Fall 2013

# Measuring and Modeling The Response Characteristics Of The Environmental Phosphate Transducer In Escherichia Coli

Chetan Sood
*Purdue University*

## Recommended Citation

Sood, Chetan, "Measuring and Modeling The Response Characteristics Of The Environmental Phosphate Transducer In Escherichia Coli" (2013). *Open Access Dissertations*. 43.
https://docs.lib.purdue.edu/open_access_dissertations/43

Graduate School ETD Form 9
(Revised 12/07)

# PURDUE UNIVERSITY
## GRADUATE SCHOOL
## Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Chetan Sood

Entitled
Measuring and Modeling the Response Characteristics of the Environmental Phosphate Transducer in Escherichia coli

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Ken P. Ritchie
Chair

Brian Todd

Ephraim Fischbach

Barry L. Wanner

To the best of my knowledge and as understood by the student in the *Research Integrity and Copyright Disclaimer (Graduate School Form 20)*, this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Approved by Major Professor(s): Ken P. Ritchie

Approved by: Mark P. Haugan 11/26/2013
Head of the Graduate Program Date

# MEASURING AND MODELING THE RESPONSE CHARACTERISTICS OF THE ENVIRONMENTAL PHOSPHATE TRANSDUCER IN *ESCHERICHIA COLI*

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Chetan Sood

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2013

Purdue University

West Lafayette, Indiana

your job was the most difficult.
i only put together the pieces you gave to me.
to mom, dad and raj.
to sharad and in his memory.

# ACKNOWLEDGMENTS

The success of this project is due to the contributions of many to whom I am indebted. I would like to acknowledge Professor Ken Ritchie for giving me the opportunity to work on this project, for the guidance and encouragement he offered through the highs and lows and for the numerous and various stimulating conversations we had along the way. I thank all the members of the Ritchie Lab for their helpfulness, especially Drs. Dongmyung Oh and Yoriko Lill who introduced me to wet lab techniques and Drs. Jeffrey Spector, Jacob Hale and Hochan Lee who introduced me to fluorescence microscopy techniques. Also, I recognize Professor Barry L. Wanner for offering guidance and discussion on matters related to biology, biological systems and two-component systems. Researchers in the laboratory of Prof. Wanner constructed and provided the *E. coli* strain and samples necessary for the experimentation; and I thank especially Drs. Yi-Ju Hsieh, Kiril Datsenko, Jun Teramoto and Katsushi Yokoyama for their contributions in this regard. I am grateful to Professor Ephraim Fischbach for his support and encouragement that started with the way he engaged me in the very first graduate-level physics class I took. And, I am grateful for the advice and feedback Professor Brian Todd provided regarding the dissertation manuscript. Finally, I am thankful to have loved ones, family and friends, who supported me through this lengthy and oftentimes difficult project.

## TABLE OF CONTENTS

APPENDICES

# LIST OF TABLES

Table Page

# LIST OF FIGURES

Figure Page

# ABSTRACT

Sood, Chetan Ph.D., Purdue University, December 2013. Measuring and Modeling the Response Characteristicsof the Environmental Phosphate Transducerin *Escherichia coli*. Major Professor: Ken Ritchie.

The PhoR/PhoB two-component system in *Escherichia coli* is a biological transducer that senses the limitation of environmental inorganic orthophosphate, the bacteria's preferred source of the essential nutrient phosphate, and transmits that information to the interior of the cell initiating a response that mitigates phosphate starvation. In the first part of this study, we present and apply a fluorescence microscopy technique to measure, *in vivo*, the dynamic response characteristics of the transducer with single-cell resolution. We report that the transience in the PhoR/PhoB TCS response is consistent with the transducer having a threshold sensitivity to the concentration of environmental phosphate, below which the transducer stochastically switches from a low to high operating point. Significantly, we find that the transducer response overshoots before settling to its final operating point. In the second part of this study, we investigate a series of minimal models, simple extensions of the birth-death process, that have response characteristics in common with the measured transducer response. We find that using the Hill equation as the functional form of the birth rate in an autoregulating birth-death process gives a bistable system with stochastic switching. This demonstrates that the model complexity necessary to reproduce the transducer's qualitative behavior is less than the complexity of the biological system.

# 1. INTRODUCTION

*Escherichia coli* live amid a fluctuating and sometimes hostile chemical environment. Drift in parameters such as temperature, pH, osmolarity, and the concentrations of foods, nutrients and toxins can impact the efficiency of the biological processes running within the bacteria. *E. coli* has evolved response mechanisms to mitigate or even exploit changing environmental conditions. The majority of these mechanisms are transcription programs that incorporate a type of biological transducer called a two-component system to couple to environmental signals [1].

The two-component system (TCS) is a protein interaction network that converts an environmental signal to a build-up in the intracellular concentration of a functional protein. The excess functional protein controls the physiological response; in the case that it is a transcription program, the functional protein usually directly promotes or inhibits the transcription of a set of genes by interaction with the DNA. At least thirty different TCSs have been identified in *E. coli* [2] and their implementation is extensive among other prokaryotes.

Of course, prokaryotes are not the only organisms to use such systems: signal transduction is a fundamental biological process found in all kingdoms of life. In general, the signal transduction systems of higher organisms tend to be more complex. Whereas a TCS is usually sensitive to one type of signal, the signal transduction systems in humans often integrate and process many, including inter-organ communications such as hormones. But the basic principle is the same. Signal transduction systems process, transmit, and convert extracellular signals to an intracellular signal.

The transduction of an extracellular signal can have interesting properties. Even the relatively simple TCSs demonstrate a rich set of response characteristics. Steady-state, large-population studies have shown that TCS signal response can be thresholded [3] and bistable [4], where transitions between states occur stochastically.

Numerous efforts have been made to construct predictive biochemical models of TCSs and other signal tranduction systems. The value of system models (and the possibility of whole-cell models) to a diverse range of disease and infection problems is vast. More so, deep understanding of cell behavior demands model abstractions; as the list of the cellular "nuts and bolts" grows to overwhelming proportion so does the need to properly modularize the components into biologically meaningful functional units. But mechanistic models remain elusive. Although component lists are long, they remain incomplete, as is knowledge of the interactions among the components. And the data against which the models are tested tends to be limited in the sense that the data cannot convincingly choose among the competing imperfect models.

It has been suggested that measuring the transient response of a TCS at the single-cell level would provide a better benchmark against which models can be tested (see, for example, [5] which reviews the value of single-cell data and [6] which suggests the necessity for measurements in the time-domain). In this study, we present and apply an experimental method to measure, *in vivo*, the single-cell transient response of a TCS. The method, a variation of the traditional fluorescence microscopy technique, illuminates key features in the transience. Then, instead of taking the standard approach of using this data to test a model constructed from knowledge of the system's "nuts and bolts", we present a series of minimal models, simple extensions of the birth-death process, that have features in common with the observed behavior. The goal in this effort is not to find the best model of the TCS, but to identify the smallest number of parts necessary to model its essential behavior.

The signal transduction system we study is the PhoR/PhoB TCS in *E. Coli* that responds to the environmental limitation of inorganic orthophosphate ($P_i$), the bacteria's preferred source of the essential nutrient phosphate. Following a short comment below, the remainder of this chapter describes the "nuts and bolts" of two-component systems, in general, and the PhoR/PhoB TCS, in particular. In the second chapter we describe the experimental methods and analysis, and, in the third chapter, we present the results of the experimentation on the PhoR/PhoB TCS and an interpretation of the results. The fourth chapter introduces the basic model building framework and reviews a comprehensive “nuts and bolts” model of the PhoR/PhoB TCS found in the literature, before presenting the series of minimal models. As the minimal models will be of general interest, time and detail is devoted to developing a framework within which to understand their behavior.

## The Big Picture

This study represents the initial stage of a long-time research project. With the exception of the basic fluorescence microscopy technique (and the *E. coli* mutant strain construction and sample culturing necessary for the experimentation, which was performed in the laboratory of our collaborator: Professor Barry L. Wanner, Biological Sciences, Purdue University), the experimental methods, analysis and interpretation and the modeling methods, analysis and interpretation have been accomplished for the first time in this laboratory. This study points the laboratory to a new line of research; it is not an iteration of an established method on a new system. As such, there did not exist a body of knowledge to motivate or guide the investigation, and, as is usually the case at the beginning of a long and difficult project, we took many steps backwards, redefining and narrowing the scope, before stepping forward. It is important to take note of this, as the primary contributions of this work is establishing the foundation and expertise necessary for the future stages of the project.

To give an example of this “moving backwards before moving forwards”, as originally conceived, this project was to use microfluidic devices as the experimentation platform on which to deposit, grow, perturb and observe the *E. coli* response. Much effort was devoted to developing the expertise to use the microfluidic devices in concert with developing the experimental method and analysis, before it became clear the use of the devices was overly burdensome for this phase in the project. The need to rapidly and repeatedly perform the experiments overruled the initial desire to use the microfluidic devices. So the scope of the project was narrowed, with the idea that migration of the experiments to the microfluidic platform would be a later goal in the project. In anticipation of this eventual goal (and because the expertise regarding microfluidic device usage represents, in itself, a significant contribution to the long-time research project), a discussion of microfluidic devices has been included in Appendix A.

## 1.1 Two-Component Systems

Signal transduction by a TCS is achieved by the interactions of two proteins: a sensory histidine kinase and a response regulator. Three primary interactions control the signal transduction [7]. In the first, an environmental signal regulates the autophosphorylation of the sensory histidine kinase (HK). In the second interaction, the response regulator (RR) catalyzes the stoichiometric transfer of the phosphoryl group from a phosphorylated HK to itself. And, in the third interaction, the phosphorylated RR dephosphorylates by transferring the phosphoryl group to water. The phosphorylated RR is the functional protein, and the build-up of the intracellular concentration of phospho-RR controls the physiological response.

Transduction is achieved in the sense that the TCS converts the environmental signal to a change in the intracellular concentration of phospho-RR. For a typical TCS, like the PhoR/PhoB TCS, the sensory histidine kinase (HK) is an integral membrane protein and the response regulator (RR) is a cytosolic protein that, when phosphorylated, binds with the DNA to regulate the transcription of a set of genes (an important TCS in which the phospho-RR is not a transcription factor is the Che system that controls the chemotaxis of *E. coli*). As we describe below for the PhoR/PhoB TCS, it is possible that interactions among proteins in addition to the HK and RR contribute to TCS signal trasduction.

As described in [7], an HK has two functional modules: a diverse sensing domain and a highly conserved kinase core. In conjunction with other proteins, the sensing domain interacts specifically with the environmental signal. This interaction causes a conformational change in the protein's structure that modulates the activity of the kinase core. In its activated conformation, the kinase core can undergo an ATP-dependent autophosphorylation at a conserved histidine residue.

An RR also has two functional modules: a highly conserved regulatory domain and a diverse effector domain. The RR catalyzes phosphoryl transfer from a phosphorylated HK to a highly conserved aspartate residue in its own regulatory domain. Phosphorylation induces a conformational change that modulates the activity of the effector domain. In its activated conformation, the effector domain can interact with other biochemical species. The modular design of the HK and RR has allowed the basic TCS template to be adapted to fit a variety of cellular signal transduction needs.

The three primary reactions that control TCS signal transduction can be summarized as [7]:

1. Autophosphorylation: HK-His + ATP $\Leftrightarrow$ HK-His~P + ADP
2. Phosphotransfer: HK-His~P + RR-Asp $\Leftrightarrow$ HK-His + RR-Asp~P
3. Dephosphorylation: RR-Asp~P + $H_2O$ $\Leftrightarrow$ RR-Asp + P + $H_2O$.

## 1.2 Phosphate Homeostasis in *Escherichia coli*

Phosphate is essential for the survival, growth, and division of *E. coli.* In fact, VanBogelen *et al.* [8] found that almost a tenth of the *E. coli* genome is differentially expressed in response to the limitation of extracellular inorganic orthophosphate ($P_i$), the bacteria's preferred phosphate source. Elemental phosphorus comprises approximately three percent of the cell's dry mass [9] mostly in the form of phosphate. A large majority of the phosphate is found in the tens of millions of phospholipids that make the cell membrane, and the entire membrane must be doubled in size each cell cycle. Phosphate is also a building block of proteins and nucleic acids. As cellular processes consume phosphate, the *E. coli* must capture more from the environment to replenish its internal stock.

**The Phosphate-specific Transport System** The uptake of $P_i$ by *E. coli* is carried out by the Pst (phosphate-specific transport) system which belongs to the superfamily of ABC (ATP-binding cassette) transporters [10]. The Pst system is made up of four proteins, PstS, PstA, PstB, and PstC. PstS specifically binds $P_i$ in the periplasm with high affinity, and PstA and PstC form the membrane channel for $P_i$ entry while PstB interacts with ATP to provide the energy necessary for transporting $P_i$ across the membrane [3].

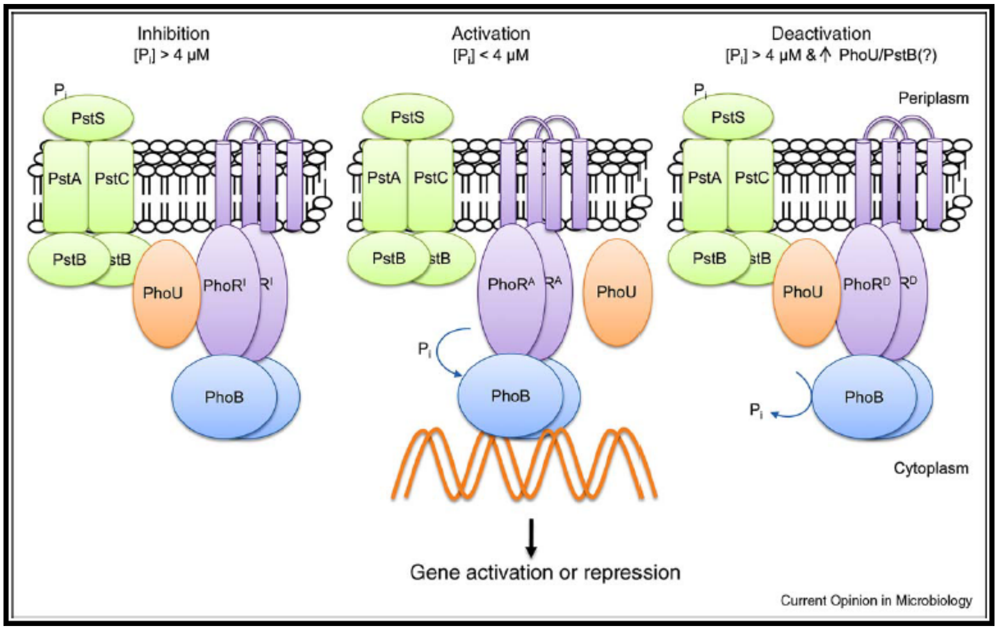
### 1.2.1 The PhoR/PhoB Two-Component System

When *E. coli* are in an environment with excess $P_i$, the Pst proteins are transcribed at a basal rate. Basal production replenishes the Pst systems that are lost to degradation or dilution via cell division (the transporters are randomly divided among the daughter cells when the cell splits).

Transcription of the Pst proteins is upregulated when $P_i$ is limited – *E. coli* responds to the starvation signal by creating additional $P_i$-specific transport systems, increasing the efficiency of $P_i$ transport into the cell. This transcriptional response is controlled by the PhoR/PhoB TCS, where PhoR is the integral membrane HK and PhoB is the DNA-binding RR. It has been previously reported [3] [10] that, from measurements in bulk cultures, the threshold delineating $P_i$ limitation is near a concentration of 4 microMolar ($[P_i] = 4\mu M$).

As suggested by Figure 1.1, a PhoR protein exists in one of three states: inhibition, activation, and deactivation [10]. The state of PhoR determines which reaction channels (autophosphorylation, phosphotransfer, dephosphorylation) are available to PhoR and PhoB. Transition between the PhoR states is controlled by the binding of $P_i$ to the Pst system (specifically to PstS) in the periplasm and facilitated by another protein called PhoU.

**Inhibition of PhoR** When the Pst system is bound with a $P_i$ molecule in the periplasm, a PhoR, PhoB, PhoU, and the Pst system form an inhibition complex



Current Opinion in Microbiology

Figure 1.1. The PhoR/PhoB Two-Component System (image reproduced from [10]). The TCS responds to $P_i$ limitation by increasing the intracellular concentration of phospho-PhoB which upregulates the production of the $P_i$ transporter system (and other proteins). PhoR exists in three states: inhibition, activation and deactivation. In the inhibition state, the binding of PhoR to a Pst system prevents PhoR from autophosphorylating, quenching the phosphorelay interaction between PhoR and the PhoB. In the activation state, PhoR is released from the inhibition complex and quickly autophosphorylates. This opens the phosphorelay reaction channel, allowing a PhoB to autophosphorylate. Phospho-PhoB can interact with the DNA to upregulate the transcription of the Pho genes. In the deactivation state, PhoR and the Pst system reform the inhibition complex and PhoR (and possibly PhoU) participates in the dephosphorylation of PhoB. At high $[P_i]$, the inhibition and deactivation states are favorable, and at low $[P_i]$, the activation state is favorable.

that prevents the autophosphorylation of that PhoR; none of the reaction channels are available. This state of PhoR is favorable when the $P_i$ concentration ($[P_i]$) is high.

**Activation of PhoR** When the Pst system unbinds with the $P_i$ molecule in the periplasm, the inhibition complex dissociates and the PhoR is able to rapidly autophosphorylate from cytoplasmic ATP. The active phospho-PhoR and a PhoB can access the phosphotransfer reaction channel. The resulting phospho-PhoB can bind to the DNA to regulate gene transcription. Along with the Pst proteins, phospho-PhoB upregulates the expression of PhoB (auto-regulation), PhoR and PhoU (and other Pho proteins). This state of PhoR is favorable when $[P_i]$ is low.

**Deactivation of PhoR** When a Pst system rebinds a $P_i$ molecule in the periplasm, it interacts with a PhoR, PhoB, and PhoU to reform the inhibition complex. In the deactivated state, PhoR and, possibly, PhoU participate in the dephosphorylation of PhoB [10]. And, once PhoB is dephosphorylated, the inhibition state has been fully recovered. The deactivation state of PhoR is favorable in high $[P_i]$.

### 1.2.2 Stochasticity in PhoR/PhoB Response

It has previously been reported [4] that the response of the PhoR/PhoB TCS has a stochastic nature. Specifically, that the TCS activity has two operating points (bistability) and the switching between the operating points is intrinsically a random process. This leads to response variability (phenotype variation) across an isogenic population subjected to identical environmental conditions.

As a fundamental property of many biological systems, stochasticity has garnered tremendous excitement and attention (see, for example, [11] for a review). Initially receiving treatment only as an interesting source of noise [12] [13] in biological systems, there is a growing consensus [14] that stochastic processes offer survival advantage at a population level. For example, [15] describes an experiment in which a genetically homogenous *E. coli* population survives a temporary antibiotic treatment, not because of an acquired resistance, but, because a subpopulation spontaneously adopts a slow-growing state that allows it to persist through the treatment.

Generally, stochastic effects arise when biochemical pathways incorporate interactions involving species that occur at small enough copy number [14] that fluctuations become important. We will see evidence suggestive of a stochastic response in the results of the experiments investigating PhoR/PhoB TCS response presented in Chapter 3. In Chapter 4, we will investigate a very simple bistable reaction system that stochastically switches between stable states.

# 2. EXPERIMENTAL METHODS

The experimental goal of this study is to measure, *in vivo*, the single-cell transient response of the PhoR/PhoB TCS to $P_i$ limitation. We define the PhoR/PhoB TCS response as the change in intracellular concentration of phospho-PhoB caused by a change in $[P_i]$. To measure the intracellular concentration of phospho-PhoB, we conduct the experimentation on an *E. coli* strain that has inserted into the chromosome a fluorescent reporter gene whose transcription in upregulated by phospho-PhoB. Then a measurement of the production rate of fluorescent protein (FP) acts as a proxy for the instantaneous concentration (or number) of phospho-PhoB in a cell. We use fluorescence microscopy and a CCD to image the emission intensity of the FP that has accumulated in single-cells over a time interval. By analysis of the image we quantify the single-cell production rate of FP during the interval.

An experiment runs as follows: after preparing a sample for observation, we take measurements at high $[P_i]$ to characterize the basal rate of FP production. Then we decrease $[P_i]$ and take further measurements for some period of time. After analysis, the data forms a time series reporting the response of the PhoR/PhoB TCS to step-decrease in $[P_i]$ (we will call the step-decrease in $[P_i]$ the $P_i$-signal). Using the time series we can investigate any transience in the TCS response to the $P_i$-signal.

In finding the results reported in Chapter 3, for the experimentation we used glass dishes as the platform for cell deposition, growth, imposition of the $P_i$-signal, and observation. We note that the glass dish platform suffers from operational weaknesses related to mismatch between the length and volume scales of the dish (centimeter and milliliter) and that of the *E. coli* (micrometer and femtoliter), including lack of precision and speed in altering the extracellular environment and lack of control of the movements of the cells within the dish. These weaknesses are pointed out, in particular, because they can be specifically addressed by migrating the experimentation to a microfluidic platform. This should be an experimental goal for the next phase of the project, and a discussion regarding microfluidic devices is included in Appendix A.

To build a more complete picture of the experimental methods we use in this study, the following sections discuss in further detail fluorescent proteins, fluorescence microscopy, the experimental protocol and image analysis.

## 2.1 Fluorescent Protein and Reporter Complex

The isolation of the gene in bioluminescent jellyfish that encodes the green fluorescent protein [16] led to the development of a library of fluorescent proteins with different useful properties, such as color, photostability, maturation time, and brightness (see, for example, [17] for a review). And, recombineering (recombination-mediated genetic engineering) techniques [18] [19] can incorporate into the chromosome the gene encoding a fluorescent protein (FP) along with an arbitrary promoter that controls the transcription of the FP gene (by the normal machinery available in a living cell). These technologies give the investigator the ability to dictate the conditions under which the cell produces FP, providing a versatile tool for probing different cellular processes and systems *in vivo*.

We chose to insert into the chromosome of the *E. coli* the fast-folding variant of the yellow fluorescent protein, called Venus [20]. Figure 2.1(a) shows the excitation and emission spectra of Venus [21], with the maximum excitation and emission at 515 and 528 nm, respectively. In the chromosomal insertion, the Venus gene is encoded downstream from the promoter of PhoA (*phoAp*), a Pho protein whose transcription is upregulated by phospho-PhoB. Also in the insertion is the gene of the membrane protein TatB, so that Venus and TatB are produced together as a complex. TatB carries the Venus to the membrane, and the fluorescent reporter complex diffuses in the membrane in two-dimensions where it is easily imaged.

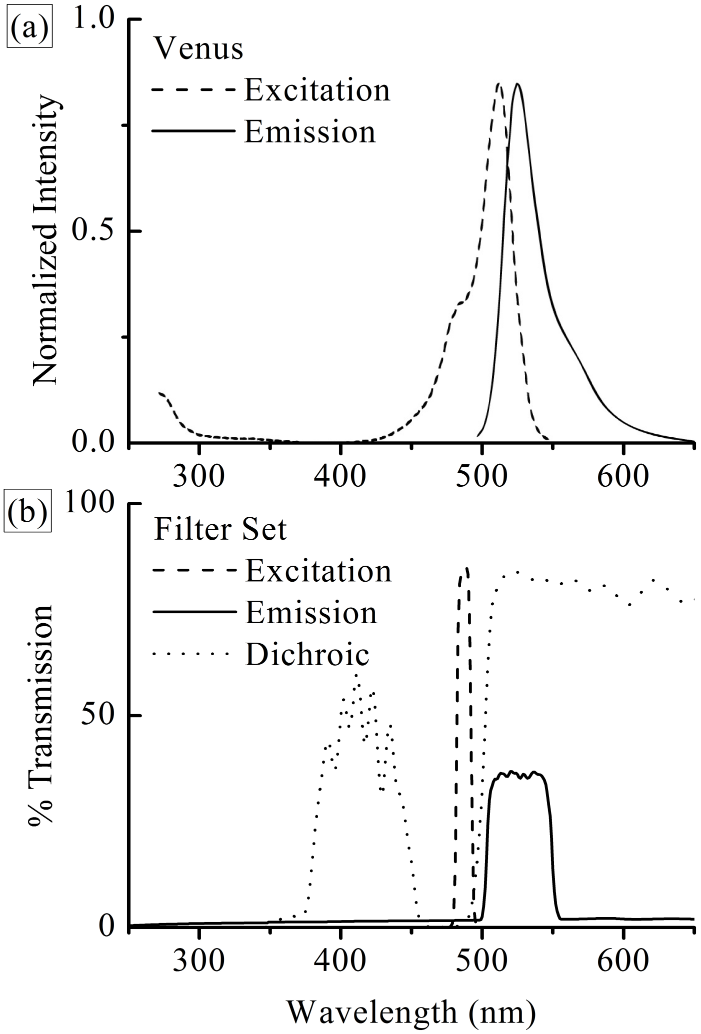


Figure 2.1. Excitation and emission spectra of the Venus fluorescent protein compared to the transmission spectra of the filter set used to image Venus. (a) The emission spectra (solid) of Venus is red-shifted from the excitation spectra (dashed). (b) The measured transmittance spectra of the optical components in the fluorescence microscopy set up used to image Venus. The excitation filter is a narrow bandpass filter used to select the 488 nm line from the Argon laser. The dichromatic mirror (dichroic) reflects the 488 nm excitation beam to the sample and transmits the resulting red-shifted Venus emission collected by the objective. The emission filter transmits the Venus emission to the camera while attenuating stray light from the excitation beam and other sources.

### 2.1.1 The Advantage of Venus Photobleaching

As discussed previously, basal transcription of the Pho genes ensures the Pho proteins are continuously produced. With our insertion of the fluorescent reporter complex downstream the PhoA promoter, Venus is also continuously produced in the cells we study. The accumulation of Venus over the entire (unknown) cell history poses a technical challenge for the single-cell emission intensity measurement, because, each cell in the field of view enters the measurement time interval at a different brightness. In order to maximize the sensitivity of the CCD to all the single-cells in the field of view, it is preferable that each cell begins the measurement interval at the same brightness. We exploit the ability to photobleach Venus as a means by which to reduce all the cells in the field of view to the same brightness, effectively erasing their unknown history.

Photobleaching is a photon-induced chemical damage that causes an FP to lose the ability to fluoresce. The incidence of photobleaching has been found to correlate approximately with the total number of excitation-emission cycles, as the FP is more likely to undergo the damage while in the excited state (or, in an operational sense, the incidence of photobleaching FP in a bulk sample increases with power of excitation beam and the duration of exposure to the beam). This number generally depends on the molecular structure of the FP and is reported as its photostability [17].

The illumination protocol for the data acquisition is as follows. Before the data acquisition, we conduct a calibration experiment in order to tune the length of the "bleaching"' interval and the power of the excitation beam during the bleaching interval so that, at the maximum rate of Venus production, after the bleaching interval the Venus accumulated in every cell in the field of view (FOV) is photobleached. During data acquisition, we precede each measurement interval with a bleaching interval during which the field of view is continuously exposed to the excitation beam, photobleaching all the cells in the FOV. At the end of the bleaching interval we illuminate the FOV with the excitation beam and record the emission intensity (as a reference for the measurement). Then the FOV is subjected to the dark, measurement interval during which Venus is produced in the cells. After that, we, again, shine the excitation beam onto the FOV and record the emission intensity (the measurement).

As described in further detail in Section 2.4, for a given cell, by taking the difference of the emission intensity before and after the measurement interval and dividing by the duration of the measurement interval we calculate a quantity that we infer as the rate of Venus production in the cell.

## 2.2 Fluorescence Microscopy

To record the single-cell Venus emission intensity in an image that we can analyze, we arrange a laser, camera, objective and supporting optical components in the standard epifluorescence microscopy setup. The setup is schematically depicted in Figure 2.2. We will generically discuss the setup before describing the specific components used to measure the single-cell Venus emission intensity.

In epifluorescence microscopy, a light source is conditioned and passed through an excitation filter, a narrow bandpass filter that isolates an excitation wavelength. This monochromatic beam travels to a dichromatic mirror (dichroic), which reflects in a band containing the excitation wavelength, but transmits at higher wavelengths. The dichroic reflects the excitation beam through the objective and through the sample containing the cells. This excitation beam causes some FP in the cells to fluoresce, and a part of the emission is collected by the objective. The collected light travels to the dichroic, where it is transmitted to the CCD for recording.

To measure the single-cell Venus emission intensity, we use, as the excitation source, an Argon-ion laser that emits at 488 nm. The excitation filter transmits the band 488±10 nm, eliminating other Argon lines, while the dichroic almost completely reflects these wavelengths. The transmittance of the dichroic increases rapidly from 495-510 nm, above which it transmits at least 75 percent. In order to capture as much of the emission spectrum as possible, the emission filter has a broader transmittance

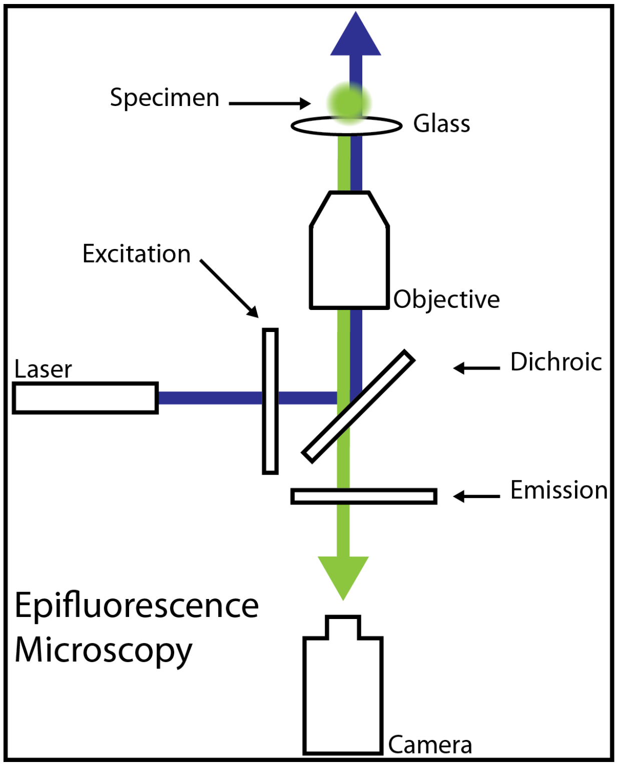


Figure 2.2. Epifluorescence microscopy illumination path. After conditioning (lenses not shown) the excitation source (laser) passes through a narrow bandpass excitation filter to select the appropriate excitation wavelength. The dichromatic mirror (dichroic) reflects the excitation beam towards the sample. The excitation beam excites fluorescent molecules in the sample. The resulting red-shifted emission from the sample is gathered by the objective and transmitted by the dichroic. The emission beam is transmitted by the bandpass emission filter and focused on the CCD for imaging.

Table 2.1
Optical components used in the fluorescence microscopy setup. The table presents product and provider information regarding major components in the setup.

| Component | Product | Provider |
|---|---|---|
| Laser | M177G Air-cooled Argon Ion | SpectraPhysics |
| Microscope | IX-71 Inverted | Olympus |
| Objective | PlanApo 40x NA0.95 $\infty$/0.11-0.23 | Olympus |
| CCD | XR/MEGA-10Z Turbo ICCD | Stanford Photonics |
| Excitation | 488/10nm single band filter | Semrock |
| Emission | 525/20nm single band filter | Semrock |
| Dichroic | 505nm edge long-pass square filter | Semrock |

band then the excitation filter, about 525±20 nm. The measured transmittance spectra of the filter set (excitation filter, dichromatic mirror, and emission filter) chosen to image Venus is shown in Figure 2.1(b). In the figure, for reference the transmittance spectra are diplayed directly below the Venus excitation and emission spectra. A 40x objective is used to collect the fluorescence emission and focus the image for recording on a cooled, dual-intensified CCD. Table 2.1 lists each of the components used in the setup by product name and provider.

### 2.2.1 Fluorescence Intensity

It is important to note that the value of the single-cell Venus emission intensity that we report can be arbitrarily changed by adjusting any of three experimental parameters: the power of the laser, the gain of the intensifier, or the length of the measurement interval. The optimal parameter settings produce the most significant signal by maximizing the number photons emitted by the FP that are recorded by the

CCD and minimizing background, while also giving a suitable dynamic range over which to compare measurements of different intensity.

To address the issue of dynamic range, the parameters are adjusted so that the measurement of the basal production rate (measured in high $[P_i]$) can “just be seen” by eye in the CCD image. This gives a large dynamic range in which to image the cells as they produce FP at higher rates, while still retaining the ability to measure the basal production rate. Measuring the basal rate in every experiment is important because it gives a standard relative to which higher FP production rates can be compared.

A working set of the parameters was found by trial-and-error, taking a few considerations into account. First, increasing the power of the laser increases the number of times each FP is excited during an image acquisition which can give a better signal. But this also increases the rate at which the FP photobleach. Though this might be desirable when the cells are intentionally bleached, laser power must not be so high that the FP are significantly photobleached during an image acquisition. Increasing the intensifier gain increases the signal as well as the background. Increasing the length of the measurement interval increases the amount of FPs produced during the interval in a cell. Though this does increase the single-cell signal without affecting the background, the interval cannot be so long that cells grows significantly during the duration of the measurement interval.

## 2.3 Sample Preparation

This section describes the method used to prepare a sample for observation.

### 2.3.1 *E. coli* Strain and Culturing

The construction of the *E. coli* strain (K-12 BW39341) with the *phoAp*=TatB-Venus chromosomal insertion was constructed in the laboratory of our collaborator, Professor Barry L. Wanner, Biological Sciences, Purdue University. This lab also cultured and prepared the *E. coli* samples for the experimentation. The goal of the culturing is to put the cell population in a reproducible state just before experimentation. The desired state corresponds to the exponential growth of the population.

For the experiments discussed in Chapter 3, the cultures were picked up from the Wanner Lab on the day of the experiment. Provided with the samples was MOPS minimal media at various $[P_i]$ for use during the experiments.

### 2.3.2 Constructing Sample for Observation

The glass dishes used in this study were formed using commercially available FLEXIperm® silicone chambers. The chambers are open on both sides, and one side can be reversible adhered to clean glass forming a dish. *E. coli* are deposited in a single layer on the clean glass inside the dish, and the MOPS media sits in the ~1 mL volume above the cells. Before cell deposition, the clean glass is treated with Poly-L-lysine (Sigma) to electrostatically adhere the cells to the glass. This keeps the cells near the glass and within the focal volume of the objective.

To keep the FLEXIperm® silicone chambers clean, they were autoclaved weekly for 55 minutes at 250 degrees Farenheit. If the chambers were to be used more than once a week, they were soaked overnight in a 50% ethanol solution before use.

We used 25 mm glass coverslips to close the chamber and form the dish. Before use the glass was cleaned by the following steps: (1) 24 hour soak in 5% solution of Contrad detergent, (2) 30 minute sonication on highest power, (3) 10x wash with distilled water, (4) 3x wash with ultra-high quality water, (5) 24 hour soak in 1 M HCl, (6) 30 minute sonication on highest power, (7) 10x wash with distilled water, (8) 3x wash with ultra-high quality water, (9) 3x wash with methanol. The cleaned glass can be stored in methanol for a few weeks.

The Poly-L-lysine treatment of the clean glass was achieved by the following steps: (1) pipette 70 $\mu$L of 0.1 mg/mL solution of Poly-L-lysine onto the surface of glass, (2) incubate 30 minutes to allow the polymer to coat the surface, (3) aspirate the solution off the glass, (4) rinse dish with ultra-high quality water, and (5) let glass air dry.

With the sample platform prepared, 100 $\mu$L of the cell culture was pipetted onto the glass and allowed to settle for 20 minutes. Then the cell solution was aspirated off the glass and the dish was washed three times with MOPS media to remove all cells that did not adhere. Finally, 1 mL of MOPS media was pipetted onto the dish to provide the cells with nutrients and food during the course of the experimentation. Change in $[P_i]$ was accomplished by manually exchanging the 1 mL of MOPS media for a MOPS media with a different $[P_i]$ using the same procedure (aspirate old media, wash dish three times with new media, pipette 1 mL of new media onto the dish).

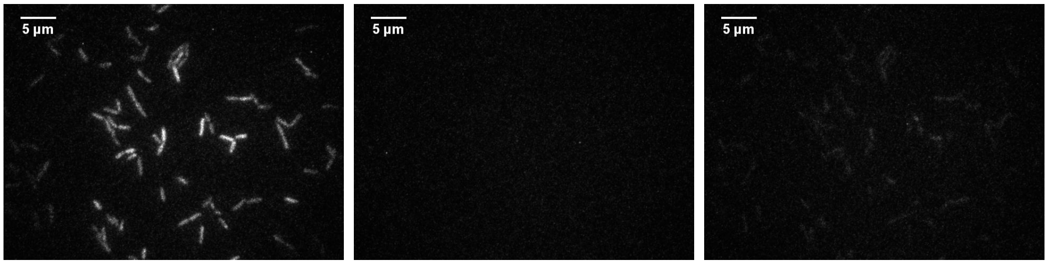
During the experimentation, the temperature of the glass dish was maintained at 34 degrees Celsius by means of heater.

We do note that the Poly-L-lysine deforms and flattens [22] cells adhered to the glass. As this might possibly effect the cell physiology, experiments have been conducted using Cell-tak (BD), in place of the Poly-L-lysine, to adhere the cells to the glass. Though not reported in the next chapter, the experiments using Cell-tak give the same qualitative results.

## 2.4 Image Analysis

The goal of the image analysis is to convert an acquired image, which includes the single-cell Venus emission intensity of many *E. coli* in the field of view, to a set of datapoints that each represents the fluorescence intensity of a single cell. The first step in the image analysis is constructing masks that coincide with the shape of a cell. Once the appropriate masks are found, it is straightforward to read off the value recorded for each pixel inside the mask.

To guide the following discussion, a representative set of images acquired during experimentation is shown in Figure 2.3. In the figure, each image shows the fluorescence intensity of cells in the same field of view at different times during the data

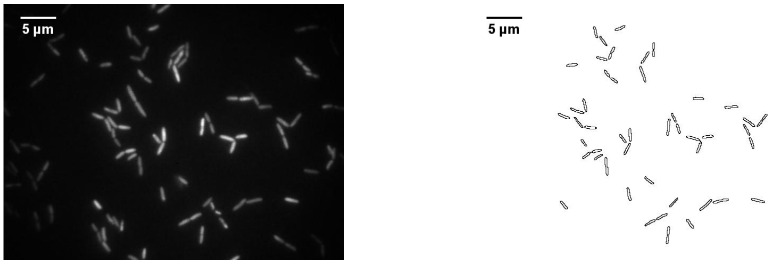


(a) (b) (c)

Figure 2.3. Single-cell fluorescence intensity imaging. The fluorescence intensity of a field of view containing several cells is imaged (a) before the photobleach interval, (b) after the photobleach interval, and (c) after the measurement interval. In this experiment, the photobleach interval was set at three minutes and the recovery interval at five minutes.

acquisition. Figure 2.3(a) is taken at the start of the photobleaching interval and shows the accumulation of Venus over the cells' history. Figure 2.3(b) is taken at the end of the bleaching interval. The cells appear dark because all the Venus has photobleached. The time Figure 2.3(b) is taken also corresponds to the start of the measurement interval. Figure 2.3(c) is taken after the dark, measurment interval, and the cells show some fluorescence intensity. For this experiment, the bleaching interval was set to three minutes and the measurement interval to five minutes. So Figures 2.3(a) and 2.3(b) were acquired three minutes apart with the laser illuminating the FOV for that entire period. And, Figures 2.3(b) and 2.3(c) were acquired five minutes apart, during which time the field of view was dark.

Cell masks are constructed using a time-series of images collected as a movie during the bleaching interval. For brighter cells the movie is taken deeper into the bleaching interval so that no CCD pixel is saturated. The images of the movie are averaged together so that noise in the background is supressed. The average image corresponding to the bleaching interval between Figures 2.3(a) and 2.3(b) is shown in Figure 2.4(a).



(a) Average image.

(b) Cell masks.

Figure 2.4. Constructing masks to coincide with the shape of the cells. (a) The image constructed by averaging together subsequent frames in a movie acquired during the bleaching interval. (b) Cell masks are constructed by processing the average image and selecting areas in the processed images that visually correspond to cells.

The average image is fed through a bandpass filter and thresholded, and then compared, by eye, to a brightfield image of the same field of view. Areas that can be trivially selected in the thresholded image and that correspond visually to single cells in the brightfield are chosen as masks. Figure 2.4 shows the masks that result from processing 2.4(a).

If $I_i(0)$ is the average pixel value (arbitrary units) inside the $i$th mask at the end of the bleaching interval, $I_i(\Delta t)$ is the average pixel value inside the same mask at the end of the measurement interval, and $\Delta t$ is the duration of the measurement interval, then we report the rate of Venus production for the $i$th cell as

$$\text{Production Rate (a.u.)} = \frac{I_i(\Delta t) - I_i(0)}{\Delta t}.$$

In the following chapter we discuss the results of the experiments that use the methods just described.

# 3. EXPERIMENTAL RESULTS

In this chapter we describe the results of experiments conducted according to the methods discussed in Chapter 2. Briefly, cells of *E. coli* strain K-12 BW39341 (with chromosomal insertion $phoAp$=TatB-Venus) were immobilized with Poly-L-lysine in clean glass dishes. The samples were then washed and incubated in 2 mM $[P_i]$ MOPS minimal media for a few generations. During the incubation period single-cell Venus emission intensity measurements were collected to determine the basal rate of protein production. After the incubation, the media was manually exchanged for MOPS with $[P_i]$ equal to either 0, 5, 10, 25 or 50 $\mu$M and further measurements were collected to determine the change in the rate of protein production induced by the PhoR/PhoB TCS.

Figure 3.1 shows the results from the image analysis of the measurements taken during the five experiments. In the figure, each data point (red circle) gives the rate of Venus production (in arbitrary units) for a single cell. The grey, dashed line shows when the $P_i$-signal was imposed (when $[P_i]$ was changed from excess, 2 mM, to the limited concentration indicated on each graph).

An immediate conclusion that we can draw from Figure 3.1 is that the threshold for the PhoR/PhoB TCS response is greater than the 4 $\mu$M $[P_i]$ reported previously; the threshold appears to be between 25 and 50 $\mu$M – in 3.1(c) there is a significant and obvious change in the observed single-cell protein production measurements after the change to 10 $\mu$M, but not so in 3.1(e) after the change to 50 $\mu$M.

Also, we note the large cell-to-cell variability in the measurements in Figures 3.1(a), (b), (c) and (d). We believe the variability is a property of the PhoR/PhoB TCS response, and that it demonstrates the stochastic nature of the TCS.

In the following sections we identify qualitative characteristics in TCS response suggested by the data in Figure 3.1. The strength of these suggestions are limited

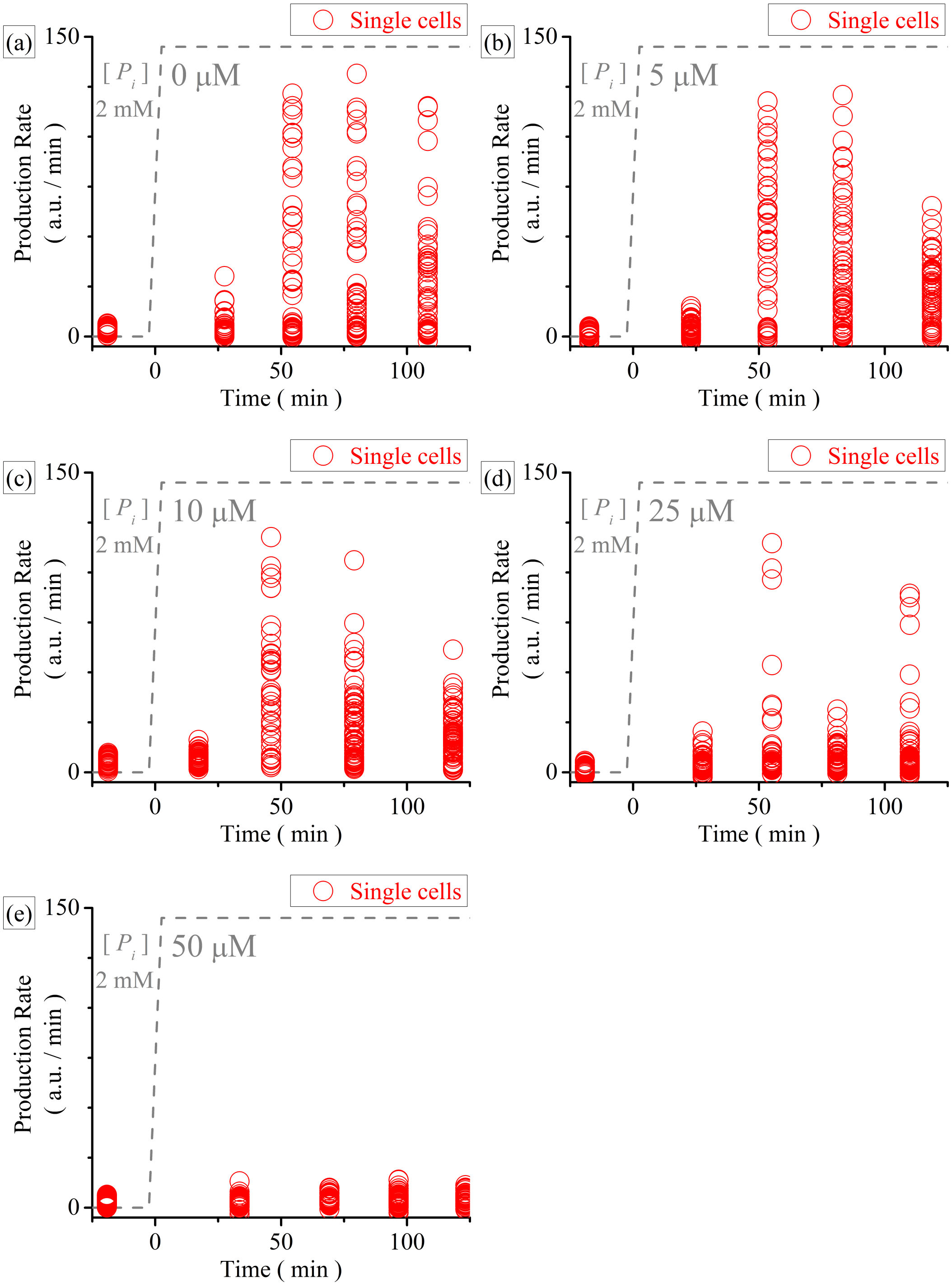


Figure 3.1. Single-cell PhoR/PhoB TCS response to various $P_i$-signals. Each graph shows the results of an experiment in which $[P_i]$ was changed from 2 mM (excess) to (a) 0, (b) 5, (c) 10, (d) 25 or (e) 50 $\mu$M. In each graph, the red circles represent a single-cell measurement of protein production as described in the text. The change in $[P_i]$ (indicated by the grey, dashed curve) occurs at time zero in all experiments.

by the number of experiments conducted for each $P_i$-signal and the coarse sampling rate of the experiments. These two concerns are currently being addressed by further experimentation performed by another investigator in our lab, Dr. Yoriko Lill, and we note that the results found so far by Dr. Lill are consistent with the data presented in Figure 3.1.

## 3.1 Bistability and Subpopulation Isolation

Figure 3.1(a), (b) and (c) show clearly that the PhoR/PhoB TCS responds to the $P_i$-signal: in these three figures it is obvious that for a fraction of cells the rate of Venus production increases above the basal rate. But, another fraction of the cells seems never to leave the basal state. This is consistent with the findings of Zhou *et al.* [4], that, using flow cytometry techniques, measured enough single cells to conclusively determine that the PhoR/PhoB TCS response was bistable. Though only suggested by the data in Figure 3.1, the measurement is consistent with the bistable hypothesis.

As an aside, we note that, in regards to the segmentation of the data, the assumption of bistability is equivalent to changing the nature of the question we ask of the data. Instead of asking "what is the character of the single-cell PhoR/PhoB TCS response?", we can ask "what is the character of the single-cell PhoR/PhoB TCS response in cells significantly outside of the basal state?". After changing the question we ask of the data, we would proceed in identical fashion from the population segmentation to the analysis and interpretation presented in the remainder of this chapter. Though the bistability hypothesis has implications for the model construction discussed in Chapter 4, it is not absolutely necessary for the data analysis and interpretation.

From the assumption that there are two distinct subpopulations represented in the data, we separate the data points into the subpopulations and analyze the subpopulations independently. In order to find a cutoff that delineates the subpopulations,

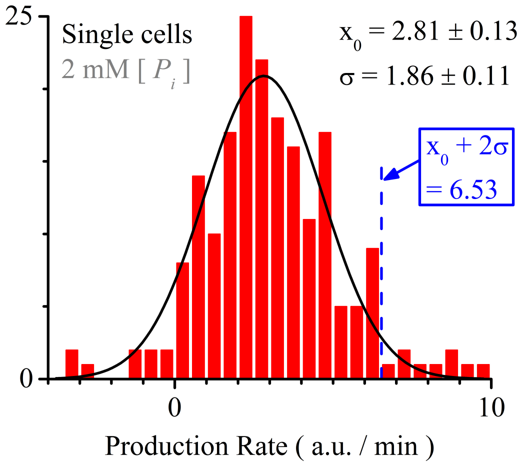


Figure 3.2. Distribution of single-cell PhoR/PhoB TCS basal activity. The single-cell measurement of protein production from multiple experiments at excess $[P_i]$ (2 mM) are combined in the histogram. The histogram is fit to a Guassian to find the center ($x_0$) and standard deviation ($\sigma$). The cutoff used to decide whether cells are in the basal state is chosen as $x_0 + 2\sigma$.

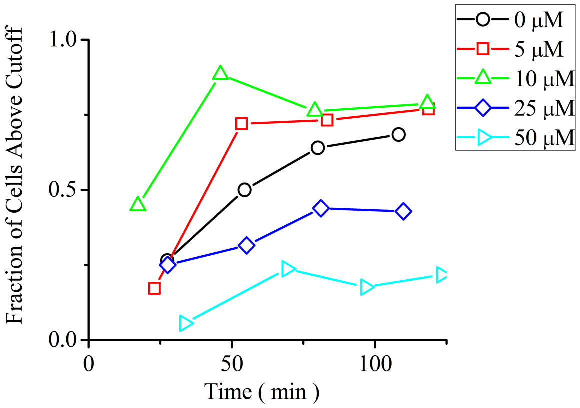


Figure 3.3. Fraction of single-cells above the cutoff. The cutoff, described in the text, isolates cells in the basal state. The plot shows, for the timepoints after the $P_i$-signal, the fraction of cells outside the basal state.

we use the data from all five experiments before the imposition of the $P_i$-signal to place bounds on the population with protein production at the basal rate. Figure 3.2

collects this data and shows the distribution of measured values of single-cell basal protein production (in excess, 2 mM, $[P_i]$) from all the experiments. We characterize the basal state by fitting the histogram in Figure 3.2 to a Gaussian distribution. Two standard deviations above the mean captures nearly 98% of the subpopulation in the basal state, and delineates a reasonable cutoff between the subpopulations. We note that reporting a “TCS response” is subject to the definition of the basal state which will depend on the experimental method and analysis.

If the PhoR/PhoB TCS is not affected by the $P_i$-signal, we would expect no more than about 2% of the population to be above the cutoff. For the timepoints after the $P_i$-signal in every experiment, Figure 3.3 shows the fraction of cells above the cutoff. Surprisingly, there exists a substantial fraction of cells outside the basal state after the 50 $\mu$M $P_i$-signal, and we must conclude that these cells are affected by that signal or affected by the changing of the media. Even so, as we can see in Figure 3.1(e), the data points in the 50 $\mu$M experiment are bunched so closely that we will not treat the results from this experiment to further analysis.

Upon manually exchanging media from MOPS with 2 mM $[P_i]$ to MOPS with 50 $\mu$M $[P_i]$ basal transcriptional activity of the PhoR/PhoB TCS is measurably perturbed, but this perturbation is significantly different in character to that measured upon changes to 0, 5, 10 and 25 $\mu$M $[P_i]$. We interpret the significant difference in character between the response to the 50 $\mu$M $P_i$-signal versus the response to other signals as indicative of the threshold sensitivity of the PhoR/PhoB TCS.

## 3.2 Summary Statistics: Median and Median Absolute Deviation

We would like understand how the subpopulations behave after the 0, 5, 10 and 25 $\mu$M $P_i$-signal. Because the sample of cells we measured at each timepoint in each experiment can contain outliers (see, in particular, Figure 3.1(d)), we do not believe that the mean and the standard deviation are the most robust characterizations of the central tendency and dispersion, respectively, of each sample. In their place, we calculate the median and the median absolute deviation. The median is the data value that separates the higher- and lower-valued halves of the sample. To calculate the median absolute deviation (MAD) we first calculate the median of the sample, $median(X_i)$, where the $X_i$ are the data points in the sample. Then we calculate the set of deviations between the data points and the median,

$$X_j - median(X_i),$$

where we have, now, indexed the data points with $j$. Then we create another set from the absolute values of the deviations, and take the median of that set,

$$MAD = median_j(|X_j - median(X_i)|).$$

## 3.3 Ramp Up after $P_i$-signal

Figures 3.4 and 3.5 show, in the left-hand column, the original raw data with the subpopulations marked with different symbols and colors. And, in the right-hand column, these figures show the median and MAD of the subpopulations at each timepoint, where the MAD is represented by the bars surrounding the symbols that show the location of the median. Figure 3.6 shows, on the same plot, the medians calculated for timepoints in the 0, 5, 10 and 25 $\mu$M $[P_i]$ experiments.

We interpret Figures 3.4 and 3.5 as suggesting that the rise time to maximal response is signal-independent, while the value at maximal response is signal-dependent. Since our measurements are coarse in time, we cannot support these claims quantitatively.

## 3.4 Overshoot before Settling

An intersting feature of the PhoR/PhoB TCS response suggested by in Figures 3.4 and 3.5, is that the TCS overshoots it final steady-state. In order to demonstrate that this data indicates an overshoot, we need to calculate the error in the estimate of the subpopulation median, and show that the median at the maximal response is

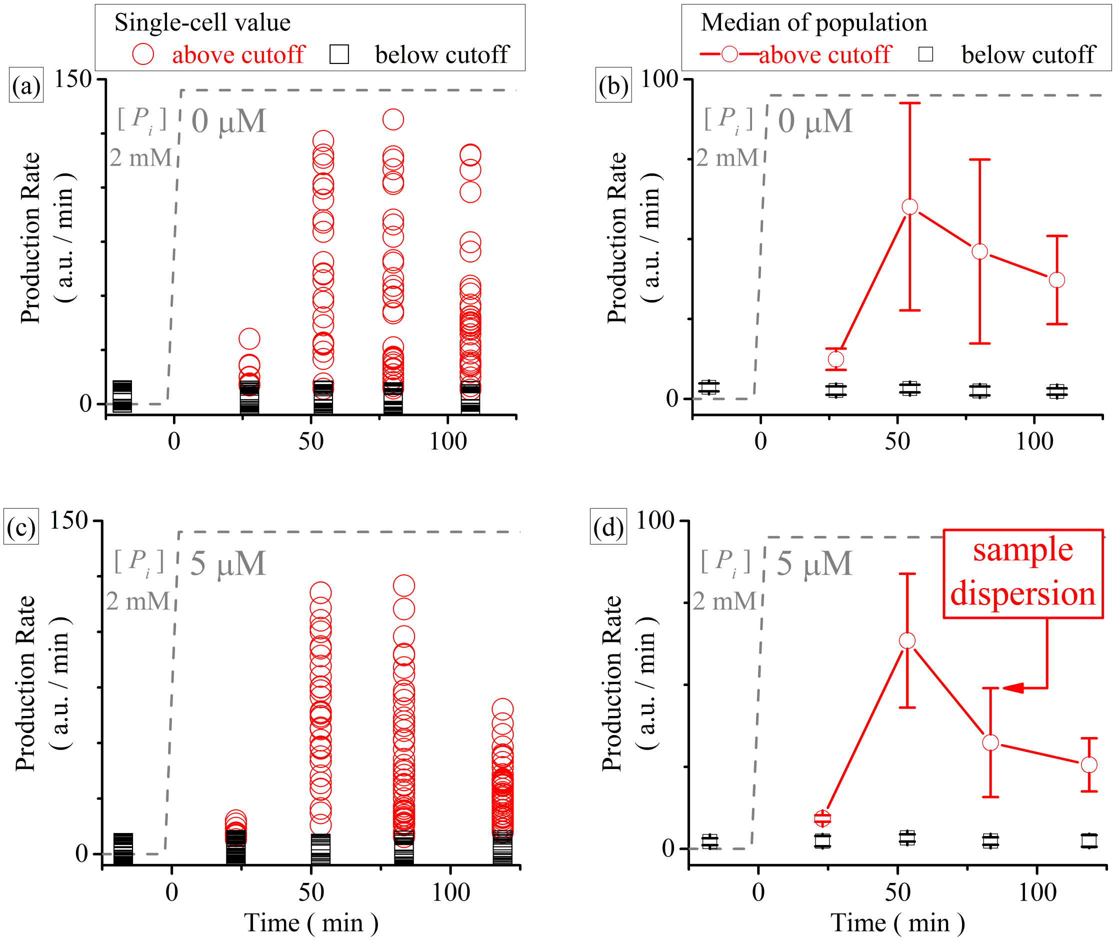


Figure 3.4. Summary statistics describing the behavior of the two subpopulations before and after the $P_i$-signal. (a) The raw data for the 0 $\mu$M experiment has been segmented into the subpopulation in the basal state (black square) and the subpopulation out of the basal state (red circle) for each timepoint. (b) The subpopulations from (a) are characterized by their median (red circle) and median absolute deviation (red bars) showing, respectively, the central tendency and dispersion of each sampling. (c) Same as (a), but for the 5 $\mu$M experiment. (d) Same as (b), but for the 5 $\mu$M experiment.

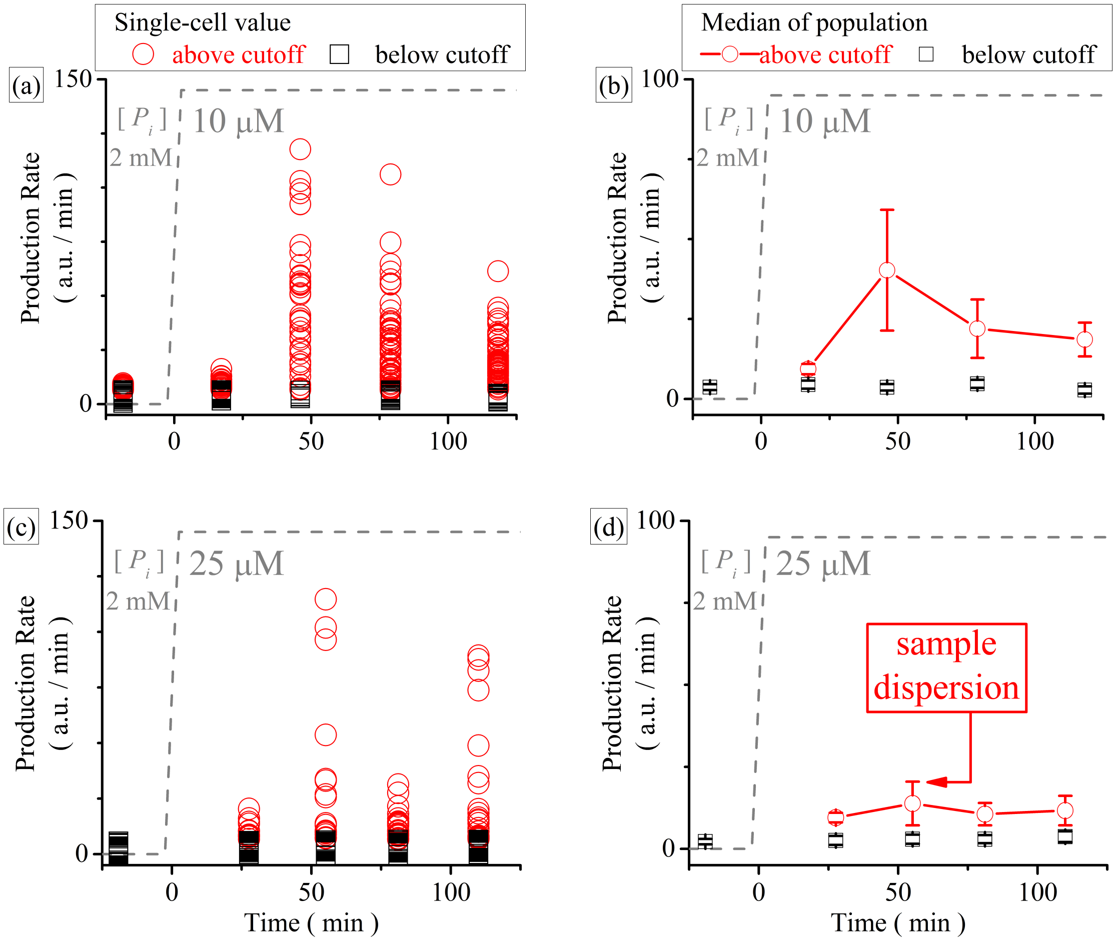


Figure 3.5. Summary statistics describing the behavior of the two subpopulations before and after the $P_i$-signal. (a) The raw data for the 10 $\mu$M experiment has been segmented into the subpopulation in the basal state (black square) and the subpopulation out of the basal state (red circle) for each timepoint. (b) The subpopulations from (a) are characterized by their median (red circle) and median absolute deviation (red bars) showing, respectively, the central tendency and dispersion of each sampling. (c) Same as (a), but for the 25 $\mu$M experiment. (d) Same as (b), but for the 25 $\mu$M experiment.

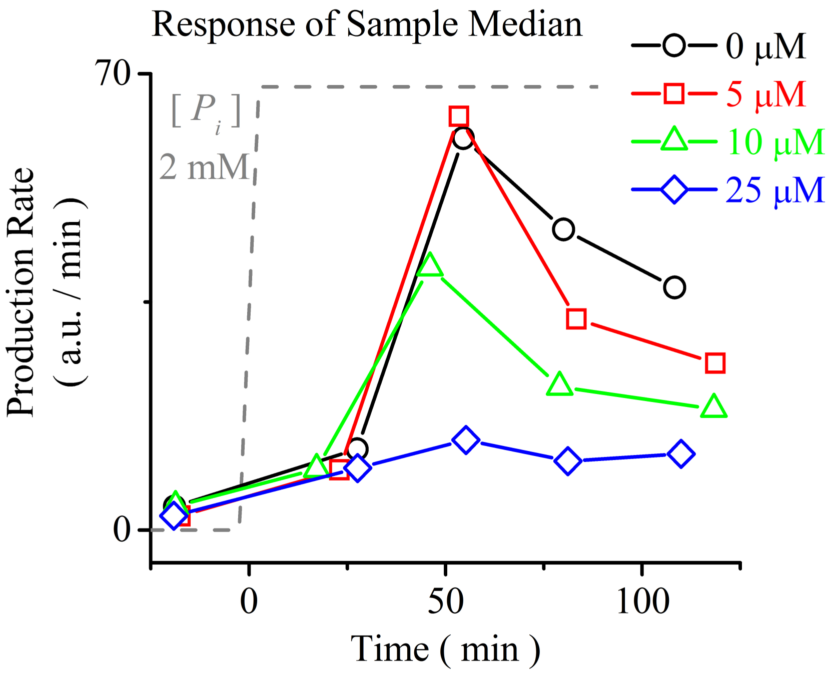


Figure 3.6. Response of subpopulation median to various $P_i$-signals. For the time points in the different experiments, the median of the subpopulation out the basal state is compared. As a reference to the basal state, the first time point in 2 mM [$P_i$] is included. The grey, dashed line indicates the arrival of the $P_i$-signal.

significantly different then the median at the long-time response. But the error in the sample median is difficult to reliably estimate by the common boot strap method since the sample size is not very large (30-50 datapoints).

To find an estimate of the error in the estimate of the central tendency, we turn to more familiar statistics. If we use the mean and standard deviation as indicators of central tendency and dispersion, respectively, then we can calculate the standard error in the mean in the usual way. For mean $\bar{X}$, standard deviation $s$ and sample size $N$, the standard error in the mean is

$$\sigma_{\bar{X}} = \frac{s}{\sqrt{N}}.$$

Figure 3.7 shows the mean and standard error in the mean calculated for the subpopulation out of the basal state for time points in 0, 5, and 10 $\mu$M $[P_i]$ experiments. We see in Figure 3.7 for the 5 and 10 $\mu$M experiments that the location of the center of the sample at the maximal response level is significantly higher then the center of the sample at the long-time response level – the TCS response overshoots before settling. The effect is also present in the 0 $\mu$M experiment, though the standard error in the mean is larger because the presence of outliers in the samples at each time point in this experiment corrupts the estimate of the mean and standard deviation.

For the 0, 5 and 10 $\mu$M experiments, Table 3.1 shows the statistics describing the subpopulation out of the basal state for time points after the $P_i$-signal.

## 3.5 Summary

We set out to draw qualitative conclusions regarding the response behavior of the PhoR/PhoB TCS from the data presented in Figure 3.1. Significantly, our measurement demonstrates that, at the population level, the transducer ramps up and overshoots before settling to its final steady-state. We also found that our measurements were consistent with the transducer having, at the single-cell level, a threshold sensitivity and a bistable response that stochastically switches between fixed points. We have been cautious in the interpretation of this data as it is limited, though recent

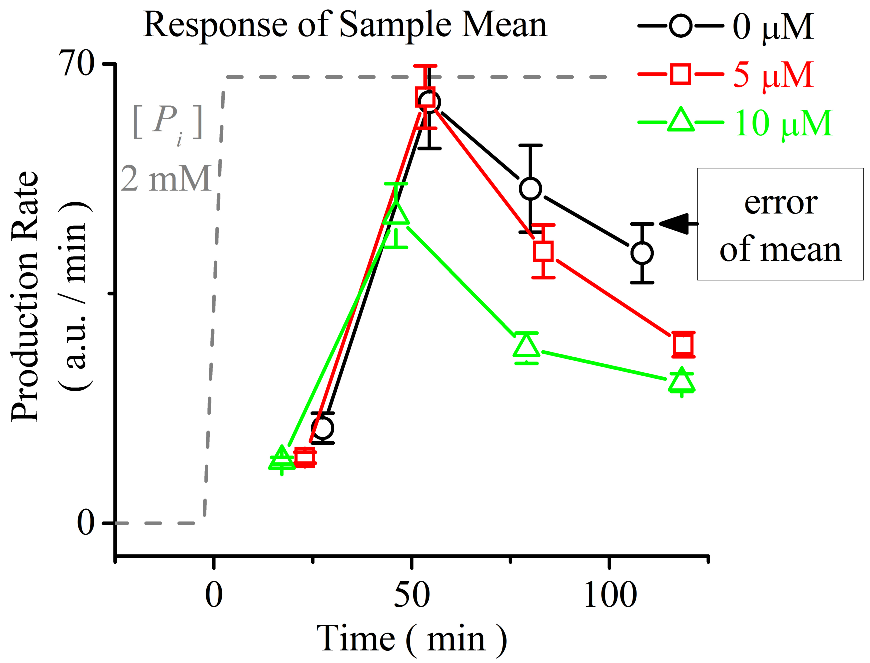


Figure 3.7. Response of subpopulation mean to various $P_i$-signals. For the time points in the different experiments, the men of the subpopulation out of the basal state is compared. The bars indicate the standard error in the mean. The grey, dashed line indicates the arrival of the $P_i$-signal.

Table 3.1
Statistics describing the subpopulation out of the basal state for time points after the $P_i$-signal in the 0, 5 and 10 $\mu$M experiments.

| **0 $\mu$M Experiment** | Data points above cutoff | | | | |
|---|---|---|---|---|---|
| minutes after $P_i$-signal | $Md$ | $MAD$ | $\bar{x}$ | $s$ | $\sigma_{\bar{x}}$ |
| 27.5 | 12.4 | 3.4 | 14.5 | 6.8 | 2.3 |
| 54.5 | 60.1 | 32.4 | 64.2 | 35.9 | 7.0 |
| 80.1 | 46.1 | 28.8 | 51.0 | 37.4 | 6.6 |
| 108.3 | 37.2 | 13.8 | 41.2 | 27.9 | 4.5 |

| **5 $\mu$M Experiment** | Data points above cutoff | | | | |
|---|---|---|---|---|---|
| minutes after $P_i$-signal | $Md$ | $MAD$ | $\bar{x}$ | $s$ | $\sigma_{\bar{x}}$ |
| 23 | 9.2 | 1.0 | 10.0 | 2.7 | 0.8 |
| 53.4 | 63.5 | 20.4 | 65.0 | 28.5 | 4.8 |
| 83.3 | 32.4 | 16.6 | 41.4 | 29.0 | 4.0 |
| 118.7 | 25.6 | 8.1 | 27.2 | 13.1 | 1.8 |

| **10 $\mu$M Experiment** | Data points above cutoff | | | | |
|---|---|---|---|---|---|
| minutes after $P_i$-signal | $Md$ | $MAD$ | $\bar{x}$ | $s$ | $\sigma_{\bar{x}}$ |
| 17.2 | 9.3 | 1.6 | 9.5 | 2.4 | 0.5 |
| 46.0 | 40.3 | 18.9 | 46.9 | 29.9 | 4.9 |
| 79.0 | 22.0 | 9.1 | 26.7 | 18.6 | 2.3 |
| 118.3 | 18.6 | 5.3 | 21.4 | 10.4 | 1.4 |

results found by another investigator in our lab are consistent with the data in Figure 3.1.

The strategy we have developed for analyzing the data of single-cell fluorescence microscopy experimentation should prove valuable as further experimentation is conducted.

# 4. MODELING

The goal of this chapter is to demonstrate that simple reaction systems can qualitatively reproduce some of the features observed in PhoR/PhoB TCS signal response. We do not try to find a comprehensive model for the TCS, that is beyond the scope of this work. Instead, we identify basic model components that give rise to the observable response features. The value in this approach is that we can build an intuition regarding how model components contribute to the overall behavior of the system; giving a toolbox of model components that can be systematically integrated into a complete model of the PhoR/PhoB TCS or other signaling systems.

The chapter is comprised of three parts. We begin the chapter by introducing, in the first part, the basic principles for building reaction systems, and the tools we will use to solve the time evolution of a system. In the second part, we review the model of the PhoR/PhoB TCS proposed by Kierzek *et al.* [23] in which a detailed and comprehensive reaction system is proposed as a model for the PhoR/PhoB TCS. The model of [23] is presented to constrast the approach to model construction we take in the final part of the chapter. In that part, we propose and investigate a series of related simple reaction systems and build a framework in order to understand their behavior.

## 4.1 Reaction System and Reaction-Rate Theory

In analogy to chemical reaction theory, we can think of a signal transduction network as a collection of reactions that, over time, changes the concentration or number of the biomolecular species present inside the cell by consuming or creating copies of the species. One of the smallest reaction systems that is biologically interesting consists of the birth and death of a single species, $X$, and in the simplest case, the birth and death of $X$ occur at constant rates, $c_1$ and $c_2$, respectively. This birth-death reaction system can be represented schematically as

$$\varnothing \xrightarrow{c_1} X \xrightarrow{c_2} \varnothing.$$

For a closed system, in the case that the number of copies of $X$ is large and the reaction volume, $V$, fixed, the classical reaction rate theory gives [24] the change of the mean number of $X$ (or, equivalently the concentration of the species) in time as

$$\dot{X} = c_1 - c_2 X. \tag{4.1}$$

Note that we use the same designation (in this case, $X$) for both the species name and the mean number of the species. The mean number $X$ evolves deterministically in time according to the solution to Equation 4.1, which is

$$X(t) = \frac{c_1}{c_2} + (X(0) - \frac{c_1}{c_2})e^{-c_2 t}.$$

From intial value $X(0)$, the mean number exponentially approaches the steady-state value of $X_{ss} = \frac{c_1}{c_2}$ with a time constant equal to the inverse of the death rate $c_2$. The appearance of the two scales, $\frac{c_1}{c_2}$ and $c_2$, suggests a scale invariant form of the solution.

Finding the scale invariant solution can reduce the number of free parameters in the system, allowing us to focus on the system behavior and not get caught up in the details of precisely defining the all the parameters. Starting with Equation 4.1, we can rescale time $c_2 t \rightarrow \tau$ and define $Z = \frac{X}{X_{ss}}$. Doing so reduces Equation 4.1 to $\dot{Z} = 1 - Z$ (where the derivative is taken with respect to $\tau$), which has solutions

$$Z(\tau) = 1 + (Z(0) - 1)e^{-\tau}.$$

The variable $\tau$ is dimensionless and represents the number of time constants that have elapsed. Figure 4.1 shows two solutions $Z(\tau)$ that start at a different initial values of $Z$.

The reaction-rate formulation can be used to study the dynamics of the mean numbers of species of a reaction system of arbitrary complexity. For example, we can add to the birth-death scheme a reversible isomerization

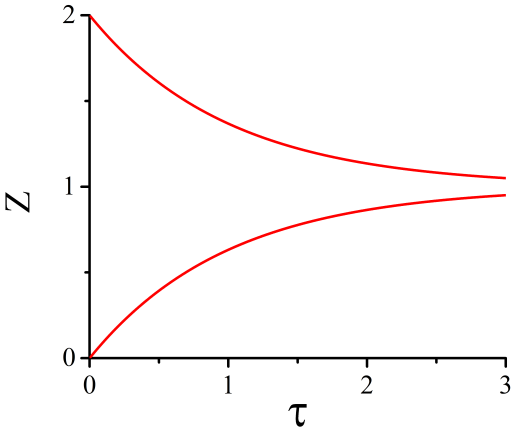


Figure 4.1. Deterministic solutions of the constant rate birth-death reaction system. The time evolution of $Z$ evolves deterministically according to $Z(\tau) = 1 + (Z(0) - 1)e^{-\tau}$. The curves begin at the different initial conditions, $Z(0) = 2$ and $Z(0) = 0$, and exponentially approach the steady-state $Z = 1$ with time constant 1.

$$X \rightleftharpoons Y,$$

or a two species reaction

$$X + Y \longrightarrow Z,$$

or any number of variations or complications of these types of reactions (three-species reactions are usually treated as a series of two-species reactions). To add further complexity we can define the reaction rates as functions of one or more species.

In a reaction system that specifies $N$ reactants interacting in $M$ reaction channels, the time derivative for the change in mean number of the $i^{th}$ species will generally depend on the mean numbers of all the species and on the reaction structure and kinetic parameters:

$$\dot{X}_i(t) = f_i(X_1(t), ..., X_N(t)).$$

The solution to this set of coupled, ordinary differential equations gives the time evolution of the mean number of every species in the system. In most cases, finding an analytic solution to this set of differential equations is impossible; but there are established numerical methods and readily available software implementations of these methods to computationally solve such reaction systems (we use the ‘ode45’ solver in MATLAB whose method is described in [25]).

### Monte Carlo Method

As noted by Daniel Gillespie [26] [27] [28] the reaction-rate theory makes the explicit assumptions that the time evolution of the system is deterministic and continuous in the $X_i$ space. For some systems, especially for systems with species that occur at small numbers, the continuum approximation fails.

In place of the reaction-rate formalism, Gillespie proposed [27] a kinetic Monte Carlo method to study the time evolution of a reaction system. The method randomly samples trajectories of the reaction system, where each trajectory is a timed path through the state space comprised of the number of each species in the system. From a collection of trajectories the time evolution of various statistics of the number of a species, including the mean number, can be determined.

In the description of his Monte Carlo method, Gillespie demonstrated that [26] [27] [28], if well-stirred, the time evolution of the reaction system can be rigorously described as a continuous-time Markov process. So each step in a trajectory depends only on the current state of the system and the structure and kinetic parameters describing the reactions. In the demonstration, he connected these kinetic parameters to those of the reaction-rate formulism – no additional information regarding the reaction system is required.

Gillespie also described a computer algorithm, commonly called the Stochastic Simulation Algorithm (SSA), to build the simulated trajectories. The SSA, and its many refinements, is currently a pervasive tool in simulating the behavior of biological reaction systems. Appendix B describes in more detail Gillespie’s SSA.

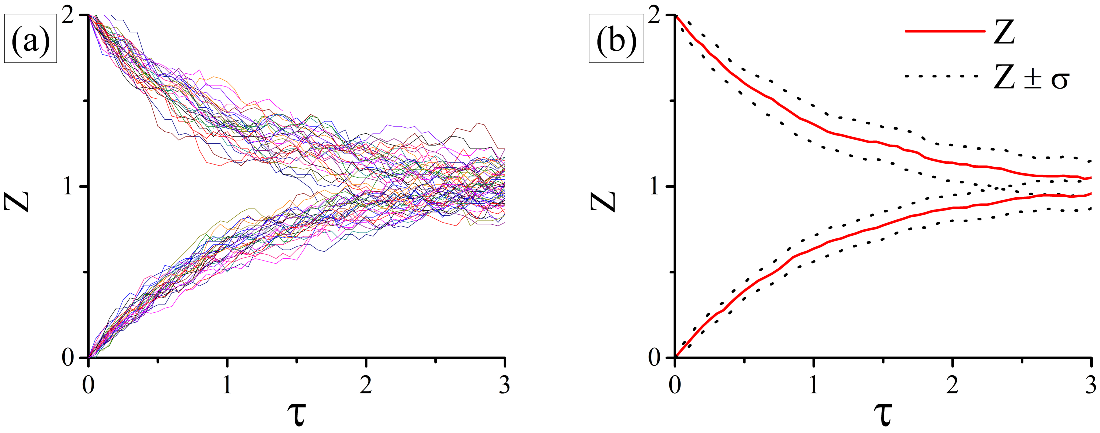


Figure 4.2. Monte Carlo (stochastic) analysis of the constant rate birth-death reaction system in scale-invariant form. (a) A collection of trajectories that simulates the time evolution of the reaction system from different intial states. (b) From the simulations, the time evolution of the sample mean, $Z$, and standard deviation, $\sigma$, are calculated.

Figure 4.2 presents a Monte Carlo (stochastic) analysis of the scale invariant form of the constant rate birth-death reaction system. Figure 4.2(a) shows sixty simulated trajectories of the reaction system, with half the trajectories starting below and half starting above the steady-state value. And Figure 4.2(b) shows the time evolution of the sample mean and standard deviation as calculated from the trajectories. The mean response is in good agreement with the deterministic solution shown in Figure 4.1.

For the various reaction systems we present later in this chapter, we will use both the deterministic and stochastic analyses as tools to study system response. The deterministic analysis is less computationally expensive and yields precise knowledge regarding the time evolution of the mean number as described by the differential equations. But this solution is not always a good predictor of the system's behavior. In particular, as we will see in Section 4.6, for a bistable system where switching between states depends on stochastic fluctuations, the long-time mean number found by the deterministic analysis incorrectly describes the behavior discovered via the stochastic analysis. Aside from giving a more accurate picture regarding the behavior of a reaction system, the stochastic analysis allows the computation of the time evolution of various sample statistics in addition to the mean.

## 4.2 Remarks Regarding Model Freedom

Now that we have a basic framework in which to build and study reaction systems, we must decide how to put together a reaction system that models the behavior of the PhoR/PhoB TCS. And, we must do this with incomplete and imperfect knowledge of the actual signal transduction network. As we can imagine, there are many reaction systems that might sufficiently reproduce the observed behavior while having network structure and kinetic parameters that might sufficiently overlap what is known of the protein signaling network. The criteria constraining the freedom in model construction are ill-defined and must be imposed, *ad hoc*, by the investigator. Before broadly commenting on the two contrasting philosophies regarding the problem of model freedom, we take the opportunity to directly quote Bolouri and Davidson [29] who elegantly characterize the problem:

> All modeling is an abstraction of reality. The only exact model of any system is the system itself. So, when we set out to build a model of a system..., we must make a choice about the level of detail and type of features that the model should represent. To a large extent, this is dictated by the characteristics of the system being studied, the type of experimental data available, and the type of questions that we wish to address through modeling.

On the one hand, the investigator can choose to faithfully use all the information currently available to build the most complete model possible and test that model against the data. Success with this approach can imply that no significant component in the real system has been left out of the model. But, inevitably, this approach suffers as the proposed model might not be structurally identifiable [30] – parts of the model might be untested by the experimental data and thus cannot be uniquely determined to be necessary for the model's behavior. Also, with the excess of parameters, by overspecifying the model, it can be difficult to build an intuition that can be applied to other modeling problems.

On the other hand, the investigator can choose simplicity as the criteria with which to constrain model freedom. This is the strictly phenomenological approach; the investigator seeks the smallest and simplest model that reproduces certain features of the measurement. Success with this approach can significantly constrain the unknown structure of the real system. And, this approach lends itself to building an intuition regarding model parts that can be widely applied to further problems. But, the approach suffers as a direct and obvious connection to the known parts of the real system can be lost. It can be difficult to connect a phenomenological model with quantitative experimental data.

Of course, these constrasting approaches are at the extremes, and every effort at model construction follows some intermediate path. The model of the reconstructionist, by necessity, must include some 'lumping' or modularization as the real system is not perfectly known. And the model of the phenomenologist needs to maintain some connection with the real system, otherwise the conclusions lose their relevance.

We will adopt the minimalist approach and investigate the behavior of simple reaction systems, drawing broad connections between the behavior of the reaction systems and that of the PhoR/PhoB TCS. Before doing that we review, in the following section, the model of the PhoR/PhoB TCS proposed by Kierzek *et al.* [23]. Kierzek *et al.* take the constrasting approach, using the best available knowledge to construct a detailed and comprehensive reaction system as a model of the PhoR/PhoB TCS.

## 4.3 A comprehensive model of the PhoR/PhoB TCS

Kierzek *et al* [23], taking the reconstructionist approach to model creation, construct a detailed and comprehensive reaction system to model the response dynamics of the PhoR/PhoB TCS. The reaction system consists of interactions among the fourteen species – variants of PhoR, PhoB, and mRNA; fluorescent reporter and RNA polymerase – listed in Table 4.1. Table 4.2 breaks the twenty nine reactions considered in the model into signaling (protein-protein interactions), production (birth reactions) and degradation (death reactions) categories. For brevity and clarity, many of the rows in Table 4.2 describe more than one of the reactions included in the Kierzek model. It is worthwhile to note that, although it is comprehensive, this reaction system is not complete, as it does not include species like PhoU or the Pst proteins believed to play some role in the signal transduction [10].

In the following discussion we point to several notable features included in the construction of the Kierzek model. The model handles protein production in a two-step process, treating transcription and translation separately. This gives the flexibility to independently vary the time scales of transcription and translation, allowing the investigators to correctly model the relatively slow translation of PhoR. In the signaling reactions of the model, the dimerization of PhoR and PhoB are explicitly treated, as the dimerization is biologically necessary for the proteins to be functioning. And, the phosphorelay channel leading to the creation of phospho-PhoB is treated as a two-species interaction between PhoB and a phosphorylated dimer of PhoR. In the model, the production of fluorescent reporter is explicitly treated so that the number of fluorescent reporter produced in a simulation can be directly compared to data from flow cytometry experiments.

Monte Carlo simulations of the Kierzek model using biologically realistic parameters showed the external signal causes the appearance of two subpopulations (bistability), a population that continued to produce reporter at a basal level and another that produced reporter at a higher level. The simulations also showed the flow of

Table 4.1
Subset of species Kierzek *et al* include in their model of the PhoR/PhoB TCS [23].

| Type | Name | Description |
|---|---|---|
| PhoR Variants | $hk$ | PhoR protein |
| | $hkP$ | phospho-PhoR |
| | $hk2$ | dimer-PhoR |
| | $hk2P$ | phospho-dimer-PhoR |
| PhoB Variants | $rr$ | PhoB protein |
| | $rrP$ | phospho-PhoB |
| | $rrP2$ | dimer-phospho-PhoB |
| Complexes | $rr - hk2P$ | PhoB/phospho-dimer-PhoR |
| | $rrP - hk2$ | phospho-PhoB/dimer-PhoR |
| – | $fr$ | fluorescent reporter |
| mRNA Variants | $mRNAhk$ | mRNA of PhoR |
| | $mRNArr$ | mRNA of PhoB |
| | $mRNAfr$ | mRNA of fluor. rep. |
| – | $RNAP$ | RNA polymerase |

Table 4.2

Reactions among the species that form the model of the PhoR/PhoB TCS constructed by Kierzek *et al* [23]. Abbreviations used in the table: dimer. is dimerization; phosph. is phosphorylation; transc. is transcription; const. transc. is constitutive transcription; transl. is translation; degrad. is degradation. $rrP2$ (dimer of phospho-PhoB) is the transcription factor necessary for active transcription. Modulation of the rate of dephosphorylation of $hk2P$ (phosphorylated dimer of PhoR) acts as a proxy for the external signal. For brevity and clarity, many of the rows in the table describe more than one of the reactions included in the model.

| Signaling | |
|---|---|
| Description | Reactions |
| Dimer. $hk$ | $hk + hk \rightleftharpoons hk2$ |
| Auto/Dephosph. $hk2$ | $hk2 \rightleftharpoons hk2P$ |
| Phosph. $rr$ | $rr + hk2P \longrightarrow rr - hk2P \longrightarrow rrP + hk2$ |
| Dephosph. $rrP$ | $rrP + hk2 \longrightarrow rrP - hk2 \longrightarrow rr + hk2$ |
| Dimer. $rrP$ | $rrP + rrP \rightleftharpoons rrP2$ |
| **Production** | |
| Description | Reactions |
| Active Transc. $rr(fr)$ | $RNAP + rrP2 \longrightarrow mRNArr(fr)$ |
| Basal Transc. $rr(fr)$ | $RNAP \longrightarrow mRNAreg(fr)$ |
| Transl. $mRNAreg(fr)$ | $mRNArr(fr) \longrightarrow mRNArr(fr) + rr(fr)$ |
| Const. Transc. $hk$ | $RNAP \longrightarrow mRNAhk$ |
| Transl. $mRNAhk$ | $mRNAhk \longrightarrow mRNAhk + hk$ |
| **Degradation** | |
| Description | Reactions |
| Degrad. $mRNA$ | $mRNArr, mRNAhk, mRNAfr \longrightarrow \varnothing$ |
| Degrad. $rr$ species | $rr, rrP, rrP2 \longrightarrow \varnothing$ |
| Degrad. $hk$ species | $hk, hkP, hk2, hk2P \longrightarrow \varnothing$ |
| Degrad. $fr$ | $fr \longrightarrow \varnothing$ |

members of one subpopulation into the other (stochastic switching). Significantly, these features of the model response are in good qualitative agreement with results from the flow cytometry experiments.

## 4.4 Phenomenological Models

In the remainder of the chapter we will examine simple models that fold together and strip away many of the reactions explicitly treated by Kierzek *et al.* Our motivation is a desire to understand what components of the Kierzek model contribute to the signal response. More generally, we desire to understand what components are necessary for a certain type of behavior.

### 4.4.1 Parameter Choice

It is not obvious how to use parameters described in the Kierzek model or elsewhere in the simple models we will study, because so many of the reactions have been lumped or ignored. So we will follow a general principle that the parameters used in our phenomenological models should not be biologically absurd nor lead to biologically absurd results. Though heuristic at best, this principle implies, for example, that the death of a species (corresponding to degradation or dilution of a protein via cell splitting) should be the slowest process in the reaction system, or that we should not have protein numbers in the billions. We offer further discussion as parameter choices are made.

Though this the connection to the real system and to real parameters is an inherent limitation in our approach, because parameters are only loosely constrained (and there are only a few parameters in the simple systems), we are able to investigate model behavior as a function of the parameters. Giving a more complete picture of system behavior allows us to build intuition regarding model components that can be intelligently integrated to build a model from the ground up.

### 4.4.2 Connecting Models to Experiments

In the experimental results discussed in Chapter 3, we measured changes in the rate of protein production, which depends indirectly on the intracellular concentration of phospho-PhoR. We assume that the transience in the measurement reflects the transience in the number of phospho-PhoR in the single-cells. So in the reaction systems we study, we look for changes in the number of an autoregulating species that qualitatively matches the changes in protein production we measured.

The features that we will demonstrate in the simple reaction systems are bistability, stochastic switching, ramp up after signal, overshoot and threshold sensitivity.

## 4.5 Autoregulation in Birth-Death System

We begin by investigating a reaction system composed of a single autoregulating species. To model an autoregulating species we can refine the birth-death system discussed previously so that the birth rate is functionally dependent on the number of the species produced in the birth reaction:

$$\varnothing \xrightarrow{c_1(X)} X \xrightarrow{c_2} \varnothing.$$

The elementary birth reaction is already a tremendous simplification of the "nuts-and-bolts" of protein production as the birth rate, $c_1(X)$, must account for the sequential processes of transcription (protein-DNA interaction creating an mRNA transcript from a DNA gene), translation (ribosome-mRNA interaction converting an mRNA transcript into a polypeptide chain) and post-translational modification (e.g., folding of polypeptide chain into functional protein).

### 4.5.1 Hill Function for Birth

Following Walczak *et al.* [31], we will use the Hill equation as the functional form of the birth rate:

$$c_1(X) = \frac{c_0 X^n + c_1 K^n}{X^n + K^n}. \tag{4.2}$$

The Hill equation causes the birth rate to toggle between constant values $c_0$ and $c_1$ depending on the concentration of $X$. At large $X$ ($X >> K$), the value of the Hill equation is near $c_0$, and at small $X$ the value is near $c_1$. The parameter $K$ sets the half-saturation point; at $X = K$ the value of the Hill equation is the average of $c_0$ and $c_1$. And, $n$ is the so-called Hill coefficient which changes the steepness of the saturation curve.

The Hill equation finds pervasive application in chemistry and biochemistry to model a process with saturating kinetics. In this context, the parameter $n$ describes cooperativity in the reaction process.

Investigators commonly use the Michaelis-Menten equation (Hill equation with $n = 1$) to model two-state (basal and active) transcription initiation that depends, in a saturating fashion, on the level of transcription factor (see, for example, [29] and [32]). As previously discussed, the Pho proteins (including PhoB) are always produced at a basal rate. Upon an increase in the intracellular concentration of phospho-PhoB the rate of production of the Pho proteins increases; phospho-PhoB is a transcription factor that binds to the DNA to upregulate the transcription initiation rate (and, thus, the production rate) of the Pho proteins. Due to the finite time of phospho-PhoB/DNA binding events, the increase in transcription rate eventually saturates. So, it would be appropriate to use a Hill equation to describe transcription initiation of a Pho gene as a function of phospho-PhoB.

But we propose to use the Hill equation to describe species production; we are, in a sense, ignoring processes other than transcription that are required for protein synthesis (namely, translation and post-translational modification). These additional processes could be added later, if needed.

As we analyze the Hill birth-death reaction system, we will take advantage of the inherent scale $K$ to rewrite the Hill equation. We introduce this form of the Hill equation now, so we can plot various Hill curves in a manner consistent with what is to come. Defining $Z \equiv X/K$, $\Gamma = c_0$ and $\alpha = \frac{c_0}{c_1}$ we rewrite Equation 4.2 as

$$c_1(Z) = \Gamma \frac{Z^n + \frac{1}{\alpha}}{Z^n + 1}. \tag{4.3}$$

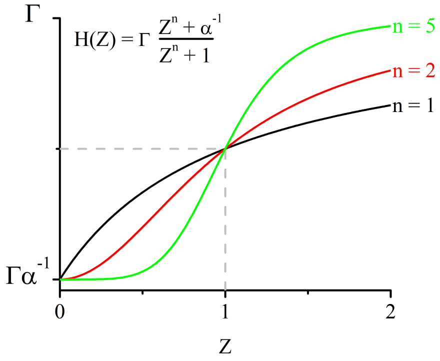


Figure 4.3. Hill curves for different values of $n$.

Figure 4.3 shows representative Hill curves for different values of $n$.

For the auto-regulating species, the time derivative of the change in mean number of that species is

$$\dot{X} = \frac{c_{10}X^n + c_{11}K^n}{X^n + K^n} - c_2X. \tag{4.4}$$

The Hill birth-death system will form the backbone of the reaction systems we introduce in later sections. More so, we will find that the form of the structure of the steady-state solutions to Equation 4.4 is a general feature of the simple extensions to birth-death reaction system that we introduce. As we can learn much that will be applicable to subsequent systems, in the following sections we will study Equation 4.4 in detail. These sections will provide tools and a strategy to our study of the simple extensions to the Hill birth-death reaction system.

To simplify our study of Equation 4.4 we rescale time $c_2t \to \tau$ and define $Z = X/K$, $\Gamma = \frac{c_{10}}{c_2K}$ and $\alpha = \frac{c_{10}}{c_{11}}$ yielding

$$\dot{Z} = \Gamma\frac{Z^n + \frac{1}{\alpha}}{Z^n + 1} - Z. \tag{4.5}$$

### 4.5.2 Steady-states of Hill Birth-Death System

A few comments are necessary to preface the mathematical treatment that follows. At the start of our modeling effort we put together several reaction systems simply related to the Hill birth-death reaction system. As we varied model parameters and studied the behavior of the reaction system (using the ode45 solver and the SSA), we were struck by how unpredictably the behavior varied as a result of small changes in model parameters. In an effort to understand this, we stripped away model parts to isolate the sources of the strong parameter dependence. Eventually, we were left with just the Hill birth-death system, and we were surprised by the richness of phenomenon coming forth from this seemingly simple reaction system. The following treatment summarizes our current understanding of the dependence of the behavior of the Hill birth-death reaction system on the parameters $n$, $\Gamma$ and $\alpha$.

In Equation 4.5 the term $\Gamma\frac{Z^n+\frac{1}{\alpha}}{Z^n+1}$ is the propensity for a birth reaction and determines the probability that the reaction will occur. Similarly, $Z$ is the propensity for a death reaction. A steady-state of $X$ occurs at levels of $X$ that balance the birth and death propensities; from such a fixed point it is equally probably to either gain or lose a copy of $X$.

The steady-state ($\dot{Z} = 0$) solutions to Equation 4.5 take the form of roots of the $(n+1)$-order polynomial

$$Z^{n+1} - \Gamma Z^n + Z - \frac{1}{\alpha}\Gamma = 0. \tag{4.6}$$

Solutions of Equation 4.6 give the steady-states or fixed points of the reaction system. In the remainder of this section, we will investigate the dependence of the structure of the steady-states on the parameters $n$, $\Gamma$ and $\alpha$. Since we have some freedom to set these parameters, we would like to build a picture of how each affects the behavior of the system.

### $n$ Sets Maximum Number of Steady-states

The Hill coefficient, $n$, sets the maximum number of solutions to Equation 4.6, and, thus, the number of fixed points of $Z$. Since $X = ZK$ is the number of species $X$ and corresponds to the number of a protein it must be greater than or equal to zero. We restrict our interest to those roots of the steady-state polynomial that are real and positive.

First, we will demonstrate that for $n = 0$ and $n = 1$ Equation 4.6 has only one solution that is real and positive. Then, we will show graphically that for $n \geq 2$ Equation 4.6 either has one or three solutions that are real and positive, depending on choice of $\Gamma$ and $\alpha$.

For $n = 0$, Equation 4.6 reduces to

$$Z - \frac{\Gamma}{2}(1 + \frac{1}{\alpha}) = 0,$$

and $Z$ has only one solution at $Z = \frac{\Gamma}{2}(1 + \frac{1}{\alpha})$. With $n = 0$ the Hill term is constant-valued and equal to $\frac{\Gamma}{2}(1 + \frac{1}{\alpha})$. So the steady-state solution of $Z$ is the ratio of this constant birth rate to the death rate (1), in agreement with what we found for the constant rate birth-death system we examined at the beginning of this chapter.

For $n = 1$, the steady-state equation reduces to

$$Z^2 + (1 - \Gamma)Z - \frac{1}{\alpha}\Gamma = 0.$$

The second-order polynomial has a positive discriminant and two solutions:

$$Z_{\pm} = \frac{\Gamma - 1}{2} \pm \frac{\sqrt{(\Gamma - 1)^2 + 4\alpha^{-1}\Gamma}}{2},$$

both of which are real, but only one of which is positive ($Z_+$).

For $n \geq 2$, we can rearrange Equation 4.6

$$Z^{n+1} - \Gamma Z^n = -Z + \frac{1}{\alpha}\Gamma,$$

and we can interpret the solutions of Equation 4.6 as the points of intersection of the polynomial $p(Z) = Z^{n+1} - \Gamma Z^n = Z^n(Z - \Gamma)$ and the line $l(Z) = -Z + \frac{1}{\alpha}\Gamma$. The intersections are best shown graphically, and Figure 4.4 shows $p(Z)$ and $l(Z)$ for some choices of $n$, $\Gamma$ and $\alpha$.

Because $l(Z)$ has positive offset and negative slope and $p(Z)$ dips into the fourth quadrant before increasing to the first quadrant, the intersections between the line

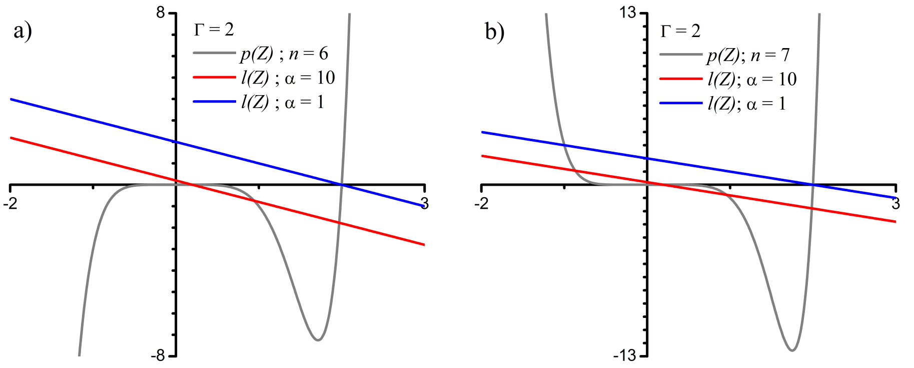


Figure 4.4. The solutions to Equation 4.6 can be visualized as the points of intersection of $p(Z) = Z^{n+1} - \Gamma Z^n$ and $l(Z) = -Z + \frac{1}{\alpha}\Gamma$. For Hill coefficient $n \geq 2$, $p(Z)$ and $l(Z)$ intersect at either one or three points for positive $Z$ depending on the values of $\Gamma$ and $\alpha$. Plot a) shows a representative graph for $n \geq 2$ and even, and b) shows a representative graph for $n \geq 2$ and odd.

and polynomial will always occurs in the fourth quadrant. More so, at minimum, there is one intersection in this quadrant, and, since it occurs for positive $Z$, this will be one of the steady-states accessed by the reaction system. As shown in Figure 4.4 the choice of $\Gamma$ and $\alpha$ can yield two additional intersections for $Z > 0$.

### Choosing $\Gamma$

The additional intersections arise when the slope of $p(Z)$ is such that it increases above $l(Z)$ then decreases below it. In order to cross $l(Z)$ from above, at some point the negative slope of $p(Z)$ must be less than than the negative slope of the line. The maximum negative slope of $p(Z)$ occurs at the inflection point $\frac{n-1}{n+1}\Gamma \equiv Z_{crit}$ (found by solving $\frac{d^2p}{dZ^2} = 0$). So the criteria sufficient for the additional intersections is $\frac{dp}{dZ}|_{Z_{crit}} < -1$. Solving this criteria we find $\Gamma > (\frac{n+1}{n-1})^{\frac{n-1}{n}} \equiv \Gamma_{crit}$, which restricts the values of $\Gamma$ that allow for the possibility of the two additional intersections.

Figure 4.5 shows the $\Gamma_{crit}$ as a function of $n$, which tends to 1 for large $n$. Remembering the definition of $\Gamma = \frac{c_{10}}{c_2 K}$, to satisfy the criteria for the two additional

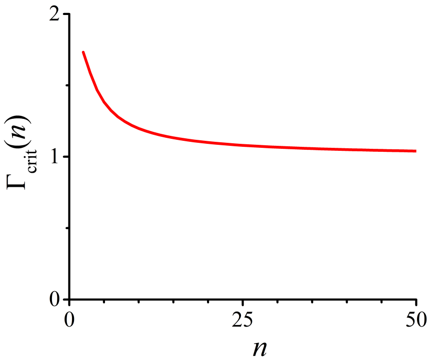


Figure 4.5. Values of $\Gamma_{crit}$ as a function of $n$. If $\Gamma < \Gamma_{crit} = (\frac{n+1}{n-1})^{\frac{n-1}{n}}$ then it is not possible for $p(Z) = Z^{n+1} - \Gamma Z^n$ and $l(Z) = -Z + \frac{1}{\alpha}\Gamma$ to intersect more than once in the fourth quadrant (for positive $Z$).

intersection the ratio between active production rate and the death rate must be greater than the half-saturation number of $X$.

### Choosing $\alpha$

We are trying to find the bounds on $\alpha$ between which the line $l(Z)$ intersects $p(Z)$ more than once in the fourth quadrant (for positive $Z$). In order for this to occur, $p(Z)$ must increase above $l(Z)$ and then decrease below the line. This situation is satisfied if the maximally negative slope of $p(Z)$ at point $(Z_{crit}, p(Z_{crit})$ is less than $-1$ and $l(Z)$ lies between the two points at which the slope of $p(Z)$ equals $-1$. So, if we draw the tangent lines through these two points at which $p(Z)$ has slope equal to $-1$ we create an envelope within which $l(Z)$ must lie in order to intersect $p(Z)$ three times. Outside of this envelope $l(Z)$ only crosses $p(Z)$ once. Figure 4.6 depicts this envelope for $n = 6$ and two different values of $\Gamma$.

If we can find the two lines tangent to $p(Z)$ of slope $-1$ that create the envelope then we will have found the range of $\alpha$ within which $l(Z)$ crosses the polynomial three times. The first step is to find the two points at which $p(Z)$ has slope equal to $-1$. We can find the $Z$-values of the points by setting the derivative of the polynomial

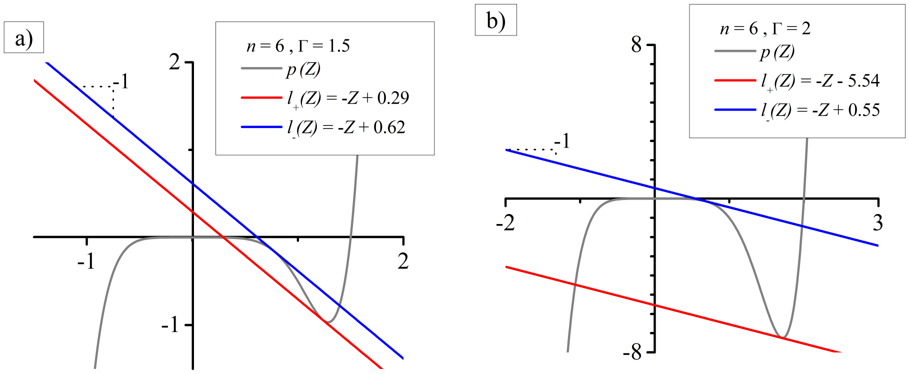


Figure 4.6. If the line $l(Z)$ lies within the envelope created by $l_+(Z)$ and $l_-(Z)$ then it will intersect the polynomial $p(Z)$ three times in the fourth quadrant.

equal to $-1$. Evaluating the derivative and adding one to both sides of the equation, we find that $Z$-values of the points will be the only real solutions to the equation

$$(n+1)Z^n - \Gamma n Z^{n-1} + 1 = 0. \tag{4.7}$$

The $Z$-values of the points, let's call them $Z_\pm$, can be found numerically. So, the points at which $p(Z)$ has slope equal to $-1$ are $(Z_\pm, p(Z_\pm))$. And, the lines that cross these points and have slope equal to $-1$ are

$$l_\pm(Z) = -Z + (Z_\pm + p(Z_\pm)). \tag{4.8}$$

In order for $l(Z) = -Z + \frac{1}{\alpha}\Gamma$ to lie within the envelope, its offset, $\frac{1}{\alpha}\Gamma$, should lie between the offsets of $l_\pm(Z)$. This criteria yields the following condition $\alpha$ must satisfy for $l(Z)$ to cross the polynomial three times in the fourth quadrant:

$$\frac{1}{\Gamma}(Z_+ + p(Z_+)) < \frac{1}{\alpha} < \frac{1}{\Gamma}(Z_- + p(Z_-)). \tag{4.9}$$

On the plots in Figure 4.6 we show the values of the offsets, found numerically according to the described procedure, of the lines $l_\pm(Z)$ for the indicated choices of $n$ and $\Gamma$. Of course, $\alpha$ cannot be negative, so the lower bound for $\alpha$ corresponding to the parameter values in Figure 4.6(b) is zero.

### Negative and Complex Roots of the Polynomial

For completeness, we note that when $n \geq 2$ and odd there is an additional intersection of $p(Z)$ and $l(Z)$ for $Z$ less than zero, as shown in Figure 4.4(b). This point is also a solution to Equation 4.6, but is not a physical solution as it requires the number of proteins to be negative.

Also, for completeness, we note that a polynomial of order $n+1$ has $n+1$ roots in the complex plane. The additional solutions to Equation 4.6 that we have not accounted for are complex numbers that occur in conjugate pairs (because the coefficients of the polynomial are real). So, for example, if $n = 7$ and $\Gamma$ and $\alpha$ are chosen such that Equation 4.6 has three real, positive solutions, then the equation also has the one real, negative solution and four solutions comprised of two pairs of complex conjugates, giving the eight solutions, as expected.

## 4.5.3 Stability of Steady-states

In this section we present a global and local analysis of the steady-states of the Hill birth-death reaction system. The global analysis allows us to understand the stability of the fixed points as a function of the parameters $\Gamma$ and $\alpha$. So this type of analysis helps us build intuition regarding the behavior of the reaction system. But, as we complicate the reaction system in later sections by incorporating additional species, we will only be able to use a local analysis to calculate the stability of the steady-states. In this discussion we will continue to use as an example the reaction system with Hill coefficient, $n$, equal to 6 (see Figures 4.4(a) and 4.6).

### Phase Plane

The phase plane plots $\dot{Z}$ against $Z$ and is a useful tool as it allows us to visualize the tendency for changes in $Z$ and identify which fixed points are attractors (stable).

On a phase plane plot, when $\dot{Z}$ is positive there is a tendency for $Z$ to increase, and when it is negative there is a tendency for $Z$ to decrease.

Fixing $n = 6$ and $\Gamma = 2$, Figure 4.7 shows the phase plane for the Hill birth-death system $\alpha = 0.3$ and 30. When $\alpha = 0.3$, there is only one steady-state solution ($\alpha = 0.3$ places $l(Z)$ outside the envelope shown in Figure 4.6). As shown by the black arrows in Figure 4.7 that emphasize the tendency for changes in $Z$, the one fixed point is stable as there is a tendency for $Z$ to move towards the point. When $\alpha = 30$, there are three steady-state solutions, two of which are stable and one of which is unstable.

In order to characterize the dependence of the structure of the steady-state solutions for all values of $\alpha$ we can collect the information of many phase plane plots in a single plot like that shown in Figure 4.8. The figure shows, for $n = 6$ and $\Gamma = 2$, how the number of fixed points, their stability and their value change with $\alpha$. The sudden appearance of the pair of fixed points with increasing $\alpha$ is called a saddle-point bifurcation. When we fix $\alpha$ and vary $\Gamma$ we find the two bifurcations shown in Figure 4.9.

From the bifurcation diagrams we can learn a few things. First, in the three-solution regime, the steady-state solution structure is not very sensitive to increasing $\alpha$. This is connected to the fact the for $n = 6$ and $\Gamma = 2$ the lower bound on $\alpha$ defining the three-solution regime is zero, and $\alpha$ must always be greater than zero.

On the other hand, for $n = 6$ and $\alpha = 30$, the solution structure is quite sensitive to increasing $\Gamma$. Notably, the value of the largest fixed points appears to increase approximately proportional to $\Gamma$, and the distance between the steady-states increases with increasing $\Gamma$.

It is worthwile to note that the types of dependencies discovered so far are general for the Hill birth-death reaction system regardless of the choice of $n$. With the exception of the negative-valued fixed point that occurs for odd $n$, the one or three positive $Z$ fixed points tend to behave in a manner similar to that found here.

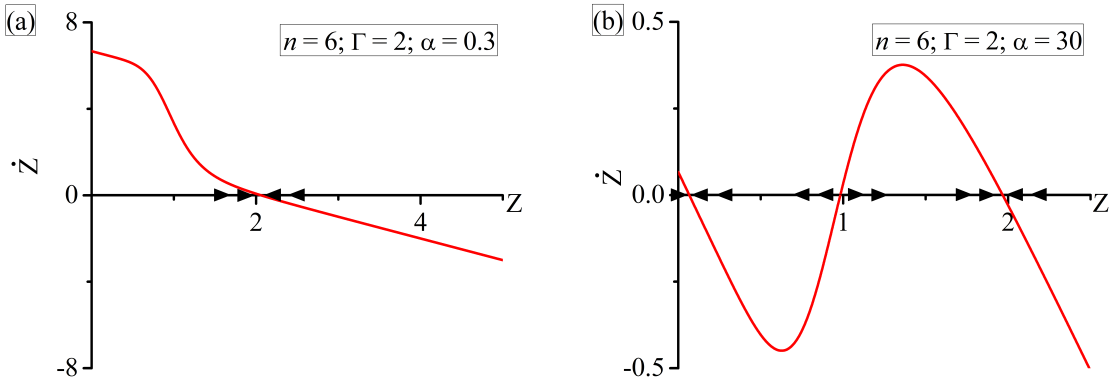


Figure 4.7. Phase plan plots for Hill birth-death system $n = 6$ and $\Gamma = 2$. When $\dot{Z}$ is positive there is a tendency for $Z$ to increase, and when $\dot{Z}$ is negative there is a tendency for $Z$ to decrease. The black arrows show the tendency. (a) $\alpha = 0.3$ yields one fixed point and it is stable and attractive, as the level of $Z$ is driven towards the fixed point. (b) $\alpha = 30$ yields three fixed points. Two are stable and attractive and third is unstable.

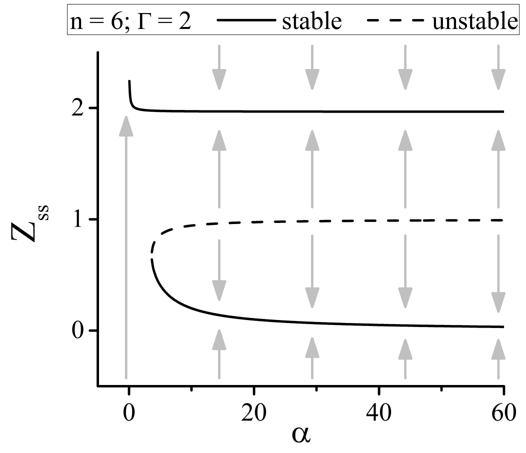


Figure 4.8. Saddle-node bifurcation with variation of model parameter $\alpha$. At small values of $\alpha$ there is a single stable fixed point. Increase in $\alpha$ causes the appearance of a pair of additional fixed points, one stable and the other unstable.

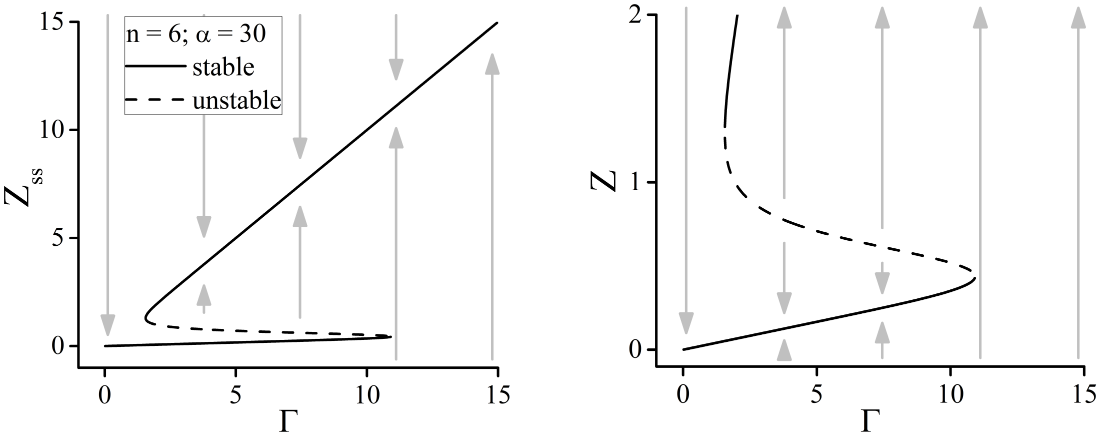


Figure 4.9. Saddle-node bifurcation with variation of model parameter Γ. At small values of Γ there is a single stable fixed point. Increase in Γ causes the appearance of a pair of additional fixed points, one stable and the other unstable. Further increase in Γ causes the disappearance of a pair of fixed points, the unstable one and the stable fixed point that exists at small values of Γ. The plot on the right is a magnification in the vertical axis.

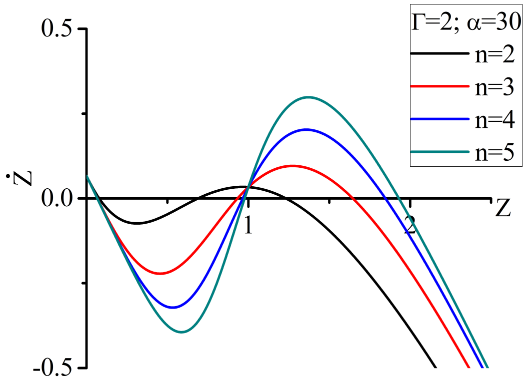


Figure 4.10. Model parameter $n$ controls strength of tendency to stay near a fixed point. For different values of $n$ the phase plot shows $\dot{Z}$ for fixed $\Gamma$ and $\alpha$. As $n$ increases, the values of the local minimum and maximums of $\dot{Z}$ also increase, implying an increase in the strength of the tendency driving the level of $Z$ towards a fixed point.

The parameter $n$ does control the strength of the tendency to stay near a fixed point. Figure 4.10 shows the phase plane for different values of $n$ (keeping $\Gamma$ and $\alpha$ fixed). As $n$ increases, the values of the local minimum and maximums of $\dot{Z}$ also increase, implying an increase in the strength of the tendency driving the level of $Z$ towards a fixed point.

### Local Stability near Fixed Points

A different approach to stability analysis is to characterize the stability of a reaction system in the vicinity of a fixed point. To do so, we find the time evolution of a small deviation from that steady-state. We begin by expanding the function that gives the time-derivative of the species to first order about a steady-state

$$\dot{Z} = f(Z) \approx f(Z_{ss}) + f'(Z)|_{Z=Z_{ss}}(Z - Z_{ss}).$$

Defining the deviation variable as $\tilde{Z} \equiv Z - Z_{ss}$ and using the fact that $\dot{Z}_{ss} = f(Z_{ss})$, we find

$$\dot{\tilde{Z}} = f'(Z)|_{Z=Z_{ss}}\tilde{Z}.$$

So the time evolution of the deviation variable in the vicinity of the fixed point is

$$\tilde{Z}(t) = \tilde{Z}(0)e^{(f'(Z)|_{Z=Z_{ss}})t}.$$

Near the steady-state the deviation variable grows exponentially in time if $f'(Z)|_{Z=Z_{ss}} > 0$ and diminishes if $f'(Z)|_{Z=Z_{ss}} < 0$; in the first case the steady-state is unstable and in the second the state is stable.

## 4.6 Stochastic Switching in the Hill Birth-Death System

We have learned that the Hill birth-death reaction system is bistable for $n \geq 2$ and particular choices of $\Gamma$ and $\alpha$. In this section we will demonstrate that the system is also capable of stochastic switching between the two stable fixed points. By using the ideas developed above, we will find a set of parameters (design a model) that gives the desired behavior.

The goal is to increase the chance that the reaction system randomly jumps between the stable steady-states, or, equivalently, to increase the chance that the level of $Z$ randomly crosses the unstable fixed point. This can be achieved by choosing a small value for $n$ (diminishing the strength of the tendency driving the level of $Z$ towards the fixed points) and a small value for $\Gamma$ (decreasing the distance between the stable steady-states and, thus, the distance that the level of $Z$ must walk before crossing the unstable fixed point).

### 4.6.1 Parameter Choice and Steady-states

We start by choosing $n = 2$. For this choice, Equation 4.6 becomes the cubic equation

$$Z^3 - \Gamma Z^2 + Z - \frac{1}{\alpha}\Gamma = 0. \tag{4.10}$$

Before we find values for $\Gamma$ and $\alpha$, we can take advantage of the fact that an analytic form of the discriminant of a cubic equation exists, and use that form to visualize the

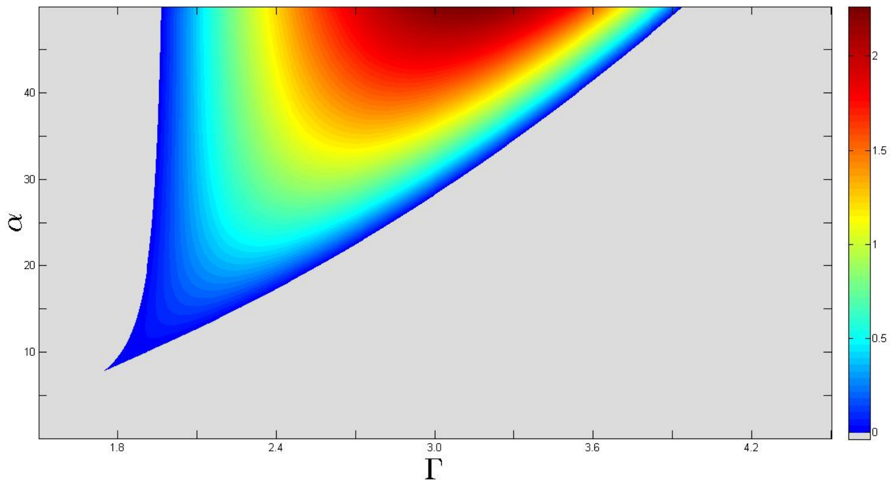


Figure 4.11. Positive values of the discriminant of $Z^3 - \Gamma Z^2 + Z - \frac{1}{\alpha}\Gamma = 0$ yields three real solutions. In the figure positive values of the discriminant are colored where as negative values are grey.

region in the $\Gamma$-$\alpha$ parameter space for which Equation 4.10 has three real solutions (all of which must be positive). The discriminant of Equation 4.10 is

$$\Delta = -\frac{4}{\alpha}\Gamma^4 + (-\frac{27}{\alpha^2} + \frac{18}{\alpha} + 1)\Gamma^2 - 4,$$

and the equation has three real solutions when the discriminant is greater than zero. Figure 4.11 plots in color the positive values of the discriminant over a region of the $\Gamma$-$\alpha$ parameter space; in the figure, all negative values of the discriminant are colored grey.

For $n = 2$, $\Gamma_{crit} = \sqrt{3} \approx 1.73$. Though a detailed discussion was previously omitted, we note that for values of $\Gamma$ near $\Gamma_{crit}$, the steady-state solution structure does depend strongly on the choice of $\alpha$. This occurs when the lower bound on $\alpha$ that delineates the bistable regime is greater than zero, so that changing $\alpha$ gives two saddle-point bifurcations (with increasing $\alpha$ the number of fixed points jumps from one to three and then back to one). This effect gives rise to the “tail” in Figure 4.11 that appears for values of $\Gamma$ between $\Gamma_{crit}$ and 2.

As part of the model design, we choose to avoid this complication, so we set $\Gamma = 2$. Solving Equation 4.7 for $n = \Gamma = 2$, we find $Z_- = \frac{1}{3}$ and $Z_+ = 1$. Plugging the $Z$-values into $p(Z) = Z^3 - 2Z^2$, we find that the points at which $p(Z)$ has slope equal to $-1$ are $(Z_-, p(Z_-)) = (\frac{1}{3}, -\frac{5}{27})$ and $(Z_+, p(Z_+)) = (1, -1)$.

Inserting the values of $Z_\pm$ and $p(Z_\pm)$ into the inequality in Equation 4.9 we find that, in order for Equation 4.10 to have two stable steady-states, $\alpha$ must satisfy

$$0 < \frac{1}{\alpha} < \frac{2}{27}.$$

As long as the choice of $\alpha$ does not lead to biologically absurd rate constants, any $\alpha$ within this range is suitable. But, since we have learned that the structure of the steady-state solutions will not depend strongly on $\alpha$ for our choice of $\Gamma = 2$ (dependence is similar to that shown in Figure 4.8), for convenience, we can try to find an $\alpha$ that yields steady-state values that have a relatively simple form. Appendix C discusses a short procedure that allows us to find steady-states with values that are tractable by, essentially, taking advantage of the freedom in $\alpha$.

Using that procedure, we find that choosing $\alpha = 16$ allows us to factor Equation 4.10

$$Z^3 - 2Z^2 + Z - \frac{1}{8} = (Z - \frac{1}{2})(Z^2 - \frac{3}{2}Z + \frac{1}{4}) = 0,$$

immediately giving us one of the steady-state values $Z_{ss} = \frac{1}{2}$. From the unfactored term, the quadratic formula gives the other two steady-state values as $Z_{ss} = \frac{3 \pm \sqrt{5}}{4} \approx 0.19; 1.31$.

We omit a detailed discussion of the stability of the fixed points, and we only remark (as can be seen in the deterministic solutions presented in Figure 4.12) that, as expected, $Z_{ss} = \frac{1}{2}$ is unstable while the two outer fixed points are stable.

### 4.6.2 Time evolution

For the choice of parameters, the differential equation defining the reaction system is

$$\dot{Z} = 2\frac{Z^2 + \frac{1}{16}}{Z^2 + 1} - Z. \tag{4.11}$$

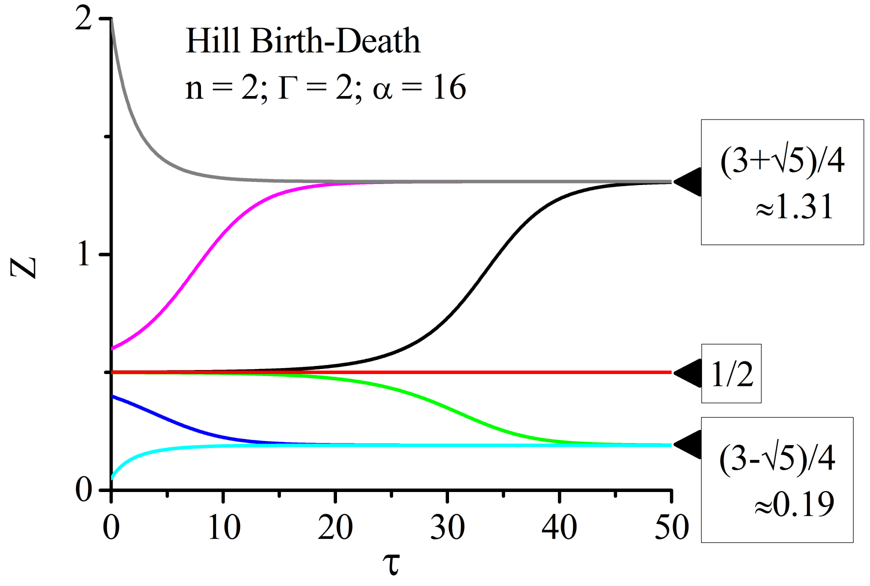


Figure 4.12. Deterministic solutions Hill Birth-Death reaction system. For all intial values except the value of the unstable fixed point, the deterministic solutions tends towards either of the outer, stable fixed points.

Figure 4.12 shows the deterministic solutions of Equation 4.11 found using the ode45 solver of MATLAB. For every initial value except that corresponding to the value of the unstable fixed point, the solutions tend towards the stable steady-states. If the initial value is set at the value of the unstable fixed point, the level of $Z$ never changes from that value (that is the definition of a fixed point). We are only able to demonstrate this in the figure because we found the precise value of the unstable fixed point (analytic forms for the roots of a cubic are usually unwieldy and must be approximated numerically).

Figure 4.13 shows trajectories of $Z$ from stochastic simulations of the system defined by Equation 4.11, where Figure 4.13(a) shows the time evolution over the first 5 time constants and Figure 4.13(b) shows the time evolution over the first 50 time constants (remember that a time constant is equal to $\frac{1}{c_2}$). In the figures all trajectories are initialized to a value ($Z(t = 0) = 0.1$) below the lowest fixed

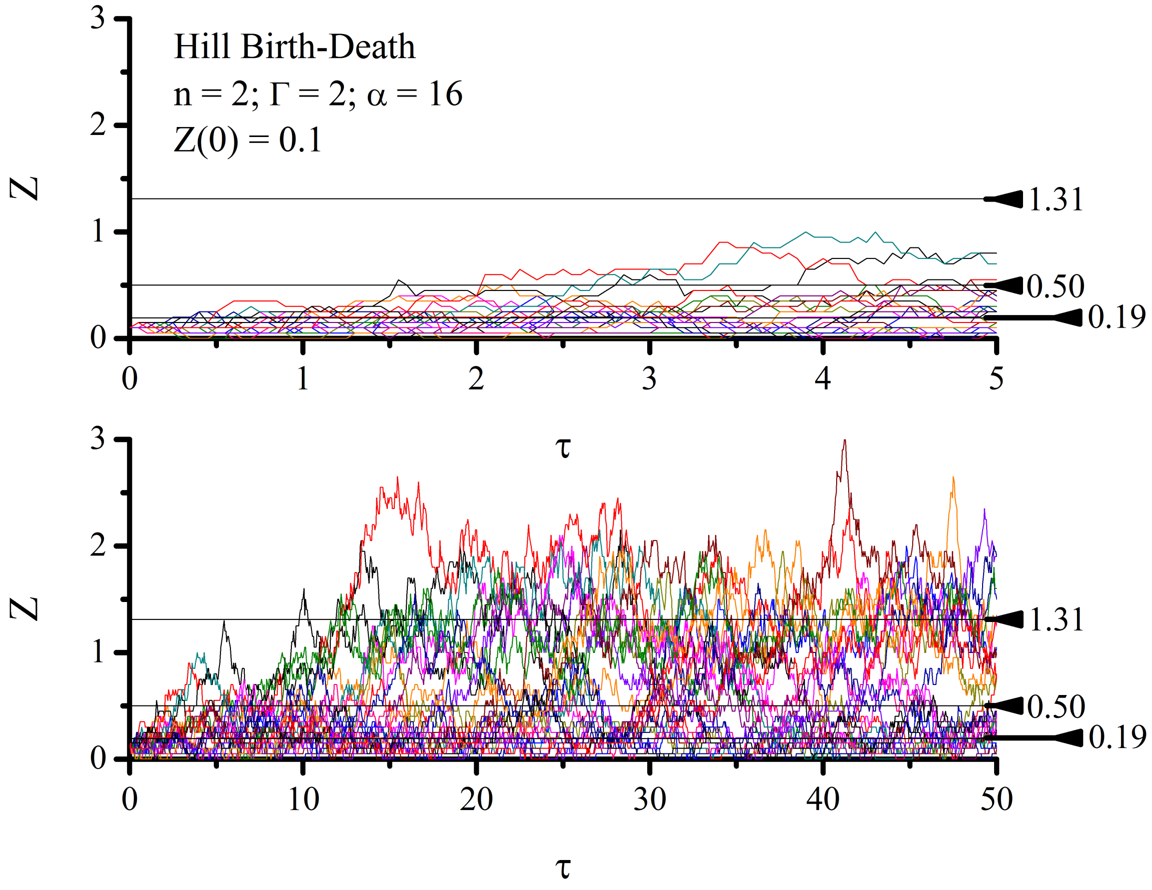


Figure 4.13. Stochastic solutions Hill Birth-Death reaction system.

point. After twenty time constants we see clear evidence that some trajectories have 'found' the higher fixed point. This reaction system demonstrates both bistability and stochastic switching.

### 4.6.3 An Example of a Realistic Set of Kinetic Parameters

Remembering the definitions $\Gamma = \frac{c_{10}}{c_2 k}$ and $\alpha = \frac{c_{10}}{c_{11}}$ we can construct a set of biologically feasible kinetic parameters that satisfies our choices $\Gamma$ and $\alpha$. We only have the freedom to choose two of the four kinetic parameters and, then, $\Gamma = 2$ and $\alpha = 16$ sets the other two.

The following considerations guide the choice of $c_2$ and $c_{10}$. The death rate, $c_2$, describing the loss of species from the system is connected to the dilution of protein due to cell division (proteins are randomly split among daughter cells), and not to the degradation of protein that occurs on a significantly slower time scale [23]. Setting $c_2 = 1/2000\ s^{-1}$ is consistent with the fact that an *E. coli* cell divides about once every thirty minutes under optimal growth conditions [9]. And, setting $c_{10} = 160/2000\ s^{-1}$ is consistent with the rate $0.0751\ s^{-1}$ used for active transcription of PhoB in the Kierzek model [23].

The choices $\Gamma = 2$, $\alpha = 16$, $c_2 = 1/2000\ s^{-1}$ and $c_{10} = 160/2000\ s^{-1}$ yield $c_{11} = 10/2000\ s^{-1}$ and $k = 80$. Using these parameters and the definition $Z \equiv X/K$ we can find the steady-state values of X; in ascending order $X_{ss} = 20(3-\sqrt{5}) \approx 15;\ 40;\ 5(3+\sqrt{5}) \approx 105$. Table 4.3 summarizes this realization of the Hill birth-death reaction system.

## 4.7 Elementary Inhibition Reaction System

In this section we investigate a reaction system that introduces simple control by an external signal into the Hill birth-death reaction system. The control is achieved by adding an isomer of the autoregulating species into the system in such a way that the isomer does not contribute to the birth reaction and the rate of isomerization

Table 4.3

Birth-death reaction system. The scheme illustrates the two reactions, the birth and death of species $X$, that occur in the system with the propensities for each reaction given in the second column of the table. The parameters in the third column have been selected so that $\Gamma = \frac{c_{10}}{c_2 k} = 2$ and $\alpha = \frac{c_{10}}{c_{11}} = 16$. With the listed parameters the reaction system has stable steady-states at $X = 20(3 \pm \sqrt{5})$.

| Scheme | Propensities | Parameters |
|---|---|---|
| $\varnothing$ | $a_1 = \frac{c_{10}X^n + c_{11}k^n}{X^n + k^n}$ | $n = 2$ |
| $_{2}\uparrow$ | | $c_{10} = 160/2000\ s^{-1}$ |
| $X$ | | $c_{11} = 10/2000\ s^{-1}$ |
| $_{1}\uparrow$ | | $k = 80$ |
| | $a_2 = c_2 X$ | $c_2 = 1/2000\ s^{-1}$ |
| $\varnothing$ | | |

is signal-dependent. In a biological sense, this corresponds to the environmentally-induced transition of the autoregulating protein to an inhibited state in which it cannot promote its own production.

We will consider the case in which the conversion frequency to the 'inhibited' isomer increases linearly with the signal. An increase in the signal increases the frequency of conversion from the free species to its inhibited isomer, effectively locking up the system as their are fewer free species to contribute to the birth propensity. A decrease in the signal releases the stock of inhibited isomer, abruptly converting many to the free species that can contribute to the birth propensity. We can schematically represent the reaction system as

$$\begin{tikzcd}
 & \varnothing & \\
X_1 \arrow[r, shift left, "c_3"] & X_2 \arrow[u, "c_2"'] \arrow[l, shift left, "c_4(t)"] & . \\
 & \varnothing \arrow[u, "c_1(X_2)"'] &
\end{tikzcd}$$

We introduce the external signal by step-modulating the parameter $c_4(t)$ in time. Our goal is to design the reaction system in such a way that the number of $X_2$ responds to the signal by ramping up and overshooting its final equilibrium value.

The time derivatives of the mean number of the two species are

$$\begin{aligned}
\dot{X}_1 &= -c_3 X_1 + c_4(t) X_2 \text{ ,and} \\
\dot{X}_2 &= \frac{c_{10} X_2^2 + c_{11} k^2}{X_2^2 + k^2} - c_2 X_2 + c_3 X_1 - c_4(t) X_2.
\end{aligned}$$

As we did previously with the Hill birth-death system, we can divide out the inherent scales in the problem and study a scale-invariant form of these equations. Rescaling time $c_2 t \to \tau$ and defining $Z_i = \frac{X_i}{k}$, $\Gamma = \frac{c_{10}}{c_2 k}$, $\alpha = \frac{c_{10}}{c_{11}}$, $\Delta = \frac{c_3}{c_2}$ and $\beta = \frac{c_4}{c_3}$ we find

$$\dot{Z}_1 = -\Delta Z_1 + \Delta \beta Z_2 \tag{4.12}$$

$$\dot{Z}_2 = \Gamma \frac{Z_2^2 + \frac{1}{\alpha}}{Z_2^2 + 1} - Z_2 + \Delta Z_1 - \Delta \beta Z_2. \tag{4.13}$$

### 4.7.1 Steady-states

Imposing the steady-state condition ($\dot{Z}_1 = \dot{Z}_2 = 0$) on Equations 4.12 we find

$$Z_{1,ss} = \beta Z_{2,ss} \tag{4.14}$$

$$Z_{2,ss} = \Gamma \frac{Z_{2,ss}^2 + \frac{1}{\alpha}}{Z_{2,ss}^2 + 1}. \tag{4.15}$$

Equation 4.15 is identical to the steady-state equation for the Hill birth-death reaction system that we solved previously – so we have all the tools necessary to understand $Z_{2,ss}$. And, $Z_{1,ss}$ is simply $\beta$ times $Z_{2,ss}$. Since $\beta$ is proportional to the signal, we see that the larger the magnitude of the signal the larger the amount of $Z_2$ in the system at steady-state.

Interestingly, the value of $\Delta$ does not affect the steady-state solutions of $Z_1$ or $Z_2$.

### 4.7.2 Parameter Choice

In order to computationally solve the time evolution of the system, we need to choose values for each of the parameters. We can set $n = 2$, $\Gamma = 2$ and $\alpha = 16$ as we did previously. The parameter $\Delta$ controls the speed of the isomerization dynamics relative to the death dynamics. In order to get a ramp up, the isomerization must occur on faster time scale then death. Also, from a biological perspective, it is reasonable to assume that species death is the slowest timescale in the system. For these reasons we demand that $\Delta = \frac{c_3}{c_2} > 1$. The time evolution of the reaction system is essentially unchanged for a large range of $\Delta > 1$, so, for simplicity, we let $\Delta = 10$.

Step-modulation of $\beta$ simulates a change in the strength of the external signal. The parameter also sets the ratio $\frac{Z_{1,ss}}{Z_{2,ss}}$ and the rate of conversion of $Z_2$ into $Z_1$. We design the initial state of the system so that $Z_2$ is at the lower fixed point and a large amount of $Z_1$ is stored in the system (large $\beta$). When $\beta$ is suddenly reduced, the stored $Z_1$ rapidly converts to $Z_2$ in such a way that $Z_2$ overshoots its higher steady-state value before settling back to it. We found that an initial value of $\beta = 50$ produces an overshoot for the range of step-modulations to $\beta =$0.05, 0.5, 1 or 2.

### 4.7.3 Time Evolution

For the choice of parameters the differential equations defining the reaction system are

$$\dot{Z}_1 = -10Z_1 + 10\beta Z_2 \tag{4.16}$$

$$\dot{Z}_2 = 2\frac{Z_2^2 + \frac{1}{16}}{Z_2^2 + 1} - Z_2 + 10Z_1 - 10\beta Z_2, \tag{4.17}$$

where $\beta$ is step-modulated from 50 to 0.05, 0.5, 1 or 2.

Figure 4.14 shows deterministic solutions of Equation 4.16 found using the ode45 solver of MATLAB. $Z_2$ is initialized to the value of the smaller stable fixed point. At $\tau = 1$, $\beta$ is step-modulated to the various final values, and each curve shows the response to one of the step-modulations. As designed, the level of $Z_2$ overshoots (to varying degrees dependent on the strength of the signal modulation) before settling to the higher of the two fixed points.

Figure 4.14 shows trajectories of $Z_2$ from stochastic simulations of the system defined by Equation 4.16. In each of the four experiments shown in the figure, $Z_2$ is initially near the lowest fixed point and $\beta = 50$ and then step-modulated to the value indicated.

Interestingly, in this case, the stochastic simulations are in good agreement with the deterministic solutions. In the current study, we were not able to design a system that shows both stochastic switching and overshoot before settling. This is currently an active area of research.

### 4.7.4 An Example of a Realistic Set of Kinetic Parameters

If we use the same parameter choices previously discussed for the Hill birth-death reaction system ($c_2 = 1/2000\ s^{-1}$ and $c_{10} = 160/2000\ s^{-1}$) and the choices of $\Gamma$, $\alpha$, $\Delta$ and $\beta$ described above, we find a complete kinetic parameter description of a realization of the inhibition reaction system: $c_{10} = 160/2000\ s^{-1}$, $c_{11} = 10/2000\ s^{-1}$,

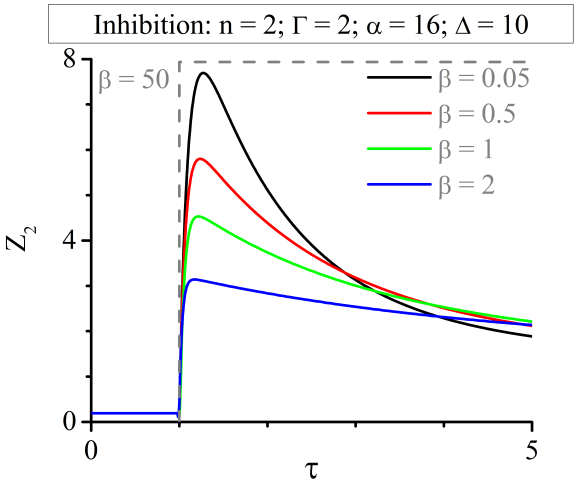


Figure 4.14. Deterministic solutions to the Elementary Inhibition reaction system for different strength signals. The value of the autoregulating species $Z_2$ is initialized to the value of the smaller stable fixed point. At $\tau = 1$ the model parameter $\beta$ is step-modulated by the amount indicated in the plot in order to simulate the arrival of an external signal. Each curve shows the solution for a different magnitude modulation. The signal strength is sufficient, in all cases, to cause $Z_2$ to transition to the higher fixed point with varying degrees of overshoot.

$k = 80$, $c_2 = 1/2000\ s^{-1}$, $c_3 = 10/2000\ s^{-1}$ and $c_4$ is modulated from $500/2000\ s^{-1}$ $(\beta = 50)$ to $0.5/2000\ s^{-1}$ $(\beta = 0.05)$.

The steady-states of $X_2 \equiv KZ_2$ in ascending order are $X_{ss} = 20(3 - \sqrt{5}) \approx 15;\ 40;\ 20(3 + \sqrt{5}) \approx 105$. Table 4.4 summarizes this realization of the elementary inhibition reaction system.

## 4.8 Elementary Activation Reaction System

In this section we present an alternative method to introduce control by an external signal into the birth-death reaction system. We do not subject this reaction system to a detailed analysis, but only highlight a remarkable characteristic the system possesses.

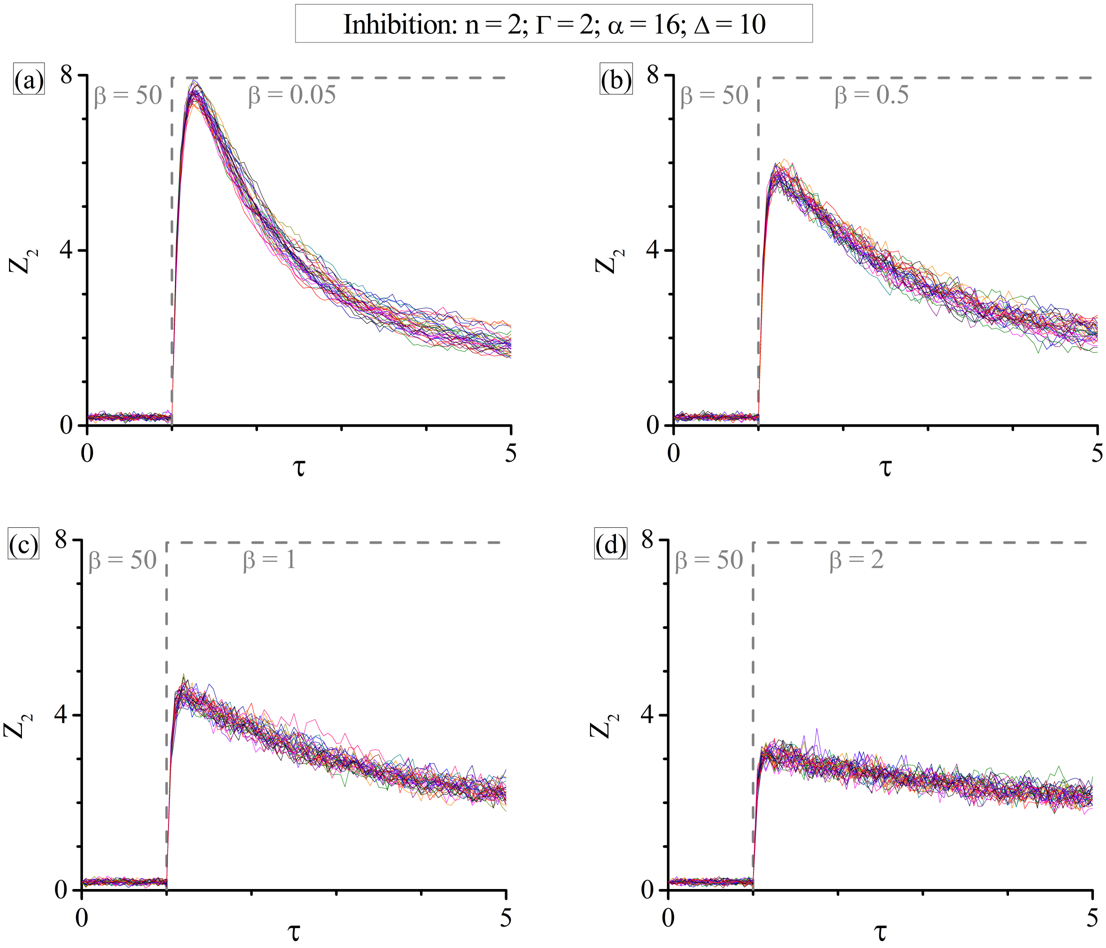
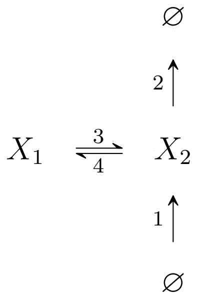


Figure 4.15. Stochastic solutions to the Elementary Inhibition reaction system for different strength signals. The value of the autoregulating species $X_2$ is initialized to the value of the smaller stable fixed point. At $\tau = 1$ the model parameter $\beta$ is step-modulated by the amount indicated in each plot in order to simulate the arrival of an external signal. Each plot shows the solution for a different magnitude modulation. The signal strength is sufficient, in all cases, to cause $X_2$ to transition to the higher fixed point with varying degrees of overshoot.

Table 4.4

Elementary inhibition reaction system. The scheme illustrates the four reactions that occur among the two species in the system. The second column of the table lists the propensities for the reactions. Species $X_2$ autoregulates its birth via a Hill function. $X_2$ converts to the inhibited isomer $X_1$ with a frequency that depends linearly on the external signal. Step modulation of the kinetic parameter $c_4$ acts as a proxy for the external signal. The parameters in the third column have been selected so that $\Gamma = \frac{c_{10}}{c_2 k} = 2$, $\alpha = \frac{c_{10}}{c_{11}} = 16$, $\Delta = \frac{c_3}{c_2} = 10$ and $\beta = \frac{c_4}{c_3}$ varies in step fashion from 50 to 0.05.

| Scheme | Propensities | Parameters |
|---|---|---|
|  | $a_1 = \frac{c_{10} X_2^n + c_{11} k^n}{X_2^n + k^n}$ | $n = 2$ |
| | | $c_{10} = 160/2000\ s^{-1}$ |
| | | $c_{11} = 10/2000\ s^{-1}$ |
| | | $k = 80$ |
| | $a_2 = c_2 X_2$ | $c_2 = 1/2000\ s^{-1}$ |
| | $a_3 = c_3 X_1$ | $c_3 = 10/2000\ s^{-1}$ |
| | $a_4 = c_4(t) X_2$ | |
| | $c_4(t) = c_{4h} -$ | $c_{4h} = 500/2000\ s^{-1}$ |
| | $(c_{4h} - c_{4l})\theta(t - t_0)$ | $c_{4l} = 0.5/2000\ s^{-1}$ |

Another way to introduce control by an external signal into the birth-death reaction system is to necessitate that the autoregulating species isomerizes before it can cause its own birth, and that the isomerization depends on the strength external signal. In a biological sense, this corresponds to the environmentally-induced activation of a protein into a self-regulating transcription factor.

We will consider the case in which the conversion frequency to the 'inactive' isomer increases linearly with the signal strength. An increase in the signal increases the frequency of conversion from the active to inactive isomer, effectively locking up the system as there are fewer active isomers to contribute to the birth propensity. A decrease in the signal abruptly converts the stock of inactive isomers to the active species that can contribute to the birth propensity. We schematically represent the activation reaction system as

$$\begin{tikzcd}
\varnothing & \\
X_1 \arrow[u, "c_2"] \arrow[r, shift left, "c_3"] & X_2 \arrow[l, shift left, "c_4(t)"] \\
\varnothing \arrow[u, "c_1(X_2)"] &
\end{tikzcd}$$

where modulation of parameter $c_4(t)$ acts as a proxy for the external signal.

Reducing the model in the same way we did previously, we find the differential equations for the reduced species $Z_i = \frac{X_i}{k}$ are

$$\dot{Z}_1 = \Gamma \frac{Z_2^2 + \frac{1}{\alpha}}{Z_2^2 + 1} - Z_1 - \Delta Z_1 - \Delta\beta Z_2 \tag{4.18}$$

$$\dot{Z}_2 = \Delta Z_1 - \Delta\beta Z_2, \tag{4.19}$$

where $\Gamma = \frac{c_{10}}{c_2 k}$, $\alpha = \frac{c_{10}}{c_{11}}$, $\Delta = \frac{c_3}{c_2}$ and $\beta = \frac{c_4}{c_3}$.

### 4.8.1 Steady-states

Imposing the steady-state condition ($\dot{Z}_1 = \dot{Z}_2 = 0$) on Equations 4.18, we find

$$Z_{1,ss} = \beta Z_{2,ss} \tag{4.20}$$

$$Z_{2,ss} = \frac{\Gamma}{\beta} \frac{Z_{2,ss}^2 + \frac{1}{\alpha}}{Z_{2,ss}^2 + 1}. \tag{4.21}$$

Other than the factor of $\beta$, Equation 4.21 is identical to the steady-state equation for the Hill birth-death system we solved previously. The parameter $\beta$ acts as the proxy for the external signal. In addition to setting the steady-state ratio $\frac{Z_{1,ss}}{Z_{2,ss}}$, $\beta$, by its placement in Equation 4.21, increasing $\beta$ is identical to decreasing $\Gamma$. This implies that changing $\beta$ will affect the steady-state solution structure.

So, remarkably, for this reaction system the external signal controls the steady-state solution structure. The implications of this property is an area of active research.

## 4.9 Summary and Concluding Remarks

Two-component systems are a large class of signal-response mechanisms (transducers) prevalent in prokaryotes. In this study we developed a method using single-cell fluorescence microscopy techniques to measure, *in vivo*, the transience in the signal response of a two-component system. We applied that technique to measuring the characteristics of the response of the environmental phosphate transducer in *Escherichia coli* (the PhoR/PhoB TCS) to phosphate limitation. Then we connected features measured in the PhoR/PhoB TCS response to simple extensions of the birth-death reaction system.

Significantly, our measurement of the PhoR/PhoB TCS transient signal-response demonstrates that, at the population level, the transducer ramps up and overshoots before settling to its final steady-state. We also found that our measurements were consistent with the transducer having a threshold sensitivity and, at the single-cell level, a bistable response that stochastically switches between fixed points.

Instead of building a model of the PhoR/PhoB TCS from first principles like that of Kierzek *et al.*, we sought a more rigorous understanding of pervasive model components, like the autoregulating species. We studied simple extensions of the birth-death reaction system that qualitatively reproduce the PhoR/PhoB TCS phenomenology found in the measurements. The simple models have a small number of parameters and, thus, allow a complete characterization of the model behavior in the parameter space.

Specifically, we developed a mathematical method to find the parameter-regime in which the Hill birth-death reaction system is bistable. Then we showed how to choose parameters within this regime that allows the system to stochastically switch between the stable states. This demonstrates how the stochasticity observed in many biological systems, including the PhoR/PhoB TCS, can arise simply from an autoregulating transcription factor.

Then we demonstrated that a simple extension of the Hill birth-death system gives a model that, in response to an external signal, ramps up and overshoots before settling to its final steady-state. We achieved this by adding to the Hill birth-death system a signal-dependent isomerization of the autoregulating species to an inhibited state. Again, this can be connected to the inhibition state of the PhoR/PhoB TCS described in Chapter 1.

Though we were not able to design a system that showed both stochastic switching and overshoot, we have identified the basic requirements a model must possess in order to have either. Further research can use the tools we developed here to construct a reaction system that combines these features in a simple model.

# APPENDICES

# A. MICROFLUIDIC DEVICES

As mentioned at the beginning of Chapter 2, the initial experiments conducted in a glass dish suffer from several weakness related to precisely controlling the chemical environment and the cell population in the relatively large ~1 mL volume of the dish. First, in the dish, Pi is controlled by manually exchanging the chemical environment surrounding the cells. Even though the environment can be changed in just a few minutes, there is no way to tell what parameters the cells "see" during the exchange. Second, though using Poly-L-Lysine is a common laboratory technique, the cell is deformed and flattened [22] as it adheres. It is not clear to what degree this affects the physiology of the cell. As the goal is to measure a physiological response, this is not an optimal condition. Third, during the course of the experiment, and especially when the cell density is high, the cells can detach from the glass, move through the volume of the dish, and land at some other spot. And, fourth, the cells have a tendency, especially when the cell density is high, to grow on top of each other. In such a case, it is impossible to reliably measure the emission intensity from a single cell. The last two effects are very difficult to account for and reduce the number of cells in the field of view that can be analyzed.

Each of these weaknesses can be addressed by using a microfluidic device instead of a glass dish. Microfluidic describes the manipulation or control of a fluid at the sub-millimeter length scale, just one or two orders of magnitude greater than the size of *E. coli* (roughly the size of a cylinder with diameter 1 micron and length 2 microns). This scale is well suited for isolating single cells for measurement. Figure A.1 shows a few *E. coli* in a microfluidic chamber. In addition to giving control of the cell population, microfluidics give precise and sophisticated control of the chemical environment surrounding the cells.

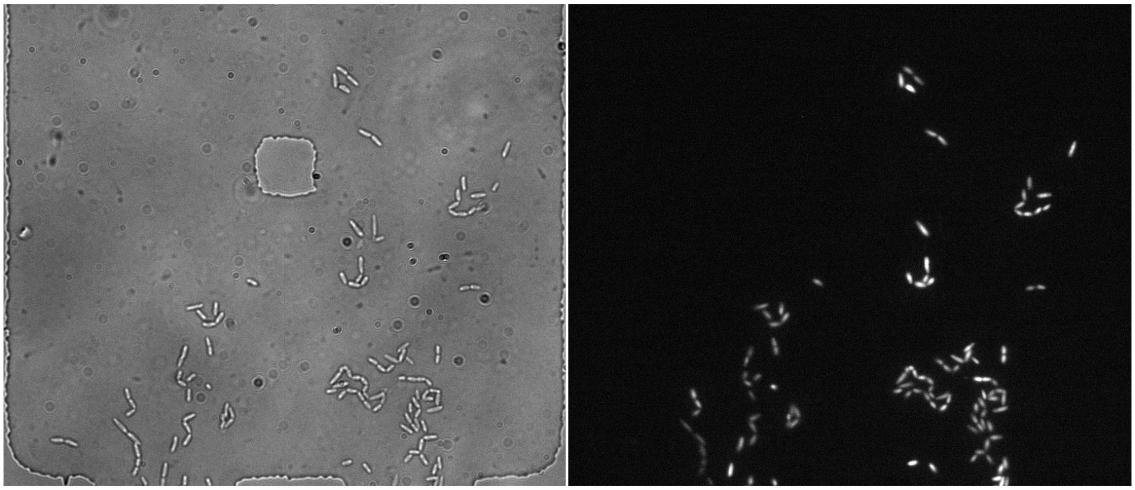


Figure A.1. Microfluidic are well suited for isolating single cells for imaging. The figure on the left shows a brightfield image at 40x magification of several *E. coli* in a microfluidic chamber. The figure on the right shows the single-cell fluorescence intensity of the same cells.

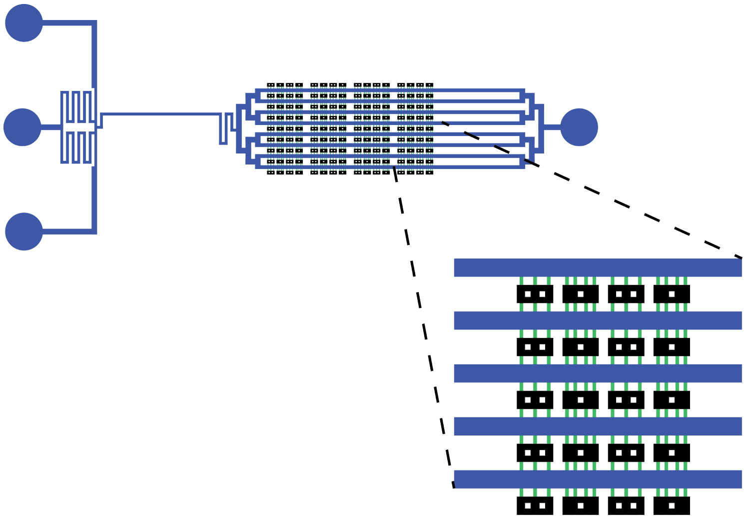


Figure A.2. 10:1 scale representation of microfluidic device. Courtesy of the Groisman Lab (UCSD).

Testing of microfluidic devices (provided by the Alex Groisman Lab, UCSD) with fluorescent beads and with *E. coli* has been conducted. Migration of the experimentation to this platform is one of the more difficult experimental goals of this proposal, but also one of the more exciting. As such, the following sections will provide an introduction to the use of microfluidic devices.

## A.1 Device Design

The microfluidic devices that we will use in this study are made from a patterning of three basic features: channels, chambers, and capillaries. Figure A.2 shows a 10:1 scale representation of a microfluidic device.

Channels carry the large-volume (microliter) flow throughout the device, and the chambers provide a small volume (femtoliters) outside the large-volume flow within which to isolate and observe single cells. The capillaries are the smallest of the three features and serve to couple the chambers to the channels. Table A.1 shows approximate sizes of each feature.

Table A.1
Approximate feature dimensions of the microfluidic devices.

| Feature | Length (micron) | Width (micron) | Height (micron) |
|---|---|---|---|
| Channel | $10^5$ | 100 | 15 |
| Chamber | 100-200 | 100 | 1.5 |
| Capillary | 25-50 | 20 | 0.6-1.5 |

## A.2 Device Construction

Currently, there are competing techniques for microfluidic device fabrication available to researchers that are well documented in the literature. For example, [33] provides a comprehensive review of materials and methods for device construction. The poly(dimethyl-siloxane) (PDMS) microfluidic devices used in this study are fabricated using soft lithography [34] to pattern the PDMS elastomer. The most difficult part of this procedure is creating a mold on which to cast the PDMS. The casting mold, upon which the PDMS prepolymer is prepared and cured, contains the negative relief structure that defines the feature geometry of the microfluidic device.

The result of the soft lithography process is a PDMS block containing the feature geometry (channels, chambers, and capillaries) open on one side. The geometry is closed by reversibly adhering the open side to a glass substrate. Prior to PDMS-glass bonding, the PDMS and glass are cleaned and prepared to promote strong adhesion. After use, the glass is separated from the PDMS, reopening the feature geometry. The PDMS block can be cleaned and rebonded to substrate for reuse.

Since device fabrication techniques are only indirectly relevant to this study and are extensively reviewed elsewhere, a more detailed discussion is omitted.

## A.3 Fluid Flow

As cells consume food and nutrients and expel waste, the chemical environment of the small-volume chamber degrades. To maintain precise control of the environmental parameters, the volume within the chamber should be exchanged on a timescale much shorter than the cell's life cycle. The following argument will demonstrate that continously flowing fresh media through the large-volume channels at a speed of order 100 micron/s is sufficient to achieve the desired volume exchange.

The Reynolds scale of the microfluidic device dictates that the fluid flow is well described by streamlines (non-crossing, steady-state particle trajectories). To find the Reynolds number we take as the characteristic length scale, $L$, the hydraulic diameter of the channel, of order 10 microns, and approximate the media density, $\rho$, and viscosity, $\mu$, as that of water. Using the channel flow speed, $V$, of order 100 micron/s we find

$$Re = \frac{\mathtt{F}_{inertial}}{\mathtt{F}_{viscous}} = \frac{\rho\, V\, L}{\mu} = \frac{10^3 \frac{\mathtt{kg}}{\mathtt{m}^3}\, 10^{-4} \frac{\mathtt{m}}{\mathtt{s}}\, 10^{-5} \mathtt{m}}{10^{-3} \frac{\mathtt{Ns}}{\mathtt{m}^2}} << 1,$$

much less than the value of 2040 [35] that marks the maximum of laminar flow and the onset of sustained turbulence in pipe flow. In this regime, energy dissipation by the relatively large viscous forces, $\mathtt{F}_{viscous}$, damps turbulent flow structure, giving creeping flow. This implies that off-streamline mass transfer can only occur by diffusion.

To put an upper limit on the timescale for the chamber volume exchange, we can ignore the small effect of active flow into the chamber and consider only the diffusive mass tranfer from the channel. For now, let us assume that the effect of the large-volume channel flow is to maintain a constant concentration of chemical species at the channel-capillary interface. Solving for this setup, the characteristic timescale for molecular diffusion into the chamber for small molecules with diffusion coefficient, $D$, of order $10^{-6}$ cm$^2$ /s can be estimated as [36]

$$\tau = \frac{w_{\mathtt{chamber}}\, h_{\mathtt{chamber}}\, l_{\mathtt{capillary}}}{D\, h_{\mathtt{capillary}}} = 125 \quad \mathtt{s},$$

where $w$, $h$, and $l$ are the width, height, and length of the features from Table A.1 chosen to maximize the calculation. Since the doubling time for a healthy population of E. coli is approximately 20-30 minutes [9], diffusive mass transfer is sufficient to cycle the chamber volume in a time much less than the cell's life cycle.

Now, the diffusive mass transfer described above will, in fact, deplete the concentration of chemical species at the channel-capillary interface. The large-volume flow within the channel must be fast enough to replenish this depletion. The volumetric flow through the channel is equal to the cross sectional area of the channel times the flow speed, $\nu$. This flow must replace the entire volume of the chamber taken from the channel in time $\tau$ found above:

$$\nu = \frac{V_{\mathtt{chamber}}}{\tau\, \sigma_{\mathtt{channel}}} << \;\mathtt{1\ micron/s}.$$

So 100 micron/s flow in the channel is more than sufficient for precise control of the chemical environment in the chamber. We can reasonably assume that the cells experience an environment in the chamber that is identical to that of channel.

## A.4 Micro/Macro Interface and Device Use

The microfluidic devices we will use have four millimeter-size wells (see Figure A.2) that interface with standard tubing and allow hydraulic access to the sub-millimeter feature geometry. Three of these wells, two inlets and a vent, are at the front-end of the device and the fourth, the waste, is at the back-end of the device. At the front-end, fluid flows in through the inlets and out through the vent (inbetween inlets). Part of the fluid from the inlets flows through a fluid manipulation construct (discussed next), through the chambers, and exits the back-end of the device through the waste.

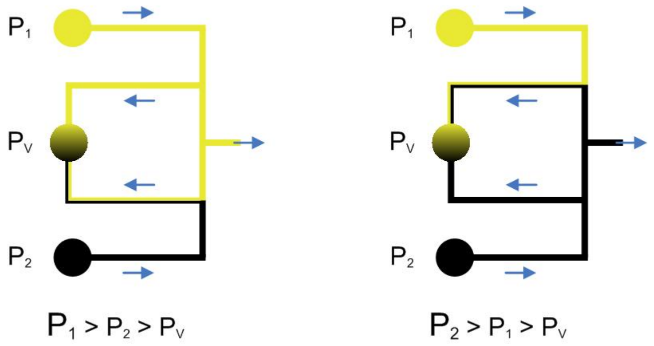
Tubing connects the four wells to 2 mL containers that provide fluid to the inlets and accepts fluid from the waste and vent. When a continous fluid fills the device, the tubing, and the containers, a pressure difference applied between any two containers will cause fluid to flow in the microfluidic device. The pressure difference necessary to obtain the small volumetric flow rate required to sustain the chambers is on the order of 10 inches of water (~0.4 psi), and is achieved by placing the containers at different heights.

## A.5 Fluid Manipulation

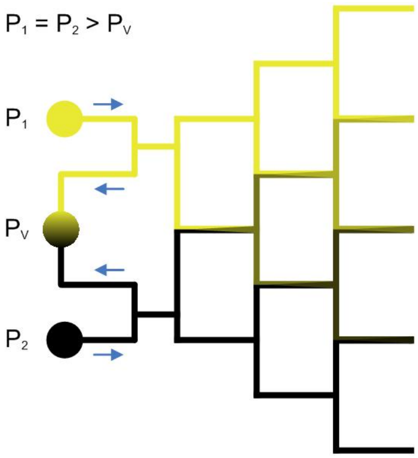
The microfluidic devices we will use implement two basic constructs to provide manipulation of the fluid from the inlets. These constructs are located downstream of the inlets and upstream of the growth chambers. Described in more detail below, the two constructs are the Fluid Switch, providing rapid and predictable fluid exchange to the growth chambers, and the Multiplex Fluid Mixer, mixing fluids from the two inlets in various exact proportions before the fluid reaches the growth chamber.

**Fluid Switch** The media switch gives the capability of very rapidly exchanging the fluid supplied to the channels that feed the chambers. Any fluid exchange must occur onboard the device, as the small volumetric flow rate through the device will not support the exchange of the ~mL volume wells. Figure A.3(a) shows a schematic of the switch in two pressure configurations. The direction of the fluid flow is shown with arrows. In the left figure, the pressure applied to the container that feeds inlet 1, $P_1$, is greater than $P_2$, the pressure applied to the container that feeds inlet 2. Both these of these pressures are larger then that applied to the container that collects from the vent, $P_V$. In this configuration, fluid from inlet 1 travels to the vent as well as downstream to the chambers and to the waste (chambers and waste not shown). Fluid from inlet 2 only travels to the vent. In the alternate pressure configuration, $P_2 > P_1 > P_V$, fluid from inlet 2 flows towards the chambers. This is shown in the right figure. Fluid exchange using the switch has been observed to occur in the chambers on the order of minutes (data not shown).

**Multiplex Fluid Mixer** The multiplex fluid mixer gives the capability of flowing a different mixture of fluid from inlets 1 and 2 down each of the channels. Figure A.3(b) shows a schematic of the mixer. Fluid from both inlets flows towards the first stage of the device. Because the pressure applied to the inlets is the same, the fluids meet in the middle of this stage and travel side-by-side down a large channel to the next stage. Before reaching the next stage, the fluids from the inlets diffusively mix in the middle channel, so that three different fluids arrive at the third stage (fluid 1, a 1:1 mixture of fluids 1 and 2, and fluid 2). A similar process occurs at the second stage. Fluid 1 is mixed in 1:1 proportion with the 1:1 mixture of fluid 1 and 2. Fluid 2 is also mixed in 1:1 proportion with the 1:1 mixture. This yields four fluids: fluid 1, fluid 1/fluid 2 mixture at 3:1, fluid 1/fluid 2 mixture at 1:3, and fluid 2. The device is designed to have eight stages; the end result is ten channels each with a fluid mixture in a different proportion. This device gives a good method for determining



(a) Fluid switch.



(b) Multiplex fluid mixer.

Figure A.3. Onboard fluid manipulation integrated into microfluidic devices. Arrows show the direction of fluid flow. (a) In the fluid switch device, changing the pressure applied to inlets 1 and 2 rapidly switches which inlet provides fluid to the chambers. The vent is kept at a pressure below both inlets. (b) In the multiplex fluid mixer, fluid from inlet 1 and 2 are mixed 1:1 at a branching point in the first stage. This gives three different fluids (the fluid from inlet 1 and from inlet 2 plus the 1:1 mixture). Two combinations of these fluids are mixed at branching points in the second stage, giving four different fluids. This procedure is iterated eight times (all not shown) yielding ten combinations of the fluids from inlet 1 and 2.

steady-state response properties as a function of concentration, like the Pi threshold of PhoR/PhoB TCS response.

# B. STOCHASTIC SIMULATION METHOD AND ALGORITHM

Daniel Gillespie rigorously demonstrated [26] [27] [28] how a single trajectory of a reaction system can be described as a continuous-time Markov process in the state space of the species numbers. From this demonstration, he described a method to simulate a trajectory.

During each step in the simulation a reaction, $\mu$, occurs during the time interval, $\tau$, where $\mu$ and $\tau$ are randomly selected from the probability density $P(\tau, \mu)$. Gillespie found the correct form of $P(\tau, \mu)$ in the following way.

Given state $(X_1, ..., X_N)$ at time $t$ we should calculate the probability, $P(\tau, \mu)d\tau$, that reaction $\mu$ occurs after time interval $\tau$.

$P(\tau, \mu)d\tau$ is equivalent to the probability, $P_0(\tau)$, that no reaction occurs in interval $\tau$ and that $\mu$ occurs immediately following interval $\tau$ and within the interval $(t+\tau, t+\tau+d\tau)$. If the probability that reaction $\mu$ occurs in an infinitesmal interval is related to the propensity of the reaction $a_\mu d\tau$, then:

$$P(\tau, \mu)d\tau = P_0(\tau)a_\mu d\tau.$$

To find $P_0(\tau)$, consider that the probability that no event occurs in an infinitesmal interval is $[1-\sum_{\nu=1}^{\infty} a_\nu d\tau]$ so the probability that no event occurs in interval $(t+\tau+d\tau)$ is

$$P_0(\tau + d\tau) = P_0(\tau)[1 - \sum_{\nu=1}^{\infty} a_\nu d\tau].$$

This means that

$$\frac{dP_0(\tau)}{P_0(\tau)} = -\sum_{\nu=1}^{\infty} a_\nu d\tau,$$

and

$$P_0(\tau) = \exp{[-\sum_{\nu=1}^{\infty} a_\nu \tau]}.$$

Defining $a_0 = \sum_{\nu=1}^{\infty} a_\nu$ gives

$$P(\tau, \mu) = a_\mu \exp\left(-a_0 \tau\right) \tag{B.1}$$

Gillespie used this formulation to propose a computational method, the Stochastic Simulation Algorithm [27] (SSA), essentially a type of kinetic Monte Carlo, to rigorously simulate chemical reactions regardless of the numbers of species. The SSA is an iterative procedure that takes at time $t$ a state $(X_1, ..., X_N)$ and steps it forward in time by $\tau$ by choosing the next reaction $\mu$, where $\tau$ and $\mu$ are randomly picked according to the probability density $P(\tau, \mu)$. This can be achieved [26] by picking two uniform random numbers, $r_1$ and $r_2$ and calculating

$$\tau = \frac{1}{a_0} \ln \frac{1}{r_1},$$

and finding the integer $\mu$ such that

$$\sum_{\nu=1}^{\mu-1} a_\nu < r_2 a_0 \leq \sum_{\nu=1}^{\mu} a_\nu.$$

The chosen reaction, $\mu$, is used to update the state of the system by subtracting the reactants used up and adding the products created to species populations. It is straight forward to write a computer program to implement the SSA.

# C. SOLUTIONS TO THE STEADY-STATE CUBIC POLYNOMIAL

We want to find solutions to the $n = 2$ steady-state polynomial

$$Z^3 - \Gamma Z^2 + Z - \frac{1}{\alpha}\Gamma = 0. \tag{C.1}$$

The general method to compute the roots of a cubic polynomial

$$ax^3 + bx^2 + cx + d = 0$$

can be found in many math textbooks. A brief summary of this method is as follows. First we compute the following quantities:

$$\begin{aligned}
\Delta_0 &= b^2 - 3ac \\
\Delta_1 &= 2b^3 - 9abc + 27a^2d \\
C &= \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}}.
\end{aligned}$$

Then, the $k^{th}$ of the three roots to the cubic polynomial is

$$x_k = \frac{-1}{3a}(b + u_k C + \frac{\Delta_0}{u_k C}), \tag{C.2}$$

where the $u_k$ are the three cubic roots of unity:

$$u_1 = 1\ ,\ u_2 = \frac{-1 + i\sqrt{3}}{2}\ ,\ u_3 = \frac{-1 - i\sqrt{3}}{2}.$$

This method yields the three roots of the cubic in an analytic form that is rather cumbersome. For certain choices of $\Gamma$ and $\alpha$ Equation C.1 has three real, positive roots, and the imaginary terms must cancel. But it is non-trivial to reduce the solutions in Equation C.2 to a form that is easy to use in calculations.

Instead of solving the cubic in this way, we can use a trick that takes advantage of the fact that $\Gamma$ and $\alpha$ are essentially free parameters (constrained only in that the underlying kinetic parameters that contribute to $\Gamma$ and $\alpha$ should be real, positive and not biologically absurd). First, we arbitrarily factor Equation C.1 using real, positive numbers $a$, $b$ and $c$:

$$\begin{aligned} Z^3 - \Gamma Z^2 + Z - \frac{1}{\alpha}\Gamma &= (Z-a)(Z^2 - bZ + c) \\ &= Z^3 - (a+b)Z^2 + (ab+c)Z - ac = 0. \end{aligned}$$

Equating the constants common to the various powers of $Z$ we can find the relationships

$$\begin{aligned} b &= \Gamma - a \\ c &= a^2 - a\Gamma + 1 \\ \frac{1}{\alpha} &= \frac{a}{\Gamma}(a^2 - a\Gamma + 1). \end{aligned}$$

In this construction the freedom in $\alpha$ has been swapped for freedom in $a$. We can choose any $\Gamma$ and $a$ subject to the constraint that $b$ and $c$ must be positive. This constraint is met as long as $a < \Gamma < a + \frac{1}{a}$.

Choosing $\Gamma$ and $a$ sets the values of $b$, $c$ and $\alpha$ and the three solutions of the cubic polynomial follow trivially as

$$Z = a \ ; \ \frac{b \pm \sqrt{b^2 - 4c}}{2}.$$

We can impose the further constraint that $b^2 - 4c > 0$ so that the three solutions will be real.

# LIST OF REFERENCES

# LIST OF REFERENCES

[1] James A. Hoch. Two-component and phosphorelay signal transduction. *Current Opinion in Microbiology*, 3(2):165–170, 2000.

[2] Takeshi Mizuno. Compilation of all genes encoding two-component phosphotransfer signal transducers in the genome of escherichia coli. *DNA Research*, 4(2):161–168, 1997.

[3] Martin G. Lemarche, Barry L. Wanner, Sebastien Crepin, and Josee Harel. The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis. *FEMS Microbiology Reviews*, 32(3):461–473, 2008.

[4] Lu Zhou, Gerald Gregori, Jennifer M. Blackman, J. Paul Robinson, and Barry L. Wanner. Stochastic activation of the response regulator phob by noncognate histidine kinases. *Journal of Integrative Bioinformatics*, 2, 2005.

[5] Byron F. Brehm-Stecher and Eric A. Johnson. Single-cell microbiology: Tools, technologies, and applications. *Microbiology and Molecular Biology Reviews*, 68(3):538–559, 2004.

[6] Bree B. Aldridge, John M. Burke, Douglas A. Lauffenburger, and Peter K. Sorger. Physicochemical modelling of cel signalling pathways. *Nature Cell Biology*, 8:1195–1203, 2005.

[7] Ann M. Stock, Victoria L. Robinson, and Paul N. Goudreau. Two-component signal transduction. *Annual Review of Biochemistry*, 69(1):183–215, 2000.

[8] Ruth A. VanBogelen, Eric R. Olson, Barry L. Wanner, and Frederick C. Neidhardt. Global analysis of proteins synthesized during phosphorus restriction in escherichia coli. *Journal of Bacteriology*, 178(15):4344–4366, 1996.

[9] John L. Ingraham, Ole Maaloe, and Frederick C. Neidhardt. *Growth of the Bacterial Cell*. Sinauer Associates Inc., Sunderland, Massachusetts, first edition, 1983.

[10] Yi-Ju Hsieh and Barry L. Wanner. Global regulation by the seven-component $P_i$ signaling system. *Current Opinion in Microbiology*, 13(2):198–203, 2010.

[11] Jonathan M. Raser and Erin K. O'Shea. Noise in gene expression: Origins, consequences, and control. *Science*, 309:2010–2013, 2005.

[12] Harley H. McAdams and Adam Arkin. Stochastic mechanisms in gene expression. *Proceedings of the National Academy of Sciences*, 94:814–819, 1997.

[13] Peter S. Swain, Michael B. Elowitz, and Eric D. Siggia. Intrinsic and extrinsic contributions to stochasticity in gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20):12795–12800, 2002.

[14] Mads Kaern, Timothy C. Elston, William J. Blake, and James J. Collins. Stochasticity in gene expression: from theories to phenotypes. *Nature Reviews Genetics*, 6:451–464, 2005.

[15] Nathalie Q. Balaban, Jack Merrin, Remy Chait, Lukasz Kowalik, and Stanislas Leibler. Bacterial persistence as a phenotype switch. *Science*, 305:1622–1625, 2004.

[16] Martin Chalfie, Yuan Tu, Ghia Euskirchen, William W. Ward, and Douglas C. Prasher. Green fluorescent protein as a marker for gene expression. *Science*, 263(5148):802–805, 1994.

[17] Nathan C. Shaner, Paul A. Steinbach, and Roger Y. Tsien. A guide to choosing fluorescent proteins. *Nature Methods*, 2(12):905–909, 2005.

[18] Kirill A. Datsenko and Barry L. Wanner. One-step inactivation of chromosomal genes in escherichia coli k-12 using pcr products. *The Proceedings of the National Academy of Sciences USA*, 97(12):6640–6645, 2000.

[19] Lynn Thomason, Donald L. Court, Mikail Bubunenko, Nina Costantino, Helen Wilson, Simanti Datta, and Amos Oppenheim. *Recombineering: Genetic engineering in bacteria using homologous recombination.* John Wiley and Sons, Inc., Hoboken, New Jersey, 2007.

[20] Takeharu Nagai, Keiji Ibata, Eun Sun Park, Mie Kubota, Katsuhiko Mikoshiba, and Atsushi Miyawaki. A variant of yellow fluorescent protein with fast and efficienct maturation for cell-biological applications. *Nature Biotechnology*, 20(1):87–90, 2002.

[21] Pabak Sarkar, Srinagesh V. Koushik, Steven S. Vogel, Ignacy Gryczynski, and Zygmunt Gryczynski. Photophysical properties of cerulean and venus fluorescent proteins. *Journal of Biomedial Optics*, 14(3):034047, 2009.

[22] Daniel Mazia, Gerald Schatten, and Winfield Sale. Adhesion of cells to surfaces coated with polylysine. *The Journal of Cell Biology*, 66(1):198–200, 1975.

[23] Andrzej M. Kierzek, Lu Zhou, and Barry L. Wanner. Stochastic kinetic model of two component system signalling reveals all-or-none, graded and mixed mode stochastic switching responses. *Molecular BioSystems*, 6(3):531–542, 2010.

[24] A. D. Mcnaught and A. Wilkinson. *IUPAC. Compendium of Chemical Terminology, 2nd ed. (the "Gold Book").* WileyBlackwell; 2nd Revised edition, August 1997.

[25] Lawrence F. Shampine and Mark W. Reichelt. The matlab ode suite. *SIAM Journal on Scientific Computing*, 18:1–22, 1997.

[26] Daniel T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, 22(4):403–434, 1976.

[27] Daniel T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*, 81(25):2340–2361, 1977.

[28] Daniel T. Gillespie. Stochastic simulation of chemical kinetics. *Annual Review of Physical Chemistry*, 58:35–55, 2007.

[29] Hamid Bolouri1 and Eric H. Davidson. Modeling transcriptional regulatory networks. *BioEssays*, 24:1118–1129, 2002.

[30] Khuloud Jaqaman and Gaundez Danuser. Linking data to models: data regression. *Nature Reviews Molecular Cell Biology*, 7:813–819, 2006.

[31] Aleksandra M. Walczak, Andrew Mugler, and Chris H. Wiggins. Analytic methods for modeling stochastic regulatory networks. In Xuedong Liu and Meredith D. Betterton, editors, *Computational Modeling of Signaling Networks*, pages 273–322. Springer Protocols, 2012.

[32] R. Khanin, V. Vinciotti, V. Mersinias, C. P. Smith, and E. Wit. Statistical reconstruction of transcription factor activity using michaelismenten kinetics. *Biometrics*, 63:816–823, 2007.

[33] Wei-Cheng Tian and Erin Finehout, editors. *Microfluidics for Biological Applications*, chapter 2. Springer Science, 2008.

[34] Younan Xia and George M. Whitesides. Soft lithography. *Annual Review of Material Science*, 28(1):153–184, 2005.

[35] Kirsten Avila, David Moxey, Alberto de Lozar, Marc Avila, Dwight Barkley, and Bjorn Hof. The onset of turbulence in pipe flow. *Science*, 333:192–196, 2011.

[36] Alex Groisman, Caroline Lobo, Hojung Cho, J. Kyle Cambell, Yann S. Dufour, Ann M. Stevens, and Andre Levchenko. A microfluidic chemostat for experiments with bacterial and yeast cells. *Nature Methods*, 2(9):685–689, 2005.

# VITA

# VITA

Chetan Sood, born on September 1, 1981 in Austin, Texas, completed a Bachelor of Science in Physics and in Mathematics at the University of Texas at Austin in May 2004. Complimenting the two degrees, while at the university, he studied extensively in the fields of computer science, philosophy and language arts.

In August 2004, he moved to West Lafayette, Indiana to begin the doctoral program in Physics at Purdue University. He was awarded the Ross Doctoral Fellowship to support his first year of coursework. After completing his coursework, in May 2006 he joined the Applied Physics Laboratory at Purdue University (APL) directed by Professor David Koltick, where he was awarded an industrial fellowship from Clarion Systems. At APL, he worked as part of a small team to build a hazardous material scanner using a neutron-based scanning technique. In the culmination of this project, he facilitated the successful blind testing of the hazardous material scanner against a series of samples that included live chemical warfare agents (e.g., Sarin) at the Edgewood area of the Aberdeen Proving Ground.

In August of 2008, he changed research tracks and joined the biophysics laboratory directed by Professor Ken Ritchie. There he employed fluorescence microscopy and computational modeling techniques to investigate the response characteristics of a biological transducer in living *Escherichia coli.* In May 2013 he was awarded a Purdue Research Foundation fellowship, and during the duration of that award he completed his dissertation describing the investigation of the biological transducer. In December 2013, his dissertation earned Ph.D. in Physics from Purdue University.

Throughout his graduate career he has been supported by research or teaching assistantships. He also held leadership positions at various organizations while in graduate school, with the primary intent of engaging in science outreach and professional development. For example, while on the executive board of the SACNAS chapter at Purdue, he designed and managed a one-hour, after-school science workshop that visited local elementary schools. And, while on the executive board of the Physics Graduate Student Association at Purdue, he created and managed a seminar series for Physics graduate students to present their research to the department.

Currently he is completing two manuscripts to submit for peer-reviewed publication regarding his dissertation work, and he anticipates starting a post-doctoral research appointment in the near future.