = b_{1p0} \quad \text{and} \quad b_{2p} = b_{2p0}.$$

8.4.2 Results

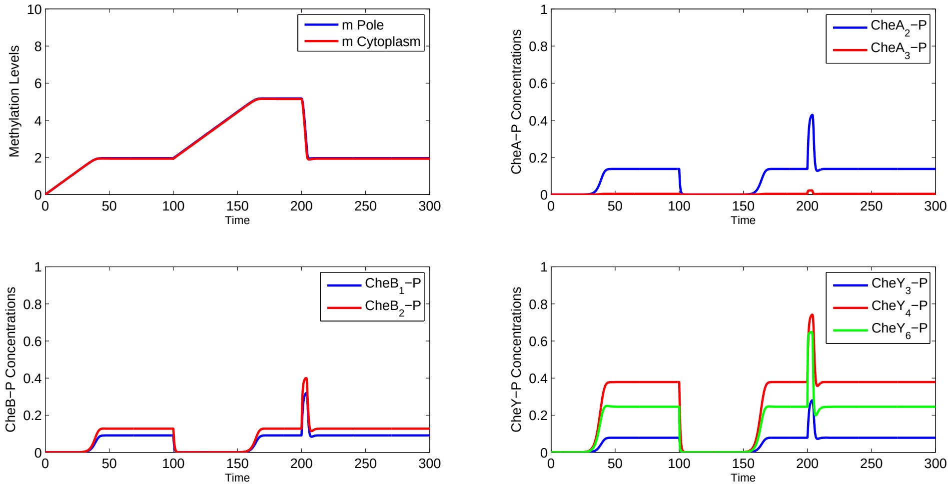
Within this section we seek to understand whether or not the model proposed in Section 8.4.1 (equations (8.18)-(8.25)) represents an improvement on that of Kojadinovic et al. [161] (see Figure 8.7). As such, here we present the results of numerical simulations representing a number of different biological scenarios. In particular, we consider numerical simulations of wild-type and ΔCheB$_2$ (CheB$_2$ deletion) mutant cells in response to both an external ligand signal and an internal metabolic signal.

In the case of an external stimulus, this is modelled as either an increase or decrease in the value of $[L]$. Results of this are displayed in Figure 8.8. Within these results it can be seen that both wild-type and ΔCheB$_2$ cell types each display similar responses. In particular, we note the similarity between results shown here and those of *E. coli* cells. This is perhaps unsurprising since the external ligand stimulus is sensed only at the polar cluster and adaptation is assumed to proceed using the same mechanism as *E. coli*.

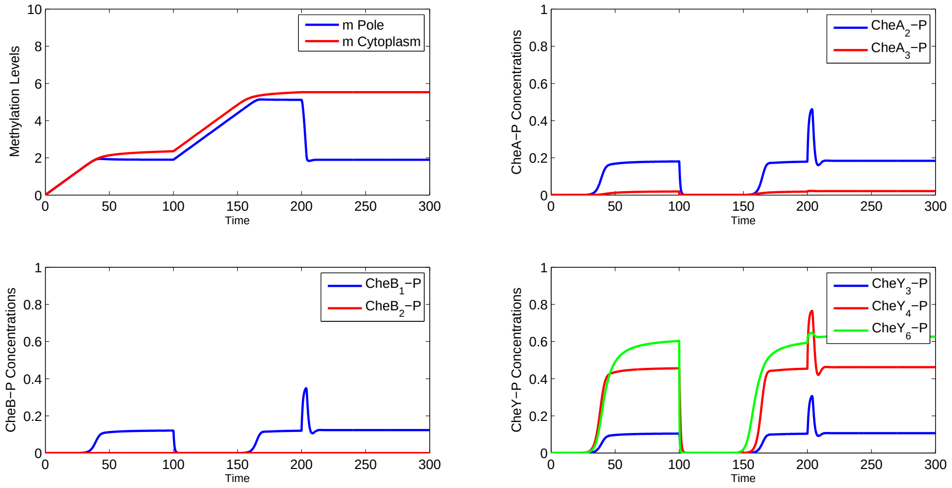
The response of a cell to some internal metabolic signal must be modelled differently to external signals. In order to do this we consider alterations to the value of Φ_C^* with an increase or decrease representing a positive or negative metabolic signal, respectively. Results obtained for wild-type and ΔCheB$_2$ cell types are given in Figure 8.9. Upon inspection of these results it is clear that both cell types display similar responses. In particular, it is noticeable that a positive metabolic signal results in an increase in the steady-state concentration of CheY$_6$-P. For the simulated wild-type cell an increase in the steady-state CheB$_2$-P concentration may also be seen.

Within the results of both Figures 8.8 and 8.9 we do not obtain the near saturation of CheY$_6$-P when protein CheB$_2$ is removed. Thus, the model considered here clearly appears to resolve this issue. In order to investigate the validity of this model yet further we consider how the steady-state concentration obtained from this model would impact upon the rotational behaviour of the cell's flagellum. Results of this are given in Figure 8.10.

We have already demonstrated that this proposed model eliminates the issue of CheY$_6$-P saturation in ΔCheB$_2$ mutant cells. In order to investigate this in greater detail we compare the results of Figure 8.10 to the experimental data of de Beyer [162] (Figure 7.2). Within this experimental data it can be seen that the flagellum of ΔCheB$_2$ mutant cells spends ∼40% of time in a stopped state. This differs only slightly from wild-type cells which spend ∼35% of time in a stopped state. Assuming that the flagellar motor control mechanism discussed in Section 8.3.3 (described by equation (8.17)) holds, the results of this model appear very positive.

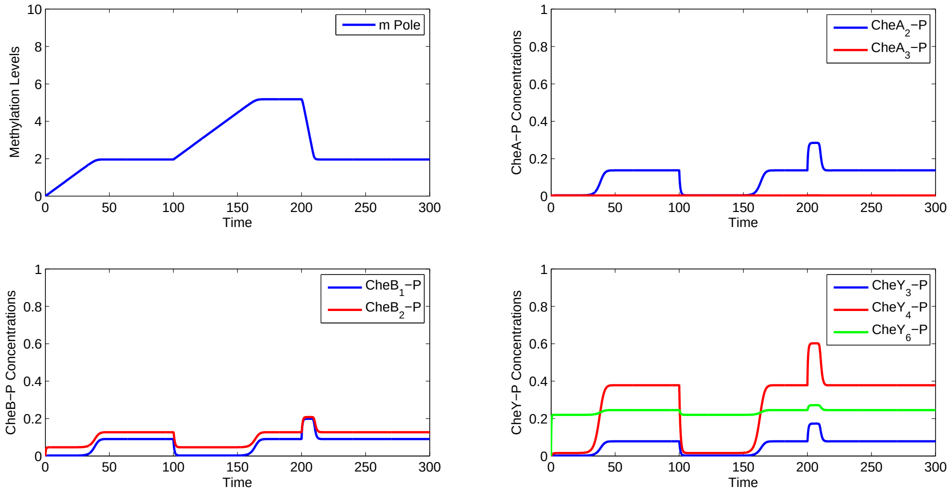


(a) Wild-Type

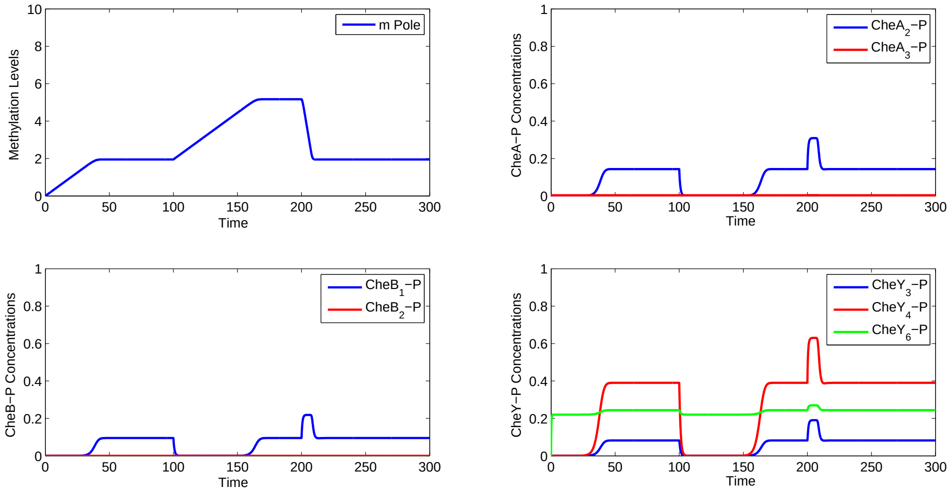


(b) ΔCheB$_2$

Figure 8.7: Plots showing numerical simulations of the Kojadinovic et al. [161] model when subjected to an increase (at $\tau = 100$) and a subsequent decrease (at $\tau = 200$) in the ligand concentration detected at both the polar and cytoplasmic clusters. Results are shown for (a) a wild-type cell and (b) a ΔCheB$_2$ mutant cell.

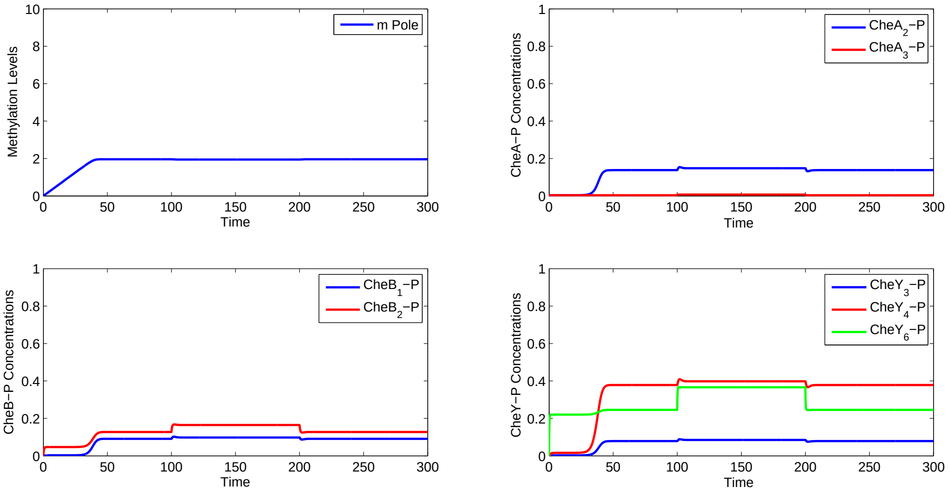


(a) Wild-Type

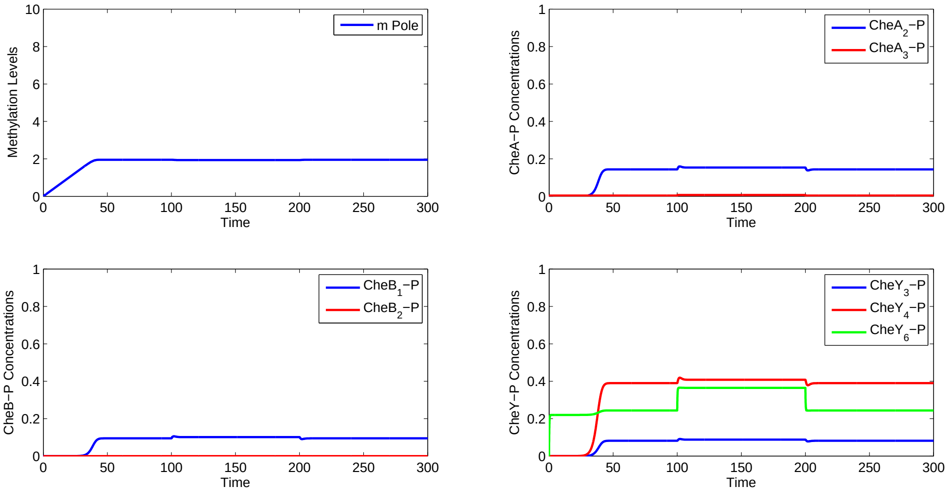


(b) ΔCheB$_2$

Figure 8.8: Plots showing numerical simulations of equations (8.18)-(8.25) when subjected to an increase (at $\tau = 100$) and a subsequent decrease (at $\tau = 200$) in the external ligand concentration. Results are shown for (a) a wild-type cell and (b) a ΔCheB$_2$ mutant cell. Here, each cell type displays a response similar to that of *E. coli* cells and does not display near saturation of CheY$_6$-P when CheB$_2$ is deleted.



(a) Wild-Type



(b) ΔCheB$_2$

Figure 8.9: Plots showing numerical simulations of equations (8.18)-(8.25) when subjected to an increase (at $\tau = 100$) and a subsequent decrease (at $\tau = 200$) in a metabolic signal. This is modelled as an increase in Φ_C^* from 0.345 to 0.545 and a subsequent decrease from 0.545 to 0.345. Results are shown for (a) a wild-type cell and (b) a ΔCheB$_2$ mutant cell. Here, each cell type displays a similar response in which the CheY$_6$-P steady-state is scaled according to the metabolic signal.

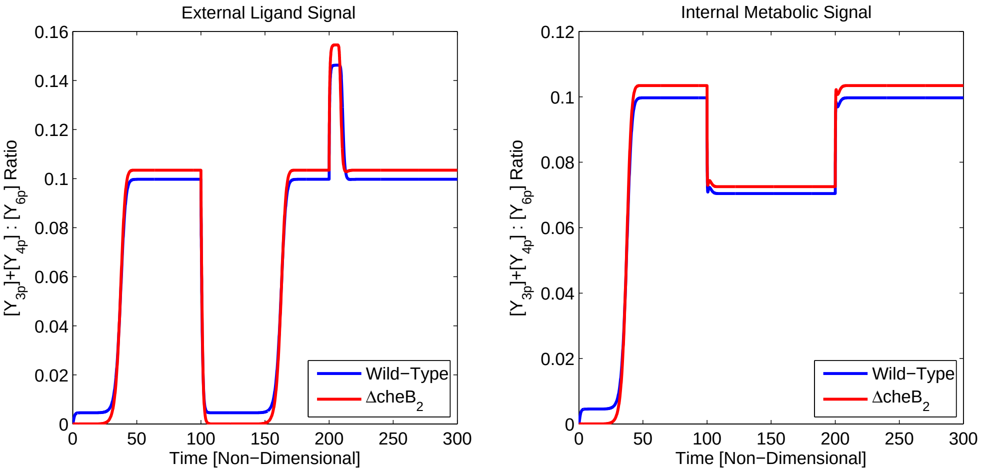


Figure 8.10: Plots showing CheY phosphorylation ratios (as per equation (8.17)) for wild-type (blue) and ΔCheB$_2$ mutant (red) cells. Values here are obtained from the numerical simulations of Figures 8.8 and 8.9.

In particular, Figure 8.10 shows that the deletion of CheB$_2$ increases the CheY phosphorylation ratio (equation (8.17)) by just a small amount. This increase appears to be of a magnitude suitable to produce the ~5% difference in time spent stopped by *R. sphaeroides* wild-type and ΔCheB$_2$ mutant cells.

The mathematical model proposed here clearly eliminates a number of issues associated with the model of Kojadinovic et al. [161]. In particular, this model produces three main improvements, namely

- the ability of the model to represent the effects of an internal metabolic signal sensed at the cytoplasmic cluster;
- elimination of CheY$_6$-P saturation when protein CheB$_2$ is deleted; and
- production of CheY phosphorylation ratios suitable to explain the ~5% difference in time spent stopped by the flagella of wild-type and ΔCheB$_2$ cells.

It would therefore appear that this proposed model represents an improvement on that of Kojadinovic et al. [161]. However, to further improve this model, more experimental work is required. In particular, it would be necessary to elucidate the metabolic signal sensed at the cytoplasmic cluster. This would then allow us to formulate an appropriate functional description of this process, thus improving the model yet further.

8.5 Summary & Discussion

Within this chapter both a reduced model of the *R. sphaeroides* cytoplasmic cluster and a simplified form of the Kojadinovic et al. [161] model, in which adaptation has been removed, have been examined. In doing so a number of features of the signalling pathway have been investigated. Specifically, steady state analyses of these two models have revealed the following.

Firstly, a simple ODE model of the *R. sphaeroides* cytoplasmic cluster was investigated. In doing so it was possible to propose roles for each of the processes that occur about this cytoplasmic cluster. Specifically it was proposed that

- autophosphorylation is essential for the cytoplasmic cluster to act in any way;
- the phosphatase action of CheA$_3$ and CheA$_3$-P on CheY$_6$-P acts to ensure the basal level of CheY$_6$ phosphorylation is maintained at a reasonable level;
- autodephosphorylation of CheY$_6$-P is not as effective as phosphatase action at limiting CheY$_6$ phosphorylation but gives an extra layer of robustness should the phosphatase action fail; and
- CheB$_2$ very slightly lowers the steady-state CheY$_6$-P concentration but is more likely to play a role in adaptation.

The simplified Kojadinovic et al. [161] model, in which adaptation is removed, was examined and steady-state results compared to experimental data. Specifically, separate versions of this model representing each mutant strain from the work of de Beyer [162] were formulated and their respective steady-states analysed. This revealed a number things.

Section 8.3.2 compared a model including reverse phosphotransfer from CheB$_2$-P onto CheA$_2$ to one from which it had been eliminated. This showed that removing the reaction from the network results in a large increase in the CheB$_2$-P concentration when the reverse phosphotransfer was removed. This would likely limit the ability of this protein to participate in any chemotactic response. Adding this to the fact that the reverse phosphotransfer reaction has been shown to occur *in vitro* led to the conclusion that it is most likely the process acting to limit the CheB$_2$-P concentration *in vivo*.

Further to this, a comparison of W-T and ΔCheY$_{3,4}$ mutant cell models was conducted. This was motivated by experimental data showing that these cell types displayed different chemotactic behaviour at steady-state [162]. Using the results in Table 8.3 we can suggest that the CheY$_3$-P and CheY$_4$-P concentrations are likely to account for the differences between their respective behaviours at steady-state.

In Section 8.3.3 a comparison of experimental data against model predictions of the steady-state behaviour was given. Using this comparison in addition to that of W-T and ΔCheY$_{3,4}$ strongly suggest that CheY$_3$-P and CheY$_4$-P must play some role in setting the flagellar rotation behaviour of *R. sphaeroides* cells. In support of this idea it was shown that the relative concentration of CheY$_6$-P against the cumulative concentration of proteins CheY$_3$-P and CheY$_4$-P displays a correlation between this ratio and the experimentally determined fraction of time cells spend in a stopped state. Fluctuations from the predicted relationship displayed in Fig. 8.5 are likely due to the simplifications considered within our modelling.

The results of Section 7.5 clearly demonstrated that the *E. coli* adaptation mechanism assumed by Kojadinovic et al. [161] to function at the cytoplasmic cluster of *R. sphaeroides* cells was not adequate. Thus, as discussed here, we investigated a number of features of *R. sphaeroides* chemotaxis signalling pathways with the aim of identifying ways in which the Kojadinovic et al. [161] model may be improved. This led us to propose a new mathematical model of *R. sphaeroides* chemotaxis. This model is based on that of Kojadinovic et al. [161] and retains the *E. coli* adaptation mechanism at the polar receptor cluster. However, the *E. coli* adaptation mechanism at the cytoplasmic cluster is substituted for a simple representation of internal metabolic sensing. Analysis of this model for wild-type and ΔCheB$_2$ cells showed CheY phosphorylation ratios consistent with the experimentally observed $\sim$5% difference in the proportion of time a cell's flagellum spends in a stopped state. It therefore appears that the model postulated here produces a better fit to experimental data than that of Kojadinovic et al. [161]. However, further work is necessary to allow a functional form to be derived for the sensing of metabolic signals at the cytoplasmic cluster of *R. sphaeroides* cells.

Chapter 9

Discussion

The main achievements of the work in this thesis have been the analyses of ODE models of *E. coli* and *R. sphaeroides* chemotaxis signalling pathways.

In particular, analysis of the *E. coli* ODE model system revealed key features associated with the phenomena of fold-change detection (FCD) and overshoot. These studies illustrated the importance of receptor dynamics, methylation and phosphorylation timescales as well as total protein concentrations. Further to this, agent-based modelling allowed an investigation in to the effects of these features/phenomena upon the overall behaviour of a cell population to be investigated.

Analysis of the *E. coli* chemotaxis signalling network equipped us with a number of tools that were used in order to study the more complex *R. sphaeroides* system. This allowed us to assign roles to a number of signalling proteins as well as identifying areas that require improvement within future work.

Here we summarise the main findings of this work in addition to giving areas for future theoretical and experimental work. This will help to further elucidate the signalling cascade characteristics of *E. coli* and *R. sphaeroides* as well as other biological systems.

9.1 Summary & Conclusions

In Chapter 2 we provided a summary of the past literature most relevant to the work contained within this thesis. In particular, we began by summarising the biological work that sought to identify key components of chemotaxis signalling cascades and their respective roles. Additionally, throughout this chapter we focused on how experimental results inspired mathematical modelling and vice versa, leading to the modern understanding of bacterial chemotaxis. Due to their importance throughout this thesis, aspects of chemotaxis signalling cascades including the roles of various signalling proteins, chemoreceptors, methylation and phosphorylation were explored in addition to their relevance to signal transduction, adaptation and the flagellar motor response.

There have been a number of mathematical models formulated that have sought to elucidate key features of chemotactic responses of single *E. coli* cells. One such model, namely that of Clausznitzer et al. [70], is analysed within Chapter 3. In particular, we began by comparing different offset energy functions to the available experimental data in order to produce the best possible model fit. Subsequent to this, the model was subjected to a rigorous mathematical analysis. In particular, it was shown that this mathematical model may display just one biologically realistic steady-state. The asymptotic stability of this equilibrium state was then conducted, demonstrating that it is asymptotically stable with eigenvalues that demonstrate that it is a stiff model system.

The mathematical model laid out in Chapter 3 was then utilised for investigations into a number of features observed within the experimental literature relating to the chemotactic response of single *E. coli* cells.

One such feature was that of FCD which has previously been analysed in both experimental and theoretical settings. In Chapter 4 we considered the sufficient conditions given by Shoval and colleagues [121, 123] alongside experimentally determined parameter values in order to demonstrate that the model from Chapter 3 does indeed display FCD. Alternative sets of sufficient conditions are then considered in order to predict that there should exist more than the two experimentally observed regimes of FCD behaviour. We went on to show that FCD is a robust property so long as all parameter values remain constant in time.

In Chapter 5 mechanisms associated with the emergence of overshoot behaviour were investigated. In order to do so we observed a negative feedback loop in which CheB-P reduces the methylation level of receptors. We thus tested the hypothesis that overshoot is the result of damped oscillations caused by this negative feedback loop. Within this investigation we identified cell-to-cell variation in total intracellular protein concentration as being essential for the emergence of this phenomenon. This cell-to-cell variation creates differences in a number of different intracellular features. In particular, we were able to identify a balance between chemoreceptor dynamics and CheB-P feedback as important in causing overshoot. Results obtained within this chapter were then compared to experimental data showing that they produce a good agreement for small to intermediate step-up ligand stimuli (i.e. $\leq 50\mu$M) and the full range of step-down stimuli (i.e. 1μM to 1mM).

Based on the investigations within Chapters 4 and 5 it is clear that the mathematical model of Clausznitzer et al. [70] is both capable of explaining significant amounts of experimental data and being used for predictive work. As such, within Chapter 6 we built this model into an ABM framework in order to investigate *E. coli* population level phenomena. Using the ABM we were able to extrapolate single cell findings to study the effect of variation in protein concentration on the efficiency of the chemotactic response within ligand gradients of varying steepness. Also considered here was the ability of cells to perform chemotaxis when two separate ligand gradients are present simultaneously. This revealed that cells will accumulate about the ligand concentration to which they are more sensitive and that this may not necessarily be the largest absolute concentration.

Chapters 3 to 6 investigated features associated with chemotaxis in *E. coli* cells. The remaining chapters focused on testing a recent single cell model of *R. sphaeroides* chemotaxis against experimental data and using the results to help elucidate new understanding of the signalling cascade and mechanistic function of the cell.

In Chapter 7 we analysed the recent mathematical model of the *R. sphaeroides* chemotaxis signalling network due to Kojadinovic et al. [161]. It was shown that the model exhibits one biologically feasible steady-state which was shown to be asymptotically stable. A number of variations of this model representing mutant cell types finally showed that this model fails to capture the steady-state behaviour observed in recent experimental data. In particular, it was clear that the cytoplasmic cluster adaptation mechanism considered here represented the main failing of the model, as evidenced by the high $\mathrm{CheY_6}$-P levels obtained when $\mathrm{CheB_2}$ is removed from the system. Within these results we could observe that the cytoplasmic cluster produces the greatest variation from the experimental data. It is clear therefore that this model has been invalidated and as such we required alternative approaches in order to gain further knowledge of this species.

Due to the invalidation of the Kojadinovic et al. [161] model in Chapter 7 we proceeded to formulate and analyse two simplified mathematical models within Chapter 8. These were utilised in order to gain further understanding of the *R. sphaeroides* chemotaxis signalling system.

The first of these simplified models was a three ODE model of the cytoplasmic cluster. This allowed us to identify the role of each reaction. In particular, it was found that in order to regulate the phosphorylated concentration of $\mathrm{CheY_6}$-P, autodephosphorylation and phosphatase action are required. This phosphatase action was shown to be far more effective than autodephosphorylation at regulating $\mathrm{CheY_6}$-P levels, although this is likely to represent an additional layer of regulation should the phosphatase action fail. In addition to this it was shown that $\mathrm{CheB_2}$ cannot act as a phosphate sink due to the small concentration of this relative to $\mathrm{CheY_6}$.

Further to the cytoplasmic cluster model, we also considered a non-adapting version of the model in Chapter 7 in order to investigate two further questions. Similar to earlier work within this thesis, a number of variants of this model were considered representing mutant cell types used within a recent experimental study (neglecting those relating to CheR and CheB proteins associated with adaptation). Within this work it was possible to show that a model without the reverse phosphotransfer reaction from $CheB_2$-P onto $CheA_2$ that has been demonstrated *in vitro* leads to near saturation of $CheB_2$-P. It is clear therefore that some regulatory process is required and due to the *in vitro* evidence for this reverse phosphotransfer, it would seem likely that this is indeed the regulatory mechanism acting *in vitro*. The same model was then used in order to prove that $CheY_3$ and $CheY_4$ must play some role in setting flagellar rotation. In addition to this it is shown that the ratio of $CheY_3$-P and $CheY_4$-P to $CheY_6$-P produced a correlation with the proportion of time flagellar motors spend in a stopped state. Using the work in Chapters 7 and 8 we postulate a new signalling model for *R. sphaeroides* and show that it helps remove some issues associated with previous models.

Using the results of Chapters 7 and 8 we then proposed a new mathematical model of chemotaxis in *R. sphaeroides* within Section 8.4. This model was based upon that of Kojadinovic et al. [161], however it incorporated the simplified cytoplasmic cluster model studied in Section 8.2.1 that allowed for the consideration of intracellular metabolic stimuli. Results obtained from this model revealed two distinct response types. In particular, for an extracellular ligand stimulus, the cell displays a response similar to that observed in *E. coli*. In contrast to this, intracellular metabolic signals acted to tune the steady-state phosphorylation levels of $CheB_2$ and $CheY_6$. Each of these responses appears in line with experimental data. The main improvement of this model however is that it removes the issue of near saturation of $CheY_6$-P upon the deletion of $CheB_2$.

Whilst it is clear that we have identified a number of features associated with *R. sphaeroides* chemotaxis signalling pathways, there is still much work to do. In particular, it is necessary to produce both extra experimental and theoretical work in order to identify the exact mechanisms associated with adaptation at the polar and cytoplasmic clusters of chemotactic *R. sphaeroides* cells.

9.2 Future Work

In this section we shall discuss a number of ideas for future work. These suggestions will cover a number of different areas, each of which would potentially complement, further or verify work contained within this thesis.

9.2.1 Modelling Assumptions

Within this thesis a number of different mathematical models have been formulated and analysed. As is commonplace in mathematics, these are based on a number of assumptions and simplifications that allow for the production of tractable models. In doing so it should be clear that to some degree the use of such simplifications/assumptions will necessarily cause differences between the model and reality. We shall now discuss the main assumptions associated with models contained in this thesis alongside methods that would allow the effects of these to be studied.

The mathematical models utilised within this work are comprised of systems of deterministic nonlinear ODEs, most of which are formulated using the law of mass action. In utilising this method for deriving mathematical models, a number of assumptions are inherently made. In particular, the two main assumptions associated with the law of mass action are that all reactions proceed in a continuous manner (i.e. concentrations and intracellular space are large enough to avoid significant stochastic effects) and that the environment is well mixed (i.e. concentrations are spatially homogeneous).

In the context of the chemotaxis signalling pathways of *E.coli* and *R. sphaeroides* cells, theoretical literature has demonstrated that models created using the law of mass action provide a good fit with available experimental data. This would suggest that the models used here are valid for comparison with experimental observations as well as for more predictive work. It would however be possible to create alternative mathematical models that are able to capture more complex features and as such examine how these affect the behaviour of chemotactic bacteria.

The continuum models utilised in this work provide a good approximation of general behaviour within the relevant phosphorylation cascades. However, it is clear upon examination of the experimental literature that reactions in these networks do not occur in a purely deterministic manner, i.e. they are stochastic. Whilst the models utilised appear suitable for the purposes considered here it would be useful to consider a stochastic model of, in particular, the *E. coli* chemotaxis signalling pathway [26, 30, 129, 155, 180]. It would then be possible to build this into our ABM framework in order to assess the importance of stochasticity within intracellular chemotaxis signalling cascades.

The other main assumption of the law of mass action is that the environment is well mixed. In the case of *E. coli* in particular, this assumption has been shown to be a good approximation. For *R. sphaeroides* cells it is less clear how important spatial inhomogeneity may be in their chemotactic response. One example of the spatial inhomogeneity observed within *R. sphaeroides* cells are the proteins CheA$_3$ and CheA$_4$ that are both localised about the polar receptor cluster. In order to understand how spatial organisation such as this affects the cell's chemotactic response, it is possible to use a PDE model. This would allow for the consideration of protein diffusion within the cytoplasm as well as restriction of certain components to specific regions of the cell, similar to the work of Rao et al. [181].

Within the formulation of models such as those considered here, it is assumed that phosphotransfer reactions occur as instantaneous events. In reality however, such reactions occur in a few stages. Initially, one phosphorylated and one non-phosphorylated protein will form a complex. This exists only transiently before they split apart, with the phosphoryl group exiting attached to the other protein. Following the approach of Tindall et al. [182] it is possible to formulate a model that considers a more complete representation of each step in phosphotransfer reactions. We could then examine the ways in which this intermediate complex formation affect the overall chemotactic response and whether there is any impact upon the overshoot mechanism discussed within this work. Conducting a similar analysis to that presented in Chapter 5 would be more difficult however since such a model would result in much higher dimensional ODE models.

The last main assumption made in formulating models used throughout this work is that we are only interested in events occurring on a relatively short timescale. This allows us to neglect protein synthesis since this occurs on a longer timescale. Supposing we wished to simulate behaviour over longer periods of time then we would be required to consider how proteins are synthesised and degraded. This would involve incorporating a genetic model describing how DNA is transcribed into mRNA and subsequently translated into protein molecules.

The models used within this thesis are clearly based on a number of assumptions. We would not anticipate that these would significantly alter the conclusions of this work, however a number of techniques that could be used in order to assess the effects of these have been proposed above. Further to this, a number of limitations associated with this work shall be discussed alongside techniques/methods that could be utilised in order to address these.

9.2.2 *E. coli*

In Chapter 5 the phenomenon of overshoot in the chemotactic response of *E. coli* cells was examined. This work noted that, in spite of producing a good fit to experimental data, our results were unable to capture the negligible overshoot amplitudes observed for large step-up ligand stimuli. This would appear to be a likely consequence of the phenomenon whereby receptor clusters break apart when subjected to large, saturating stimuli. As previously stated, this has not featured within the theoretical literature and as such it would be an interesting area for future modelling work. We would anticipate that this would extend upon the work of Endres et al. [62] who gave a functional form describing the size of a receptor cluster in different ligand environments. Upon doing so, this mechanism could then be incorporated into the methodology used within this thesis in order to re-examine the proposed overshoot mechanism. Results could then be compared against the experimental literature.

Chapter 6 formulated an ABM and utilised this in order to study population level behaviour. This followed a similar methodology to a number of hybrid models within the literature [109, 183, 184, 185]. Within this work it was assumed that the ligand profile was temporally homogeneous. In experimental settings this is unlikely to be the case since ligand is likely to diffuse as a result of spatial heterogeneity and certain ligands would be metabolised by cells. In order to formulate a more detailed ABM we could define a PDE describing the time evolution of the ligand profile.

It would seem simple enough to state that for ligand diffusion we must solve a PDE of the form

$$\frac{\partial [L]}{\partial t} = D\nabla^2 [L], \tag{9.1}$$

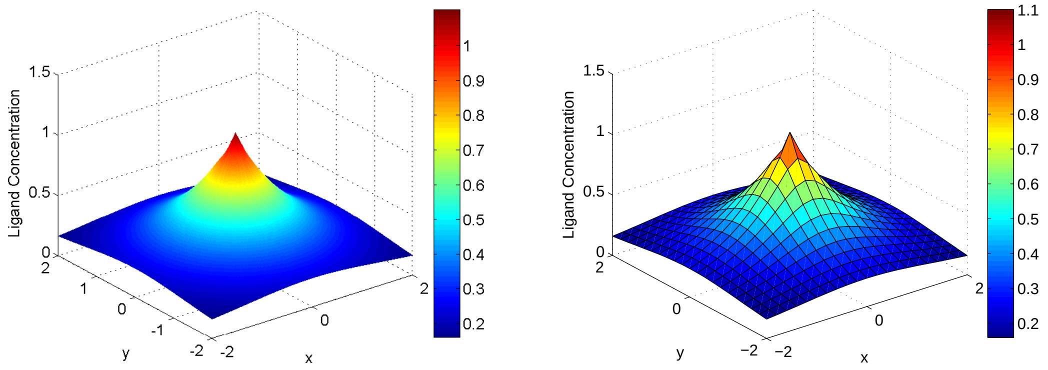
however this does not do justice to the computational complexities that must be addressed in order to do so. In the first instance it would be necessary to define a mesh over which the numerical solution of the ligand PDE may be calculated. As such we no longer have a ligand concentration defined at each point within the domain. Thus we would need to use some interpolation method in order to approximate the ligand concentration at the location of each simulated cell. Some initial results show that for an exponential ligand profile (see equation (6.2)), application of bilinear interpolation (over a 21×21 grid) results in a maximal error of $<2\%$ as shown in Figure 9.1. Allowing an extension of this type would allow modellers to create more accurate simulations of experimental situations.

When seeking to produce a model with full biological detail, the addition to our ABM of ligand diffusion is not necessarily sufficient. In particular, certain chemoattractants have been shown to be metabolised by cells and some may naturally degrade over time, thus reducing the ligand concentration. Each of these two processes (metabolism and degradation) could feasibly be incorporated into the ABM. By generalising equation (9.1) to a reaction-diffusion equation of the form

$$\frac{\partial [L]}{\partial t} = D\nabla^2 [L] - \rho [L] - \zeta, \tag{9.2}$$

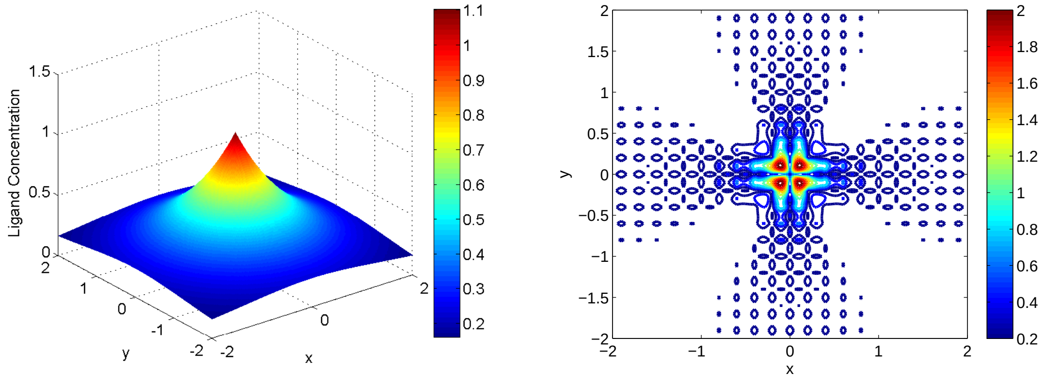
where $\rho(\eta)$ is the rate of ligand degradation, $\zeta(\eta)$ is the rate of ligand metabolism by cells and η is the number of cells at/close to a given spatial location. Once again, this would require interpolation methods as well as experimental data on the rates of ligand degradation and metabolism in order to produce a realistic model.

The final limitation associated with the ABM in this work is the amount of computational time required to produce each simulation. In order to illustrate this issue, a simulation of 50



(a) Exponential ligand profile

(b) 21×21 grid representation



(c) Interpolated profile

(d) Percentage error

Figure 9.1: An example of results obtained from application of bilinear interpolation. (a) The exponential ligand profile that we wish to capture using bilinear interpolation. (b) A representation of the ligand profile in (a) created on a 21×21 mesh. (c) The ligand profile resulting from application of bilinear interpolation to the rough mesh representation in (b). (d) The percentage error resulting from application of bilinear interpolation. This is calculated as the difference between results in panels (a) and (c) expressed as a percentage of the exact value in panel (a).

cells in a fixed, exponential ligand field over a time period of ∼12 minutes requires ∼45 minutes to run across a pool of three processor cores (referred to as workers in the MATLAB parallel computing toolbox). As can clearly be seen in Chapter 6, for each result we generally require a reasonable number of separate simulations, thus the time required can begin to add up rather significantly. It would therefore be sensible, if seeking to use this ABM for more complex work, to first optimise the code in order to reduce the necessary computational time.

There are a number of adjustments that could potentially make the ABM code more time efficient. A number are summarised below.

- **ODE solution:** The numerical solver ode15s is used in order to progress the signalling cascade ODE model within our ABM model code. For a ∼12 minute simulation of 50 cells, the ABM currently requires information be passed in and out of the ode15s solver function ∼2.5 million times. Since the solver has been restricted to taking one time step on each function call, it may be faster to directly include a finite-difference (or other method) representation of the ODE model within the code itself. As such, the model would no longer need to pass values in and out of the ode15s function, potentially saving some time, especially when scaled up over a large simulation.

- **Save sparingly:** In many numerical codes, passing information to and from system memory can be a time consuming process. As such, within the ABM code utilised here it would be possible to reduce the number of pieces of data that are passed into and out of system memory as well as reducing the amount of data that is permanently stored. At present the ABM code stores each of the system variables at each time step. Thus a significant reduction in running time could possibly be achieved by storing just one previous time step and using this to update results. In addition to this, permanently storing only the output variable of interest at each time step would significantly reduce the total amount of data that is permanently stored.

Whilst the above alterations could increase the efficiency of our ABM, there may be other slightly less conventional methods that could be considered. One such example is that of Flegg and colleagues [186, 187] who utilise a two-regime method. Within this method a spatial domain is split into two halves, each of which uses a different solution. One half of the domain uses an individual-based Brownian motion simulation while the other makes use of a lattice based compartmental model. The two solutions are then matched according to a set of rules at the domain interface. These two methods each produce different levels of detail and differ significantly in their computational costs. An approach such as this could potentially simplify our ABM by allowing a less detailed model to be used in regions of less interest while a full

ABM simulation is completed for some region of interest. This has the potential to significantly reduce computational costs; however a significant amount of work would be required to allow the matching of results between the two different simulation regions/regimes.

9.2.3 *R. sphaeroides*

In Chapters 7 and 8 of this thesis, the chemotaxis signalling pathways of *R. sphaeroides* cells were our focus. In spite of being able to show that the ODE model of the phosphotransfer network appears, in the most part, to re-produce experimental data the adaptation mechanisms included in the modelling work were shown to be inadequate. As such, a new modelling approach is required in order to fully explain this data. There are a number of ways in which a new mathematical model of adaptation could be formulated, a few of which shall be briefly summarised here.

The first approach that could be taken mimics early models of *E. coli* ligand binding and adaptation. This would aim to create a mechanistically complete description of receptor state for each receptor cluster. In order to do this a model would be formulated that takes into account ligand binding and the relevant number of methylation states available for each receptor cluster. In the case of the polar receptor cluster where there are known to be 4 methylation sites this approach would yield a model consisting of 20 possible states and therefore 20 ODEs. The total cluster activity would also need to be calculated and this is likely to take the form of a sum of each receptor state balanced by the relative contributions of each. As discussed earlier in this thesis, mathematical models of this type can be difficult to analyse due to their sheer size. Another issue is models such as this require large numbers of parameter values. These must either be obtained through mathematical parameter fitting techniques or experimentation, both of which can take significant amounts of time. Once such a model has been formulated and parameterised, it could be possible to apply a number of different model reduction techniques. In many cases such methods are capable of reducing large ODE systems into much smaller systems that maintain the input-output profile within a small error tolerance.

In addition to the mechanistic description described above it would also be possible, given further experimental data, to produce a description of the average methylation of each receptor cluster in a similar manner to that considered by Clausznitzer et al. [70] for *E. coli* cells. In particular it would be possible to observe that

$$\frac{d\Phi}{dt} = \frac{\partial\Phi}{\partial m}\frac{dm}{dt} + \frac{\partial\Phi}{\partial[L]}\frac{d[L]}{dt} = \frac{\partial\Phi}{\partial m}\frac{dm}{dt}, \tag{9.3}$$

at the time immediately after a step change in the ligand concentration is applied (i.e. where

$d[L]/dt = 0$). Using currently available experimental techniques it is possible to experimentally determine both $d\Phi/dt$ and $\partial\Phi/\partial m$. As such, we could then formulate a functional description of the methylation/demethylation ODE (dm/dt) that provides a good fit to the experimental data.

9.2.4 Experimental Work

We have so far discussed a number of assumptions and limitations associated with the work in this thesis. Another key area in which further work would help to improve our knowledge of chemotaxis signalling is through comparison of results from this work with theory and new experimental work. For the remainder of this section we shall briefly mention some key areas in which further work would be beneficial and outline proposals as to how this work may be undertaken.

A number of predictions made within this work could be tested using experimental methods. Some of these have been discussed within the relevant chapters/sections however here we shall briefly summarise some of these predictions and the methods that could be used to verify (or indeed invalidate) them.

- **FCD regimes:** It was predicted in Chapter 4 that a system of n receptor types with well separated ligand dissociation constants would display $2n + 1$ regimes of FCD behaviour. In the case of *E. coli*, it was shown that just three of these five predicted regimes would be observed due to the cell's maximum detectable ligand concentration. Our work has shown good agreement with the two experimentally observed FCD regimes, leaving just one unverified. Using the same FRET approach as Lazova et al. [122] it should be possible to assess whether this third regime exists.

- **Overshoot:** A number of theoretical predictions have been made in this thesis relating to the occurrence of overshoot behaviour. Section 5.8 discussed experimental methods that could be used to test these. In particular, for Figure 5.6 it is possible to create cells under/overexpressing certain proteins; Figure 5.7 needs the *meche* and *mocha* operons to be under/overexpressed whilst Figures 5.12 and 5.13 would require cells under/overexpressing all chemotaxis signalling proteins simultaneously. In each example analysis of flagellar rotation time-courses, for the under/overexpression mutant cells, in response to step-changes in ligand concentration should reveal the magnitude of overshoot exhibited by cells (if any).

- **Protein concentration effects:** Using our ABM a number of population level phenomena were observed as a result of variation in total concentration of chemotaxis signalling proteins between cells. In order to test these results it would first be necessary to create cells in which the total concentration of all signalling proteins may be controlled. In addition to this it would be necessary to create a well-controlled ligand gradient that very closely matches those in our ABM simulations. These newly created cells could then be placed into the controlled ligand gradient. Observing their behaviour over time, it should then be possible to compare the migration of cells in this experimental work to our ABM simulations, thus either confirming or invalidating this work.

In conclusion, a number of phenomena associated with chemotaxis of *E. coli* cells have been investigated. It is hoped that the modelling framework/methodology, observations and predictions within this work will be useful to future work on bacterial chemotaxis and other biological systems. In the case of *R. sphaeroides* a number of areas have been identified that would benefit from further theoretical and experimental work. In addition to this, a new model was postulated that helps to remove some of the issues of previous models. As such it is hoped that this work proves a useful step in providing a more complete understanding of chemotaxis in *R. sphaeroides* as well as signalling networks (chemotaxis or otherwise) of other bacterial species.

Appendix A

Law of Mass Action

The law of mass action is an example of a kinetic formula. It states that a reaction will proceed at a rate that is proportional to the product of the concentrations of participating species. This is often used in the formulation of mathematical models in the field of mathematical biology and is based upon a number of key assumptions that will be discussed shortly.

Here two examples are considered in order to demonstrate how the law of mass action works to represent a set of reactions as a system of ODEs. The first such example considers the case whereby an enzyme interacts with a substrate resulting in the creation of a product. This reaction ($\longrightarrow$) may be represented schematically as

$$[A] + [B] \ \xrightarrow{k} \ [C], \tag{A.1}$$

within which [...] indicates the concentration of the relevant species with A representing a substrate, B an enzyme, C the product and k the constant kinetic rate parameter associated with this reaction. Now, using the law of mass action, we may represent this reaction as a set of ODEs of the form

$$\frac{d[A]}{dt} = -k[A][B], \tag{A.2}$$

$$\frac{d[B]}{dt} = -k[A][B], \tag{A.3}$$

$$\frac{d[C]}{dt} = k[A][B]. \tag{A.4}$$

Further to the simple example above, we now consider the example of Michaelis-Menten reaction kinetics [188]. In this example an enzyme interacts with a substrate to form a complex. This complex then separates giving back the enzyme and converting the substrate into a product. This reaction mechanism is a more complex example than that given above since it requires the formation of a complexed state as well as a reversible reaction ($\rightleftharpoons$). Schematically this may be written as

$$[S] + [E] \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} [SE] \xrightarrow{k_2} [P] + [E], \tag{A.5}$$

within which S denotes a substrate, E an enzyme, SE is a complex formed of S and E, P is a product and k_i $(i = 1, -1, 2)$ are the constant kinetic rate parameters associated with these reactions. The law of mass action may then be applied to this reaction mechanism in order to produce a set of ordinary differential equations of the form

$$\frac{d[S]}{dt} = -k_1[E][S] + k_{-1}[SE], \tag{A.6}$$

$$\frac{d[E]}{dt} = -k_1[E][S] + (k_{-1} + k_2)[SE], \tag{A.7}$$

$$\frac{d[SE]}{dt} = k_1[E][S] - (k_{-1} + k_2)[SE], \tag{A.8}$$

$$\frac{d[P]}{dt} = k_2[SE]. \tag{A.9}$$

The law of mass action was formulated in order to model reactions that are subject to a specific set of conditions. These conditions correspond to the associated assumptions and relate to the speed of reaction, the environment in which reactions occur as well as the amounts of each reactant. Each of these assumptions imposes some restrictions which must be considered before the law of mass action is used in modelling any biological system.

Firstly, we must consider the time taken for each reaction to occur. The law of mass action generally corresponds to a reaction that occurs instantaneously. An example of this would be that an enzyme and substrate act to form a product instantaneously when a collision between the two occurs. This simple case corresponds to example (A.1).

The law of mass action as stated above applies to reactions that occur in a well mixed environment. Specifically this refers to the fact that the environment in which a reaction occurs contains homogeneous concentrations of all relevant species, i.e. the concentration of each reactant is equal at every point within the environment, hence any spatial features may be neglected.

Finally, it is assumed within the definition of the law of mass action that reactions proceed in a continuous, deterministic manner. Reactions of the type modelled by the law of mass action generally occur via collisions between reactant components. These can be considered as deterministic processes in situations whereby the concentrations of each reactant are sufficiently high, i.e. the continuum limit holds. In cases whereby the concentrations of reactants are low, stochastic effects become more prevalent (i.e. the continuum limit breaks down) and as such the law of mass action becomes invalid.

Appendix B

Descartes' Rule of Signs

Descartes' rule of signs may be used to investigate solutions of polynomial equations. More specifically it may be used to find the number of positive, negative and complex roots that can possibly be obtained from a polynomial equation.

In order to utilise this rule it is first necessary to arrange the polynomial equation in order from highest to lowest power of the relevant variable, for example

$$p(x) = a_1 x^3 \pm a_2 x^2 \pm a_3 x \pm a_4 = 0, \tag{B.1}$$

within which x is the variable of interest and a_i $(i = 1, 2, 3, 4)$ are the polynomial coefficients expressed such that every $a_i > 0$, i.e. the relevant signs should be expressed in the equation rather than hidden in the coefficient definitions. It is then possible to apply Descartes' rule of signs in two main forms (Rule 1 and Rule 2), which are as follows.

- **Rule 1:** For a polynomial equation $p(x)$ with real coefficients, the maximum number of positive real roots of this equation will be equal to the number of sign changes in the coefficients (with zero coefficients not constituting a sign change). Then, the number of positive real roots will either be equal to this maximum or will be less than this by some multiple of two depending upon the number of complex conjugate pairs of roots of the polynomial.

- **Rule 2:** For a polynomial equation with real coefficients the maximum number of negative real roots may be determined by examining the number of sign changes in the polynomial equation $p(y)$ where $y = -x$. Then the number of negative, real roots of the polynomial $p(x)$ is equal to the number of coefficient sign changes of the polynomial $p(y)$ or less than this by some multiple of two depending upon the number of complex conjugate pairs of roots of the polynomial. Once again, a zero coefficient does not constitute a sign change.

- **Rule 3:** It is also sometimes possible to obtain some information about the minimum number of complex roots of a polynomial equation of the form $p(x)$. If this polynomial does not have a zero root then there will be at least $n - (q + r)$ complex roots, where q is the maximum number of positive roots and r is the maximum number of negative roots, which may be determined using the two forms of Descartes' rule of signs explained above.

B.1 Example 1

In order to demonstrate the application of Descartes' rule of signs we here consider an example of the form

$$p(x) = a_1x^3 - a_2x^2 - a_3x + a_4 = 0. \tag{B.2}$$

Making use of Rule 1 we can clearly see that there exist two changes of sign in the polynomial coefficients. As such there will be a maximum of two positive real roots. Thus, here we may have either two positive, real roots and zero complex roots or alternatively there may be zero positive, real roots and two complex roots.

In order to use Rule 2 we then re-write equation (B.2) such that it takes the form

$$p(y) = p(-x) = -a_1x^3 - a_2x^2 + a_3x + a_4 = 0. \tag{B.3}$$

In this case we can clearly see that there exists just one sign change in the polynomial coefficients meaning that there will be one negative, real root to the polynomial equation (B.2).

B.2 Example 2

Suppose we now consider the same example as that in Example 1, but with the polynomial coefficients altered such that $a_2 = a_3 = 0$, i.e. we have

$$p(x) = a_1x^3 + a_4 = 0. \tag{B.4}$$

In this example we can clearly see that there exist no sign changes in the polynomial coefficients and so there will be zero positive, real roots.

If the polynomial equation (B.4) is then re-written such that it becomes

$$p(y) = p(-x) = -a_1x^3 + a_4 = 0. \tag{B.5}$$

Inspecting this expression it is possible to see one change in sign in the polynomial coefficients and so there will exist one negative, real root to the polynomial equation (B.4).

In order to use Rule 3 we may then see that $q = 0$ and $r = 1$, from above. We may then calculate that there will be two complex roots since $n - (q + r) = 3 - (0 + 1) = 2$.

Appendix C

Asymptotic Stability Analysis

Asymptotic stability analysis is a technique utilised in order to examine the stability of solutions to a system of differential equations and the relevant trajectories of such systems under small perturbations. In particular, asymptotic stability analysis may be used to investigate the nature of a dynamical system's behaviour in the vicinity of a steady-state.

In order to conduct an analysis of this type we define a system of ODEs in the form

$$\frac{d\boldsymbol{x}}{dt} = \dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}), \tag{C.1}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, with n denoting the dimensionality of the system and

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{and} \quad \boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} f_1(x_1, ..., x_n) \\ \vdots \\ f_n(x_1, ..., x_n) \end{pmatrix}. \tag{C.2}$$

Before we seek to investigate the system behaviour in the vicinity of some steady-state (equilibrium), we must obtain an expression for the steady-state of interest which is denoted $\boldsymbol{x}^*$. Once this equilibrium point has been obtained we may take a small perturbation to this steady-state of the form $\boldsymbol{\xi}(t) = (\xi_1, ..., \xi_n)^T$, which can be used to linearise about the equilibrium point. This is carried out by expanding $\boldsymbol{f}$ in a vector-valued Taylor series of the form

$$\dot{\boldsymbol{x}}^* + \dot{\boldsymbol{\xi}} = \boldsymbol{f}(\boldsymbol{x}^* + \boldsymbol{\xi}) = \boldsymbol{f}(\boldsymbol{x}^*) + \boldsymbol{J}\boldsymbol{\xi} + ... \ , \tag{C.3}$$

in which ... denotes the addition of higher order terms and $\boldsymbol{J}$ is the $n \times n$ Jacobian matrix of partial derivatives $\partial f_i / \partial x_j$, evaluated at the equilibrium point $\boldsymbol{x}^*$.

In order to assess the asymptotic stability of the relevant equilibrium point, we solve for the eigenvalues by evaluating

$$\det |\boldsymbol{J} - \lambda \boldsymbol{I}| = 0, \tag{C.4}$$

in which λ are the eigenvalues of the linearised system and $\boldsymbol{I}$ indicates the $n \times n$ identity matrix. The nature of these eigenvalues tell us about the qualitative behaviour of our dynamical system in the vicinity of the relevant equilibrium state.

All eigenvalues of this Jacobian matrix may be considered as the addition of a real and an imaginary part. The magnitude of the real component of the eigenvalues are used to determine the stability of the equilibrium point. In particular there are three main possibilities:

- $\lambda < 0$: In cases whereby all of the eigenvalues possess strictly negative real parts then the system will be asymptotically stable. This will be the case since the small perturbation will, in time, shrink to zero, i.e. $|\boldsymbol{\xi}(t)| \to 0$ as $t \to \infty$.

- $\lambda > 0$: Supposing the system possesses one (or more) eigenvalue(s) with a real part strictly greater than zero then the system will be asymptotically unstable. In such cases the small perturbation will grow in time, i.e. $|\boldsymbol{\xi}(t)| \to \infty$ as $t \to \infty$.

- $\lambda = 0$: The final possibility here is that the real components of the eigenvalues may equal zero. In this case we possess a state that is neutral, i.e. it is neither asymptotically stable or unstable. Here we would have a case in which the small perturbation will maintain the same value for all time, i.e. $|\boldsymbol{\xi}(t)| = |\boldsymbol{\xi}(0)|$ for all $t \geq 0$.

Whilst the real parts of the eigenvalues clearly denote the asymptotic stability of the relevant equilibrium point, the imaginary components are also of interest in many cases. In particular, non-zero eigenvalues may occur but only in complex conjugate pairs, i.e.

$$\lambda_1 = a + bi \qquad \text{and} \qquad \lambda_2 = a - bi.$$

The occurrence of one or more such pairs of eigenvalues indicate the existence of oscillatory behaviour. As before, the real component of the eigenvalues determine the stability of the system which in the context of oscillatory behaviour may be considered as damping (stable) or amplification (unstable) of the amplitude of oscillation. The magnitude of the imaginary components are also of interest since they are indicative of the frequency of oscillation within the solution.

Appendix D

Protein Synthesis

Within Chapter 5 it was discussed that there exists a significant amount of variation in the concentrations of signalling proteins in *E. coli* cells. One of the causes of this is the stochastic nature of gene expression which has been studied in some depth for *E. coli* cells. Here we briefly summarise the process by which proteins are synthesised within cells.

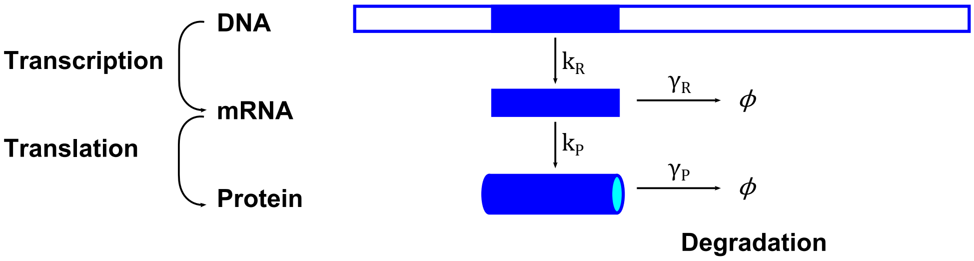


Figure D.1: Protein synthesis is a process comprising multiple stages. In the first instance, a DNA sequence is copied to produce an mRNA via a process called transcription. This occurs at a rate k_R and the resultant mRNAs are degraded at a rate γ_R. Once an mRNA has entered the cytoplasm of the cell, translation acts to produce protein molecules based on the mRNA instructions. This occurs at some rate k_P whilst the protein molecules produced will be degraded at a rate γ_P.

Protein synthesis begins within the cell nucleus where DNA (deoxyribonucleic acid) is located. DNA contains all of the information the cell needs in order to correctly produce the proteins and other components it requires for healthy function. The first stage of this process is called transcription. This is the process whereby upon receiving a signal which the cell needs to produce more of a particular protein, the relevant gene or section of a DNA nucleotide sequence is copied in order to produce RNA (ribonucleic acid). This is a one-to-one process in so much as each base on the DNA corresponds to one base on the resultant RNA. In the case of the *E. coli* chemotaxis signalling pathway this process will form an mRNA (messenger RNA) since the gene being transcribed encodes a protein.

Every different mRNA may be considered as a template for the production of a particular protein required for healthy cellular function. Once the appropriate mRNA has been transcribed from a DNA nucleotide sequence, the mRNAs move from the nucleus into the cytoplasm of the cell. At this point they interact with large RNA and protein complexes known as ribosomes in order to build proteins from the mRNA template, a process known as translation. In particular, the ribosome moves along an mRNA strand and builds the relevant amino acid polymer (also known as a protein molecule) one amino acid at a time. Different to the process of transcription, translation is not a one-to-one process. During translation a protein is created piece by piece as per the instructions encoded in the mRNA. For every three bases in the mRNA, there will be one particular amino acid added to the strand. This will continue until a particular set of three bases known as a stop codon gives the signal that the protein is complete.

Once the translation process is halted by the stop codon on the mRNA, the resultant amino acid strand is released into the cytoplasm. At this point the final stage of protein synthesis occurs. This is known as protein folding, which is the process whereby the elements of a linear strand of amino acids interact with one another in order to assume a particular three-dimensional shape or conformation.

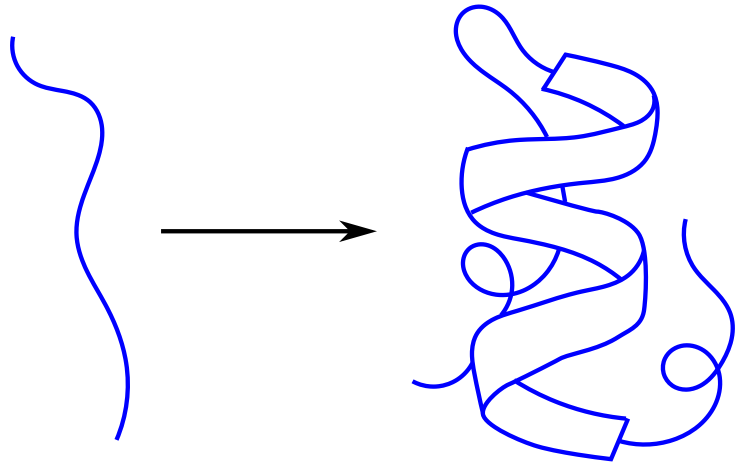


Figure D.2: Once the ribosome is finished with translating the mRNA into a strand of amino acids (left) there is one final stage that must be undergone in order to have a fully functional protein molecule. This is known as protein folding and is the process whereby the linear strand of amino acids assume the necessary three-dimensional conformation, which often appears similar to the right hand panel.

Within this section, a summary of the main processes involved in protein synthesis has been given. In addition to this there are a number of regulatory processes that act to control the rates at which each of these steps proceeds. Within this thesis we consider the effects of protein variation upon the chemotactic response of the cell rather than mathematically modelling all of the genetic processes. Whilst the genetic processes summarised here are not included within our modelling work it is still important to consider how these processes lead to the signalling cascade analysed within this work, especially since they are partially responsible for the variation in protein concentration considered in Chapter 5.

Appendix E

FCD in a Model with Multiple Receptor Types

Within Chapter 4 we have demonstrated that, subject to a number of assumptions and simplifications, the property of FCD holds for a model of the form in equations (3.16)-(3.19). However, as stated in Section 4.4.1 it is discussed that *E. coli* cells produce a response to aspartate using both Tar and Tsr receptors. Here we outline a proof that a mathematical model including the effects of Tsr binding aspartate is also able to exhibit FCD. This proof was provided in a personal communication by A. Hamadeh and E. Sontag.

In order to examine the multiple receptor case, it is necessary to consider the generalised homogeneity conditions of Shoval et al. [121]. These are given by

$$f\left(\varphi x(m), \Phi, p[L]\right) = \varphi f(x(m), \Phi, [L]), \tag{E.1}$$

$$g\left(\varphi x(m), \Phi, p[L]\right) = g(x(m), \Phi, [L]). \tag{E.2}$$

These sufficient conditions are considered here in the context of a multiple receptor model similar to that in equations (3.16)-(3.19). However, in order to capture the effects of multiple receptors, the receptor free-energy expression of equation (3.10) is now utilised. As in Section 4.3, assuming that $K_i^{off} \ll [L] \ll K_i^{on}$ for each receptor type (i.e. for $i = a, s$). This simplifies equation (3.10) such that it takes the form

$$F = N\left[\left(1 - \frac{m}{2}\right) + \nu_a \ln\left(\frac{[L]}{K_a^{off}}\right) + \nu_s \ln\left(\frac{[L]}{K_s^{off}}\right)\right], \tag{E.3}$$

where all symbols retain their earlier definition. Substituting this into equation (3.9) gives

$$\Phi = \frac{1}{1 + \left(\frac{[L]^R}{x(m)}\right)^N}, \tag{E.4}$$

where $R = \nu_a + \nu_s$ and $x(m) = K_a^{off} K_s^{off} e^{(m-2)/2}$ is the effective dissociation constant of this model. In order to assess the homogeneity conditions here we consider $\varphi = p^R$.

In order to examine the first homogeneity condition (equation (E.1)), it is necessary to compute the temporal derivative of $x(m)$. Upon use of the chain rule this is found to be

$$\dot{x} = \frac{\partial x}{\partial m}\frac{dm}{d\tau} = \frac{1}{2}x(m)\frac{dm}{d\tau} = f(x(m), \Phi, [L]). \tag{E.5}$$

As such it is possible to satisfy the first homogeneity condition as follows

$$f\left(\varphi x(m), \Phi, p[L]\right) = \frac{1}{2}\varphi x(m)\frac{dm}{d\tau} = \varphi\left(\frac{1}{2}x(m)\frac{dm}{d\tau}\right) = \varphi f\left(x(m), \Phi, [L]\right). \tag{E.6}$$

The second homogeneity condition may be satisfied in a similar manner since

$$g\left(\varphi x(m), \Phi, p[L]\right) = \frac{1}{1 + \left(\frac{\varphi[L]^R}{\varphi x(m)}\right)^N} = \frac{1}{1 + \left(\frac{[L]^R}{x(m)}\right)^N} = g\left(x, \Phi, [L]\right). \tag{E.7}$$

Thus, it is clear that a model including the effects of multiple receptors binding the same ligand will exhibit FCD when $K_i^{off} \ll [L] \ll K_i^{on}$.

Glossary

- **Adaptation**

  The process by which cells return to a pre-stimulus state. In the case of *E. coli* this occurs via balanced receptor methylation and demethylation.

- **Amidation**

  Conversion of a specific protein domain into an amide a compound derived from carboxylic acids.

- **Aspartate**

  An amino acid acting as a chemoattractant for *E. coli*. The most common form of aspartate considered in the literature is MeAsp (α-methylaspartate) since it is non-metabolisable.

- **Autophosphorylation**

  Phosphorylation of a protein as a result of its own activity.

- **Chemoattractant (Ligand)**

  A chemical substance that causes cells to display a positive chemotactic response. The terms chemoattractant, attractant and ligand are used interchangeably within this thesis.

- **Chemoreceptor**

  A type of protein that detects certain chemical substances within the environment.

- **CW/CCW Bias**

  The fraction of time that the flagella of a cell spend rotating in a clockwise (CW) or counter-clockwise (CCW) direction.

- **Cytoplasm**

  The contents of a cell, i.e. everything inside of the cytoplasmic membrane (for prokaryotes) but outside of the nucleus (for eukarotes).

- ***Escherichia coli* (*E. coli*)**

  A rod-shaped bacterium commonly found in the lower intestine of warm-blooded organisms. Most strains are harmless and are part of the regular flora of the gut. There are however a number of strains that may cause illnesses such as gastroenteritis, urinary tract infections and neonatal meningitis.

- **Eukaryote**

  A type of cell that contains a membrane-bound nucleus.

- **Flagellum (Plural: Flagella)**

  A long, slender, whip-like extension of a cell used for movement. The rotational movement of each flagellum is controlled by a flagellar motor.

- **FliM**

  A protein controlling the rotational behaviour of flagella. In the case of *E. coli*, a phosphorylated protein (CheY-P) binds to FliM, resulting in an increase in clockwise flagellar rotation.

- **Gain**

  This refers to the amplification of an external stimulus. In particular, here this refers to the fact that a small change in receptor occupancy can produce a much larger difference in intracellular signalling activity.

- **Histidine Protein Kinase (HPK)**

  A protein playing a vital role in passing extracellular signals into the cell. For *E. coli* and *R. sphaeroides* cells, this is achieved by changing the rate at which the histidine kinase protein (CheA or CheA$_2$) produces phosphoryl groups (autophosphorylates).

- *In vitro*

  Literally translates to “in glass”. Here this is used to signify that an experiment was conducted outside of a living cell.

- *In vivo*

  Literally translates to “within the living”. This is used to indicate that an experiment was conducted within the context of a living cell.

- **Linker Protein**

  A protein that is used as an intermediate in the binding of two other items. In the case of *E. coli* chemotaxis, the linker protein CheW allows CheA to bind to receptors. It may essentially be thought of as an binding adapter.

- **Methyl Group**

  A type of chemical compound derived from methane. In the chemotaxis signalling pathway of *E. coli* cells these regulate the activity at which the chemoreceptors of the cell act.

- **Methylation (Demethylation)**

  The addition of methyl groups onto, for example, a chemoreceptor is referred to as methylation. Demethylation relates to the removal of methyl groups.

- **Methylesterase**

  A protein that is able to remove methyl groups from chemoreceptors. This plays a key role in the process of adaptation.

- **Methyltransferase**

  Also known as a methylase. This type of protein is able to pass a methyl group from one location to another. In the case of *E. coli* the methyltransferase (CheR) passes methyl groups onto chemoreceptors forming a key part of the adaptation process.

- **Mutant Cell**

  This refers to a cell that has been altered from its natural state, often resulting in altered cell behaviour.

- **Operon**

  A functioning unit of DNA (Deoxyribonucleic Acid) that produce strands of mRNA (messenger Ribonucleic Acid). This mRNA is then translated into a number of gene products that in this case are the chemotaxis proteins.

- **Phosphatase**

  A protein that is able to to remove phosphoryl groups from other proteins. In the case of *E. coli*, the phosphatase (CheZ) acts to remove phosphoryl groups from the phosphorylated CheY (CheY-P).

- **Phosphoryl Group**

  A chemical compound of phosphoric acid. In terms of the chemotaxis signalling pathway, phosphoryl groups may be thought of as changing/enhancing the behaviour of a chemotaxis protein.

- **Phosphorylation (Dephosphorylation)**

  The addition of phosphoryl groups onto a protein is known as phosphorylation. Dephosphorylation refers to the loss of phosphoryl groups from a protein.

- **Phosphotransfer**

  A transfer of phosphoryl groups between two molecules. Here, this refers to the passing of a phosphoryl group from one chemotaxis protein to another.

- **Prokaryote**

  A single-cell organism lacking a distinct nucleus.

- **Receptor Cluster**

  A number of chemoreceptors that are closely packed in an area. This is often necessary to enhance the sensitivity of a cellular response.

- **Response Regulator**

  The protein controlling the response of the cell. For *E. coli* the response regulator (CheY) controls the swimming behaviour of the cell.

- *Rhodobacter sphaeroides* (*R. sphaeroides*)

  Another example of a rod-shaped bacterium and is commonly found in soil, mud and sludge. Multiple areas of research into this bacterium are ongoing including bioremediation (using biological processes to overcome environmental problems) and biofuel production [189].

- **Serine**

  Another example of an amino acid representing a chemoattractant for *E. coli* cells.

- **Signalling Team**

  A group of chemoreceptors that act together. A stimulus sensed by one chemoreceptor affects the activity of each receptor in the signalling team.

Bibliography

[1] U. Kaupp, N. Kashikar, and I. Weyand. Mechanisms of sperm chemotaxis. *Annual Review of Physiology*, 70:93–117, 2008.

[2] Y. Rao, K. Wong, M. Ward, C. Jurgensen, and J. Wu. Neuronal migration and molecular conservation with leukocyte chemotaxis. *Genes & Development*, 16(23):2973–2984, 2002.

[3] P. Nuzzi, M. Lokuta, and A. Huttenlocher. Analysis of neutrophil chemotaxis. In *Adhesion Protein Protocols*, pages 23–35. Springer, 2007.

[4] E. Roussos, J. Condeelis, and A. Patsialou. Chemotaxis in cancer. *Nature Reviews Cancer*, 11(8):573–587, 2011.

[5] G. Wadhams and J. Armitage. Making sense of it all: Bacterial chemotaxis. *Nature Reviews Molecular Cell Biology*, 5(12):1024–1037, 2004.

[6] R. Singh and M. Olson. Application of bacterial swimming and chemotaxis for enhanced bioremediation. In *Emerging Environmental Technologies*, pages 149–172. Springer, 2008.

[7] J. Kaper, J. Nataro, and H. Mobley. Pathogenic *Escherichia coli*. *Nature Reviews Microbiology*, 2(2):123–140, 2004.

[8] R. Bentley and R. Meganathan. Biosynthesis of vitamin K (menaquinone) in bacteria. *Microbiological Reviews*, 46(3):241, 1982.

[9] G. Cooper. The cell: A molecular approach, 2000.

[10] H. Berg. How to track bacteria. *Review of Scientific Instruments*, 42(6):868–871, 2003.

[11] H. Berg and D. Brown. Chemotaxis in *Escherichia coli* analysed by three-dimensional tracking. *Nature*, 239(5374):500–504, 1972.

[12] M. Baker, P. Wolanin, and J. Stock. Signal transduction in bacterial chemotaxis. *Bioessays*, 28(1):9–22, 2006.

[13] E. Kort, M. Goy, S. Larsen, and J. Adler. Methylation of a membrane protein involved in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 72(10):3939–3943, 1975.

[14] M. Springer, M. Goy, and J. Adler. Sensory transduction in *Escherichia coli*: Two complementary pathways of information processing that involve methylated proteins. *Proceedings of the National Academy of Sciences USA*, 74(8):3312–3316, 1977.

[15] M. Silverman and M. Simon. Chemotaxis in *Escherichia coli*: Methylation of *che* gene products. *Proceedings of the National Academy of Sciences USA*, 74(8):3317–3321, 1977.

[16] H. Kondoh, C. Ball, and J. Adler. Identification of a methyl-accepting chemotaxis protein for the ribose and galactose chemoreceptors of *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, 76(1):260–264, 1979.

[17] A. Boyd, A. Krikos, and M. Simon. Sensory transducers of *E. coli* are encoded by homologous genes. *Cell*, 26(3):333–343, 1981.

[18] M. Manson, V. Blank, G. Brade, and C. Higgins. Peptide chemotaxis in *E. coli* involves the Tap signal transducer and the dipeptide permease. *Nature*, 321(6067):253–256, 1986.

[19] S. Bibikov, R. Biran, K. Rudd, and J. Parkinson. A signal transducer for aerotaxis in *Escherichia coli*. *Journal of Bacteriology*, 179(12):4075–4079, 1997.

[20] A. Rebbapragada, M. Johnson, G. Harding, A. Zuccarelli, H. Fletcher, I. Zhulin, and B. Taylor. The Aer protein and the serine chemoreceptor Tsr independently sense intracellular energy levels and transduce oxygen, redox, and energy signals for *Escherichia coli* behavior. *Proceedings of the National Academy of Sciences USA*, 94(20):10541–10546, 1997.

[21] S. Clarke and D. Koshland. Membrane receptors for aspartate and serine in bacterial chemotaxis. *Journal of Biological Chemistry*, 254(19):9695–9702, 1979.

[22] L. Lin, J. Li, J. Brandts, and R. Weis. The serine receptor of bacterial chemotaxis exhibits half-site saturation for serine binding. *Biochemistry*, 33(21):6564–6570, 1994.

[23] H. Biemann and D. Koshland. Aspartate receptors of *Escherichia coli* and *Salmonella typhimurium* bind ligand with negative and half-of-the-sites cooperativity. *Biochemistry*, 33(3):629–634, 1994.

[24] R. Mesibov, G. Ordal, and J. Adler. The range of attractant concentrations for bacterial chemotaxis and the threshold and size of response over this range: Weber law and related phenomena. *The Journal of General Physiology*, 62(2):203–223, 1973.

[25] J. Hess, K. Oosawa, N. Kaplan, and M. Simon. Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell*, 53(1):79–87, 1988.

[26] K. Lipkow, S. Andrews, and D. Bray. Simulated diffusion of phosphorylated CheY through the cytoplasm of *Escherichia coli*. *Journal of Bacteriology*, 187(1):45–53, 2005.

[27] A. Bren, M. Welch, Y. Blat, and M. Eisenbach. Signal termination in bacterial chemotaxis: CheZ mediates dephosphorylation of free rather than switch-bound CheY. *Proceedings of the National Academy of Sciences USA*, 93(19):10090–10093, 1996.

[28] M. Welch, K. Oosawa, S. Aizawa, and M. Eisenbach. Phosphorylation-dependent binding of a signal molecule to the flagellar switch of bacteria. *Proceedings of the National Academy of Sciences USA*, 90(19):8787–8791, 1993.

[29] P. Cluzel, M. Surette, and S. Leibler. An ultrasensitive bacterial motor revealed by monitoring signaling proteins in single cells. *Science*, 287(5458):1652–1655, 2000.

[30] K. Lipkow. Changing cellular location of CheZ predicted by molecular simulations. *PLoS Computational Biology*, 2(4):e39, 2006.

[31] W. Springer and D. Koshland. Identification of a protein methyltransferase as the CheR gene product in the bacterial sensing system. *Proceedings of the National Academy of Sciences USA*, 74(2):533–537, 1977.

[32] J. Stock and D. Koshland. A protein methylesterase involved in bacterial sensing. *Proceedings of the National Academy of Sciences USA*, 75(8):3659–3663, 1978.

[33] B. Mello and Y. Tu. Effects of adaptation in maintaining high sensitivity over a wide range of backgrounds for *Escherichia coli* chemotaxis. *Biophysical Journal*, 92(7):2329–2337, 2007.

[34] R. Macnab and D. Koshland. The gradient-sensing mechanism in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 69(9):2509–2512, 1972.

[35] D. Koshland. A response regulator model in a simple sensory system. *Science*, 196(4294):1055–1063, 1977.

[36] S. Block, J. Segall, and H. Berg. Impulse responses in bacterial chemotaxis. *Cell*, 31(1):215–226, 1982.

[37] S. Block, J. Segall, and H. Berg. Adaptation kinetics in bacterial chemotaxis. *Journal of Bacteriology*, 154(1):312–323, 1983.

[38] M. Delbrück and W. Reichardt. System analysis for the light growth reactions of *Phycomyces. Princeton University Press*, 1956.

[39] A. Goldbeter and D. Koshland. Simple molecular model for sensing and adaptation based on receptor modification with application to bacterial chemotaxis. *Journal of Molecular Biology*, 161(3):395–416, 1982.

[40] S. Asakura and H. Honda. Two-state model for bacterial chemoreceptor proteins: The role of multiple methylation. *Journal of Molecular Biology*, 176(3):349–367, 1984.

[41] L. Segel, A. Goldbeter, P. Devreotes, and B. Knox. A mechanism for exact sensory adaptation based on receptor modification. *Journal of Theoretical Biology*, 120(2):151–179, 1986.

[42] D. Hauri and J. Ross. A model of excitation and adaptation in bacterial chemotaxis. *Biophysical Journal*, 68(2):708–722, 1995.

[43] M. Arocena and L. Acerenza. Necessary conditions for a minimal model of receptor to show adaptive response over a wide range of levels of stimulus. *Journal of Theoretical Biology*, 229(1):45–57, 2004.

[44] N. Barkai and S. Leibler. Robustness in simple biochemical networks. *Nature*, 387(6636):913–917, 1997.

[45] X. Xin and H. Othmer. A “trimer of dimers” based model for the chemotactic signal transduction network in bacterial chemotaxis. *Bulletin of Mathematical Biology*, 74(10):2339–2382, 2012.

[46] J. Segall, S. Block, and H. Berg. Temporal comparisons in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 83(23):8987–8991, 1986.

[47] V. Sourjik and H. Berg. Receptor sensitivity in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 99(1):123–127, 2002.

[48] D. Bray, M. Levin, and C. Morton-Firth. Receptor clustering as a cellular mechanism to control sensitivity. *Nature*, 393(6680):85–88, 1998.

[49] D. Bray and T. Duke. Conformational spread: The propagation of allosteric states in large multiprotein complexes. *Annual Review Biophysics and Biomolecular Structure*, 33:53–73, 2004.

[50] T. Duke, N. Le Novère, and D. Bray. Conformational spread in a ring of proteins: A stochastic approach to allostery. *Journal of Molecular Biology*, 308(3):541–553, 2001.

[51] Y. Shi and T. Duke. Cooperative model of bacterial sensing. *Physical Review E*, 58(5):6399, 1998.

[52] T. Duke and D D. Bray. Heightened sensitivity of a lattice of membrane receptors. *Proceedings of the National Academy of Sciences USA*, 96(18):10104–10108, 1999.

[53] T. Shimizu, S. Aksenov, and D. Bray. A spatially extended stochastic model of the bacterial chemotaxis signalling pathway. *Journal of Molecular Biology*, 329(2):291–309, 2003.

[54] Y. Shi. Adaptive ising model and bacterial chemotactic receptor network. *Europhysics Letters*, 50(1):113, 2000.

[55] B. Mello and Y. Tu. Quantitative modeling of sensitivity in bacterial chemotaxis: The role of coupling among different chemoreceptor species. *Proceedings of the National Academy of Sciences USA*, 100(14):8223–8228, 2003.

[56] B. Mello, L. Shaw, and Y. Tu. Effects of receptor interaction in bacterial chemotaxis. *Biophysical Journal*, 87(3):1578–1595, 2004.

[57] Y. Shi. Effects of thermal fluctuation and the receptor-receptor interaction in bacterial chemotactic signaling and adaptation. *Physical Review E*, 64(2):021910, 2001.

[58] J. Monod, J. Wyman, and J. Changeux. On the nature of allosteric transitions: A plausible model. *Journal of Molecular Biology*, 12(1):88–118, 1965.

[59] Y. Tu. Quantitative modeling of bacterial chemotaxis: Signal amplification and accurate adaptation. *Annual Review of Biophysics*, 42:337, 2013.

[60] J. Keymer, R. Endres, M. Skoge, Y. Meir, and N. Wingreen. Chemosensing in *Escherichia coli*: Two regimes of two-state receptors. *Proceedings of the National Academy of Sciences USA*, 103(6):1786–1791, 2006.

[61] B. Mello and Y. Tu. An allosteric model for heterogeneous receptor complexes: Understanding bacterial chemotaxis responses to multiple stimuli. *Proceedings of the National Academy of Sciences USA*, 102(48):17354–17359, 2005.

[62] R. Endres, O. Oleksiuk, C. Hansen, Y. Meir, V. Sourjik, and N. Wingreen. Variable sizes of *Escherichia coli* chemoreceptor signaling teams. *Molecular Systems Biology*, 4(1), 2008.

[63] Y. Tu, T. Shimizu, and H. Berg. Modeling the chemotactic response of *Escherichia coli* to time-varying stimuli. *Proceedings of the National Academy of Sciences USA*, 105(39):14855–14860, 2008.

[64] Y. Kalinin, L. Jiang, Y. Tu, and M. Wu. Logarithmic sensing in *Escherichia coli* bacterial chemotaxis. *Biophysical Journal*, 96(6):2439–2448, 2009.

[65] R. Endres, J. Falke, and N. Wingreen. Chemotaxis receptor complexes: From signaling to assembly. *PLoS Computational Biology*, 3(7):e150, 2007.

[66] Y. Meir, V. Jakovljevic, O. Oleksiuk, V. Sourjik, and N. Wingreen. Precision and kinetics of adaptation in bacterial chemotaxis. *Biophysical Journal*, 99(9):2766–2774, 2010.

[67] M. Li and G. Hazelbauer. Adaptational assistance in clusters of bacterial chemoreceptors. *Molecular Microbiology*, 56(6):1617–1626, 2005.

[68] R. Endres and N. Wingreen. Precise adaptation in bacterial chemotaxis through assistance neighborhoods. *Proceedings of the National Academy of Sciences USA*, 103(35):13040–13044, 2006.

[69] C. Hansen, R. Endres, and N. Wingreen. Chemotaxis in *Escherichia coli*: A molecular model for robust precise adaptation. *PLoS Computational Biology*, 4(1):e1, 2008.

[70] D. Clausznitzer, O. Oleksiuk, L. Løvdok, V. Sourjik, and R. Endres. Chemotactic response and adaptation dynamics in *Escherichia coli*. *PLoS Computational Biology*, 6(5):e1000784, 2010.

[71] M. Skoge, R. Endres, and N. Wingreen. Receptor-receptor coupling in bacterial chemotaxis: Evidence for strongly coupled clusters. *Biophysical Journal*, 90(12):4317–4326, 2006.

[72] M. Tindall, S. Porter, P. Maini, G. Gaglia, and J. Armitage. Overview of mathematical approaches used to model bacterial chemotaxis I: The single cell. *Bulletin of Mathematical Biology*, 70(6):1525–1569, 2008.

[73] D. Bray, R. Bourret, and M. Simon. Computer simulation of the phosphorylation cascade controlling bacterial chemotaxis. *Molecular Biology of the Cell*, 4(5):469–482, 1993.

[74] P. Spiro. *Mathematical studies of cell signal transduction*. PhD thesis, The University of Utah, 1997.

[75] P. Spiro, J. Parkinson, and H. Othmer. A model of excitation and adaptation in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 94(14):7263–7268, 1997.

[76] U. Alon, M. Surette, N. Barkai, and S. Leibler. Robustness in bacterial chemotaxis. *Nature*, 397(6715):168–171, 1999.

[77] T. Yi, Y. Huang, M. Simon, and J. Doyle. Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proceedings of the National Academy of Sciences USA*, 97(9):4649–4653, 2000.

[78] M. Levin, C. Morton-Firth, W. Abouhamad, R. Bourret, and D. Bray. Origins of individual swimming behavior in bacteria. *Biophysical Journal*, 74(1):175–181, 1998.

[79] S. Andrews and D. Bray. Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biology*, 1(3):137, 2004.

[80] C. Rao, J. Kirby, and A. Arkin. Phosphatase localization in bacterial chemotaxis: Divergent mechanisms, convergent principles. *Physical Biology*, 2(3):148, 2005.

[81] M. Tindall, P. Maini, S. Porter, and J. Armitage. Overview of mathematical approaches used to model bacterial chemotaxis II: Bacterial populations. *Bulletin of Mathematical Biology*, 70(6):1570–1607, 2008.

[82] J. Adler. Chemotaxis in *Escherichia coli*. *Cold Spring Harbor Symposium on Quantitative Biology*, 30:289–292, 1965.

[83] J. Adler. Chemotaxis in bacteria. *Science*, 153(3737):708–716, 1966.

[84] M. Beijerinck. Über atmungsfiguren beweglicher bakterien. *Zentralblatt Bakteriologie Parasitenkunde*, 14:827–845, 1893.

[85] O. Baracchini and J. Sherris. The chemotactic effect of oxygen on bacteria. *The Journal of Pathology and Bacteriology*, 77(2):565–574, 1959.

[86] J. Sherris, N. Preston, and J. Shoesmith. The influence of oxygen and arginine on the motility of a strain of *Pseudomonas sp.* *Journal of General Microbiology*, 16(1):86–96, 1957.

[87] E. Keller and L. Segel. Initiation of slime mold aggregation viewed as an instability. *Journal of Theoretical Biology*, 26(3):399–415, 1970.

[88] E. Keller and L. Segel. Model for chemotaxis. *Journal of Theoretical Biology*, 30(2):225–234, 1971.

[89] D. Horstmann. From 1970 until present: The Keller-Segel model in chemotaxis and its consequences I. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 105(3):103–165, 2003.

[90] D. Horstmann. From 1970 until present: The Keller-Segel model in chemotaxis and its consequences II. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 106(2):51–69, 2003.

[91] R. Erban and H. Othmer. From signal transduction to spatial pattern formation in *E. coli*: A paradigm for multiscale modeling in biology. *Multiscale Modeling & Simulation*, 3(2):362–394, 2005.

[92] M. Zhu and J. Murray. Parameter domains for generating spatial pattern: A comparison of reaction–diffusion and cell-chemotaxis models. *International Journal of Bifurcation and Chaos*, 5(06):1503–1524, 1995.

[93] P. Maini, M. Myerscough, K. Winters, and J. Murray. Bifurcating spatially heterogeneous solutions in a chemotaxis model for biological pattern generation. *Bulletin of Mathematical Biology*, 53(5):701–719, 1991.

[94] E. Keller and L. Segel. Traveling bands of chemotactic bacteria: A theoretical analysis. *Journal of Theoretical Biology*, 30(2):235–248, 1971.

[95] M. Brenner, L. Levitov, and E. Budrene. Physical mechanisms for chemotactic pattern formation by bacteria. *Biophysical Journal*, 74(4):1677–1693, 1998.

[96] L. Segel. Incorporation of receptor kinetics into a model for bacterial chemotaxis. *Journal of Theoretical Biology*, 57(1):23–42, 1976.

[97] L. Segel. A theoretical study of receptor mechanisms in bacterial chemotaxis. *SIAM Journal on Applied Mathematics*, 32(3):653–665, 1977.

[98] H. Othmer, X. Xin, and C. Xue. Excitation and adaptation in bacteria–a model signal transduction system that controls taxis and spatial pattern formation. *International Journal of Molecular Sciences*, 14(5):9205–9248, 2013.

[99] C. Xue. Macroscopic equations for bacterial chemotaxis: Integration of detailed biochemistry of cell signaling. *Journal of Mathematical Biology*, 70(1-2):1–44, 2015.

[100] R. Erban and H. Othmer. From individual to collective behavior in bacterial chemotaxis. *SIAM Journal on Applied Mathematics*, 65(2):361–391, 2004.

[101] C. Patlak. Random walk with persistence and external bias. *The Bulletin of Mathematical Biophysics*, 15(3):311–338, 1953.

[102] E. Codling, M. Plank, and S. Benhamou. Random walk models in biology. *Journal of the Royal Society Interface*, 5(25):813–834, 2008.

[103] W. Alt. Biased random walk models for chemotaxis and related diffusion approximations. *Journal of Mathematical Biology*, 9(2):147–177, 1980.

[104] K. Chen, R. Ford, and P. Cummings. The global turning probability density function for motile bacteria and its applications. *Journal of Theoretical Biology*, 195(2):139–155, 1998.

[105] H. Othmer, S. Dunbar, and W. Alt. Models of dispersal in biological systems. *Journal of Mathematical Biology*, 26(3):263–298, 1988.

[106] M. Rivero, R. Tranquillo, H. Buettner, and D. Lauffenburger. Transport models for chemotactic cell populations based on individual cell behavior. *Chemical Engineering Science*, 44(12):2881–2897, 1989.

[107] R. Ford and P. Cummings. On the relationship between cell balance equations for chemotactic cell populations. *SIAM Journal on Applied Mathematics*, 52(5):1426–1441, 1992.

[108] T. Emonet, C. Macal, M. North, C. Wickersham, and P. Cluzel. Agentcell: A digital single-cell assay for bacterial chemotaxis. *Bioinformatics*, 21(11):2714–2721, 2005.

[109] L. Zonia and D. Bray. Swimming patterns and dynamics of simulated *Escherichia coli* bacteria. *Journal of The Royal Society Interface*, 6(40):1035–1046, 2009.

[110] D. Bray, M. Levin, and K. Lipkow. The chemotactic behavior of computer-based surrogate bacteria. *Current Biology*, 17(1):12–19, 2007.

[111] S. Setayeshgar, C. Gear, H. Othmer, and I. Kevrekidis. Application of coarse integration to bacterial chemotaxis. *Multiscale Modeling & Simulation*, 4(1):307–327, 2005.

[112] M. Li and G. Hazelbauer. Cellular stoichiometry of the components of the chemotaxis signaling complex. *Journal of Bacteriology*, 186(12):3687–3694, 2004.

[113] N. Francis, M. Levit, T. Shaikh, L. Melanson, J. Stock, and D. DeRosier. Subunit organization in a soluble complex of Tar, CheW, and CheA by electron microscopy. *Journal of Biological Chemistry*, 277(39):36755–36759, 2002.

[114] R. Stewart, K. Jahreis, and J. Parkinson. Rapid phosphotransfer to CheY from a CheA protein lacking the CheY-binding domain. *Biochemistry*, 39(43):13157–13165, 2000.

[115] R. Stewart, A. Roth, and F. Dahlquist. Mutations that affect control of the methylesterase activity of CheB, a component of the chemotaxis adaptation system in *Escherichia coli*. *Journal of Bacteriology*, 172(6):3388–3399, 1990.

[116] J. Smith, J. Latiolais, G. Guanga, S. Citineni, R. Silversmith, and R. Bourret. Investigation of the role of electrostatic charge in activation of the *Escherichia coli* response regulator CheY. *Journal of Bacteriology*, 185(21):6385–6391, 2003.

[117] D. Bray. Research group data. http://www.pdn.cam.ac.uk/groups/comp-cell/Data.html. Accessed: 08/05/2012.

[118] T. Shimizu, Y. Tu, and H. Berg. A modular gradient-sensing network for chemotaxis in *Escherichia coli* revealed by responses to time-varying stimuli. *Molecular Systems Biology*, 6(1), 2010.

[119] A. Vaknin and H. Berg. Physical responses of bacterial chemoreceptors. *Journal of Molecular Biology*, 366(5):1416–1423, 2007.

[120] P. Dunten and D. Koshland. Tuning the responsiveness of a sensory receptor via covalent modification. *Journal of Biological Chemistry*, 266(3):1491–1496, 1991.

[121] O. Shoval, L. Goentoro, Y. Hart, A. Mayo, E. Sontag, and U. Alon. Fold-change detection and scalar symmetry of sensory input fields. *Proceedings of the National Academy of Sciences USA*, 107(36):15995–16000, 2010.

[122] M. Lazova, T. Ahmed, D. Bellomo, R. Stocker, and T. Shimizu. Response rescaling in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 108(33):13870–13875, 2011.

[123] O. Shoval, U. Alon, and E. Sontag. Symmetry invariance for adapting biological systems. *SIAM Journal on Applied Dynamical Systems*, 10(3):857–886, 2011.

[124] A. Barnakov, L. Barnakova, and G. Hazelbauer. Efficient adaptational demethylation of chemoreceptors requires the same enzyme-docking site as efficient methylation. *Proceedings of the National Academy of Sciences USA*, 96(19):10667–10672, 1999.

[125] M. Li and G. Hazelbauer. Adaptational assistance in clusters of bacterial chemoreceptors. *Molecular Microbiology*, 56(6):1617–1626, 2005.

[126] L. Jiang, Q. Ouyang, and Y. Tu. Quantitative modeling of *Escherichia coli* chemotactic motion in environments varying in space and time. *PLoS Computational Biology*, 6(4):e1000735, 2010.

[127] V. Sourjik and H. Berg. Functional interactions between receptors in bacterial chemotaxis. *Nature*, 428(6981):437–441, 2004.

[128] C. Hansen, V. Sourjik, and N. Wingreen. A dynamic-signaling-team model for chemotaxis receptors in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, 107(40):17170–17175, 2010.

[129] C. Morton-Firth, T. Shimizu, and D. Bray. A free-energy-based stochastic simulation of the Tar receptor complex. *Journal of Molecular Biology*, 286(4):1059–1074, 1999.

[130] H. Berg and P. Tedesco. Transient response to chemotactic stimuli in *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, 72(8):3235–3239, 1975.

[131] T. Min, P. Mears, I. Golding, and Y. Chemla. Chemotactic adaptation kinetics of individual *Escherichia coli* cells. *Proceedings of the National Academy of Sciences USA*, 109(25):9869–9874, 2012.

[132] T. Min, P. Mears, L. Chubiz, C. Rao, I. Golding, and Y. Chemla. High-resolution, long-term characterization of bacterial motility using optical tweezers. *Nature Methods*, 6(11):831–835, 2009.

[133] G. Lan, S. Schulmeister, V. Sourjik, and Y. Tu. Adapt locally and act globally: Strategy to maintain high chemoreceptor sensitivity in complex environments. *Molecular Systems Biology*, 7(1), 2011.

[134] S. Pigolotti, S. Krishna, and M. Jensen. Oscillation patterns in negative feedback loops. *Proceedings of the National Academy of Sciences USA*, 104(16):6533–6537, 2007.

[135] E. Snoussi. Necessary conditions for multistationarity and stable periodicity. *Journal of Biological Systems*, 6(01):3–9, 1998.

[136] J. Gouzé. Positive and negative circuits in dynamical systems. *Journal of Biological Systems*, 6(01):11–15, 1998.

[137] M. Elowitz, A. Levine, E. Siggia, and P. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[138] S. Kalir, J. McClure, K. Pabbaraju, C. Southward, M. Ronen, S. Leibler, M. Surette, and U. Alon. Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science*, 292(5524):2080–2083, 2001.

[139] E. Flach and S. Schnell. Use and abuse of the quasi-steady-state approximation. *IEE Proceedings-Systems Biology*, 153(4):187–191, 2006.

[140] J. Murray. Mathematical biology I: An introduction, 2002.

[141] C. Morton-Firth and D. Bray. Predicting temporal fluctuations in an intracellular signalling pathway. *Journal of Theoretical Biology*, 192(1):117–128, 1998.

[142] A. Lamanna, G. Ordal, and L. Kiessling. Large increases in attractant concentration disrupt the polar localization of bacterial chemoreceptors. *Molecular Microbiology*, 57(3):774–785, 2005.

[143] K. Wu, H. Walukiewicz, G. Glekas, G. Ordal, and C. Rao. Attractant binding induces distinct structural changes to the polar and lateral signaling clusters in *Bacillus subtilis* chemotaxis. *Journal of Biological Chemistry*, 286(4):2587–2595, 2011.

[144] J. Goldman, M. Levin, and D. Bray. Signal amplification in a lattice of coupled protein kinases. *Molecular BioSystems*, 5:1853–1859, 2009.

[145] E. Korobkova, T. Emonet, J. Vilar, T. Shimizu, and P. Cluzel. From molecular noise to behavioural variability in a single bacterium. *Nature*, 428(6982):574–578, 2004.

[146] J. Raser and E. O'Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309(5743):2010–2013, 2005.

[147] M. Elowitz, A. Levine, E. Siggia, and P. Swain. Stochastic gene expression in a single cell. *Science*, 297(5584):1183–1186, 2002.

[148] H. Park, C. Guet, T. Emonet, and P. Cluzel. Fine-tuning of chemotactic response in *E. coli* determined by high-throughput capillary assay. *Current Microbiology*, 62(3):764–769, 2011.

[149] M. Levin, C. Morton-Firth, W. Abouhamad, R. Bourret, and D. Bray. Origins of individual swimming behaviour in bacteria. *Biophysical Journal*, 74(1):175–181, 1998.

[150] J. Spudich and D. Koshland Jr. Non-genetic individuality: Chance in the single cell. *Nature*, 262(5568):467–471, 1976.

[151] T. Emonet and P. Cluzel. Relationship between cellular response and behavioral variability in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 105(9):3304–3309, 2008.

[152] R. Steuer, S. Waldherr, V. Sourjik, and M. Kollmann. Robust signal processing in living cells. *PLoS Computational Biology*, 7(11):e1002218, 2011.

[153] N. Barkai and B. Shilo. Variability and robustness in biomolecular systems. *Molecular Cell*, 28(5):755–760, 2007.

[154] M. Kollmann, L. Løvdok, K. Bartholomé, J. Timmer, and V. Sourjik. Design principles of a bacterial signalling network. *Nature*, 438(7067):504–507, 2005.

[155] C. Morton-Firth and D. Bray. Predicting temporal fluctuations in an intracellular signalling pathway. *Journal of Theoretical Biology*, 192(1):117–128, 1998.

[156] E. Purcell. Life at low Reynolds number. *American Journal of Physics*, 45(1):3–11, 1977.

[157] N. Darnton, L. Turner, S. Rojevsky, and H. Berg. On torque and tumbling in swimming *Escherichia coli*. *Journal of Bacteriology*, 189(5):1756–1764, 2007.

[158] L. Turner, W. Ryu, and H. Berg. Real-time imaging of fluorescent flagellar filaments. *Journal of Bacteriology*, 182(10):2793–2801, 2000.

[159] R. Macnab and D. Koshland. The gradient-sensing mechanism in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 69(9):2509–2512, 1972.

[160] N. Vladimirov, L. Løvdok, D. Lebiedz, and V. Sourjik. Dependence of bacterial chemotaxis on gradient shape and adaptation rate. *PLoS Computational Biology*, 4(12):e1000242, 2008.

[161] M. Kojadinovic, J. Armitage, M. Tindall, and G. Wadhams. Response kinetics in the complex chemotaxis signalling pathway of *Rhodobacter sphaeroides*. *Journal of The Royal Society Interface*, 10(81):20121001, 2013.

[162] J. de Beyer. *Roles of the two chemotaxis clusters in Rhodobacter sphaeroides*. PhD thesis, University of Oxford, 2013.

[163] J. Armitage and R. Macnab. Unidirectional, intermittent rotation of the flagellum of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, 169(2):514–518, 1987.

[164] S. Porter, A. Warren, A. Martin, and J. Armitage. The third chemotaxis locus of *Rhodobacter sphaeroides* is essential for chemotaxis. *Molecular Microbiology*, 46(4):1081–1094, 2002.

[165] D. Harrison, J. Skidmore, J. Armitage, and J. Maddock. Localization and environmental regulation of MCP-like proteins in *Rhodobacter sphaeroides*. *Molecular Microbiology*, 31(3):885–892, 1999.

[166] P. Hamblin, B. Maguire, R. Grishanin, and J. Armitage. Evidence for two chemosensory pathways in *Rhodobacter sphaeroides*. *Molecular Microbiology*, 26(5):1083–1096, 1997.

[167] A. Martin, G. Wadhams, and J. Armitage. The roles of the multiple CheW and CheA homologues in chemotaxis and in chemoreceptor localization in *Rhodobacter sphaeroides*. *Molecular Microbiology*, 40(6):1261–1272, 2001.

[168] A. Martin, G. Wadhams, D. Shah, S. Porter, J. Mantotta, T. Craig, P. Verdult, H. Jones, and J. Armitage. CheR-and CheB-dependent chemosensory adaptation system of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, 183(24):7135–7144, 2001.

[169] G. Wadhams, A. Warren, A. Martin, and J. Armitage. Targeting of two signal transduction pathways to different regions of the bacterial cell. *Molecular Microbiology*, 50(3):763–770, 2003.

[170] S. Porter and J. Armitage. Chemotaxis in *Rhodobacter sphaeroides* requires an atypical histidine protein kinase. *Journal of Biological Chemistry*, 279(52):54573–54580, 2004.

[171] A. Ferré, J. De La Mora, T. Ballado, L. Camarena, and G. Dreyfus. Biochemical study of multiple CheY response regulators of the chemotactic pathway of *Rhodobacter sphaeroides*. *Journal of Bacteriology*, 186(15):5172–5177, 2004.

[172] M. Roberts, E. August, A. Hamadeh, P. Maini, P. McSharry, J. Armitage, and A. Papachristodoulou. A model invalidation-based approach for elucidating biological signalling pathways, applied to the chemotaxis pathway in *R. sphaeroides*. *BMC Systems Biology*, 3(1):105, 2009.

[173] A. Hamadeh, M. Roberts, E. August, P. McSharry, P. Maini, J. Armitage, and A. Papachristodoulou. Feedback control architecture and the bacterial chemotaxis network. *PLoS Computational Biology*, 7(5):e1001130, 2011.

[174] M. Tindall, S. Porter, P. Maini, and J. Armitage. Modeling chemotaxis reveals the role of reversed phosphotransfer and a bi-functional kinase-phosphatase. *PLoS Computational Biology*, 6(8):e1000896, 2010.

[175] M. Amin, S. Porter, and O. Soyer. Split histidine kinases enable ultrasensitivity and bistability in two-component signaling networks. *PLoS Computational Biology*, 9(3):e1002949, 2013.

[176] A. Hamadeh, B. Ingalls, and E. Sontag. Transient dynamic phenotypes as criteria for model discrimination: Fold-change detection in *Rhodobacter sphaeroides* chemotaxis. *Journal of The Royal Society Interface*, 10(80):20120935, 2013.

[177] S. Porter and J. Armitage. Phosphotransfer in *Rhodobacter sphaeroides* chemotaxis. *Journal of Molecular Biology*, 324(1):35–45, 2002.

[178] S. Porter, M. Roberts, C. Manning, and J. Armitage. A bifunctional kinase-phosphatase in bacterial chemotaxis. *Proceedings of the National Academy of Sciences USA*, 105(47):18531–18536, 2008.

[179] M. Gould. *Chemotaxis gene expression in Rhodobacter sphaeroides WS8N*. PhD thesis, University of Oxford, 2006.

[180] C. Morton-Firth. *Stochastic simulation of cell signalling pathways*. PhD thesis, University of Cambridge, 1998.

[181] C. Rao, J. Kirby, and A. Arkin. Design and diversity in bacterial chemotaxis: A comparative study in *Escherichia coli* and *Bacillus subtilis*. *PLoS Biology*, 2(2):e49, 2004.

[182] M. Tindall, S. Porter, G. Wadhams, P. Maini, and J. Armitage. Spatiotemporal modelling of CheY complexes in *Escherichia coli* chemotaxis. *Progress in Biophysics and Molecular Biology*, 100(1):40–46, 2009.

[183] B. Franz and R. Erban. Hybrid modelling of individual movement and collective behaviour. In *Dispersal, Individual Movement and Spatial Ecology*, pages 129–157. Springer, 2013.

[184] H. Othmer and C. Xue. The mathematical analysis of biological aggregation and dispersal: Progress, problems and perspectives. In *Dispersal, Individual Movement and Spatial Ecology*, pages 79–127. Springer, 2013.

[185] B. Franz, C. Xue, K. Painter, and R. Erban. Travelling waves in hybrid chemotaxis models. *Bulletin of Mathematical Biology*, 76(2):377–400, 2014.

[186] M. Flegg, J. Chapman, L. Zheng, and R. Erban. Analysis of the two-regime method on square meshes. *SIAM Journal on Scientific Computing*, 36(3):B561–B588, 2014.

[187] M. Flegg, J. Chapman, and R. Erban. The two-regime method for optimizing stochastic reaction–diffusion simulations. *Journal of the Royal Society Interface*, 9(70):859–868, 2012.

[188] L. Michaelis and M. Menten. Die kinetik der invertinwirkung. *Biochemische Zeitschrift*, 49(333-369):352, 1913.

[189] D. Kim, J. Lee, S. Kang, P. Hallenbeck, E. Kim, J. Lee, and M. Kim. Enhanced photo-fermentative H_2 production using *Rhodobacter sphaeroides* by ethanol addition and analysis of soluble microbial products. *Biotechnology for Biofuels*, 7(1):79, 2014.