REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (From - To)
05-09-2014	MS Thesis	-

4. TITLE AND SUBTITLE
Optimizing a synthetic signaling system, using mathematical modeling to direct experimental work

5a. CONTRACT NUMBER
W911NF-09-1-0526

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
206023

6. AUTHORS
Keira Havens

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES AND ADDRESSES
Colorado State University - Ft. Collins
2002 Campus Delivery

Fort Collins, CO 80523 -2002

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS (ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSOR/MONITOR'S ACRONYM(S)
ARO

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
56814-CH.9

12. DISTRIBUTION AVAILIBILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.

14. ABSTRACT
Synthetic biology uses engineering principles and biological parts to probe existing biological networks and build new biological systems. As biological components become better characterized, synthetic biology can make use of predictive mathematical models to analyze the activity of biological systems. This thesis demonstrates the utility of modeling in optimizing a synthetic signaling system for a bacterial testing platform and advances the use of model-based bacterial systems as an effective tool of plant synthetic biology.
Using models in combination with experimental data, I showed that increasing the concentration of a single

15. SUBJECT TERMS
synthetic biology, histidine kinase, modeling

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			June Medford
UU	UU	UU	UU		19b. TELEPHONE NUMBER: 970-491-7865

Report Title

Optimizing a synthetic signaling system, using mathematical modeling to direct experimental work

ABSTRACT

Synthetic biology uses engineering principles and biological parts to probe existing biological networks and build new biological systems. As biological components become better characterized, synthetic biology can make use of predictive mathematical models to analyze the activity of biological systems. This thesis demonstrates the utility of modeling in optimizing a synthetic signaling system for a bacterial testing platform and advances the use of model-based bacterial systems as an effective tool of plant synthetic biology.

Using models in combination with experimental data, I showed that increasing the concentration of a single component of the synthetic signaling system, the PBP, results in a 100 fold increase in sensitivity, and an order of magnitude increase in fold change response in the response of the bacterial testing platform. Additional mathematical exploration of the system identified another component, the number of PhoB inducible promoters, which could be adjusted to further increase maximum signal. In addition, our model has suggested additional avenues of research, including the potential to introduce new functions, such as memory, to the existing circuit.

In this way the prototype synthetic signaling system developed by the Medford Lab has been refined to improve detection and generate substantial response, moving the technology closer to real-world use. Once validated, this modeling based protocol, using a microbial platform for developing and optimizing plant synthetic systems, will serve as a foundation for engineering advanced plant synthetic systems.

THESIS

OPTIMIZING A SYNTHETIC SIGNALING SYSTEM, USING MATHEMATICAL MODELING TO DIRECT EXPERIMENTAL WORK

Submitted by

Keira Havens

Department of Biology

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2014

Master's Committee:

Advisor: June Medford
Co-Advisor: Ashok Prasad

Mauricio Antunes
Olve Peersen

ABSTRACT

OPTIMIZING A SYNTHETIC SIGNALING SYSTEM, USING

MATHEMATICAL MODELING TO DIRECT EXPERIMENTAL WORK

Synthetic biology uses engineering principles and biological parts to probe existing biological networks and build new biological systems. As biological components become better characterized, synthetic biology can make use of predictive mathematical models to analyze the activity of biological systems. This thesis demonstrates the utility of modeling in optimizing a synthetic signaling system for a bacterial testing platform and advances the use of model-based bacterial systems as an effective tool of plant synthetic biology.

Using models in combination with experimental data, I showed that increasing the concentration of a single component of the synthetic signaling system, the PBP, results in a 100 fold increase in sensitivity, and an order of magnitude increase in fold change response in the response of the bacterial testing platform. Additional mathematical exploration of the system identified another component, the number of PhoB inducible promoters, which could be adjusted to further increase maximum signal. In addition, our model has suggested additional avenues of research, including the potential to introduce new functions, such as memory, to the existing circuit.

In this way the prototype synthetic signaling system developed by the Medford Lab has been refined to improve detection and generate substantial response, moving the technology closer to real-world use. Once validated, this modeling based protocol, using a microbial platform for developing and optimizing plant synthetic systems, will serve as a foundation for engineering advanced plant synthetic systems.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 From biology to synthetic biology

A new biological discipline has emerged at the intersection of engineering and biology. Synthetic biology uses engineering principles and biological parts to probe existing biological networks and build new biological systems. These goals are achieved by repurposing wild-type systems or designing biological circuitry *de novo* [1]. Crucial to the successful implementation of synthetic designs is a thorough comprehension of the governing interactions.

The modern understanding of gene networks and regulation evolved from investigations into *E. coli* sugar metabolism by Jacob and Monod over fifty years ago [2]. At this time, the molecular nature of biology was being unraveled. DNA had been recognized as the 'hereditary' informational molecule by Avery et al in 1943, and the molecular structure of DNA had been discovered less than a decade earlier [3]. That same year, messenger RNA was proposed as the mechanism of information transfer between DNA and protein [4]. However, even without explicit molecular characterization, careful analysis of generated and spontaneous mutants suggested a similar control mechanism operating in diverse areas of biology, *e.g.*, phage immunity, tryptophan synthesis, and lactose metabolism [2].

Jacob and Monod described this control mechanism as 'repression of gene synthesis'. In their words, a cytoplasmic 'agent' interacts with an 'operator locus' to control the expression of protein coding genes. The 'agent' can combine with a specific small molecule such that the constitutive activity of the 'agent' is modified, *e.g.*, a repressing agent no longer represses expression [2]. Today we understand these 'agents' to be transcription factors and the 'operator

locus' to be the DNA binding sites of these repressors. For example, the *lac* repressor binds to the *lac* operator and prevents expression of the lac operon in the absence of the small molecules allolactose and IPTG [5]. In the presence of either allolactose or IPTG, the *lac* repressor binds to the small molecule and changes conformation, releasing the *lac* operator and allowing gene expression [6]. Despite their inability to precisely characterize these components, Jacob and Monod were able to accurately characterize repression of gene expression [2].

Genetic expression and regulation was explored by mathematical analysis shortly thereafter when Goodwin et al described protein translation as a function of messenger RNA [7]. However, mathematical exploration of biological systems was soon limited by a lack of information. A model is only as useful as the assumptions upon which it is built, and in the early days of molecular biology there was an insufficient understanding of the network architecture, interactions, and components to accurately characterize the 'agents' and 'operators' at work in these systems.

Knowledge of the proteins, DNA binding sites and genes involved in biological networks has expanded as molecular biology has advanced, but the underlying framework of activation and repression proposed by Jacob and Monod remains essentially intact. Investigations into gene regulation now characterize not only the type of activity (repression or activation) of a transcription factor and the particular DNA sequence it binds to, but quantitative details of the interaction such as rates of dimerization, DNA binding, phosphorylation, degradation and other biochemical activities [8]. Now that biological components have begun to be characterized, synthetic biology can use predictive mathematical models to analyze the activity of biological systems.

1.2 Why is modeling important to biology?

The use of predictive models to design a biological system capable of performing a particular function was pioneered in two genetic circuits. Elowitz et al developed a genetic circuit consisting of three genes, each repressed in turn. In this repressible oscillator, or 'repressilator', each protein concentration oscillates as a result of periodic inhibition of expression [9]. The second system, a toggle switch, was developed by Gardner et al[10]. In this system, mutually repressible promoters can be 'toggled' between two states by inducing expression of the appropriate repressor. In each case, genetic interactions were modeled to produce a behavior, then the model predictions were validated experimentally.

To implement model predictions, the repressilator and the toggle switch were built with genetic pieces that had been thoroughly characterized, including the lac repressor. Though these components were well studied, the synthetic systems contructed with these parts did not exactly replicate model predictions. The repressilator oscillated, but it did so with irregular periodicity and amplitude height[9], and the toggle switch 'switched' at various concentrations of inducer for various cells in a population, rather than the cell population switching in synchronization as predicted[10]. Nonetheless, these two papers established that predictive modeling is a useful tool in developing synthetic gene networks with designed functions.

Modeling enables researchers to predict system-level changes as a result of molecular-level modifications, and as systems grow in complexity, modeling becomes progressively more important. DNA synthesis no longer limits the size of synthetic systems, and our increased understanding of biological regulatory mechanisms allows synthetic biologists to draw from a larger pool of network architectures [1]. Researchers are now designing increasingly complex systems with novel functions, engineering bacteria to seek out tumors and destroy them [11],

developing genetic programs that initiate apoptosis in cancer cells [12], and re-engineering multicellular organisms to detect substances of interest to humans[13]. Anderson, Xie, and many others have made use of modeling to understand the behavior of their complex system, predict outcomes, and direct experimental efforts. In this thesis I apply these same tools to the synthetic signaling system developed by the Medford lab.

1.3 The messiness of biology

Biological systems are highly complex, with multiple layers of control, and often, our understanding of the system is incomplete. *E. coli,* for example, an organism central to molecular and microbial biology, has over 1000 uncharacterized proteins [14]. In addition to the sheer number of components that can interact, multiple layers of regulatory control obscure the link between an individual genetic component and the overall behavior of a system [15].

When investigating a particular system, we first develop a theoretical model using the known components (*e.g.,* Protein X and Protein Y), and known interactions (*e.g.* Protein Y phosphorylates Protein X and Protein X regulates the expression of Protein Y). One way to determine the validity of the model is to develop, mutants as Jacob and Monod did. However, the knockout or overexpression of a gene often yields unexpected results – perhaps there is a redundant system or a previously undiscovered interaction. The lac operon itself is not only repressed by the lac repressor, but activated by the catabolite gene activator protein (CRP). CRP binds to DNA in the presence of cyclic AMP, but this metabolite is only present at high concentrations in the absence of glucose [16]. Therefore the effect of a mutant CRP that cannot activate the lac operon would not be apparent in a glucose based media. As biologists, we pursue discrepancies between model and the experiment, resolve them, and approach the original system with a new and improved model.

One such well-known model is the ABC model of flower organogenesis in the plant *Arabidopsis thaliana*. The ABC model describes the spatial and biochemical interaction of the gene products of three gene families (A, B, and C) in the formation of floral organ identity. Mutants obtained in each of the ABC gene families demonstrated abnormalities in flower organ formation – for example, single mutants of the *agamous* gene resulted in flowers with petals where the stamens and carpels should be [17]. However, the development of the ABC model required double and triple mutants to characterize the interplay between the three gene families [18]. While single mutants like *agamous* had been developed more than 50 years before the 1991 work, the framework for considering mutations in the context of gene regulation did not [19].

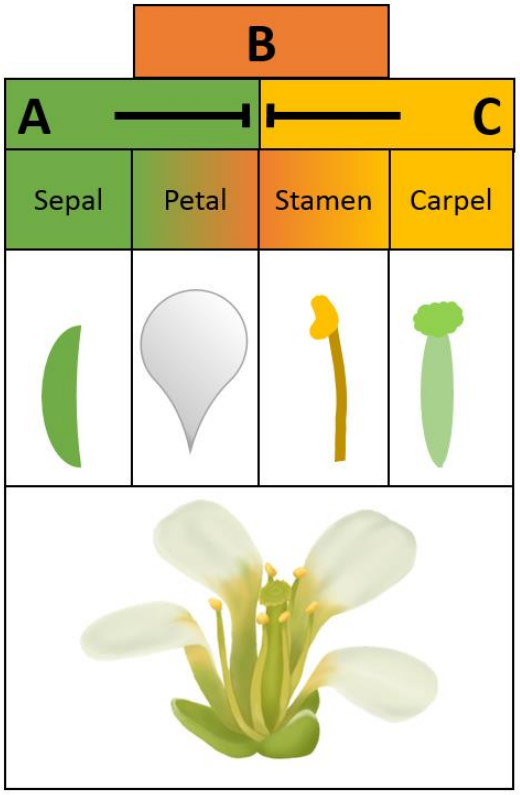


Figure 1.1. Arabidopsis ABC model of floral organogenesis. *In this model, organogenesis is the product of three interacting gene families as the flower forms. A (green) and C (yellow) are mutually antagonistic. On the innermost and outermost whorls of the forming flower, only A and C are produced, resulting in sepals on the outside of the flower and carpels on the interior. B is produced only in the middle two regions of the forming flower, and interacts with A and C to produce two other distinct organs, the petals (A + B) and the stamens (B + C). Bottom panel adapted from Virginia Polytechnic Institute and State University PREP Online Lab Notebook, Arabidopsis Anatomy (VPI-PREP).*

Like Jacob and Monod, Bowman et al used mutant phenotypes to abstract a regulatory network. Specifically, the A gene family is antagonistic to the C gene family, C is antagonistic to A, and B may be co-expressed with either (Figure 1.1). These molecular interactions result in four distinct zones as the flower matures, each corresponding to a particular floral organ: A expression only

(sepals), AB expression (petals), BC expression (stamens), and C expression only (pistils)[18]. A, B, and C interact in this system – these components can be considered as 'nodes' of the gene network. The way A, B, and C interact with one another, for example, the mutual antagonism of A and C, are the 'edges' of the network.

Synthetic biology takes a different approach. Rather than use mutants to establish nodes (molecules) and edges (interactions) in a native biological process (*i.e.*, Protein X interacts with Protein Y to generate response Z), synthetic biology analyzes proposed networks mathematically and uses agreement between experimental data and model predictions to identify likely network architectures (*e.g.*,[20,21]). Investigation into the mathematical properties of these motifs enables synthetic biologists to connect the architecture of the proposed network to biological activity [22]. Experiments (changing the concentration of a gene product or introducing a new interaction) can be carried out *in silico,* thereby reducing the number of experimental studies and focusing the time and effort spent on experimental work. In addition to analyzing endogenous systems, synthetic biologists can construct nodes and edges *de novo,* building genetic networks to achieve a particular goal (*e.g.*, Anderson et al [11], Xie et al [12], Antunes et al [23]).

1.4 From models to implementation

A predictive model is a useful tool, but must be validated by experimental work to establish biological relevance. In combination with experimental results, models can predict the behavior of a synthetic network, probe the effect of perturbations, identify areas of the system for optimization, suggest new avenues of research, and reveal gaps in our understanding of a biological system [15]. However, validating theoretical work experimentally is not always a straightforward task.

Predictive models capable of efficiently investigating complex systems rely on a level of abstraction. In other disciplines, like electrical engineering, models can be implemented using individual components known to interact with each other in a specific manner. These components are organized into complex circuitry in order to perform desired functions: transmit and receive information, perform computations, etc. Using well characterized components, electrical circuitry can be developed and fine-tuned *in silico.*

At the heart of this approach is the idea of modularity, the ability to use an individual part in combination with any other part, and orthogonality, freedom from unwanted interactions [24,25]. In the context of biological systems, modularity can be thought of as a component that has the same activity in any context, *e.g.*, a promoter with a constant transcription rate, regardless of the gene it is expressing or the cell line it is expressed in. Orthogonality can be viewed as freedom from unwanted endogenous interaction as well as freedom from interference by other synthetic system parts, *e.g.*, a protein that is only phosphorylated by its cognate, and not any other protein.

Today, synthetic biology draws heavily from the precepts of electrical engineering when designing and constructing *de novo* circuits, but there are some crucial differences between electrical and biological components. Electrical components are hardwired together, strictly defining their interactions, but we often have an incomplete understanding of all the interactions of a particular biological component. Electrical components are well characterized and have predictable behavior, whereas biological components and systems are highly context dependent and difficult to predict.[1] The complexities of living systems make achieving modularity and orthogonality challenging, and implementation of a synthetic biology model can sometimes be more difficult than the design. [26]

1.5 Thesis statement and overview

1.5.1 Thesis statement

This thesis demonstrates the utility of modeling in optimizing a synthetic signaling system for a bacterial testing platform and advances the use of model-based bacterial systems as an effective tool of plant synthetic biology. Of particular interest is the ability to optimize a circuit computationally in a bacterial testing platform and apply those predictions to systems intended for stable integration in plants.

1.5.2 Optimizing the synthetic signaling system

This thesis focuses on the optimization of a synthetic signaling system, as shown in Figure 1.2. First, I developed a dynamic model of the synthetic signaling system and characterized the behavior of the synthetic signaling system in a bacterial context. After establishing a baseline of activity, I used the model to optimize the system, maximizing fold-change response. My model predictions were experimentally validated in the bacterial model, and subsequently tested in plants.

1.5.3 Approach and summary of contributions

1.5.3.1 Characterization of Synthetic Signaling Circuit

Engineering of synthetic circuits relies on accurate, detailed component characterization[24]. Synthetic circuits are a function of the parameters internal to the system (e.g., concentrations of proteins, binding and unbinding rates) and interactions with external conditions (e.g., media, temperature, cell line). In this work I first established the response of the synthetic signaling

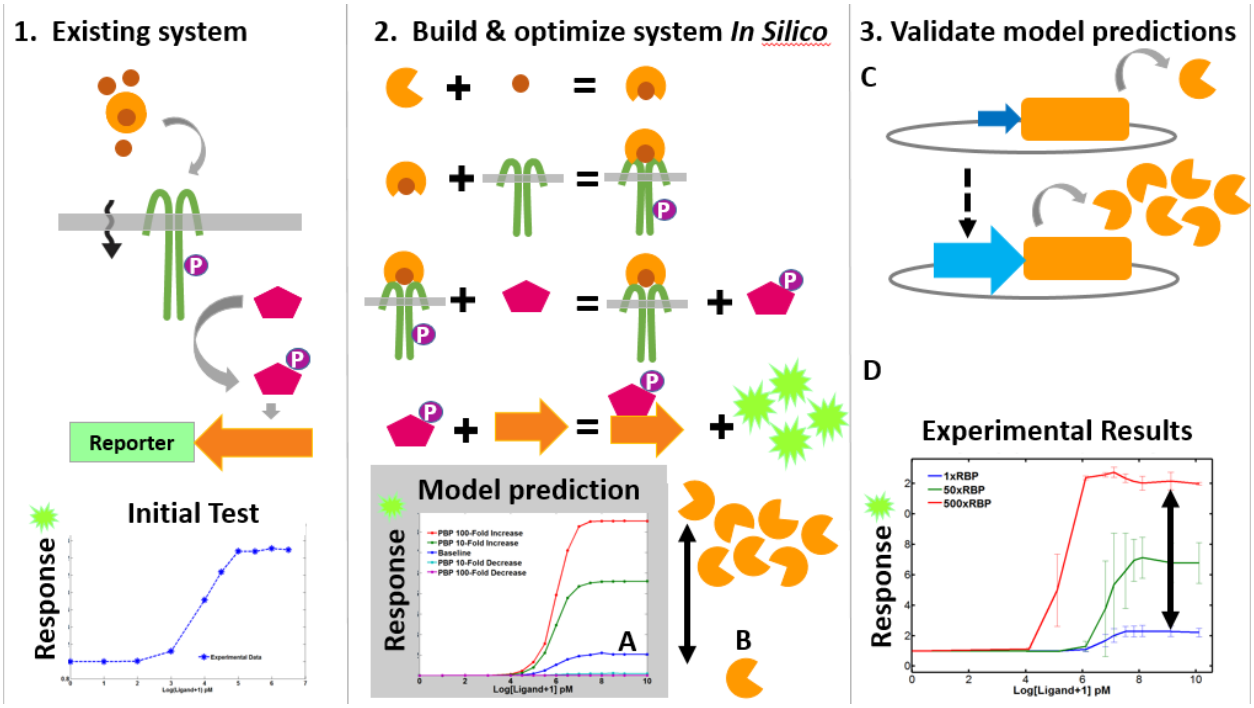


Figure 1.2. Schematic of strategy to optimize a synthetic signaling system for use in plants through* in silico *modeling and experimental validation in a microbial testing platform. *1) Existing experimental system was used to establish the baseline activity of the synthetic signaling system in a bacterial platform. Various growth conditions and measurements were tested. 2)* In silico *development of the experimental system, using literature values for protein concentrations and interactions whenever possible. Parameter space is explored to identify the characteristics of the initial system (A) and areas for optimization, in this case the concentration of PBP (B). 3) Genetic circuits are constructed per model predictions (C). Predictions are validated experimentally (D). If the results validate the model, optimized system can then be introduced into plants. The model can then be used for further optimizations, or the development of new functions. If the results do not validate the model, the model must be revisited to account for previously unknown system behaviors.*

system under a variety of external conditions. I then modeled changes to the internal parameters, seeking out broad trends in system behavior that would increase the maximum signal of the system, increase system threshold of detection, and increase 'switch-like' behavior, to develop a sharp transition from the OFF state to the ON state.

1.5.3.2 Validation of Model in Bacterial and Plant Systems.

I developed this model from a previous model of the synthetic signaling system developed in collaboration with Samy Lyons, a PhD student of Dr. Ashok Prasad. The detailed model described in this thesis was built by applying the Law of Mass Action to the known reactions involved in the synthetic signaling system. Baseline parameter values were assigned according to experimentally determined values, using *in vitro* work or research on closely related protein families when necessary. A global parameter analysis was conducted, in which every parameter and initial protein concentration was scanned over 5 orders of magnitude to determine the sensitivity of the system output to that parameter.

As a result of scanning the parameter space, the concentration of periplasmic binding protein (PBP) was identified as the parameter with the greatest impact on the maximum signal, sensitivity and 'switch-like behavior' achievable by the system. To validate these model predictions, I constructed bacterial vectors to express PBP at varying levels, and tested the system over a range of ligand concentrations. Initial tests of the system were conducted with a plate reader, and flow cytometry was later used to obtain data on individual cells. Upon validation of the bacterial model, I constructed plant vectors to express PBP at varying levels to establish whether insights gained from the bacterial model are indeed applicable to plants.

1.5.3.2 Modeling Additional Parameters and Parameter Combinations of Interest.

I then investigated parameters with more complex relationships, for example the number of 'active' response regulator (RR) binding sites, *i.e.,* RR binding sites in a promoter that induce the expression of the reporter gene when bound. Prior work suggested that multiple promoter-reporter cassettes resulted in a larger output signal in response to a given input[13]. To explore the relationship between phosphorylated response regulator and the number of promoter-reporter

cassettes more fully, I developed a truncated model in collaboration with Katherine Schaumberg, a PhD student in the Prasad lab. Model results supported the existing experimental data, and further indicated that an increase in the number of promoter-reporter cassettes would increase maximum signal as the concentration of phosphorylated response regulator increased. This hypothesis was validated experimentally in bacteria.

1.5.4 Thesis outline

After a description of the background and significance of this project in Chapter 2, in Chapter 3 I describe the construction of the model used to make the initial predictions of synthetic signaling system response. Chapter 4 presents the plasmids used in this work and discusses the experimental characterization of the synthetic signaling system. Chapter 5 illustrates how I experimentally validated model predictions in bacteria. Chapter 6 describes the process of building a model for a specific application and the experimental validation of that model. Chapter 7 concludes this thesis.

CHAPTER 2: BACKGROUND AND SIGNIFICANCE

Plants are unable to escape changes in their environment, and must instead respond quickly to environmental changes. As a result, they have evolved complex signaling mechanisms to rapidly detect and respond to environmental conditions. Among other things, plants detect light, nutrient availability, biotic and abiotic stresses, and volatiles and integrate the signals from various pathways to fine tune control of their developmental processes and transcriptional responses [27]. Plant signaling pathways are themselves highly complex networks, and signal integration adds an additional layer of complexity.

The sophisticated nature of these regulatory pathways, the lack of well characterized plant components, and the long generation time of most plants has stunted plant synthetic biology. Far from building new systems for use in plants, groups have only just begun to use modeling in combination with traditional molecular biology to tease apart endogenous genetic networks (*e.g.*, Rausenberger et al[21], Middleton et al [28], Muraro et al [29]). The Medford lab has not only constructed a synthetic signaling system for use in plants, it has advanced the technology by using mathematical models to identify key parameters and predict outcomes.

2.1 Synthetic Biology: from prokaryotes to eukaryotes

Synthetic biology started in prokaryotes, in part because bacteria are better characterized than their multicellular counterparts. The input-output functions of a large number of inducible and repressible promoters are known, transcription and translation are relatively well characterized, and tools exist to modify the rates of either or both. In addition, prokaryotic organisms grow

rapidly and are amenable to high throughput testing. Novel synthetic circuits can be implemented in prokaryotic chassis to demonstrate proof-of-concept.

Eukaryotes pose additional engineering challenges: it took four years after publication of the toggle switch to implement it in mammalian cells [30], and a synthetic toggle switch has not yet been demonstrated in plants. The differentiated tissues, organelles, and various developmental stages of eukaryotic organisms make developing synthetic biological circuits more complicated, but at the same time provide an opportunity to build spatially and temporally regulated circuits [31]. Unfortunately, the components of eukaryotic systems are not as well characterized as their prokaryotic counterparts, and eukaryotic transcription and translation is more complex [32]. The potential of eukaryotic synthetic biology can be realized through precise control of the genetic elements used in a synthetic eukaryotic circuit.

To this end, the scientific community has adapted well-characterized prokaryotic parts for use in eukaryotic cells to establish the necessary transcriptional control. Synthetic mammalian promoters have been developed to take advantage of known prokaryotic transcription factor and DNA binding pairs. For example, a synthetic mammalian promoter incorporating lac repressor binding elements was co-expressed with the lac repressor [33]. As in bacteria, the lac repressors bound to DNA binding elements and prevented expression of the downstream gene in the absence of the inducer IPTG [33]. Similarly, the *tet* repressor system includes a bacterial transcription factor, tetR, that represses expression by binding to a promoter *tetO* DNA binding site. A synthetic version of this system is used extensively in eukaryotic cells to conditionally express genes, and has been further engineered to develop new eukaryotic transcriptional activators and repressors by fusing tetR to eukaryotic transcriptional activator and repressor

domains [34]. Not only are these components well-defined, they are derived from another organism and are therefore likely to be orthogonal to the host of the synthetic circuit (*e.g.*, Wang et al [35]).

In addition to transcriptional control, there are many ways to tune the behavior of a particular synthetic circuit, for example RNA interference[12], phosphorylation cascades [36], or protein-protein interactions[37]. These components may be derived from prokaryotes to maintain orthogonality, but proteins and RNA-based technologies can also be developed through directed engineering or rational design[38]. Using these tools, synthetic biologists can address a complicated pathology like cancer, facilitating therapeutics development and drug discovery [39]. The promise of medically relevant genetic circuits has proved a powerful motivation, spurring synthetic biology research in mammalian cells (reviewed extensively in [40,41], and as described in [11,12,30,32,42,43] and many more).

2.2 Plant synthetic biology

In contrast to the many examples of prokaryotic and mammalian synthetic biology, plant synthetic biology is still in its infancy. One barrier to development of synthetic biology in plants is the long timescale required. The standard model organism for plants, *Arabidopsis thaliana*, has a small diploid genome, is easily transformed via *Agrobacterium*, and has a deep community of researchers from which to draw protocols, mutants, and tools [44]. However, after *A. thaliana* flowers are transformed via *Agrobacterium*, it takes approximately three weeks for seed formation, 5-10 days to select for transformed first generation plants (T0), and six to eight weeks after that to get seed from the first generation – three to four months in total. Moreover, Arabidopsis has one of the shortest plant lifecycles. Transformation, selection, and growth of a monocot such as rice can take up to 9 months, and trees require an even longer timespan. Despite these challenges, plants are a promising platform for synthetic biology. Plants are essential for

food, feed, fiber, and fuel and improvements in these areas have the potential to dramatically impact people's lives worldwide. In addition, plants have intricate metabolic pathways that convert light to chemical energy and can produce high value compounds like medicine and fuel[45].

A recent review on plant synthetic biology highlights the sparseness of the field. In the review, there are three published applications of synthetic biology in whole plants: generating a stress tolerant plant by introducing a gene from cyanobacteria; using bacterial enzymes in conjunction with the complicated and often incompletely described metabolic pathways of plants to yield novel metabolites; and engineering a synthetic plant signaling system as developed by the Medford laboratory[45]. Of the three, only the synthetic signaling system is designed to be orthogonal [23] – the other two examples rely on 'black box' interactions with uncharacterized host components to generate their response (stress tolerance in the first case, and non-specific metabolite production in the second). Furthermore, of the three, only the synthetic signaling system was engineered to encompass three aspects of a synthetic system: a sensor, a circuit, and a response [31]. With this synthetic system, plants have the potential to serve as simple and inexpensive detectors of pollutants, explosives, terrorist agents, and other ligands of interest[13].

2.3 Moving plant synthetic biology forward

The synthetic signaling system has the potential to demonstrate the promise of plant synthetic biology. Using synthetic components to design plant based genetic circuits enables researchers to engineer control of gene expression beyond overexpression or knockouts. These circuits can then be fine-tuned, using mathematical models to simulate the necessary control mechanisms to achieve a particular goal.

2.3.1 The benefits of synthetic components

Synthetic promoters and other synthetic genetic components offer several advantages in plant synthetic biology. First, synthetic genetic components are often derived from other organisms or are rationally designed by scientists with the explicit goal of reducing interference by endogenous networks and making the synthetic system orthogonal[35,38]. Second, genetic components from well-studied prokaryotes like *E. coli* and *B. subtilis* expand our pool of characterized components. Better characterization of components and their interactions result in better, more predictive models[26]. Third, with a synthetic signaling system it is possible to engineer a specific transcriptional response to a particular extracellular input.

Most plant signaling systems are highly interconnected. Mitogen-activated protein kinases, (MAPKs), which control diverse developmental and defense responses in plants, form a complicated, layered web, with one MAPK interacting with many substrates. There are 90 MAPKs in Arabidopsis, and the function of many of these proteins has yet to be determined [46]. Auxin, a plant hormone involved in flowering and fruit setting, among other complex developmental pathways, involves twenty-nine Aux/IAA proteins (short-lived proteins which interact with auxin response factors (ARFs)), twenty-two ARF proteins (DNA binding proteins), and five homologous degradation proteins [47]. Plant responses to three families of photo-receptors, which tie into diverse transcriptional responses (*e.g.*, de-etiolation, flowering, and circadian clock entrainment)[48] are also highly interconnected, both through plant hormone pathways like auxin and through the multitude of proteins they interact with. Engineering a signaling system using genetic components from endogenous plant pathways like these invites interaction with convoluted endogenous signaling systems.

There have been several efforts to develop synthetic transcriptional control systems in plants. One example is a transcriptional control method based on rat glucocorticoid receptors (GR) [49]. To prevent endogenous interaction, the hormone-binding domain of the GR was fused to the yeast Gal4 binding domain and a eukaryotic activator from the herpes virus, VP16. Upon binding the small molecule glucocorticoid, the chimeric GRs effectively induced transcription of luciferas. Similarly, gene transcription can be induced using synthetic chimeric ecdysone receptors that form homodimers in the presence of ecdysone [50], or tetracycline de-repressible systems [51]. While these systems generate a ligand-dependent response and use orthogonal components, they cannot translate an external signal into transcriptional response. Instead, they depend on ligand freely diffusing into the plant cell and binding to transcription factors.

The Medford synthetic signaling circuit transmits a signal from the exterior of a cell to the interior, using a 'biological wire' through which the signal propagates [13]. Both the input to the system and the transcriptional response can be engineered for an intended application – the PBP through protein re-design [52,53], and the transcriptional response by placing a gene under the control of the inducible promoter. The synthetic signaling system is therefore a flexible tool for detecting a variety of extracellular ligands and generating transcriptional responses. The expression of a reporter such as luciferase or green fluorescent protein is a convenient quantifiable response in a research setting, but agricultural applications may require the production of a transcription factor capable of inducing a developmental or metabolic change (e.g. flowering, ripening, or sugar production)[54]. In a detection capability, the synthetic signaling system could instead initiate transcription of a visual cue, alerting observers to the presence of a compound of interest. It was with this goal in mind that the Medford Lab developed this cutting-edge synthetic signaling system for plants.

2.3.2 Importance of system optimization

While small synthetic circuits can sometimes be fine-tuned through trial and error, this method becomes time-consuming and costly for complex genetic circuits. Computational tools, therefore, are a powerful tool for facilitating experimental work on complex systems [26]. Nonfunctional systems can be examined *in silico* to identify modifications that will result in functionality (e.g., [43]), and functional systems can be analyzed to determine where a genetically encoded change will have the most impact on a synthetic system, and what changes to network architecture will yield a particular behavior [26].

The synthetic signaling system used for testing induces signaling in the presence of ribose, and has been shown to function in both bacteria and plants [13]. However, in order to build a field-functional detector plant, the system should be able to detect small quantities of ligand and quickly induce a robust response. To refine the synthetic signaling system for use in real-world scenarios, we must first establish a comprehensive understanding of the dynamics of the system, then identify areas for optimization. In order to streamline the optimization process, I used mathematical models to dissect this complicated system and suggest avenues of investigation.

2.3.3 Goals of synthetic signaling system optimization

The synthetic signaling system may be used as a detection circuit. The optimization parameters investigated here are particular to this application, *i.e.*, a detection circuit should sense low quantities of ligand and move from OFF to ON with a very small change in ligand quantity. While these traits may be useful for other synthetic systems as well, other applications may also require different optimization parameters.

In this case, I am interested in three parameters: the maximum signal in the presence of ligand when compared to background signal in the absence of ligand (fold-change response), threshold of a response, and the 'switch-like' behavior of the circuit, as shown in Figure 2.1. A robust detector will produce a significant response in the presence of ligand, so I will consider parameters that increase the fold-change response. In addition, a detector should be sensitive to the ligand, so I will also consider parameters that decrease the threshold of the response (*i.e.*, generate a response at a lower level of ligand). The final parameter increases the 'switch-like' behavior of the circuit, by shrinking the change in ligand concentration needed to effect a change from the OFF state to the ON state is also investigated.

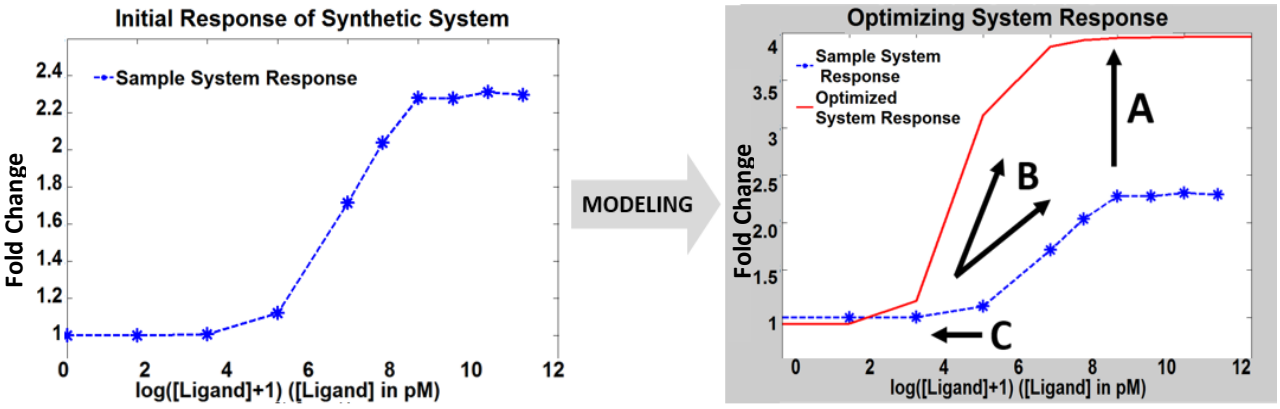


Figure 2.1 Goals of synthetic signaling system optimization. *Left: The response of the synthetic signaling system to ribose, experimental data. The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the response as fold change over background. I used modeling to identify areas of optimization (right): x and y axes remain the same, blue dashed line is experimental data and red solid line is a model prediction. The goals of optimization are A) to increase maximum signal; B) to decrease the change in ligand concentration required to transition the system from off to on; C) to decrease the threshold of detection.*

2.2 Understanding the Synthetic Signaling System

2.2.1 The synthetic signaling system in brief

The Medford Lab has developed a synthetic signal transduction system that functionally links a particular extracellular input to a specific transcriptional response, allowing for ligand-dependent control of a transcriptional response. Figure 2.2 details this system, where an extracellular ligand is bound by a periplasmic binding protein (PBP) secreted into the extracellular space. Upon binding, the PBP changes conformation and develops high affinity for the extracellular domain of a transmembrane protein. These elements are adapted from a bacterial chemotactic system.

The transmembrane protein (HK) acts as the first half of a 'biological wire', transmitting information from the outside of the cell to the inside, with the signal taking the form of a high energy phosphoryl bond [13]. The phosphate group is transferred from the membrane bound HK to a cytoplasmic response regulator (RR). The RR changes conformation into an active form upon phosphorylation and intiates a transcriptional response. These elements are adapted from a bacterial two component system. The phosphorylated response regulator binds to an inducible promoter and activates transcription of a reporter gene. The specific proteins used in the synthetic signaling system, the wild-type systems they are derived from, potential for modularity, and particulars of the phospho-relay are discussed in further detail in this section.

2.2.2 Input and signal transduction

2.2.2.1 Two component systems

A two component system (TCS) is at the center of the synthetic signaling system. TCS are relatively linear phospho-relays between histidine kinases (HKs) and response regulators (RRs) found in bacteria, plant, and fungal systems[55]. In these systems, a signal is recognized by the

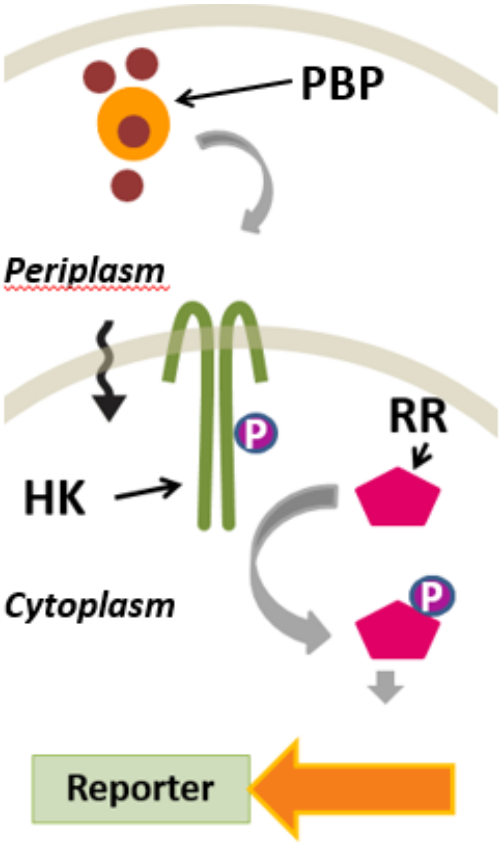


Figure 2.2. The synthetic signaling system. *The input is a periplasmic binding protein (PBP, yellow), a protein that changes conformation upon binding its ligand, developing a high affinity for the extracellular sensor domain of a transmembrane protein (HK, green). The transmembrane protein acts as a 'biological wire' transmitting information from the external binding event through the inner cell membrane and into the cytoplasm as a high-energy phosphate group (P). A response regulator (RR, pink) accepts the phosphate group and undergoes a conformational change. Thus activated, the RR is then able to bind to a promoter and induce transcription of a gene of interest.*

extracellular portion of the histidine kinase and transduced through the membrane, where it results in phosphorylation of a histidine residue in the cytoplasmic portion of the histidine kinase. Unlike serine-threonine kinases, the phosphate group of a phosphorylated histidine is quickly transferred [55]. In plants, the phosphate is generally transferred intramolecularly, to an aspartate residue in the receiver domain of the same histidine kinase. It is then transferred to the accepting histidine of a cognate effector protein. This effector protein translocates to the nucleus, where it transfers the phosphate to the aspartate of a response regulator which then induces a transcriptional response (Figure 2.3 A)[55]. TCS have been identified as ideal candidates for orthogonal gene expression in synthetic systems due to the specificity of a RR for its cognate HK.[38]

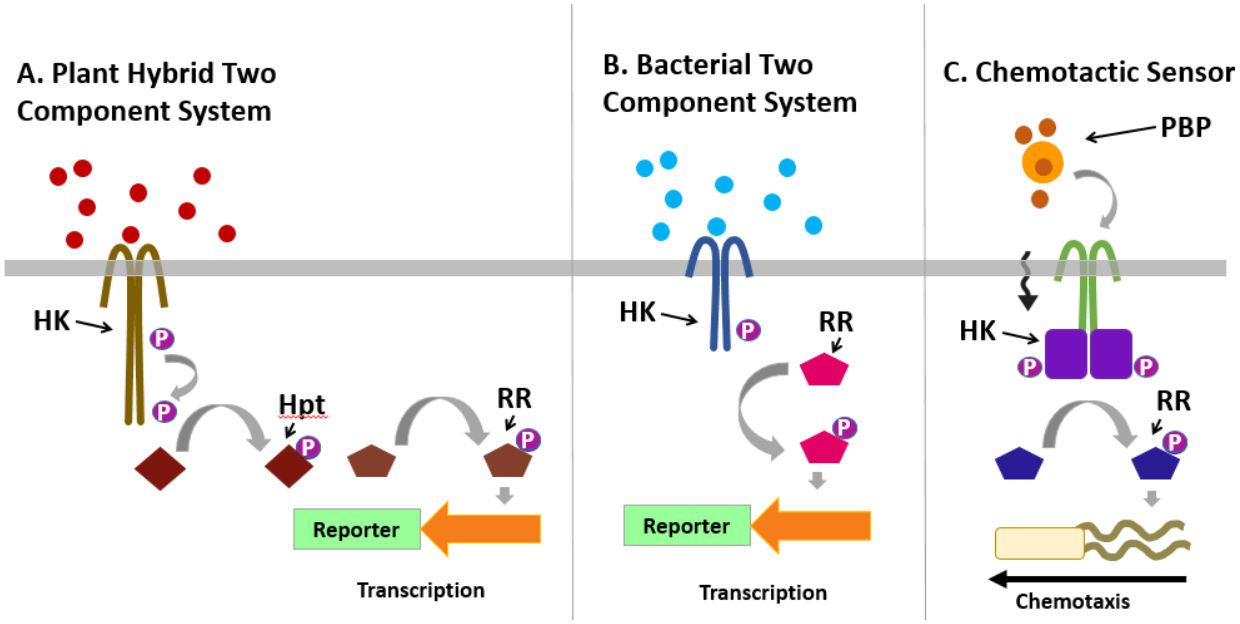


Figure 2.3. Schematic of endogenous plant and bacterial two component systems. *A) A typical plant hybrid two component system. The histidine kinase (HK) binds an extracellular ligand directly, phosphorylating the intracellular or cytoplasmic histidine in response. The phosphate group (purple P) is transferred several times: first to an aspartate in the receiver domain of the histidine kinase, then to a histidine on a histidine phosphotranferase (HPT), and finally to a response regulator which effects a transcriptional response. B) A typical bacterial two component system. The bacterial histidine kinase (HK) also binds extracellular ligand directly and is phosphorylated at the cytoplasmic histidine phospho-acceptor. The phosphate is transferred once to a response regulator which effects a transcriptional response. C) A sample chemotactic sensor, a specialized type of two component system. The chemotactic receptor binds a ligand indirectly, via a periplasmic binding protein. The PBP-ligand complex binds to the chemotactic receptor (green) which transduces the signal through the membrane and phosphorylates an associated histidine kinase (purple). This phosphate group undergoes several phosphotransfers until reaching a response regulator which effects the final phophotransfer to a flagellar motor protein, modifying bacterial motion.*

It is important to note that many histidine kinases have dual functions: they can both phosphorylate the unphosphorylated response regulator and act as a phosphatase for the phosphorylated response regulator[56]. The equilibrium of the autophosphorylation reaction favors the unphosphorylated protein, and so only a small fraction of the total HK will be phosphorylated in the absence of signal[55]. External signals can impact the equilibrium of the system towards

autophosphorylation or dephosphorylation [57], and it is the ratio of these activities that determines the state of the system rather than maximum kinase activity [58].

TCS are found in both prokaryotic and eukaryotic cells. While eukaryotic TCS are often integrated into more complex signaling pathways such as MAPK phospho-cascades, bacterial TCS can use as few as two components to transmit a signal from outside the cell to inside and effect a response[55]. In bacteria, as in plants, an extracellular signal is transduced through the membrane by the histidine kinase and phosphorylates the histidine in the cytoplasmic portion of the protein (Figure 2.3, B). Unlike plants, the phosphate group is then transferred directly to the aspartate of the cognate response regulator[59]. The phosphorylated response regulator can then initiate a response. This response may be transcriptional in nature, as in the case of the EnvZ/OmpR TCS, which senses osmolarity and regulates the production of porins, or it may be a post-translational effector, as in chemotaxis (*e.g.*, the Trg/CheA system), which initates flagellar motion in response to chemoattractants.

2.2.2.2 Chemotaxis

Chemotaxis, the movement of bacteria towards or away from a chemical, is one of the most well studied TCS [60]. The relationship between some extracellular signals and the corresponding response is characterized, and has been mathematically modeled [61]. Bacteria sense chemoattractants either directly, as in the case of aspartate binding to Tar, or indirectly, via a periplasmic binding protein (PBP) which first forms a complex with the substrate and then binds to the receptor histidine kinase (Figure 2.3, C) [60].

PBPs are known to be robust scaffolds for protein engineering[53]. If PBPs can be engineered to bind new substrates with high specificity without disrupting the binding of the PBP-ligand

complex to the extracellular portion of chemotactic receptor, this system can be repurposed to detect substances of interest. This has been demonstrated for the ribose binding chemotactic system in particular, where ribose binding protein (RBP) has been modified to bind other substrates, including but not limited to zinc[62], vanillin[63], lactate, serotonin, and TNT [52].

The modularity of the chemotactic system is a desirable trait for developing detector plants that respond to particular compounds of interest. However, the chemotactic TCS does not result in a transcriptional response. Instead, it modifies the movement of the flagella via phosphorylation of effector proteins[60]. In plants, therefore, a heterologously expressed bacterial chemotactic TCS will not yield a detectable output in the presence of ligand. Hence, there is a need to identify an alternative, transcriptional, response.

2.2.2.3 Combining the modularity of the PBP with the specificity of the TCS

To build a plant detector, we combined the modularity of the PBP and chemotactic system with the specificity of the TCS [13]. Fusions between chemotactic systems and TCS were first demonstrated by Utsumi et al, who fused the extracellular portion of the chemotactic Tar transmembrane protein to the cytoplasmic portion of the EnvZ/OmpR TCS [64]. The fusion HK induced a transcriptional response in a small-molecule dependent manner; however, it responded to the direct binding of aspartate by the chemotactic receptor, as opposed to the indirect binding of a substrate by a PBP and that complex subsequently binding to the chemotactic receptor [64]. Therefore it lacked the desired modularity of a detector system.

Expanding on the idea of a chimeric transmembrane histidine kinase, the Medford Lab further analyzed chemotactic and other TCS histidine kinases for structural and functional similarities. Figure 2.4 shows the development of a novel fusion histidine kinase between the extracellular

portion of the chemotactic receptor Trg and the PhoR histidine kinase, hereafter referred to as HK72 (Morey et al, in prep). Most experiments described in this thesis use RBP as the input and HK72 to transduce the signal across the membrane and transmit it to the response regulator. However, the synthetic signaling system is not restricted to those components and could be re-designed to accommodate different inputs (glucose/galactose binding protein for example) or different outputs (e.g., EnvZ/OmpR phosphorylation).

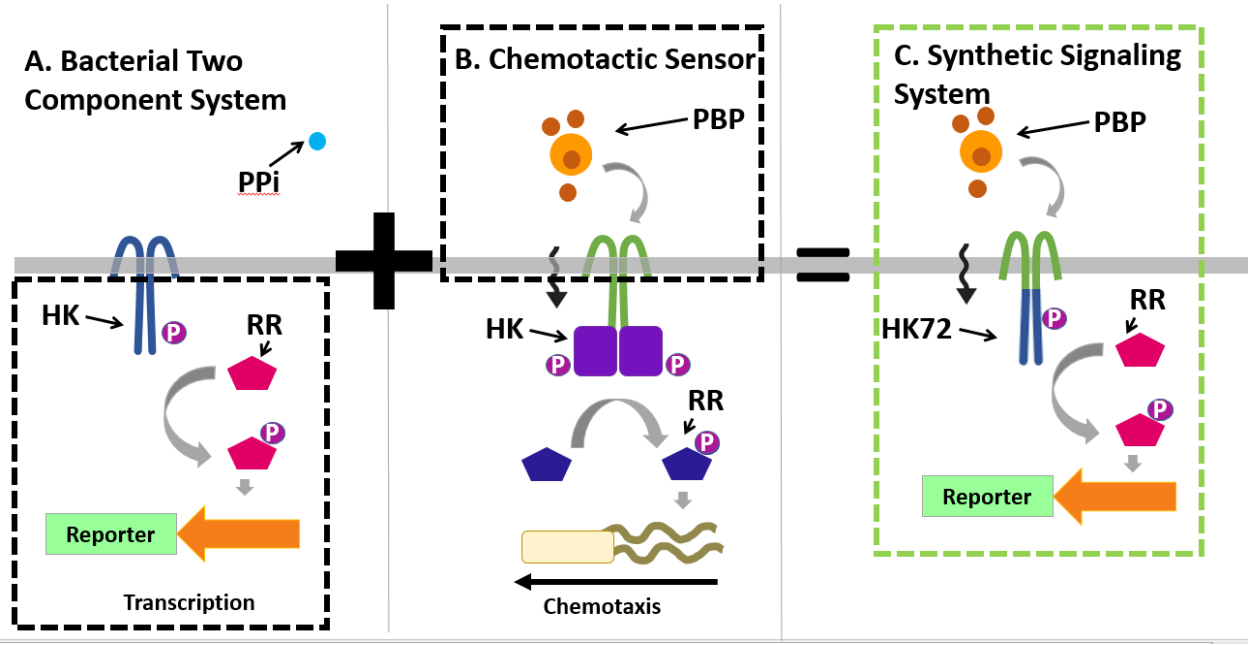


Figure 2.4. Generation of the fusion HK. *A. Bacterial TCS. In the PhoR/PhoB system, the HK is PhoR, the RR is PhoB, and the system responds to a lack of inorganic phosphate. The transcriptional response generated by this system is a desired property for a plant detector (black dashed box) B. Ribose chemotactic system. In this system, the PBP is Ribose Binding Protein (RBP), the protein transducing the signal through the membrane is Trg, the HK is the associated histidine kinase CheA and the RR in this case represents the phospho-relay from HK to response regulator that effects a change in flagellar motion. The specificity of the PBP for its ligand, and the potential modularity of this input are desirable properties for a plant detector (black dashed box). C. The synthetic signaling system. The extracellular sensing domain of Trg and the cytoplasmic portion of PhoR are fused so that the binding event initiated by the chemotactic system outside of the cell is translated into a transcriptional response by the TCS inside the cell. In this system, ligand is bound by the PBP and the PBP-ligand complex then binds to the extracellular sensing domain of the Trg chemotactic receptor. This signal is then transduced through the membrane by the Trg transmembrane domain. The extracellular sensing domain of Trg is now fused to the cytoplasmic domain of PhoR (HK72), and so the physical binding event on the exterior of the cell is translated into a chemical signal (autophosphorlyation*

of the PhoR derived domain of HK72) inside the cell through an unknown mechanism. Phosphorylated HK72 then transfers the phosphate to the RR PhoB, which undergoes a conformational change upon phosphorylation. The conformational change results in dimerization and subsequent activation of PhoB inducible promoters.

2.2.3 Transcriptional response through PhoB activation

The end result of PBP-ligand binding and activation of the synthetic HK is the phosphorylation of the response regulator PhoB. In wild-type *E. coli*, PhoB is phosphorylated in the absence of inorganic phosphate through the following process[65]. First, a repression complex dissociates from the HK, PhoR. PhoR autophosphorylates, and the phosphate group is transferred to the phospho-accepting aspartate of the cognate response regulator, PhoB. Phospho-PhoB undergoes a conformational change, exposing the DNA binding domain of the response regulator[55] and dimerizing. The PhoB dimer goes on to regulate the expression of over 40 genes in the PHO regulon[66].

Only certain parts of this wild-type TCS are used in the synthetic signaling system. For example, the PhoR domain that interacts with the repression complex is not present in the fusion protein (Morey et al, in prep). However, the synthetic signaling system uses PhoB to effect a transcriptional response, thereby making use of both the specificity of the PhoR/PhoB interaction, and the promoters known to be induced by phospho-PhoB.

PhoB contacts the major groove of the binding site with a recognition helix and the minor groove with flanking loops [55]. In bacteria, the response regulator interacts directly with DNA (Figure 2.5, *Bacteria),* but in plants, response regulators must first cross the nuclear membrane, then activate transcription (Figure 2.5, *Plants*). The Medford Lab established that PhoB is capable of

translocating to the nucleus in a signal-dependent manner, and when appropriately modified with a eukaryotic transcription activation domain (VP64), can activate eukaryotic transcription[23].

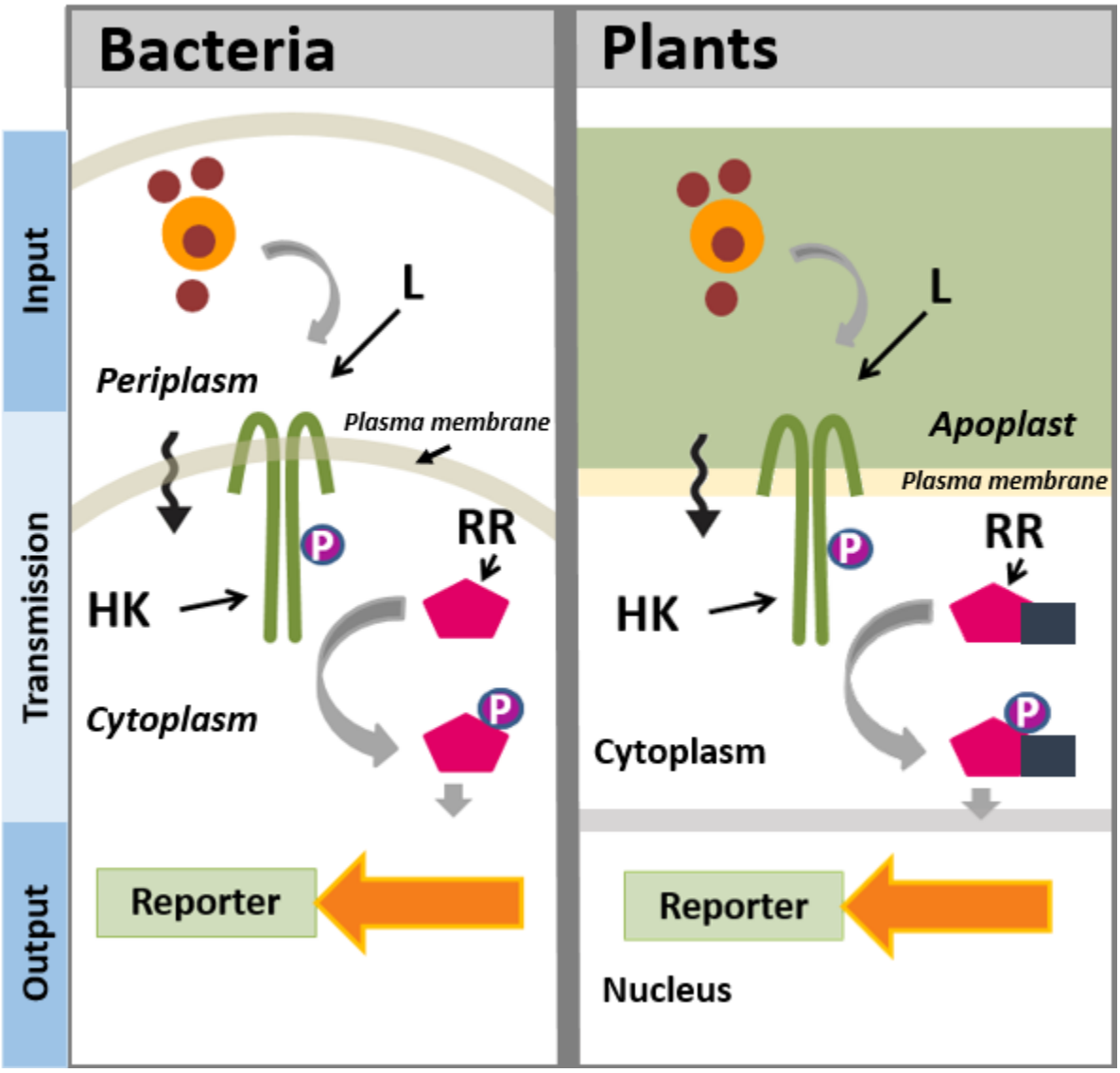


Figure 2.5. The synthetic signaling system functions in both bacteria and plants. *INPUT: The periplasmic binding protein (PBP, yellow) is derived from a bacterial chemotactic system, and directed to the plant's apoplast by the addition of a secretory signal [13]. In both plants and bacteria, this protein binds to its ligand (ribose, red dot), changing conformation and developing a high affinity for the extracellular sensor domain of Trg. SIGNAL TRANSMISSION: HK72 (HK, green), a chimeric membrane-bound HK, is a fusion protein containing the extracellular sensor domain of Trg and the cytoplasmic portion of PhoR. When the PBP-ligand complex binds to the extracellular sensor domain of HK72 the signal is transduced through the membrane via the fusion histidine kinase, thereby activating the cytoplasmic PhoR domain and resulting in*

autophosphorylation of the PhoR phospho-accepting histidine. The phosphate group is then transferred to a response regulator (RR, pink), which changes conformation, and in bacteria, dimerizes [66]. In both bacteria and plants, this activated response regulator binds to a PhoB inducible promoter and induces expression of the downstream gene. In bacteria, this promoter can be an endogenous PhoB inducible promoter like PstS. In plants, PhoB first translocates into the nucleus, then binds a synthetic PhoB inducible promoter developed by the Medford Lab (the Plant Pho promoter). The plant-adapted PhoB also has a eukaryotic activation domain, VP64 (square, navy blue).

2.2.4 Synthetic signaling system summary

In summary, the synthetic signaling system used in this thesis is a fusion of two well-studied two component systems: the extracellular components of a chemotactic system, fused to the *E. coli* PhoR/PhoB TCS. The input, a periplasmic binding protein (PBP) binds a ligand and changes conformation, developing high affinity for the extracellular domain of a chemotactic receptor histidine kinase, Trg. The extracellular domain of Trg has been fused to the cytoplasmic portion of a bacterial histidine kinase PhoR, resulting in TrgPhoR72 (HK72), a fusion HK which detects the extracellular binding of the PBP-ligand complex at the Trg domain and results in autophosphorylation of the cytoplasmic PhoR domain (Morey et al, in prep). This phosphate group is then transferred to the cognate response regulator, PhoB, inducing a conformational change in the RR. In bacteria, this results in dimerization, and subsequent transcriptional response (Figure 2.5, *Bacteria*).

The synthetic signaling system was designed to be orthogonal in plants. All of the system components are derived from bacteria, reducing but not eliminating the probability of interactions with wild-type plant systems. In plants expressing the cytokinin inducible plant HK AHK4, PhoB was shown to translocate into the nucleus in upon exposure to cytokinin [23,67]. This suggests cross talk between the wild-type AHK4 and PhoB. However, this cytokinin-dependent

translocation occurred in the absence of PhoR, PhoB's cognate HK. When the synthetic signaling system is expressed in plants, experiments have shown that the system is unperturbed by common plant stressors under laboratory conditions (unpublished data), a promising sign for the orthogonal nature of this synthetic signaling system. To mimic this orthogonality in our bacterial test system, I used an *E. coli* strain with cross-talk components deleted. This strain is discussed in more detail in Chapter 4.

2.2.5 Functional conservation and the Bacterial Testing System

Each part of the synthetic signaling system, PBP, HK, and RR, has been shown to function across kingdoms. Mizuno et al demonstrated conservation of function of HK proteins by expressing the plant hormone receptor AHK4 in *E. coli,* where it detected the plant hormone cytokinin as it would in *Arabidopsis*[68]. The Medford Lab work complements this finding, showing that bacterial parts maintain their function when expressed in plants and targeted to the appropriate plant cell compartments[13]. The individual rates of binding and concentrations of proteins may differ, but the synthetic signaling system maintains the link between specific extracellular input and ligand-dependent transcriptional output in both bacteria and plants[13]. It is this functional conservation that permits the use of *E. coli* as a test platform for our synthetic signaling system.

Like all new technologies, the synthetic signaling system has room for improvement. Instead of optimizing the system in plants, the Medford Lab has adopted a unique approach: using a bacterial testing platform for synthetic signaling systems destined for use in plants[13]. The synthetic signaling system is complex, incorporating multiple binding steps, a phospho-relay, dimerization, and transcriptional and translational events. In addition, it makes use of synthetic proteins (the fusion HK) and, in plants, synthetic genetic components (the Plant Pho promoter),

both of which are novel and not fully characterized. Traditional plant engineering methods require three months or more to test a single genetic modification, a prohibitively slow timeline for a desirable technology.

One notable advantage of a bacterial testing platform is the ease of transformation and rapid growth rate of *E. coli.* Instead of three months from transformation to experiment in *Arabidopsis*, new components can be implemented quickly and evaluated in a matter of days. Other benefits of using a bacterial testing platform include the well-characterized genetic components and predictive tools available to fine-tune networks in bacteria. Native chemotactic and two component systems are both well studied, with a wealth of literature on kinetic rates of binding, phosphorylation, phosphotransfer and in certain cases, promoter activation (*e.g.*, numerous publications, a very few selected for reference in this thesis [55,61,66,69-72]).

Mathematical analysis can make use of established literature values as a basis for teasing apart the complexities of the synthetic signaling system and making predictions that can be experimentally tested. The contributions in this thesis revolve around the experimental testing of mathematical predictions and the development of a simplified mathematical model characterizing the behavior of the system. Improvements to the system can then be integrated into the plant platform, theoretically streamlining the engineering of the synthetic signaling system in plants.

2.3 Modeling

2.3.1 Models of plant molecular kinetics

Kinetic molecular models developed using the Law of Mass Action have only recently been applied to plant genetic networks. The ABC model described earlier is indeed necessary for floral organogenesis,

but was later found to be insufficient to describe that process[73]. Using recently acquired molecular and biochemical information, mathematical models have been used to identify additional constraints on the ABC model and build a predictive framework for subsequent experimental testing [74,75]. Other plant gene networks, for example the production of protein in various regions of the meristem, are also being modeled [76].

In addition to protein concentration and cell fate, mathematical modeling can be applied to plant signaling systems. These predictive models seek to understand both the global nature of a signaling pathway and the particulars of a given signaling network. Portions of the auxin and cytokinin responses involved in lateral root initiation in *Arabidopsis* have been modeled and experimentally validated in plants[28,29]. The kinetics of the light sensing protein PhyA, to include nuclear shuttling of the protein, have also been modeled and experimentally validated in plants[21]. All of these models use the Law of Mass Action to describe the dynamics of the molecular network.

2.3.2 The Law of Mass Action

The Law of Mass Action states that the rate of a chemical reaction is proportional to the concentrations of the reactants involved[77]. This can be intuitively understood by picturing molecules of the reactants in an enclosed space. The probability that reactant A will interact with reactant B increases with the concentrations of A and B because the molecules are more likely to come into contact with one another (Box 1). The likelihood that this contact will result in a reaction is described by the rate, 'k'. The 'speed' of the reaction (the rate at which the reaction progresses when reactants are in excess) is proportional to this rate, 'k'.

The Law of Mass Action can be applied to a biological reaction as follows: A binds B at rate k, forming complex C. The rate of change of A (*dA/dt*) is therefore equivalent to the concentration of A times the concentration of B and the rate of formation, minus the rate of dissociation of C. In an ordinary

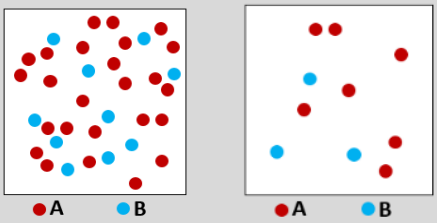
differential equation such as this one, the concentrations of A, B, and C change as a function of one independent variable, time.

In this visualization, it becomes apparent that at very low concentrations of A and B, it is essential to describe the system in terms of discrete elements instead of relying on the mean activity of the components. For example, if there are two moles of reactants, one of A and one of B, you can make a half mole of product. However, if there are only two molecules, one of A, and one of B, you cannot make half of a molecule of product – the product is either made or it is not. While the discrete, stochastic, nature of genetic elements can be important in some contexts (*e.g.*, investigation into cell heterogeneity [78]), stochastic simulations will converge on the deterministic mean as the number of particles involved increases. I do not explore stochastic mechanisms in this thesis as the synthetic signaling system is estimated to operate in the regime of hundreds of proteins, not ten.

Crucial to describing chemical reactions in terms of the Law of Mass Action is the assumption that the reactions are 'elementary', *i.e.*, involving no more than one mechanistic step with a single transition state [79]. The binding of an enzyme to a single substrate meets the criteria, (substrate is either bound or not bound), while the activation of a kinase by multiple phosphorylation steps does not (substrate can be phosphorylated once, multiple times, or not at all). However, each of the phosphorylation reactions in question can be expressed as an individual elementary reaction, and the Law of Mass Action can be applied to each one. In fact, this is how our initial model was built, applying the Law of Mass Action to each of 33 elementary reactions.

Box 1. Ordinary Differential Equations & the Law of Mass Action

Consider the reaction $A + B \leftrightarrow C$



This reaction is dependent on two factors:
1) The amount of free A and B in the system,
2) The rate at which a molecule of A and a molecule of B are converted to C upon interaction.

The speed at which this reaction forms C (that is, the change in C as a function of time) can be written as:

$$\frac{dC}{dt} = k_1[A][B]$$

When there is more reactant available (left), the rate of product formation will increase. The reverse of this process, the breakdown of C into A and B, can be written as:

$$\frac{d[A]}{dt} = k_{-1}[C]$$

Where k_1 is the rate at which two interacting molecules of A and B are converted into C and k_{-1} is the rate at which one molecule of C splits into A and B. At long reaction times, the system comes to equilibrium and the concentrations of both reactants and products are at a steady state: they no longer change with time.

$$0 = \frac{dC}{dt} = k_1[A][B] \qquad\qquad 0 = \frac{d[A]}{dt} = k_{-1}[C]$$

As both equations are now equal to zero, they can be set equal to one another

$$k_1[A][B] = k_{-1}[C]$$

It becomes clear that at equilibrium, the ratio of the rates of dissociation and association are equal to the ratio of the products to the reactants. This ratio is also known as the equilibrium constant. If more reactants or products are added to the system, the system will move towards the equilibrium (*i.e.*, if more C is added, the rate of breakdown will increase proportionally, while the rate of formation will not change)

$$K_{eq} = \frac{k_{-1}}{k_1} = \frac{[A][B]}{[C]}$$

This becomes important in biological systems where the individual $k_{on}$and k_{off} rates may be unknown, but the ratio of the two, the equilibrium constant, can be determined empirically (*e.g.*, through enzyme binding assays)

CHAPTER 3: INITIAL MODEL AND VALIDATION

3.1 Detailed Model

The complexities of the synthetic signaling system make it difficult to deduce the impact of changing a single protein concentration or adjusting a parameter like binding affinity. Simulating the entire system *in silico* allows us to examine global behavior as a function of individual parameters and protein concentrations. Once constructed, a model can simulate many different parameter combinations, thoroughly exploring the biologically relevant space occupied by this model[26]. An in depth investigation of the parameter space directs efficient experimental work, identifying areas for investigation that are likely to significantly impact the system. Experimental results then feed back into the model, either by validating model predictions and lending support to the model, or conflicting with them and prompting a re-evaluation of the mechanisms used to model the system[1].

3.1.1 Establishing a mathematical model of a biochemical system

To model the synthetic signaling system, I developed a mathematical description of the biological reactions involved. Each reaction features at least one reactant and one product, known as species in this model. For example, PhoB and phospho-PhoB are two species involved in this model. The phospho-PhoB dimer is a third species. The Law of Mass Action was applied to the detailed, step by step reactions composing the synthetic signaling system, from extracellular ligand binding to transcriptional response. Phosphotransfer from the HK to the RR, for example, is described as an initial binding event, a subsequent phosphotransfer, and a dissociation step, rather than using a Michaelis-Menten approximation.

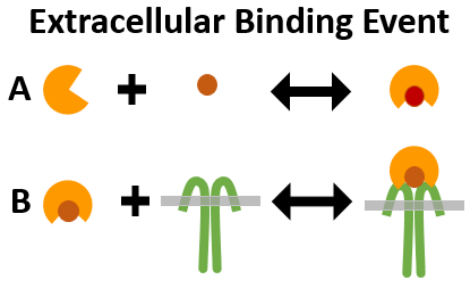
My model builds upon an early kinetic model of the synthetic signaling system developed in collaboration with Samy Lyons of the Prasad Lab. Important biochemical features of the synthetic signaling system (RR dimerization, multiple RR binding sites on the promoter controlling expression of the reporter, pathways for non-ligand dependent phosphorylation, and RR dephosphorylation pathways) are introduced in this model. In addition, the theoretical parameter values used in the first model (derived from a dynamic model of the wild-type PhoR/PhoB signaling system[69] and a model exploring alternative architectures of TCS[57]) have been updated with experimentally determined parameter values.

Parameter values are primarily drawn from published literature on wild-type components used in the synthetic signaling system[66,70-72,80]. Where these values were unavailable for the wild-type components the synthetic signaling system was derived from, parameters were instead estimated from values determined for similar systems[56,81]. In a model, any parameter or protein concentration can be varied over any range of values, and complex biochemical models with large numbers of parameters can be made to fit myriad phenomena if all of the parameters can be varied without restriction[82]. Experimentally determined parameter values constrain the model, reducing the degrees of freedom. If the model based on published mechanisms and values cannot replicate experimental results, this indicates an incomplete understanding of the system. If instead, the experimental results corroborate such a model, it supports the mechanisms incorporated into the model and also lends strength to conclusions drawn from additional *in silico* experiments.

This model uses 20 species in 23 different reactions governed by 33 parameters. At the end of this chapter, Table 3.1 has a complete list of species and initial conditions is found, model parameters, sources, and reactions are listed in Table 3.2, and the explicit equations used in this model are listed in Table 3.3. The signaling system model can be broken down into four categories: the extracellular binding event (1), the transduction of signal across the membrane (2), the transmission of signal from the HK to the RR (3) and the activation of a transcriptional response (4). Each of the four categories is described in this section in terms of the assumptions, parameters, and biological mechanisms represented in this model.

3.1.1.1 Extracellular Binding Event

The extracellular binding event is composed of two distinct reactions, represented graphically and mathematically in Figure 3.1. First, the PBP (yellow) binds to the ligand (red), forming a PBP-Ligand complex (Figure 3.1 A). This complex then binds to the extracellular portion of the histidine kinase (green, Figure 3.1 B).



$$\frac{d[Ligand]}{dt} = \frac{d[PBP]}{dt} = k_{pb} * [PBP][Ligand] - k_{pd}[L]$$

$$\frac{d[HK:L]}{dt} = k_{lb} * [HK][L] - k_{lu}[HK:L] \ldots$$

Figure 3.1. Graphic and mathematical representations of the extracellular binding event. *A. Reversible binding of PBP (yellow) to ligand (red) to form the PBP-Ligand complex. This is mathematically represented on the right hand side of the figure using Law of Mass Action kinetics. The equation can be read as follows: the rate of change of the ligand ($\frac{d[L]}{dt}$) is equivalent to the rate of change of the PBP ($\frac{d[PBP]}{dt}$). Both rates of change are equivalent to the binding rate (k_{pb}) multiplied by the concentrations of PBP and Ligand, minus the dissociation rate (k_{pd}) of the PBP-ligand complex ($[L]$). B. Reversible binding of PBP-Ligand complex to*

HK (green) to form HK bound to PBP-Ligand complex ([HK:L]). This reaction is mathematically represented on the right hand side of the figure using Law of Mass Action kinetics. The equation can be read as follows: the rate of change of the amount of HK bound to the PBP-ligand complex ($\frac{d[HK:L]}{dt}$) can be described in part by the rate of binding (k_{1b}) of the PBP-ligand complex to the HK multiplied by the concentration of PBP-ligand complex ($[L]$) and HK ($[HK]$), minus the dissociation rate of the HK bound to the PBP-ligand complex ([HK:L]) multiplied by the concentration of HK bound to the PBP-ligand complex ([HK:L]). Note the ellipsis – there are additional components to this equation, to be discussed in the next section of the model. The full list of ODEs comprising the model can be found in Table 3.3

There are several assumptions made in this portion of the model. The first is that the total concentration of PBP is conserved. Neither synthesis nor degradation of the PBP are incorporated into this model, a reasonable assumption since the PBP is constitutively produced by P_{lacI} in this system and synthesis should be in equilibrium with degradation. Furthermore, in this model, the PBP-ligand complex does not degrade. If there is a difference between the degradation rate of unbound PBP and the degradation rate of the PBP-ligand complex, there may be an accumulation of total PBP in the system over the course of the experiment, and this assumption may be invalid. In addition, I do not explicitly describe any protein loss from misfolding, nor do I describe export of the PBP to the periplasm. To simplify the model, all reactions occur in the same cell compartment.

I established a baseline of 3.5 µM for the concentration of PBP. This concentration is based on the following information and assumptions: Wild-type P_{lacI} produces approximately 40 LacI molecules, enough to form ten tetramers of LacI repressor[83]. However, this is not a direct analogy to the number of PBP in our synthetic signaling system. RBP does not tetramerize, and each one of the approximately 15 copies of the signaling system plasmid, pACYC177[84], contains a P_{lacI}::RBP cassette, suggesting a PBP concentration of around 600 molecules.

However, I must consider translation as well as transcription. The Salis laboratory has built a predictive model of the effective translation rate of ribosome binding sites (RBS) in a particular organism and in a particular genetic context[85]. In this model, the energetic binding of 16S RNA of the bacterial ribosome to the proposed RBS sequence is analyzed to determine the likely rate of translation. They have also built a calculator for use by the scientific community to analyze and design RBS with a specific translation rate. This calculator predicted that the P_{lacI} RBS is ¼ as efficient in the context of expressing RBP as it is in expressing P_{lacI}. Under these assumptions, total RBP is not 600 molecules, but 150.

While I do not describe export to the periplasm, I account for the difference in volume between the periplasm and the cytoplasm. Cayley et al reported the cytoplasmic volume of *E. coli* grown in M9 media supplemented with glycerol as 1.69 µl H_20 per mg dry weight[86], and a single *E. coli* bacterium has a dry weight of approximately 0.25 picograms[8]. This yields an approximate cytoplasmic volume of 0.42 fL. In contrast, the periplasm of a fully expanded cell was reported by Cayley et al to be 0.31 µl H_20 per mg dry weight[86]. This is approximately equivalent to 0.07 fL, one sixth of the cytoplasmic volume. 150 molecules of total PBP in a volume of 0.07 fL yields a PBP concentration of 3.5 µM.

It is important to note that the steady state concentration of a protein is a function of the rate of synthesis and the degradation of that protein. Consider a protein, X, produced at some constant

Box 2 – Protein concentration at steady state

Synthesis: $\emptyset \rightarrow X$ *at rate* $k_{synthesis}$
Degradation: $X \rightarrow \emptyset$ *at rate* $k_{degradation}$

Rate of change:

$$\frac{d[X]}{dt} = k_{synthesis} - k_{degradation}[X]$$

At steady state, $\frac{d[X]}{dt} = 0$; $[X] = [X_{ss}]$

$$0 = k_{synthesis} - k_{degradation}[X_{ss}]$$

$$[X_{ss}] = \frac{k_{synthesis}}{k_{degradation}}$$

rate, and degraded at another constant rate. The concentration of protein X at steady state ($[X_{ss}]$) is in this case determined solely by the ratio of the synthesis rate to the degradation rate (Box 2) . In the case of LacI, this simple mechanism appears to govern protein concentration: if P_{lacI} is replaced with a stronger promoter, the concentration of LacI repressor increases proportionally [87]. However, we do not know the degradation rate of RBP, and the steady state of RBP under the control of the P_{lacI} promoter may differ from that of LacI due to differences in degradation rate. Therefore, the baseline of 3.5 μM PBP established above should be considered a starting point for simulations and may be adjusted in accordance with experimental data.

The second assumption in this step is that the ligand concentration remains constant over the course of the experiment. The lowest concentration of ligand applied to cell culture is 13 nM, and cultures are assayed at an OD_{600} below 0.01 (less than 10^7 cells per mL) after a 180 to 360 minute incubation. Strains of *E. coli* grown at similar concentrations of ribose (0.2% ribose) in LB media were found to take up ribose at a rate of 23 pM per minute per 10^8 cells[88]. By that calculation, 10^7 cells would use approximately 2.3 pM per minute, up to 828 pM over a 360 minute incubation period, or less than 10% of the lowest ribose concentration tested. Given the small change in ligand concentration, I consider ligand concentration constant over the course of the experiment. In addition, we do not consider cell growth and division in this model.

The third assumption is that the PBP and the extracellular binding portion of the HK function in the same way as their wild-type counterparts. The PBP used in the test synthetic signaling system is identical to the PBP used in ribose chemotaxis by wild-type *E. coli.* Therefore, it should have the same affinity for ribose as its wild-type counterpart. The extracellular binding portion of HK72 is also derived from the *E. coli* ribose chemotactic system, however it has been fused to the cytoplasmic portion of another HK which may have affected its function. The chimeric chemotactic receptor Taz, for example, was shown to respond differently to amino acids than the wild-type chemotactic receptor it is derived from, Tar [89]. For this model, we have defined the dissociation constant of the extracellular binding portion of HK72 for the PBP-ribose complex as the wild-type dissociation constant, but this parameter may require adjustment as experimental data is obtained.

3.1.1.2 Transduction of the signal across the membrane

The synthetic signaling system uses the extracellular binding event to initiate a phospho-relay in the cytoplasm of the cell. It does this by transducing the signal from the exterior of the cell to the interior via the novel fusion protein HK72 (Figure 3.2).

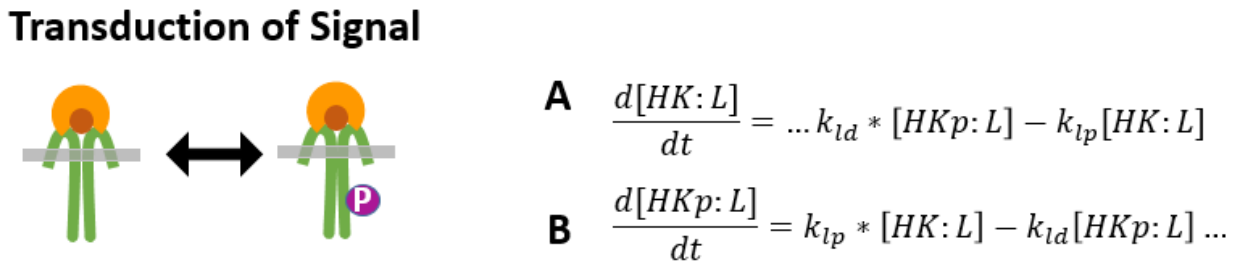


Figure 3.2. Graphic and mathematical representation of reversible phosphorylation of HK72-L complex. *HK:L (left) autophosphorylates to become HKp:L (right, purple 'P'). This is mathematically represented on the right hand side of the figure using Law of Mass Action kinetics. The equations can be read as follows: A) the rate of change of the concentration of HK bound to the PBP-ligand complex ($\frac{d[HK:L]}{dt}$) can be described in part by the dephosphorylation rate of HKp:L (k_{ld}) multiplied by the concentration of HKp:L, minus the phosphorylation rate (k_{lp}) of the HK bound to the PBP ligand complex ($[HK:L]$). Because the HK bound to the PBP-*

ligand complex is involved in the extracellular binding event and this autophosphorylation event, the concentration of HK bound to the PBP-ligand complex is dependent both the reactions shown in Figure 3.2 and Figure 3.3. (hence the ellipsis). B) The rate of change of phosphorylated HK bound to the PBP-ligand complex, ($\frac{d[HKp:L]}{dt}$) is dependent on the autophosphorylation rate of the HK bound to the PBP-ligand complex (k_{lp}) multiplied by the concentration of HK bound to the PBP-ligand complex ($[HK:L]$) minus the autodephosphorylation rate of phosphorylated HK bound to the PBP-ligand complex ([HKp:L]). Because HKp:L is involved in the autophosphorylation reaction, as well as a RR binding event the rate of change of this molecule is dependent on additional terms (discussed in Figure 3.2) The full list of ODEs comprising the model can be found in Table 3.3

There are no experimentally determined values available for either signal transduction or autophosphorylation of HK72. To model the transmission of signal from exterior to interior of the cell, I first assumed that the extracellular domain of HK72 regulates autophosphorylation of the cytoplasmic domain. This assumption is supported by data showing HK72 signaling in a ligand dependent manner (Morey et al, in prep). In their model of the wild-type PhoR/PhoB system, Van Dien et al set the HK autophosphorylation constant to such a large value (167 sec^{-1}) that it could never limit downstream reactions[69] – essentially ensuring that 100% of activated HK become phosphorylated in the model. Given that HK72 is a fusion between two different proteins, it is unlikely that autophosphorylation as a result of ligand binding is as efficient as the wild-type component from which it is derived. Therefore, autophosphorylation of HK72 bound to the PBP-ligand complex is set to a lower rate (10 sec^{-1}) in the model of the synthetic signaling system.

Conversion of the physical binding event to activation and subsequent phosphorylation of the cytoplasmic portion of the chimeric histidine kinase requires additional assumptions. The chemical reactions governing autophosphorylation of some histidine kinases favor the unphosphorylated state[55]. In order to achieve ligand-dependent signaling, this equilibrium must

shift in the presence of the ligand and favor the phosphorylated state. This can be achieved by decreasing the autodephosphorylation rate in the presence of ligand, increasing the autophosphorylation rate, or both.

Parkinson et al suggested that an extracellular binding event results in a physical, conformational change of both chemotactic sensors and sensor histidine kinases[90]. In addition, Manson and Coles have hypothesized that chemotactic sensors propagate the conformational change of the extracellular binding event into the cytoplasm, changing the conformation of the portion of the chemotactic receptor or sensor kinase internal to the cell[91,92]. In HK72, the extracellular sensing domain and transmembrane sequence of Trg are identical to those of the wild-type Trg chemotactic system. This raises the possibility that the extracellular binding event between the PBP-ligand complex and HK72 may also result in a conformational change that is propagated through the membrane and into the cytoplasmic portion of the fusion histidine kinase. HK72 fuses Trg to PhoR at a cytoplasmic residue. Therefore, a conformational change propagated into the cytoplasmic portion of the histidine kinase would likely affect the conformation of the PhoR domain.

The cytoplasmic portion of PhoR has been dissected and each domain analyzed individually[80]. The isolated dimerization and histidine phosphotransfer (DHP) domain of PhoR has been shown to dephosphorylate PhoB; the isolated DHP domain incubated with the isolated catalytic and ATP-binding (CA) domain results in phosphorylation of the DHP domain and, in the presence of PhoB, transfer of that phosphoryl group to PhoB. This suggests that kinase activity is a function of the CA and DHP domains, while phosphatase activity is the function of the DHP domain alone.[80]

If the DHP and CA domains of HK72 were capable of autophosphorylating the DHP histidine residue in presence of ligand as well as the absence of ligand, there would be no difference between the two states. However, HK72 has been shown to signal in a ligand-dependent manner (Morey et al, *in prep)*. Therefore, it is reasonable to assume that DHP and CA domains are more likely to autophosphorylate HK72 in the presence of ligand. Since both domains are part of the same molecule, it seems likely that a conformational change could separate the two domains or bring them into close proximity with one another.

Thus, it may be possible that the extracellular binding event between the PBP-ligand complex and HK72 results in a conformational change that is propagated through the membrane and into the cytoplasmic portion of the histidine kinase, as suggested by Coles et al[92]. This conformational change has the potential to bring the CA domain into contact with the DHP domain and enabling kinase activity in the presence of ligand. This could make phosphatase activity less likely in that conformation, as the DHP domain is involved in both kinase and phosphatase activity. If the DHP domain is participating in kinase activity in the presence of the CA domain it is unlikely to at the same time participate in phosphatase activity, which is facilitated by the DHP domain alone.[80] By the same reasoning, in the absence of ligand the conformation of HK72 would prevent the DHP and CA domains from effectively autophosphorylating the DHP histidine residue, therefore allowing the DHP domain to function independently as a phosphatase, limiting kinase activity in this conformation.

Based on this description of histidine kinase activity, I assume that both autophosphorylation and dephosphorylation activities are modified as a result of HK72 binding the PBP-ligand complex in this model. In addition, I assume that HK72 is a histidine kinase which is primarily unphosphorylated in the absence of ligand. I implemented the above assumptions in the model,

establishing an autophosphorylation rate ten times the rate of autodephosphorylation for HK72 bound to PBP-ligand, and an autodephosphorylation rate 100 times the rate of autophosphorylation of HK72. Because all simulations are run to steady state, the ratio of the phosphorylation and dephosphorylation rates govern the behavior of the system; the specific values used in this model were selected to achieve a ratio of 100:1.

Autophosphorylation is modeled here as a first order process, solely dependent on the concentration of HK72 bound to the PBP:Ligand complex (HK:L). This assumption treats the concentration of ATP as a constant [56]. If the ATP concentration in the cell changes, the model will no longer capture the behavior of the synthetic signaling system.

3.2 Transmission of the signal from HK to RR

In order for the synthetic signaling system to effect a transcriptional response, the phosphoryl group must be relayed from the HK to a transcription factor. Upon transduction of the extracellular binding event, HK72 autophosphorylates at the phospho-accepting histidine in the PhoR portion of the fusion. PhoB then interacts with PhoR, and accepts the phosphate at its phospho-accepting aspartate residue. Phosphorylated PhoR changes conformation and initiates a transcriptional response (Figure 3.4)[13,23].

In my model, the total amount of HK is assumed to be conserved between all of the states, an assumption shared by previous models investigating TCS [56]. The total amount of PhoB, however, is also assumed to be conserved. However, this is not the case in the wild-type system, where PhoB is under a positive feedback loop, promoting its own expression in the absence of phosphate, and has not been assumed in other models [56,70]. PhoB is expressed at a constitutive rate in the synthetic signaling system via P_{lacI} as part of the RR/HK operon.

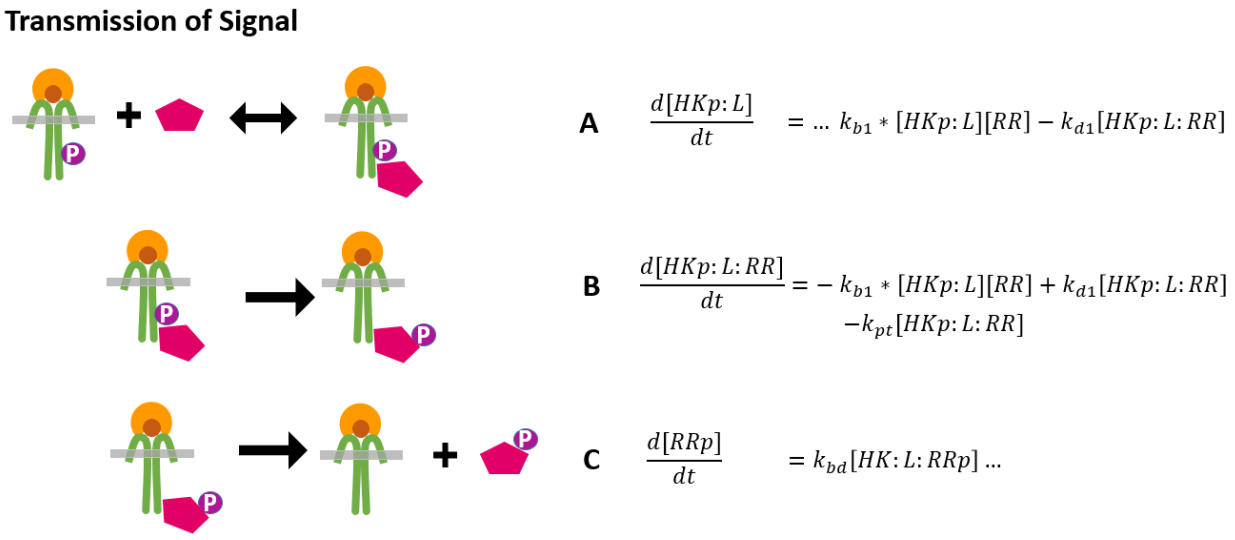


Figure 3.3. Graphic and mathematical representation of transmission of phospho-signal from HK to RR. *A) Reversible binding of phosphorylated HKp:L to RR. This is mathematically represented on the right hand side of the figure using Law of Mass Action kinetics. The equation can be read as follows: The rate of change of the concentration of phosphorylated HK bound to the PBP-ligand complex ($\frac{d[HKp:L]}{dt}$) can be described in part by the binding rate of RR to the phosphorylated HK bound to the PBP-ligand complex (k_{b1}) multiplied by the concentration of phosphorylated HK bound to the PBP-ligand complex ($[HKp:L]$) and the concentration of unphosphorylated RR ($[RR]$) minus the rate of dissociation of the RR bound, phosphorylated HK bound to the PBP-ligand complex (k_{d1}) multiplied by the concentration of RR bound, phosphorylated HK bound to the PBP-ligand complex ($[HKp:L:RR]$). Because the phosphorylated HK bound to the PBP-ligand complex is involved in an autophosphorylation event, and an RR binding event, the concentration of phosphorylated HK bound to the PBP-ligand complex is dependent on additional terms described in Figure 3.2 (hence the ellipsis). B) Irreversible phospho-transfer from HK to RR in the HKp:L:RR complex. This is mathematically represented on the right hand side of the figure using Law of Mass Action kinetics. The equation can be read as follows: The rate of change in the concentration of the RR bound, phosphorylated HK bound to PBP-ligand complex, decreases at the the binding rate of RR to the phosphorylated HK bound to the PBP-ligand complex (k_{b1}) multiplied by the concentration of phosphorylation of HK bound to the PBP-ligand complex ($[HKp:L]$) and the concentration of unphosphorylated RR ($[RR]$), increases at the rate of dissociation of the RR bound, phosphorylated HK bound to the PBP-ligand complex (k_{d1}) multiplied by the concentration of RR bound, phosphorylated HK bound to the PBP-ligand complex ($[HKp:L:RR]$), and decreases at the rate of autophosphorylation (k_{pt}) multiplied by the concentration of RR bound, phosphorylated HK bound to the PBP-ligand complex ($[HKp:L:RR]$). C) Irreversible unbinding of RRp from HK:L. This is mathematically represented on the right hand side of the figure using Law of Mass Action kinetics. The equation can be read as follows: The rate of change in the concentration of phosphorylated RR ([RRp]) is described in part by the dissociation rate of the RR bound, phosphorylated HK bound to the PBP-ligand complex (k_{bd}) multiplied by the concentration of RR bound, phosphorylated HK bound to the PBP-ligand*

complex ($[HK:Lp:RR]$). Because RRp is involved in a number of downstream reactions, the rate is dependent on additional terms not shown here.

Due to the increasing complexity of the equations, downstream reactions will only be graphically represented. The full list of ODEs comprising the model can be found in Table 3.3

There is one RR/HK operon per plasmid, and I again assumed an average copy number of 15 [84]. RR is predicted to be expressed by P_{lacI} at 2/3 the rate of LacI (calculated per Salis RBS calculator [85]). I again assumed the same degradation rate for LacI and PhoB, with a final, calculated concentration of 495 molecules/cell, or approximately 2 uM of PhoB. The HK however, is translated after a hairpin intergenic region in the operon mRNA, and is likely expressed at lower levels than the RR. Quantification of PhoB and PhoR proteins in wild-type cells indicated a 10:1 ratio of RR to HK[70], a ratio we have maintained in this model.

An additional assumption is made regarding the affinity of the various forms of RR for the HK in the system. Experimental work on the EnvZ/OmpR TCS by Inouye et al determined the affinity of all combinations of phosphorylated and unphosphorylated OmpR for the phosphorylated and unphosphorylated cytoplasmic portions of EnvZ[81]. These experiments found that phosphorylated cytoplasmic EnvZ was more than twice as likely to bind OmpR, whether the response regulator was phosphorylated or not[81]. While this work was accomplished in vitro, and in a different TCS, it serves as a starting point for the binding constants associated with the synthetic signaling system.

Phosphotransfer in two component systems is exceedingly rapid and is therefore difficult to measure effectively. Recently, Stock et al probed the kinase and phosphatase activities of TCS using a 40 minute timecourse to experimentally determine the dephosphorylation rate of PhoB by PhoR. At 20 minutes, approximately 75% of the 4 μM PhoB-P used in the *in vitro* assay was

dephosphorylated. In the same work, a phosphorylation assay showed that 2 µM of PhoB was 95% phosphorylated by 3µM PhoR in under 20 seconds[70]. While it is impossible to calculate the true phosphorylation rate from the data in this paper (there are no data points under 20 seconds for the phosphorylation assay), it is apparent that phosphorylation is a much more rapid process than dephosphorylation. To that end, the phosphorylation rate of RR by HK in this model is 10 fold greater than the dephosphorylation rate.

3.3 Ligand-dependent signaling versus ligand-independent signaling.

These three events, extracellular ligand-binding, transduction of the signal across the membrane, and transmission of phospho-signal to RR comprise the ligand-dependent signaling pathway. A graphic representation of the model of ligand-dependent signaling is shown in Figure 3.5 and the explicit form of the equations is given in Table 3.3

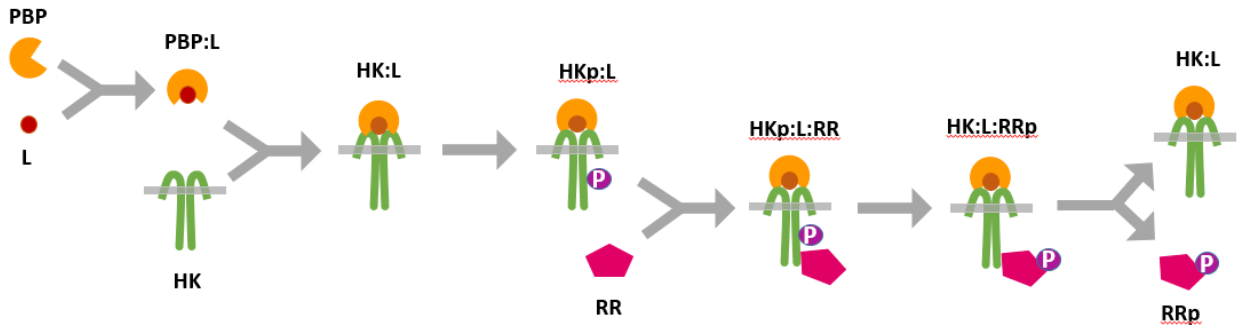


Figure 3.4 Ligand-Dependent Signaling model*. From left to right: Ligand (brown) binds to free PBP(yellow) to form the ligand bound complex PBP:L (yellow). The HK (green) is bound by L, to form HK:L. HK:L autophosphorylates (HKp:L), then binds free response regulator (pink) to form HKp:L:RR. That complex then transfers a phosphate from the HK to the RR (HK:L:RRp), at which time the complex dissociates into phosphorylated RR (RRp, magenta), and HK:L.*

In addition to ligand dependent signaling, it is possible for the HK to autophosphorylate in the absence of ligand. The model assumes that subsequent to ligand-independent

autophosphorylation of HK72, phosphorylation of the RR and initiation of transcriptional response will occur as they would with ligand-dependent signaling (Figure 3.5).

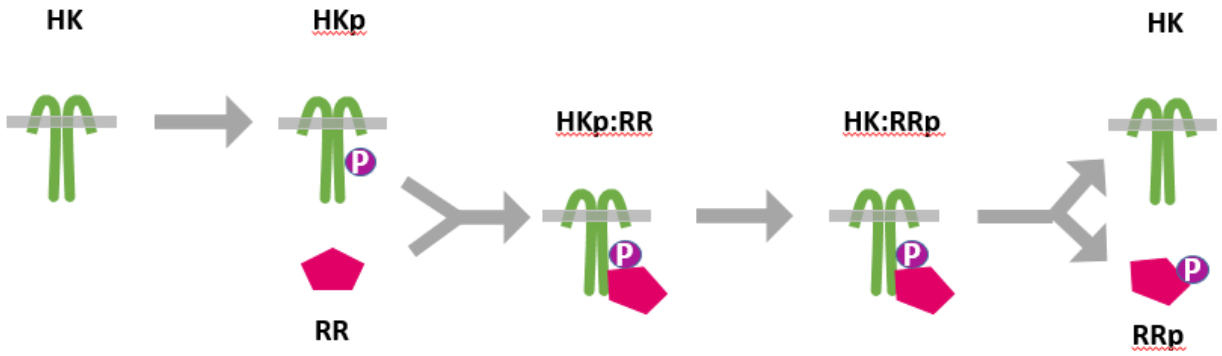


Figure 3.5. Ligand-Independent Signaling model.

Figure 3.5. Ligand-Independent Signaling model. *From left to right: Free HK autophosphorylates (HKp), then binds free response regulator (pink) forming HKp:RR. HKp:RR transfers a phosphate from the HK to the RR (HK:RRp), at which time the complex dissociates into phosphorylated RR (RRp, magenta), and HK (green).*

Ligand-independent signaling may also result from phosphorylation of the RR by wild-type *E. coli* proteins[93]. However, phosphorylation from these sources should effectively be constant if the cells are always tested in the same abiotic context (media, temperature, and carbon source) and genetic background. The end result of this presumably constant RR phosphorylation is more RRp, so this model incorporates additional sources of ligand-independent phosphorylation as an increase in the HK ligand-independent autophosphorylation rate.

3.4 Inducing a transcriptional response

The final step of the synthetic signaling system is transcriptional activation by phospho-PhoB (RRp). The production of RRp via HK phosphorylation is countered by the dephosphorylation of RRp through two mechanisms. First, PhoB-P autodephosphorylates via hydrolysis (Figure 3.6

A), a rate experimentally determined by Stock et al [70]. Second, RRp may be dephosphorylated by the HK[70]. These values, like many of the experimentally determined parameters in this model, were identified through *in vitro* experiments. It is important to keep in mind that *in vitro* experiments may not always accurately reflect *in vivo* activity of these proteins, and that these parameter values are only a starting point for the model.

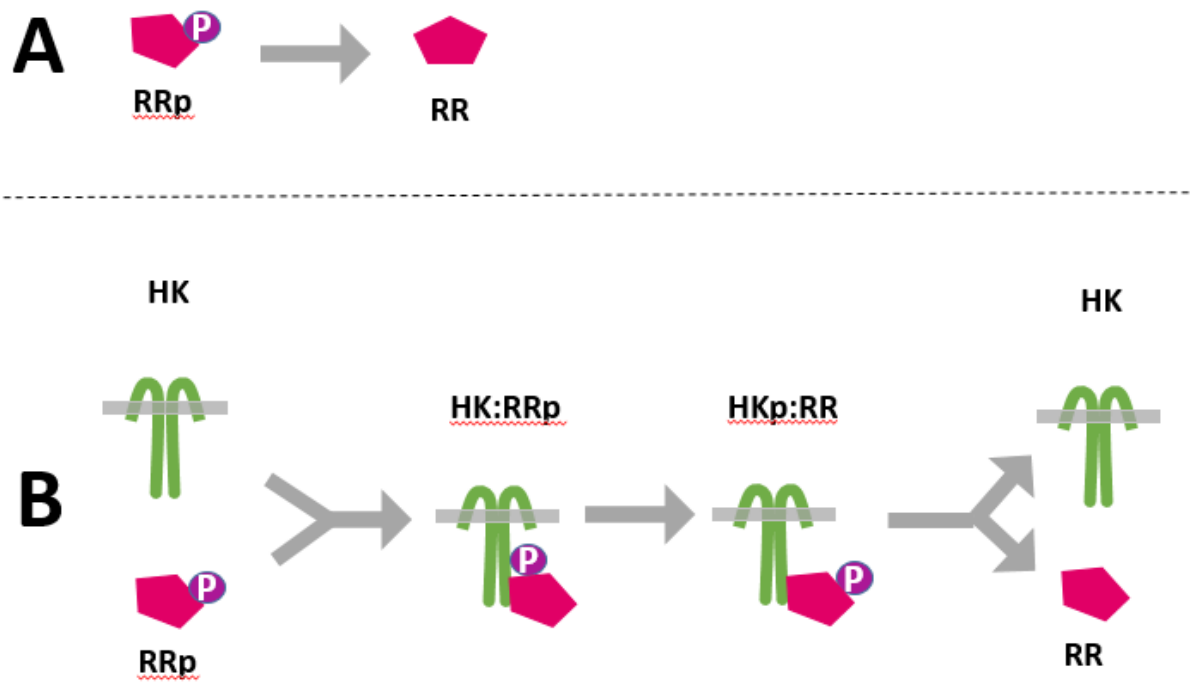


Figure 3.6. Dephosphorylation of the response regulator. *A) Phosphorylated RR (RRp, magenta) autodephosphorylates (RR). B) RRp (magenta) binds to HK (green) forming HK:RRp. The HK dephosphorylates RRp (HK:RR) and the complex dissociates into RR, and HK.*

Dephosphorylation of the RR by a bifunctional HK has been investigated in several models, and has been shown to have distinctly different behavior from a system where RR dephosphorylation is primarily independent of the TCS[56,57]. There is broad experimental support for the role of HK dephosphorylation of the RR in the regulation of wild-type TCS [55], and we assume in this model

that HK72 retains the phosphatase activity of PhoR as well as the kinase activity. The rate of dephosphorylation of PhoB by the wild-type PhoR has been experimentally determined [70], however the individual binding affinities of PhoB for PhoR are still obscure. Published work on the EnvZ/OmpR TCS suggests that there is no difference in affinity for the HK between the phosphorylated and unphosphorylated forms of the response regulator [81], however the unphosphorylated form of the HK is half as likely to bind OmpR. I have extended that observation to the PhoR/PhoB system used in the synthetic signaling system and assumed that free RRp is half as likely to bind to HK as RR is to bind to HKp, and used these experimentally determined binding affinities as a starting point for defining PhoR/PhoB binding and dissociation.

Phospho-PhoB acts as a dimer in the bacterial cell. Unphosphorylated PhoB has a dimerization dissociation constant of 378 μM (*i.e.*, at 378 μM , one half of the PhoB present in the cell will be incorporated into dimers)[66]. The intracellular concentration of PhoB in our synthetic signaling system is predicted to be approximately 2 μM, making dimers of unphosphorylated PhoB highly unlikely. Therefore, this reaction is omitted from the model and dimerization only occurs upon RR phosphorylation (Figure 3.7).

In wild-type *E. coli* PhoB-P dimers bind to the Pho-inducible promoters at binding sites known as PhoBoxes. Each PhoBox is 18 base pairs long, consisting of two 7 base pair binding sites (one for each half of the dimer), separated by a 4 base pair spacer. In the synthetic signaling system, reporter expression is controlled by the wild-type, PhoB-inducible PstS promoter. While most PhoB-inducible promoters have one PhoBox, the PstS promoter has two. The PstS promoter is a useful candidate for modeling because it has been well studied. In addition, it is tightly regulated, has strong activity, and is the PhoB inducible promoter with the highest affinity for PhoB-P [71].

The model assumes that only PhoB-P dimers (Figure 3.7, RRpD) bind to the PstS promoter and induce transcription. This assumption is supported by published work demonstrating that the affinity of PhoB-P dimers for the PstS promoter is an order of magnitude greater than the affinity of unphosphorylated PhoB[94], as well as experiments in this work demonstrating very low reporter expression in the absence of PhoB. However, the unbound PstS promoter will produce some level of mRNA. Although PstS is tightly regulated, there a basal level of expression in biological systems regardless of induction. Experimental work has shown that promoter activity increases 100 fold in the presence of PhoB-P[71], therefore in this model the basal rate of transcription has been set to 1/100th of the transcription rate of a PstS promoter bound with two PhoB-P dimers.

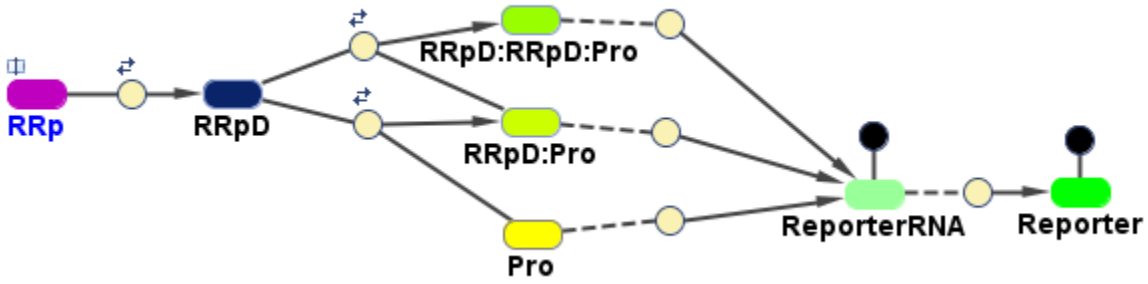


Figure 3.7. PhoB inducible expression of reporter modeled. *This is a sample of SimBiology's interface, which 'draws' nodes and edges. Phosphorylated RR (RRp, magenta) dimerizes to form RRpD. RRpD can then bind to the PstS promoter (Pro, yellow) at one of two locations. For visual simplicity, both locations are represented by RRpD:Pro in this diagram. However, the binding affinities of the two sites are different, and binding to the upstream site produces almost no reporter RNA, while binding to the downstream site produces reporter RNA at a near maximal rate. All promoter configurations result in some level of reporter RNA production. Reporter RNA is translated into Reporter (green). The beige circles represent a reaction between the two proteins, and the black circles represent a degradation reaction.*

The affinity of the PhoB-P dimer for the upstream binding site is four times that of the downstream site[71]. However, the PhoB binding to the upstream PhoBox results in only minimal

transcription, while binding to the downstream PhoBox results in near-maximal transcription[72]. This suggests that PhoB binds to the PstS PhoBoxes cooperatively – that is, the difference in promoter activity between zero bound PhoB-P dimers and one bound PhoB-P dimer is not the same as the change in activity between one and two bound PhoB-P dimers. The binding affinities and rates of transcription of each promoter state reflect these experimental results (Table 3.2).

Translation rates are not unique to the synthetic signaling system, so I applied the translation rates used in previous kinetic models to this one as well[69]. In addition, I made use of general mRNA and protein degradation rates to describe the degradation of 'reporter RNA' and 'reporter' protein. These rates are reasonable for biological systems, with reporter mRNA degrading 10 times faster than the protein itself [69]. In later experimental work, the reporter is green fluorescent protein (GFP). GFP maturation is incorporated into the translation rate in this model.

3.5 Technical considerations

The model was built with MATLAB's SimBiology plug-in, a graphic user interface which allows you to 'draw' the connections between system components (see Figure 3.7) and apply the appropriate description of the biochemical reaction (*i.e.,* PBP binds to Ligand in accordance with the Law of Mass Action). Simbiology then solves the derived system of equations numerically. Simulations were run until steady state.

3.6 Initial conditions and literature values

Initial concentrations of HK, RR, and PBP and all species are set as described in Table 3.1

Table 3.1. Initial amounts of species involved in reactions.

Species	*Concentration*	*Units*
HK	*0.2*	*µM*
HKp	*0*	*µM*
L	*0*	*µM*
RR	*2*	*µM*
RRp	*0*	*µM*
HK:L	*0*	*µM*
HKp:L	*0*	*µM*
HKp:RR	*0*	*µM*
HK:RRp	*0*	*µM*
RR:HKp:L	*0*	*µM*
RRp:HK:L	*0*	*µM*
PBP	*3.5*	*µM*
Ligand	*0.1*	*µM* *
Pro	*0.059*	*µM*
RRpD:Pro	*0*	*µM*
RRpD:RRpD:Pro	*0*	*µM*
RRpD	*0*	*µM*
Reporter	*0*	*µM*
ReporterRNA	*0*	*µM*
HK:RR	*0*	*µM*

** Ligand concentrations are varied from 13 nM to 13 mM in accordance with the in silico experimental design.*

As described above, the species associate with one another in defined ways: PBP binds to ligand, phosphorylated HK binds to unphosphorylated PhoB, etc. While the overall rate of formation of a particular product (PBP-Ligand complex (L), or phospho-PhoB (RRp)) may change with increased concentration of certain species, the affinity of PBP for ligand, and the rate at which phosphates are transferred from the HK to PhoB does not change in this model. These rate constants are derived from literature values and described fully in Table 3.2.

Table 3.2

Role	*Reaction*	*Parameter*	*Value*	*Units*		*Ref*
BINDING OF LIGAND TO PBP						
PBP dissociates from ligand	***PBP + Ligand <-> L***	*kpd*	*0.588*	*1/s*	*	61
PBP binds to ligand		*kpb*	*1*	*1/(uM*s)*	*	61
BINDING OF PBP-LIGAND COMPLEX TO HK						
PBP-Ligand complex binds HK	***HK + L <-> [HK:L]***	*klb*	*1*	*1/(uM*s)*	*	61
PBP-Ligand complex unbinds HK		*klu*	*417*	*1/(uM*s)*	*	61
LIGAND DEPENDENT SIGNAL TRANSDUCTION						
HK:L Autophosphorylation	***[HK:L] <-> [HKp:L]***	*klp*	*10*	*1/s*	**	
HK:L Autodephosphorylation		*kld*	*1*	*1/s*	**	
HKp:L binds to RR	***[HKp:L] + RR <-> [RR:HKp:L]***	*kb1*	*1*	*1/s*	**	81
HKp:L dissociates from RR		*kd2*	*0.78*	*1/s*	*◊	81
HKp:L transfers phosphate to RR	***[RR:HKp:L] -> [RRp:HK:L]***	*kpt*	*0.87*	*1/s*	**	
HKp:L dissociates from RRp	***[RRp:HK:L] -> [HK:L] + RRp***	*kd1*	*0.78*	*1/s*	*◊	81
LIGAND INDEPENDENT SIGNAL TRANSDUCTION						
HK Autophosphorylation	***HK <-> HKp***	*kap*	*1*	*1/s*	**	
HK Autodephosphorylation		*kad*	*100*	*1/s*	**	
HKp binds RR	***HKp + RR <-> [HKp:RR]***	*kbla*	*1*	*1/(uM*s)*	*◊	81
HKp dissociates from RR		*kd2a*	*0.78*	*1/s*	*◊	81
HKp transfers phosphate to RR	***[HKp:RR] -> [HK:RRp]***	*kpt2*	*0.87*	*1/s*	**	70
HKp dissociates from RR	***RRp + HK <-> [HK:RRp]***	*kd1a*	*0.78*	*1/s*	*◊	81
RRp DEPHOSPHOSPHORYLATION AND NONPRODUCTIVE BINDING						
RRp autodephosphorylates	***RRp -> RR***	*kdp*	*0.00024*	*1/s*		70
RRp binds to HK		*kbb*	*1*	*1/(uM*s)*	*	81

RRp DEPHOSPHOSPHORYLATION AND NONPRODUCTIVE BINDING						
HK dephosphorylates RRp	***[HK:RRp] -> [HK:RR]***	*kdp1*	*0.0087*	*1/s*		70
RR dissociates from HK		*kbd*	*1.96*	*1/s*	*	81
RR binds to HK	***[HK:RR] <-> HK + RR***	*kbb*	*1*	*1/(uM*s)*	*	81
TRANSCRIPTIONAL ACTIVATION						
RRp dimerization	***RRp + RRp <-> RRpD***	*kdim*	*1*	*1/(uM*s)*	*	66
RRp dimer dissociates		*kdis*	*5.1*	*1/s*		66
RRpD binds to upstream PhoBox	***RRpD + Pro <-> [RRpD:Pro]***	*ktfb*	*1*	*1/(uM*s)*	*	95
RRpD unbinds upstream PhoBox		*ktfd*	*7*	*1/s*	*	95
RRpD binds downstream PhoBox	***[RRpD:Pro] + RRpD <-> [RRpD:RRpD:Pro]***	*ktfb2*	*1*	*1/(uM*s)*	**	95
RRpD unbinds downstream PhoBox		*ktfd2*	*3.4*	*1/s*	**	95
TRANSCRIPTION AND TRANSLATION OF REPORTER						
Downstream PhoBox Transcription	***[RRpD:RRpD:Pro] -> [RRpD:RRpD:Pro] + ReporterRNA***	*ktr*	*0.00107*	*uM/s*	**	72
Upstream PhoBox transcription	***[RRpD:Pro] -> [RRpD:Pro] + ReporterRNA***	*ktr1*	*0.000054*	*uM/s*	**	72
Basal Transcription Rate	***Pro -> Pro + ReporterRNA***	*ktr0*	*1.1E-05*	*uM/s*		72
Degradation of reporter mRNA	***ReporterRNA -> null***	*kz1a*	*0.0077*	*1/s*		61
Translation of Reporter	***ReporterRNA -> Reporter + ReporterRNA***	*ktl1*	*0.228*	*1/s*		61
Degradation of Reporter	***Reporter -> null***	*kz3*	*0.0012*	*1/s*		61

** Derived from dissociation constant*

*** Assumed*

◊ Derived from experimental data on EnvZ/OmpR TCS

3.7 Initial test of the model

My model was simulated at the above conditions and generated a response profile for the synthetic signaling system at various concentrations of ligand. Figure 3.8 shows the predicted reporter concentrations generated by simulating the synthetic signaling system from 0mM to 13mM ribose. The model was simulated until it reached steady state, at which time the predicted values of reporter concentration were plotted against the ligand concentration used in the model. The experimentally determined parameters used in this model predict an OFF and an ON state for the system, a prediction corroborated by data showing that HK72 signals in a ligand-dependent manner (Morey et al, *in prep*).

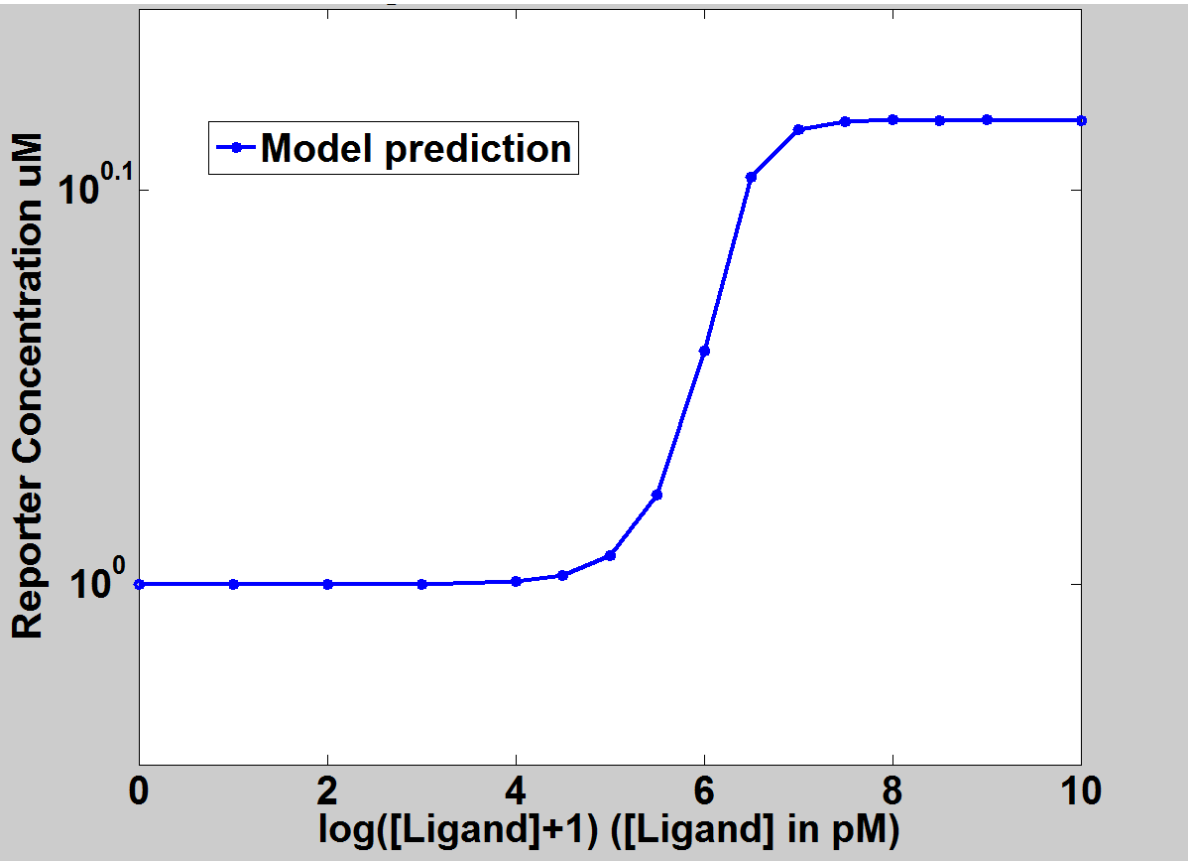


Figure 3.8. Modeled synthetic signaling system response using experimentally published parameter values. *The x-axis is log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. The ON state (presence of ligand) is predicted to generate 1.3 times as much reporter as the OFF state (absence of ligand).*

CHAPTER 4: CHARACTERIZATION OF SYNTHETIC SIGNALING SYSTEM AND PARTS USED IN OPTIMIZATION

Just as the baseline values for the model were researched thoroughly, the baseline activity of the synthetic signaling system must be well characterized. In this section I detail the plasmids used in subsequent experiments and the cloning strategies used to generate them. The response of the original synthetic signaling system, as well as the determination of optimal growth conditions is characterized in this chapter.

4.1 Construction of synthetic signaling system and variants.

The synthetic signaling system was divided into two plasmids, a signaling plasmid containing the PBP, HK, and RR, and a reporter plasmid. This allowed for flexible and combinatorial testing as necessary. Cloning was accomplished using Qiagen miniprep kits (Qiagen, Valencia, CA) to purify plasmid DNA from cultures, Zymoclean gel purification kits (Zymo Research, Irvine, CA) to extract DNA fragments from agrose gels, Herculase II DNA polymerase (Agilent, Santa Clara, CA) to generate new DNA sequences via PCR, GoTaq Green Master Mix (Promega, Madison WI) for PCR screening of putative colonies, and restriction enzymes from the FastDigest catalog (Thermo Scientific, Waltham, MA). Ligations were accomplished with T4 DNA ligase (New England Biolabs, Ipswich, MA), blunting and dephosphorylating as necessary using Clone JET PCR Cloning Kit (Thermo Scientific, Waltham, MA) and Antarctic Phosphatase (New England Biolabs, Ipswich, MA).

4.1.1 Signaling Plasmid

The signaling plasmid contains a PBP (yellow), a response regulator (magenta), the fusion HK (green), and a selectable marker gene (orange) in the pACYC177 vector, a low copy expression

plasmid (grey). Figure 4.1 depicts the standard signaling system, with particular PBP, wild-type RBP, a particular RR, PhoB, and particular fusion HK, HK72, in addition to ampicillin resistance. PhoB and the fusion HK are arranged in an operon, with the wild-type intergenic sequence found in the PhoB/PhoR operon inserted between PhoB and HK72. This intergenic sequence forms a hairpin, reducing the amount of HK72 translated per operon mRNA. The synthetic signaling system plasmid depicted here is KJM114, and was developed prior to this work.

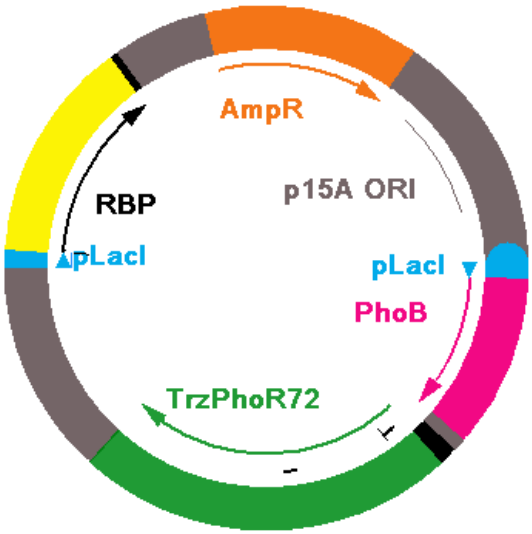


Figure 4.1. Vector map of the original synthetic signaling system plasmid. *RBP and the PhoB/PhoR operon are both controlled by the weak constitutive promoter* P_{lacI}*, in the pACYC177 vector. All signaling system plasmids were built off of this backbone.*

Expression of both the PhoB operon and the RBP open reading frame is controlled are driven by the LacI promoter, a well characterized, weak constitutive promoter. Promoters and components were modified as required for experiments using restriction sites including but not limited to Fermentas Fast Digest NheI, XhoI, NdeI, HindIII, BsteII, AfeI, NcoI, EcoRI and PsiI. Dettailed cloning strategies can be found in the appendix of this thesis.

Table 4.1 Signaling plasmids used in this work

Construct Name	Forward Primer	Reverse Primer	Components	Description
Original synthetic signaling system construct				
KJM114	n/a	n/a	P_{lacI}::RBP, P_{lacI}::PhoB::HK72	'standard' or '1x' RBP expression
Removing extra genetic material				
KLH545	31	60	P_{lacI}::RBP, P_{lacI}::PhoB::HK72	FLAG tag 3' of PBP removed
KLH551	63	64	P_{lacI}::RBP, P_{lacI}::PhoB::HK72	FLS tag 5' of HK removed
KLH581	53	89	P_{lacI}::RBP, P_{lacI}::PhoB::HK72	PhoB N227T point mutation repaired
RBP signaling systems with various levels of RBP				
KJM120	*	*	KJM114 w/ P_{LacIQ1}::RBP	Putative 50-fold increase in RBP expression
KLH532	43	37	KJM114 w/ P_{LacIQ1}:: 10xRBS::RBP	Putative 500-fold increase in RBP expression
Partial synthetic signaling systems				
KLH546	54	33	P_{lacI}::RBP, P_{lacI}::PhoB	Contains only PhoB and RBP
KLH524	n/a	n/a	P_{lacI}::PhoB::HK72	Contains only PhoR and PhoB
KLH575	53	84	P_{lacI}::RBP, P_{lacI}:: HK72	Contains only PhoR and RBP
GBP synthetic signaling system				

KJM109	54	33	P_{lacI}::GBP, P_{lacI}::PhoB::HK72	'1x' GBP expression
NAB251	*	*	P_{lacI}::GBP, P_{lacI}::PhoB::HK72	Putative 50-fold increase in GBP expression
Synthetic signaling systems with additional features				
KLH614	n/a	n/a	P_{lacI}::GBP, P_{lacI}::PhoB::HK72, 21 PhoB binding sites	RBP signaling system with additional PhoB binding sites
KLH592	95,53	94,94	P_{lacI}::GBP, P_{lacI}::PhoB::HK72, P_{lacI}::EGA	RBP signaling system with additional dephosphorylation pathway
Synthetic Signaling System for use in plants				
KLH803	70	71	P_{FMV}::RBP $P_{CaMV35S}$::PhoB::VP64 P_{NOS}::HK72 $P_{CaMV35S}$::Luciferase	Plant synthetic signaling system with strong constitutive FMV promoter controlling RBP expression
KLH804	70	71	P_{NML1}::RBP $P_{CaMV35S}$::PhoB::VP64 P_{NOS}::HK72 $P_{CaMV35S}$::Luciferase	Plant synthetic signaling system with novel Medford Lab plant promoter (P_{NML1}) controlling RBP expression
KLH805	70	71	P_{NML1}::RBP $P_{CaMV35S}$::PhoB::VP64 P_{NOS}::HK72 $P_{CaMV35S}$::Luciferase	Plant synthetic signaling system with novel Medford Lab plant promoter (P_{NML2}) controlling RBP expression

** indicates plasmid constructed by another lab member, NAB251 constructed by Nikolai Braun, PhD, KJM120 and KJM114 constructed by Kevin Morey, PhD, KLH804, 805 constructed by Mauricio Antunes, PhD; 'n/a' indicates plasmid was constructed without PCR amplification, using either a commercially synthesized DNA fragment or a previously constructed plasmid*

4.1.2 Reporter Plasmid

The reporter plasmid (TAA002, previously generated) contains a reporter gene (green) under the control of a PhoB-P responsive promoter (blue) as well as a resistance gene (orange) in the pBR322 vector (light grey), a low copy expression plasmid. The reporter plasmid shown in Figure 4.2 has the PstS promoter controlling the expression of GFP. The pMB1 (pBR322) and

p15 ORIs are compatible and can be co-expressed in the cell (pACYC177)[84]. The pACYC177 plasmid is maintained at approximately 15 copies per cell, and the pBR322 vector is maintained at approximately 20 copies per cell [84]. Promoter and resistance genes were modified as required using restriction sites including but not limited to, Fermentas Fast Digest AatII, EcoRI, NcoI, SacI, and isoschizomers of SacI.

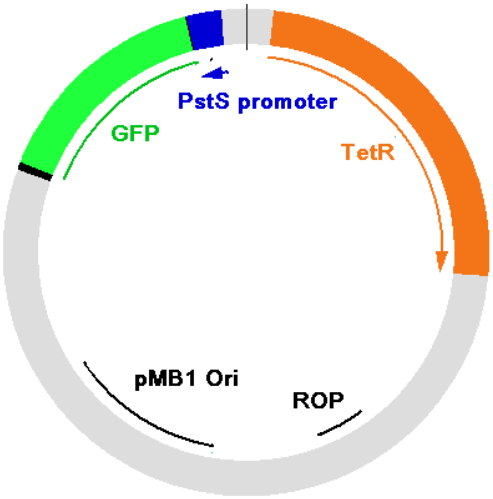


Figure 4.2. Vector map of the reporter plasmid. *GFP is controlled by the PhoB responsive PstS promoter in the pBR322 vector. All reporter plasmids were built off of this backbone. For a full list of changes made to the vector for each reporter plasmid used in this work, see Table 4.2*

Table 4.2 Reporter plasmids used in this thesis

Construct Name	Forward Primer	Reverse Primer	Components	Description
TAA002	n/a	n/a	PstS::GFP,	PhoB inducible GFP
KLH557	57	58	TAA002 + additional PstS:GFP cassettes	Reporter construct with two PstS:GFP cassettes
KLH213	62	61	PstS.Consensus::GFP	Synthesized PstS promoter with two consensus PhoBoxes, PCR to replace reporter gene.
KLH214	62	61	PstS.PhoH::GFP	Synthesized PstS promoter with two consensus PhoBoxes, PCR to replace reporter gene.
KLH215	62	61	PstS.PhoA::GFP	Synthesized PstS promoter with two consensus PhoBoxes, PCR to replace reporter gene.

KLH216	62	61	PstS.ugpB::GFP	Synthesized PstS promoter with two consensus PhoBoxes, PCR to replace reporter gene.
KLH217	62	61	PstS.PhoB::GFP	Synthesized PstS promoter with two consensus PhoBoxes, PCR to replace reporter gene.
KLH218	62	61	PstS.PhoE::GFP	Synthesized PstS promoter with two consensus PhoBoxes, PCR to replace reporter gene.

4.2 Cell lines used

All experiments use bacterial cell lines derived from BW23423, which contains a deletion in the PhoBR operon, Δ*PhoBR580* [37].This 2kbp deletion eliminates endogenous expression of PhoB and PhoR, creating a cell chassis that will not interfere with testing of the synthetic signaling system. In addition, BW23423 has mutations that eliminate other known sources of spurious PhoB phosphorylation, including CreC (a histidine kinase known to interact with PhoB [65]) via Δ*CreBCD153*, and acetyl phosphate (a small molecule shown to autophosphorylate PhoB [94]), by knocking out genes responsible for acetyl phosphate production (*i.e.*, Δ(*pta ackA*)$_{TA3516}$). Also, histidine transport complexes and supplementary *pho* regulon deletions are found in this line (Δ(*hisQ hisP),* Δ*phn).* The BW line also contains two fusions to inducible promoters: *araBAD*$_{AH37}$*::* P_{araB}-NFLAG-*'phoR*$_{AH41,}$ an L-arabinose inducible cassette controlling expression of the cytoplasmic portion of PhoR, and P_{phnC}-*lacZ*$_{WJ19}$, PhoB controlled expression of *β*-D-galactosidase inserted at the *lac* locus. BW23423 is derived from BW13711 (fully described in Haldimann, 1996 [37]), and further details on this strain can be found at the Coli Genetic Stock Center.

From BW23423, two additional cell lines were derived using the lambda red protocol[96]. Cell line LT002, developed by Lindsey Triplett, PhD, is a knockout of endogenous RBP expression in a BW23423 background. It was developed to ensure accurate measurement of the impact RBP concentration has on the synthetic signaling system. Wild-type RBP expression is known to increase with ribose concentration[97], and therefore cannot be considered a constant in these experiments. Similarly, GBP concentrations also increase with galactose concentration [97]. Therefore a strain of GBP knockouts (MJB015) was also established in a BW23423 background by Matt Barrow, PhD.

4.3 Experimental Protocol

In order to optimize the system, I wanted to experimentally validate predictions regarding increases in maximum signal over fold change, decreases in threshold of response (*i.e.*, response at a lower level of ligand), and increase the 'switch-like' behavior of the circuit (*i.e.*, reducing the change in ligand concentration needed to effect a change from the OFF state to the ON state). To accomplish these goals, it was necessary to understand the response of the synthetic signaling system across a gradient of ligand concentrations.

4.3.1 Experimental design

Two days prior to each experiment, cells were co-transformed with the appropriate synthetic signaling system and reporter plasmids and grown on Luria Broth (LB) plates with appropriate antibiotics (50 µg/ml carbenicillin (Teknova, Hollister, CA) and 50 µg/ml tetracycline (Alfa Aesar, Ward Hill, MA)) as required, overnight at 30° C. Individual colonies were randomly selected for testing. Colonies were inoculated in 10 mL M9CA+ media (Amresco, Solon, OH) and grown overnight for approximately 16 hours. Cultures were then diluted an OD600 of less

than 0.01 (20 to 50 µL culture in 12 ml fresh M9CA+). One ml of diluted culture was aliquoted into 16 ml borosilicate culture tubes.

Ribose serial dilutions were prepared from a 4M Ribose stock and 3.25 uL of serial dilution were added to each 1mL aliquot to establish ligand concentrations from 13 mM to 13nm (See Appendix A.4 for details). Galactose and glucose serial dilutions were prepared from a 2M galactose or glucose stock, respectively, and 6.25 uL of serial dilution were added to 1ml aliquot to establish ligand concentrations from 13 mM to 13 nM. Cultures were incubated with ligand from three to six hours and measured using a Avalon flow cytometer (Propel Labs, Fort Collins, CO) with excitation wavelength 488 nm and emission filter 525/25.

4.3.2 Quantification of Response

Flow cytometry is quantitative and accurate, and is therefore used by numerous synthetic biology laboratories(*e.g.*, Smolke et al[98], Arkin et al[11], Weiss et al[99], Voigt et al[85], and many more). In brief, flow cytometry passes a stream of uniformly sized and spaced droplets through a laser. Each droplet is illuminated individually as it passes through the laser, and the resulting fluorescence is recorded[100]. The machine is calibrated to expose only one cell or event to the laser per droplet and so has the added benefit of tracking individual cells within a population. In addition, flow cytometry requires only small volumes of cell culture, allowing a single culture to be tracked over time with multiple reads. Flow cytometry has been used with great success in characterizing functional and nonfunctional circuits (*e.g.*,Nevozhay [43], Hooshangi et al [99]), and has become established as a high throughput method of validating computationally designed genetic components (*e.g.*, Liang et al[98], Salis et al[85]).

In my experiments, each datapoint is the average of at least three independent experiments on three individual colonies, generated from three separate transformation events, on three different

days. Each colony was measured at each ligand concentration (from 13 mM to 13 nM, and in the absence of ligand). To establish the fold change over background, the mean of 50,000 events was normalized to background levels of fluorescence in the absence of ribose. Raw fluorescence values are shown in arbitrary units.

4.3.3 Response of the original synthetic signaling system

When co-transformed with a synthetic signaling plasmid, the fluorescence signal at a given concentration of ribose is representative of the strength of the signal transmitted by the synthetic signaling system being tested. The original signaling system was tested under the described experimental conditions, and the amount of fluorescence was measured by flow cytometry to establish a benchmark for a functional system. Figure 4.3 shows the response of the original signaling system, KJM114, over a ribose concentration gradient ranging from 0 to 13 mM.

After sequencing the original synthetic signaling system plasmid, I discovered that the plasmid contained a Flag Tag 3' of the RBP sequence, and a membrane-localization tag 5' of the histidine kinase. In addition PhoB was not wild-type, a point mutation had altered amino acid 227 from threonine to asparagine. The tags were removed, and the PhoB point mutation was reverted using polymerase chain reation (PCR) and appropriate primers to modify the regions of interest (see Appendix A.1, Strategy I for details). All changes were sequenced verified, and I retested the signaling system and compared the new synthetic signaling system variants against the tagged synthetic signaling system. I found that none of these changes had a significant impact on system response (Figure 4.4).

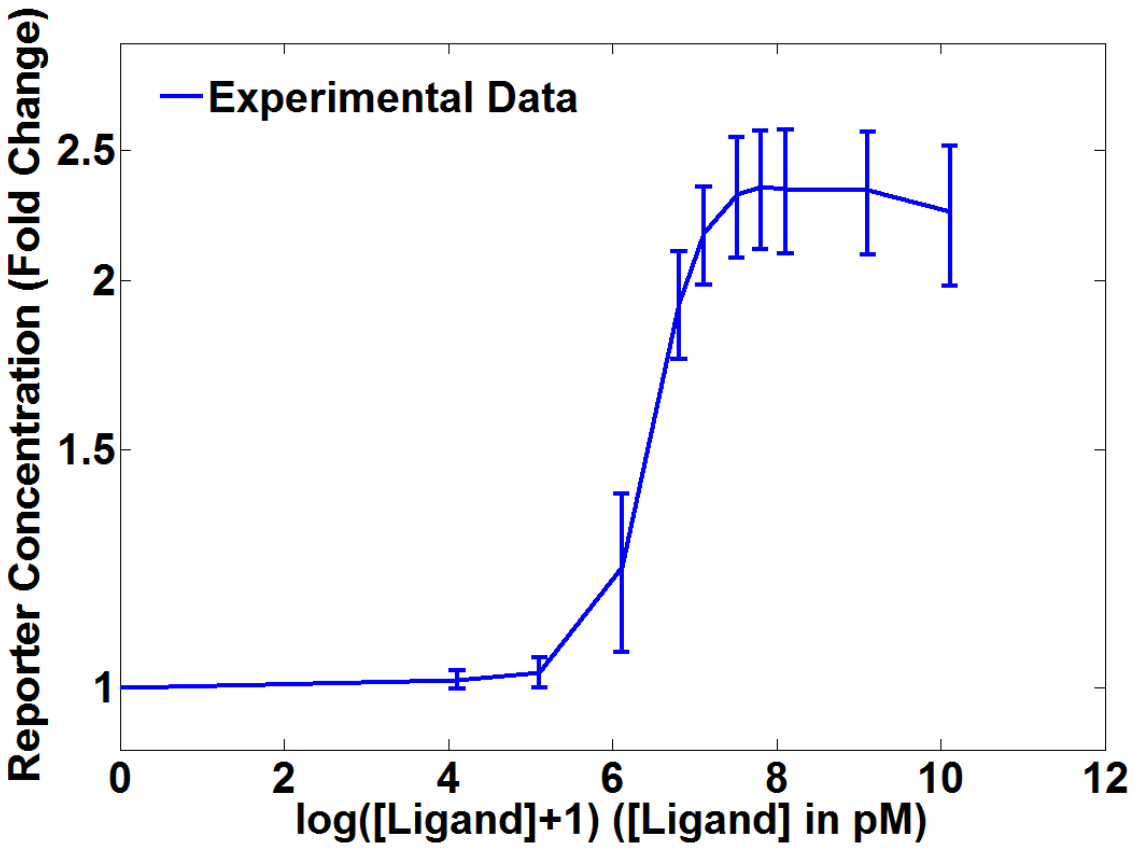


Figure 4.3. Response of original RBP synthetic signaling system. *LT002 cells were cotransformed with KJM114, (*P_{lacI}*::RBP,* P_{lacI}*::PhoB::HK72), and TAA002(PstS::GFP) and exposed to a gradient of ribose concentrations. Mean values and standard deviations of 13 synthetic signaling system responses generated over six months. The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the ON state of the cell normalized to background levels of fluorescence in the absence of ribose. Error bars show the standard deviation of fluorescence measurements from twelve different experiments.*

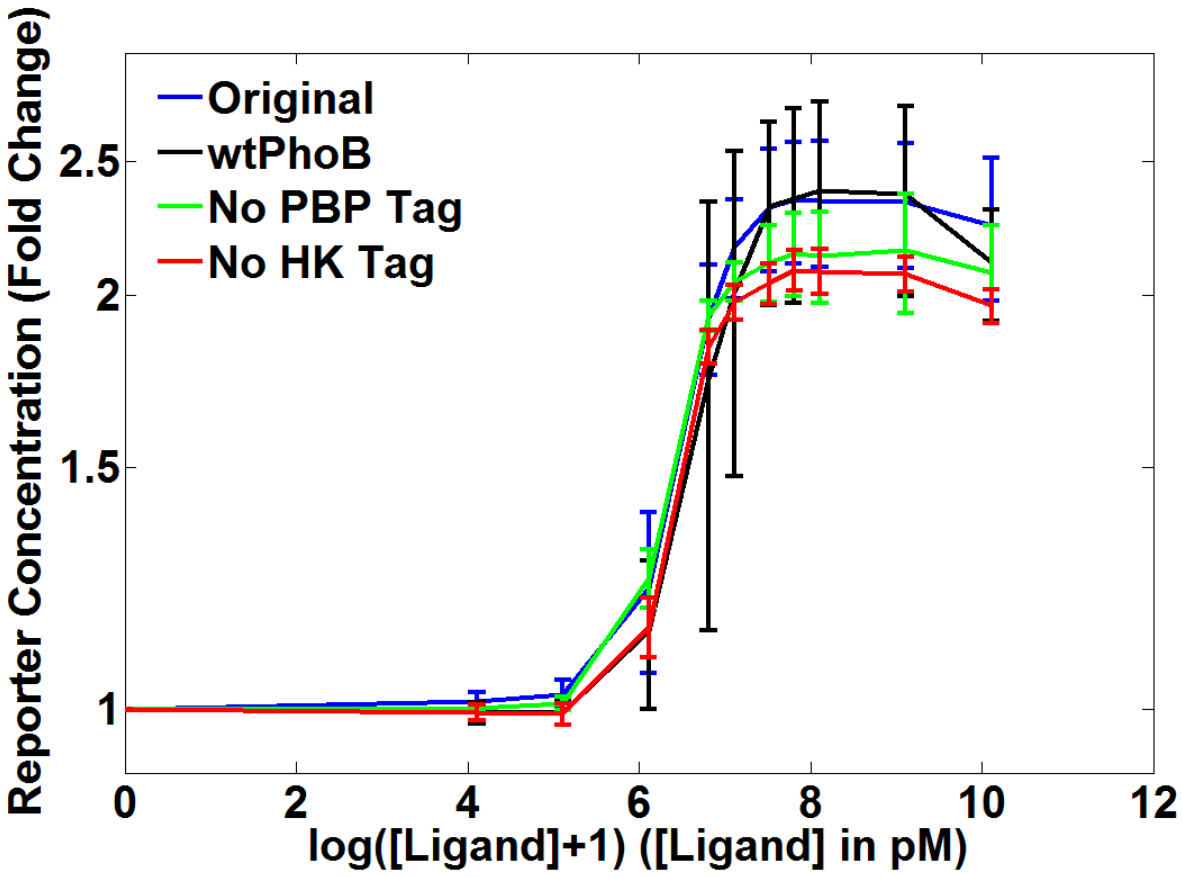


Figure 4.4. Comparison of synthetic signaling system response with and without genetic modifications to signaling system components. *Red denotes the original synthetic signaling system response (containing tags), blue shows the response of the signaling system with the PBP Flag Tag removed, and green shows the response of the signaling system with the HK leader removed. Each datapoint is the average of at least three colonies generated from three individual transformation event on three different days. Each colony was measured at each ligand concentration, and the mean of 50,000 events was normalized to background levels of fluorescence in the absence of ribose. Tags do not have a significant impact on the function of the synthetic signaling system ($p>0.05$ for all ligand concentrations with No HK Tag construct (KLH551) and the original synthetic signaling system (KJM114), $p>0.05$ for all ligand concentrations of wtPhoB (KJM584) and the original synthetic signaling system, $p>0.05$ for most ligand concentrations and $p<0.05$ for low ligand concentrations of No PBP Tag construct (KLH545) and original synthetic signaling system.*

4.4 Establishing standard growing conditions

4.4.1 Media

Initial tests of the signaling system were accomplished in the minimal media M9CA. For flow cytometry experiments, the cells can be grown in any media so the 1x synthetic signaling system was tested in both rich and minimal media. Figure 4.5 shows that cells grown in rich media made with yeast extract (LB and 2XYT) dampened system response significantly, for reasons that are unclear. Further investigation demonstrated that media containing only 10% yeast extract was sufficient to achieve signal dampening (data not shown), but the particular compound responsible has not been identified.

4.4.2 Carbon source

The M9CA media used in this workcontains 0.4%, or 5. 4 mM, glycerol. An additional carbon source can be added to M9CA after autoclaving, and glucose is often the sugar of choice (e.g. Cayley et al [86]). In this case, glucose could have an impact on the signaling system, as *E. coli* glucose and galactose binding protein (GBP) is also capable of interacting with the extracellular sensor domain of Trg. In addition, cells metabolize sugars in different ways, and the endogenous pathways involved in sugar metabolism may impact the synthetic signaling system. In order to determine the ideal carbon source for measuring the response of the synthetic signaling system, I added 13mM of various sugars to basic M9CA (see Appendix A.3.1 for media composition). I tested two 5 carbon ring sugars, L-arabinose and D-Xylose, two disaccharides, D –maltose and D-sucrose, along with glycerol, a linear three carbon molecule. M9CA+ without an additional carbon source served as a negative control.

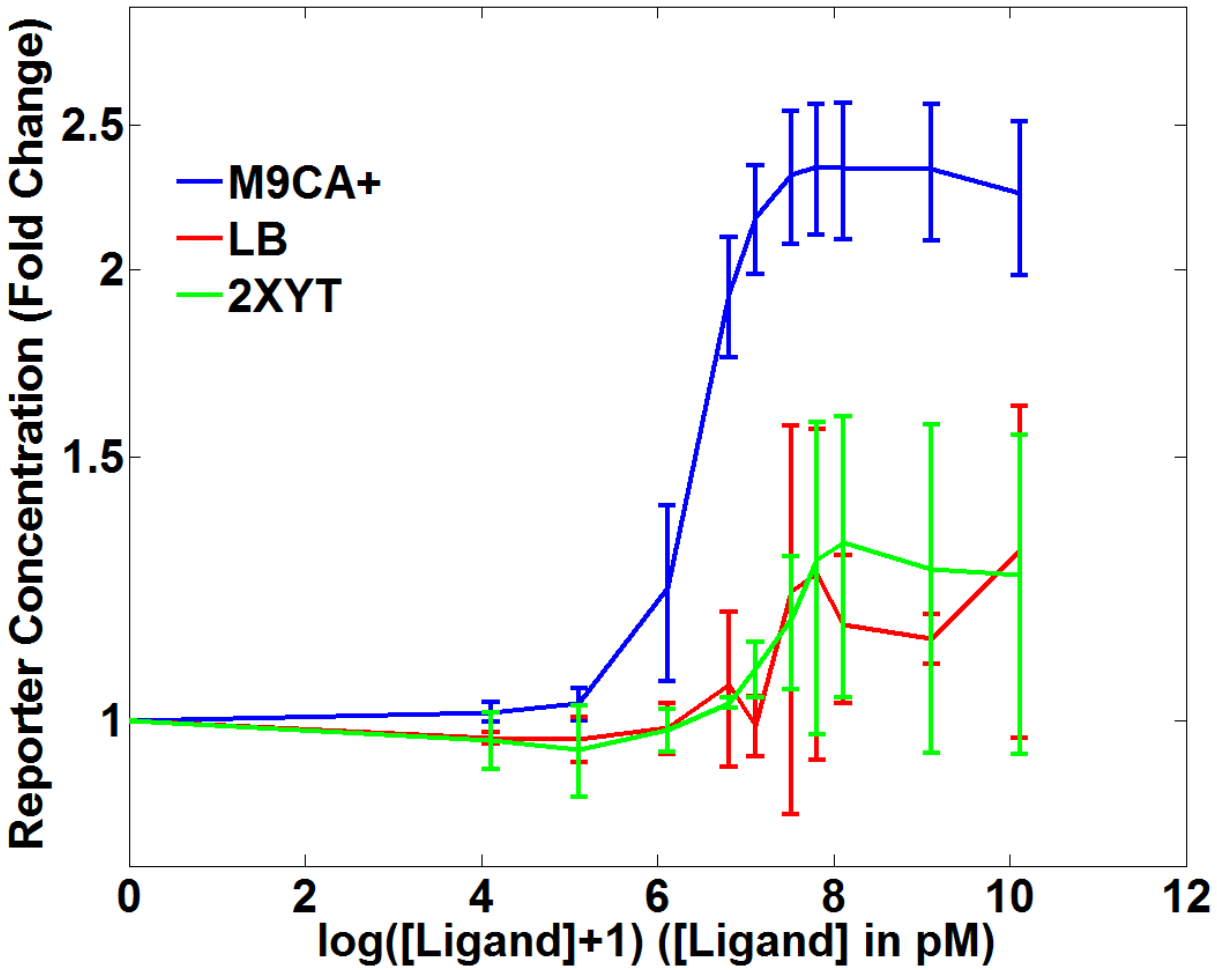


Figure 4.5. Comparison of 1x synthetic signaling system response in various culture media. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows fold change over background of arbitrary fluorescence units as captured by the flow cytometer. Cells were cotransformed with the 1x synthetic signaling system (KJM114,* P_{lacI}*::RBP,* P_{lacI}*::PhoB::HK72), and TAA002 (PstS::GFP). Cells grown in M9CA+ (blue) show a clear ribose dependent response. Responses of cells grown in 2xYT and LB (green and red, respectively) are dampened. Error bars show the standard deviation of fluorescence measurements from three different experiments. These responses are significantly different, with* $p < 0.001$ *for ligand concentrations greater than 1.3 µM of LB and M9CA responses, and 2XYT and M9CA responses.*

Figure 4.6 shows that cells grown in M9CA+ with additional glycerol (green) performed nearly identically to the no additional sugar control (blue) ($p > 0.060$ for all ligand concentrations). The disaccharides (sucrose, pink; maltose, cyan) and xylose (black) increased reporter levels in terms of raw fluorescence, but when the ON state was normalized to the fluorescence of the OFF state,

as shown in Figure 4.6, the performance of the system was degraded. L-arabinose was the poorest performer, reducing system response by more than half. From this point on, all experiments were run in M9CA+ glycerol (M9CA+).

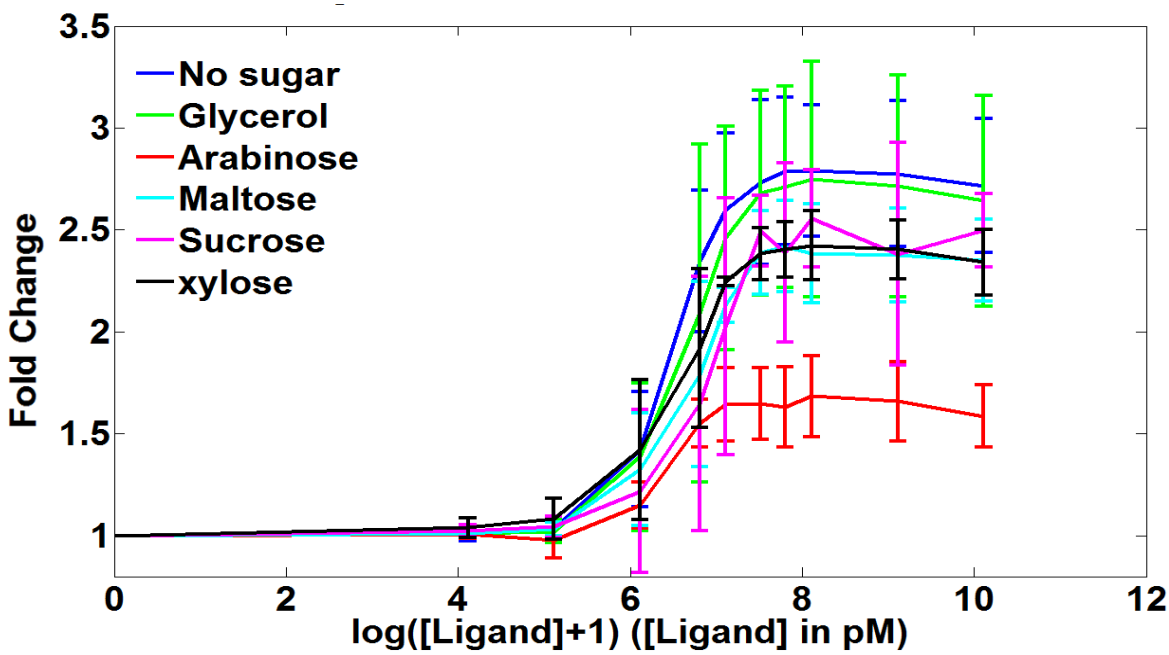


Figure 4.6. Comparison of synthetic signaling system response in the presence of various carbon sources. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows fold change over background of arbitrary fluorescence units as captured by the flow cytometer. Cells are cotransformed with the 1x synthetic signaling system (KJM114,* P_{lacI}*::RBP,* P_{lacI}*::PhoB::HK72), and TAA002 (PstS::GFP). Cells were grown in M9CA containing 0.2% glycerol and 0.2% additional sugar was added to the media. Cells grown in no additional sugar (blue) and additional glycerol (green) showed the best response, with very little deviation from one another (p>0.60 for all ligand concentrations). L-arabinose (red) showed the least response to ribose.*

4.4.2 Temperature

Temperature has the potential to impact the function of synthetic circuitry. Temperature sensitive components have been explicitly engineered into circuits [10], but temperature can also impact the activity of synthetic circuits using 'standard parts'. Wang et al characterized three promoters commonly used in synthetic circuitry, the P_{BAD} promoter, which is de-repressed by L-arabinose

application, the P_{lux} promoter, which responds to acyl-homoserine lactone, and the P_{lac} promoter, which is inducible by IPTG[35]. Each of the promoters responded to temperature differently: P_{BAD} showing very little difference in expression between 30˚C and 37 ˚C, while the basal transcription rate of P_{lux} increased at the higher temperature (the promoter became "leakier" in the absence of inducer). In contrast, P_{lac} performed more robustly at higher temperature.

In order to determine the sensitivity of P_{LacI} to temperature, I tested the 1x signaling system at both 30˚C and 37 ˚C (Figure 4.7). In this experiment, two colonies from a transformation were inoculated in M9CA+. One was incubated at 37˚C overnight and the other incubated at 30˚C. The next day the cultures were diluted, aliquoted, and inoculated with ligand as described in section 4.3.1, and grown at 37 or 30 degrees respectively during the incubation period.

My data show that temperature does not affect the fold change system response at the two points tested. Cells grown at a 37˚C show the same two-fold change and transition from OFF to ON as cultures grown at 30˚C, as seen in Figure 4.7. There is no statistical difference between the two sets of data, when the ON state is normalized to the GFP fluorescence level in the absence of ligand.

However, Figure 4.8 shows that the raw fluorescence values, *i.e.*, the absolute expression of GFP by the synthetic signaling system at 37˚C are higher over the range of ribose concentrations tested. Both the ON state and the OFF state generated much more fluorescence at 37 ˚C, but as shown in Figure 4.7 above, there is no change in the maximum fold change over background, threshold, or sensitivity, of the synthetic signaling system. It is possible that temperature impacts the expression of P_{LacI}, but because both the RBP and the PhoB::HK72 operons are controlled by this promoter, the ratios remain constant, thereby generating the comparable fold changes shown

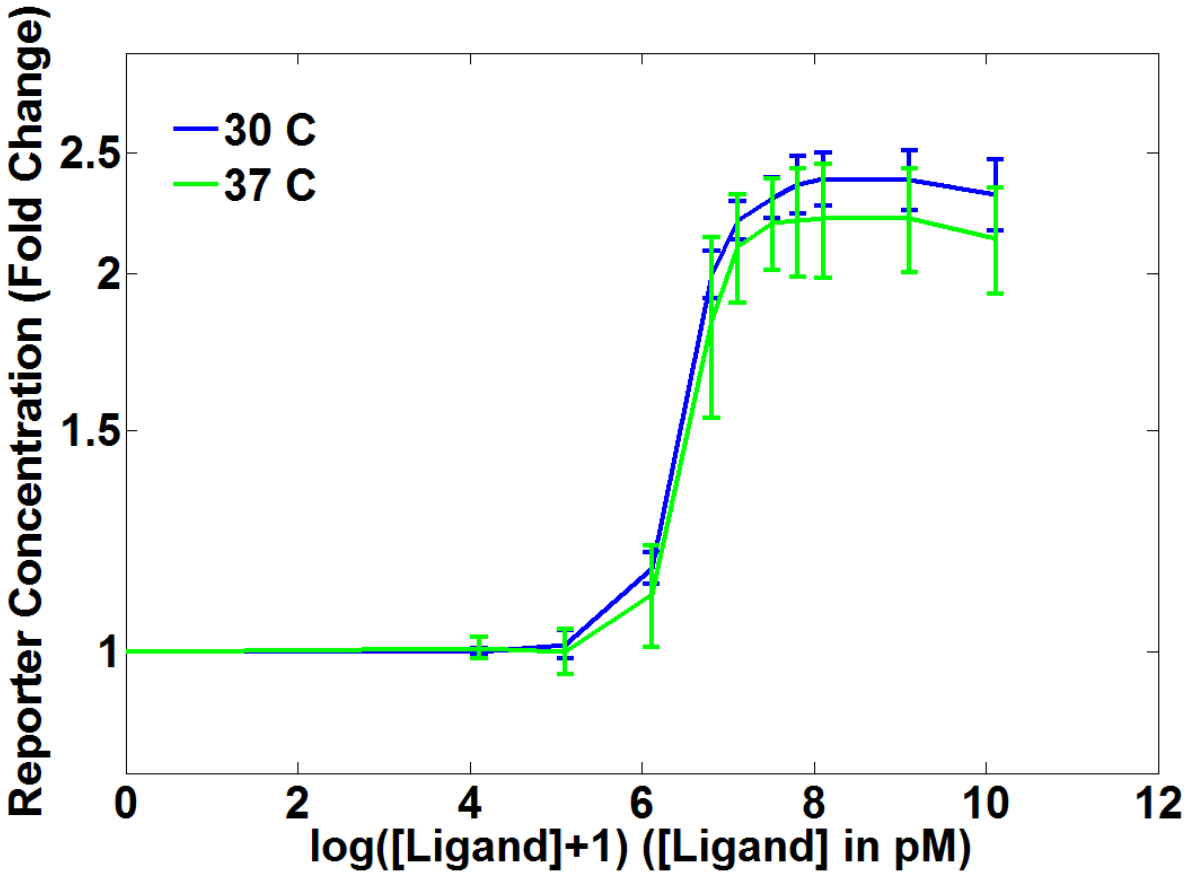


Figure 4.7. Test of synthetic signaling system fold change response at 30˚ and 37˚ C. *The ON state fluorescence was normalized to background levels of fluorescence in the absence of ribose. For each replicate of the experiment, two colonies co-transformed with the 1x synthetic signaling system (KJM114,* P_{lacI}*::RBP,* P_{lacI}*::PhoB::HK72), and TAA002 (PstS::GFP) were selected from each transformation event. One was grown at 30 and one at 37°C in M9CA+. The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows fold change over background of arbitrary fluorescence units as captured by the flow cytometer. There is no statistical difference between the sets. (p value > 0.25 for all ligand concentrations of cultures grown at 30 and 37 degrees)*

in Figure 4.7. Because there was no significant difference between the temperatures with regard to either fold change between OFF and ON states or the transition between OFF and ON states, this parameter does not optimize system performance. Therefore, I maintained our established protocol and continued to grow all cultures at 30˚C.

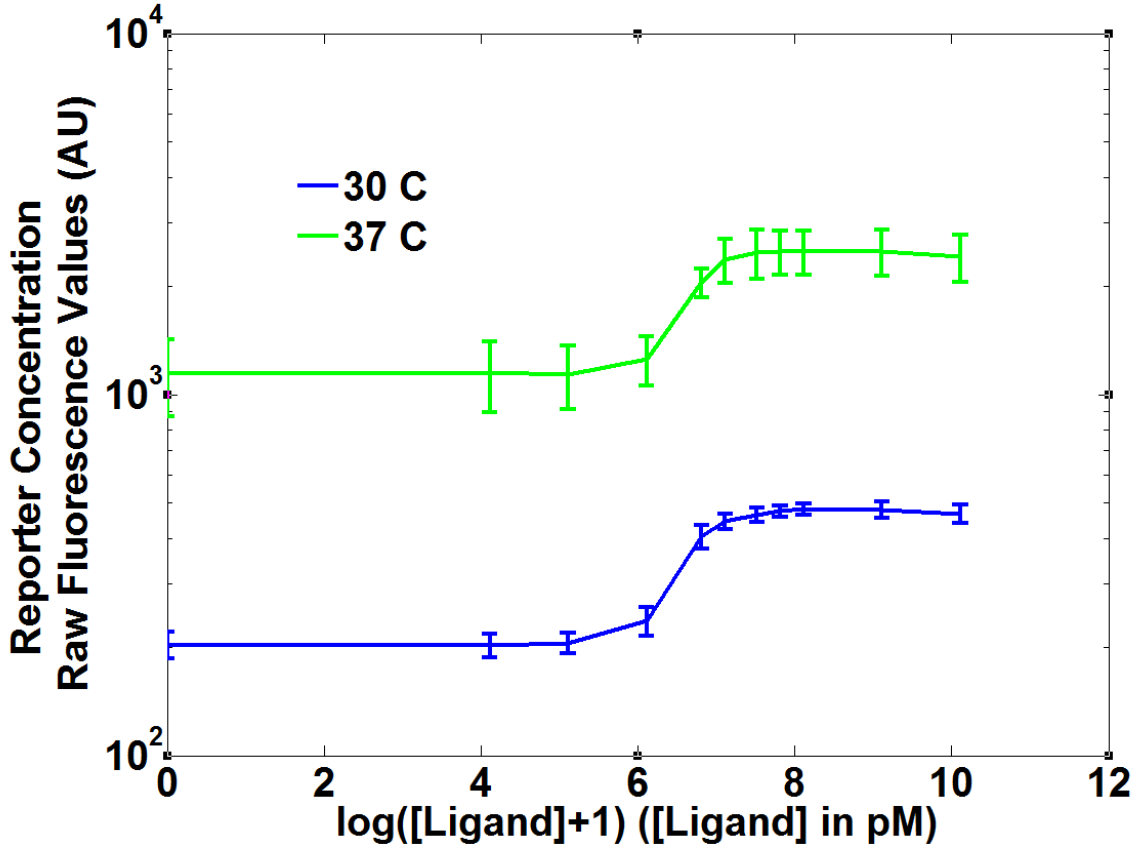


Figure 4.8. Test of synthetic signaling system at 30˚ and 37˚ C, Raw fluorescence values. *Fluorescence values are not normalized to background. For each replicate of the experiment, two colonies co-transformed with the 1x synthetic signaling system (KJM114,* P_{lacI}*::RBP,* P_{lacI}*::PhoB::HK72), and TAA002 (PstS::GFP) were selected from each transformation event. One was grown at 30 and one at 37°C in M9CA+. The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows fold change over background of arbitrary fluorescence units as captured by the flow cytometer. The difference between the sets is statistically significant. (p value < 0.02).*

4.4.3 Fold change dependence on OD_{600} and time of induction

Cell cultures were grown overnight prior to dilution for flow cytometry measurements, as described in section 4.3.2. Below an OD_{600} of 0.04 or at a dilution factor greater than 64, culture density does not appear to impact fluorescence measurement [101]. All flow cytometry experiments in this work were conducted at an OD_{600} below 0.01, and a dilution factor of greater than 100.

While OD_{600} was not a factor in these experiments, the time after induction at which fluorescence measurements were taken did have an impact on the observed signal. Figure 4.9 shows that at relatively short timescales (3 hours, blue) responses can be detected at lower levels of ribose. A distinct response above background levels of fluorescence is generated at 33 nM of ribose at three hours, but this sensitivity decreases as time after induction increases. At six hours, there is no detectable signal above the background at 33 nM of ribose. However, at higher levels of ribose, the maximum signal generated by the system reaches a peak at 6.5 hours (green). Most readings took place between three and six hours of inoculation, depending on experimental set up.

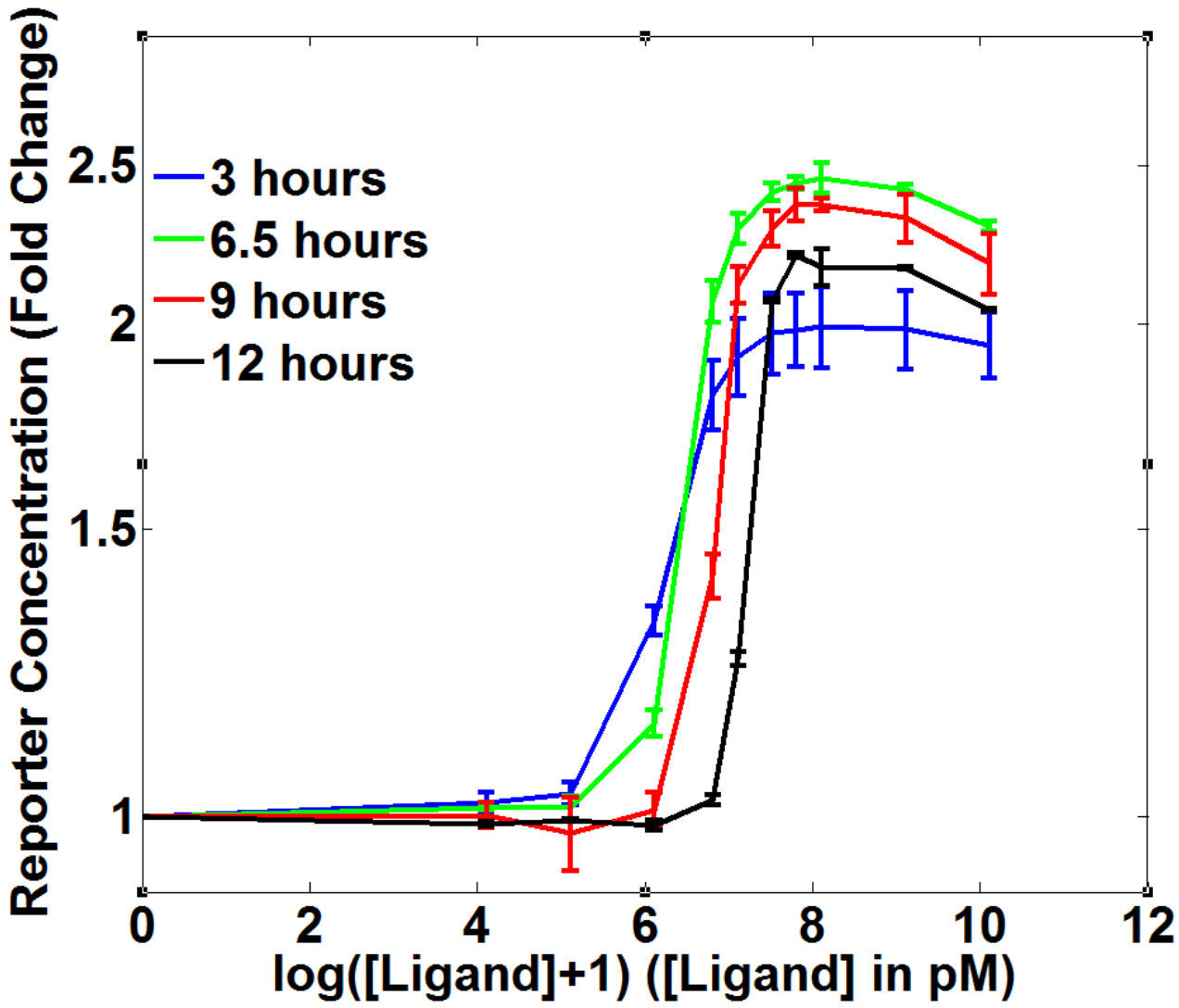


Figure 4.9. Synthetic signaling system response as a function of time. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the fold change of fluorescent response over background. Cultures grown in M9CA were diluted to an OD600 below 0.01, inoculated with the standard series of ribose concentrations and measured every three hours on the flow cytometer. The average geometric mean of the area and standard deviations of two biological replicates were plotted against the ribose concentration and responses were compared. Increased sensitivity appears at short timescales (3 hours), while maximum signal is generated at intermediate timescales (6 to 9 hours)*

CHAPTER 5: EXPERIMENTAL VALIDATION OF SYNTHETIC SIGNALING SYSTEM MODEL

I have established a functional model and a well characterized baseline for the synthetic signaling system. Figure 5.1 compares the two initial results of the model (based on experimentally determined parameter values of wild-type systems) with the experimentally determined synthetic signaling system response. The results are qualitatively similar in that the OFF state occurs in the absence of ligand, the ON state occurs in the presence of ligand, and the transition between the two states occurs at approximately the same ligand concentration (less than 1.3 μM ribose).

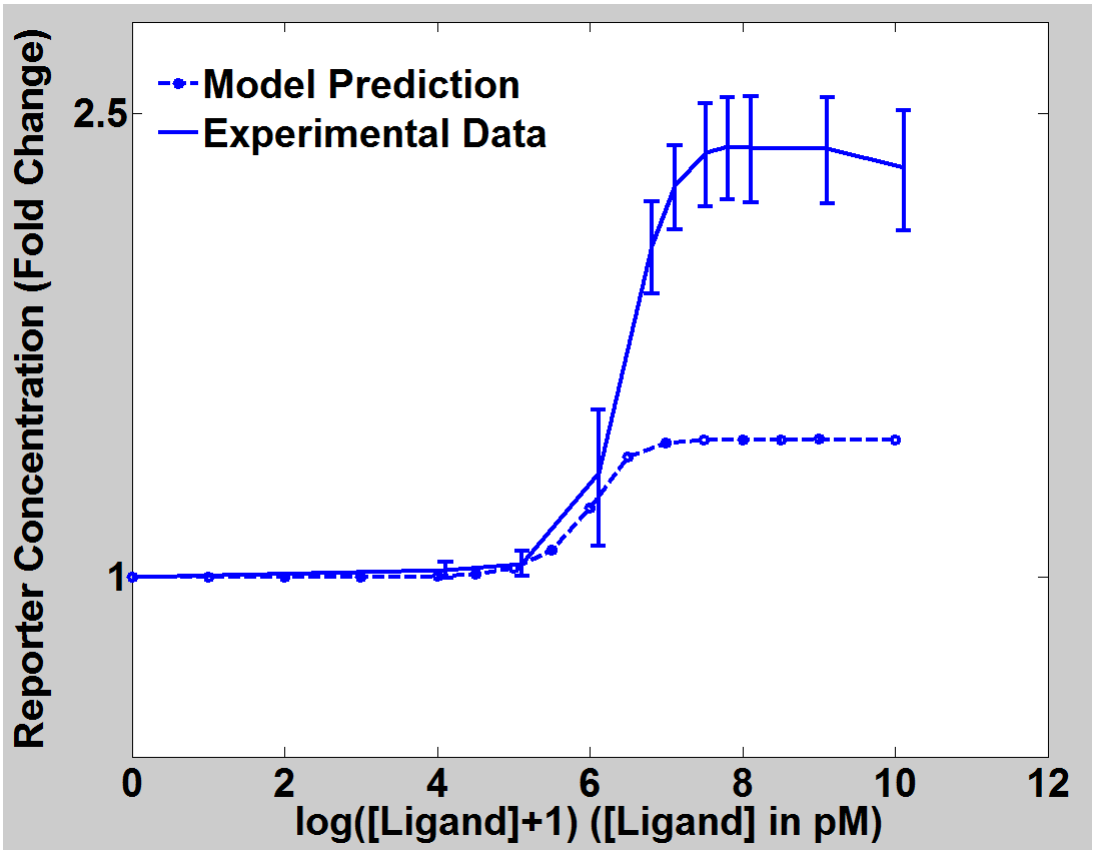


Figure 5.1 Comparison of experimentally determined synthetic signaling system response with model predictions. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the ON state of the cell normalized to background levels of fluorescence in the*

absence of ribose. EXPERIMENTAL VALUES: Mean values and standard deviations of 13 synthetic signaling system responses generated over six months. Error bars show the standard deviation of fluorescence measurements from twelve different experiments. MODEL PREDICTIONS: Blue dashed line, predicted system response in the presence and absence of ribose using initial conditions as described in Table 4.1 and 4.2. Responses are normalized to the predicted response in the absence of ribose.

While the model prediction and the experimental results are qualitatively similar, they are obviously different with regards to the maximum signal generated in each case. The model can be modified to fit the data at this time, but there are over 30 parameters and 4 species concentrations that can be changed, and several combinations of these parameters could force the model to fit the data. In essence, the model could be fit to the data regardless of biological reality. However, it is interesting to note that it is possible to get better agreement with the experimental data by changing a single parameter. Figure 5.2 shows that by increasing the concentration of RR in the model by 50%, the model better matches the experimental data. This parameter was identified after a global parameter search, which is described more fully in section 5.3.

The new model agreement may point to a false assumption about the concentration of RR. However, it is unlikely that this is the only assumption in the entire model that is incorrect. For example, if we assume the translation rate of the fully induced PstS promoter is greater than the original model definition, a model prediction fits the experimental data even more closely, as seen in Figure 5.3. However it is still important to establish if this model captures the behavior of the synthetic signaling system by experimentally validating a model prediction. All subsequent model predictions are generated assuming an increased concentration of RR, and an increased PstS transcription rate upon full induction by PhoB.

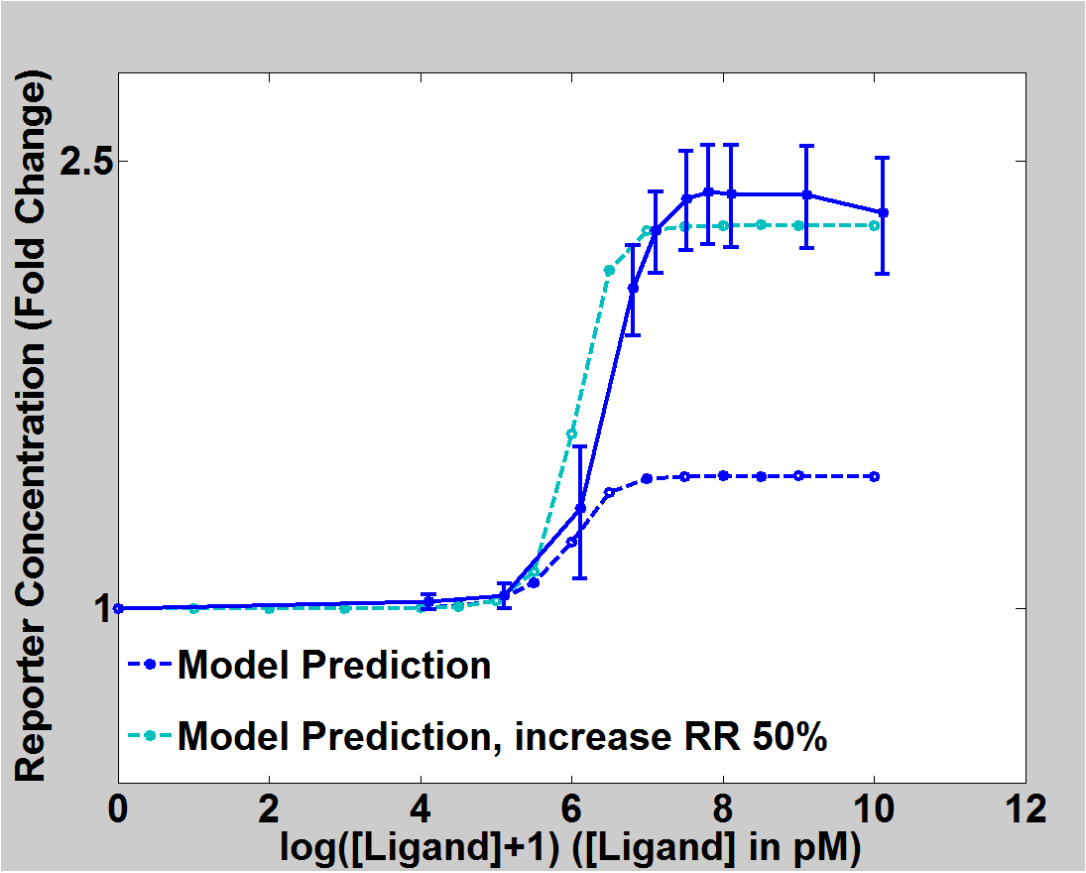


Figure 5.2 Comparison of experimentally determined synthetic signaling system response with model predictions. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the ON state of the cell normalized to background levels of fluorescence in the absence of ribose. EXPERIMENTAL VALUES: Mean values and standard deviations of 13 synthetic signaling system responses generated over six months. Error bars show the standard deviation of fluorescence measurements from twelve different experiments. MODEL PREDICTIONS: Blue dashed line, predicted system response in the presence and absence of ribose using initial conditions as described in Table 4.1 and 4.2, with the exception of the initial concentration of RR, which has been increased by 50%. Responses are normalized to the predicted response in the absence of ribose.*

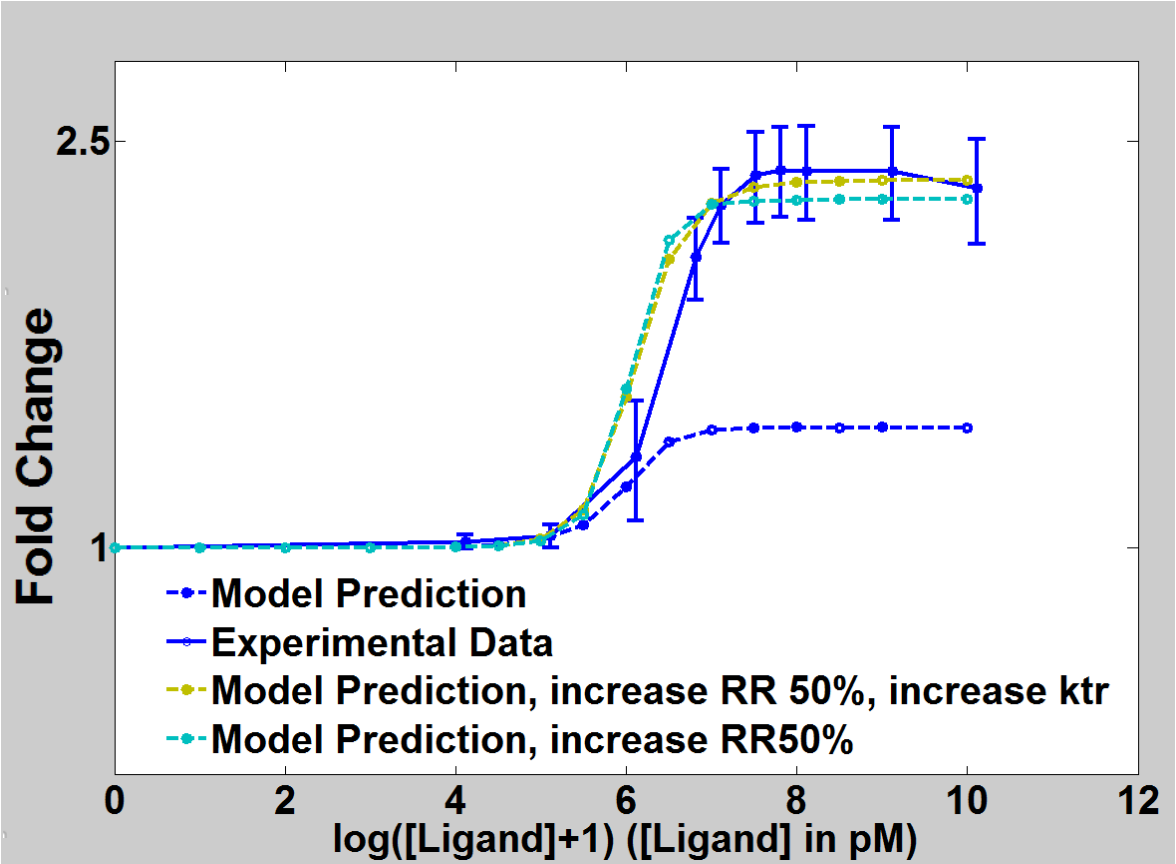


Figure 5.3 Comparison of experimentally determined synthetic signaling system response with updated model predictions. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the ON state of the cell normalized to background levels of fluorescence in the absence of ribose. EXPERIMENTAL VALUES: Mean values and standard deviations of 13 synthetic signaling system responses generated in 13 independent experiments. Error bars show the standard deviation of fluorescence measurements from twelve different experiments. MODEL PREDICTIONS: Blue dashed line, predicted system response in the presence and absence of ribose using initial conditions as described in Table 4.1 and 4.2, with the exception of the initial concentration of RR, which has been increased by 50% (teal) and ktr, which has been increased 5-fold (mustard). Responses are normalized to the predicted response in the absence of ribose.*

To determine the validity of this initial representation of the synthetic signaling system I had to identify a parameter that had an impact on the synthetic signaling system response when modified and could be engineered without impacting other components of the model. In order to identify the critical components of the system – values which had a large impact on the system response over this dynamic range – I conducted a global analysis of the parameter space, running

the model over 5 orders of magnitude for each parameter and species concentration. Ideally, the parameter selected from this analysis could not only be modified independently, it would also move us towards an optimized system by increasing the maximum fold change between the OFF and ON states, decreasing the threshold of detection, or increasing the sensitivity of the response. For each parameter, the simulation was run to steady state at 16 ligand concentrations and the concentration of reporter was recorded. The results can be grouped into three categories: Those which do not have a large impact on the system, those which change the behavior of the system, but cannot easily be engineered without modifying any other parameters, and those which change the behavior of the system and can be engineered individually.

5.1 Parameters that do not impact system behavior

Parameter *kz3*, which represents the rate of reporter degradation, is an example of the parameters which do not significantly impact the system. Some of the parameters in this category have no significant impact in this parameter space (*i.e.*, a 100-fold decrease in the parameter value generates the same response as a 100-fold increase in the parameter value). However, *kz3* acts as an indiscriminant 'volume' knob in this space, increasing expression of the reporter in both the OFF state and the ON state. Figure 5.4 shows that increasing kz3 decreases the total amount of reporter at steady state (blue), and decreasing kz3 increases the total amount of reporter at steady state (purple). Regardless of the absolute concentration of reporter, the fold change in reporter concentration between the OFF and the ON state remains the same (in this example, approximately two fold more reporter is produced in the ON state in all cases). Therefore, the

degradation of the reporter can be a rate anywhere over five orders of magnitude without affecting the fold change, threshold, or sensitivity of the system response.

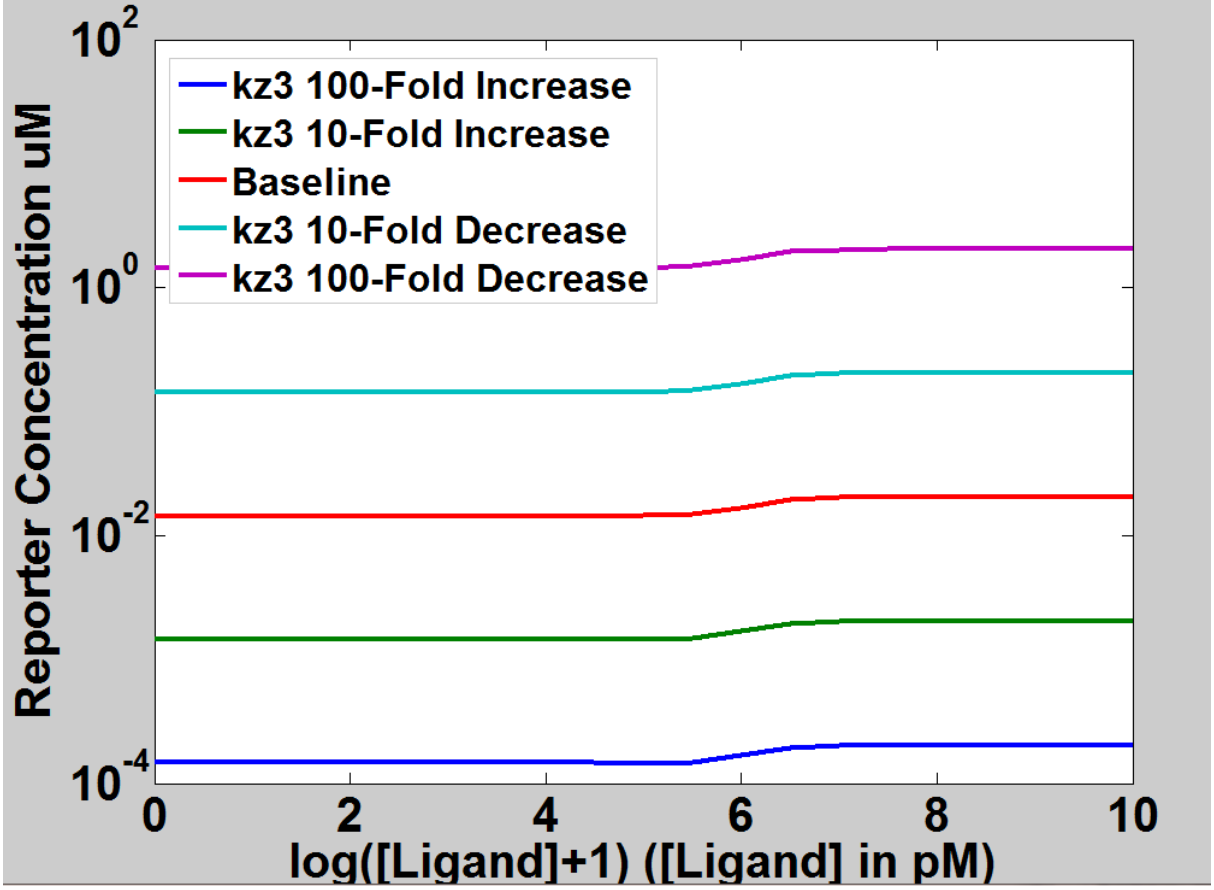


Figure 5.4. Effect of reporter degradation on system response. *The x-axis is log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. The same pattern of behavior is observed whether the rate of degradation was one hundredth (purple) or one hundred times (blue) the baseline value. Therefore the degradation of the reporter can be a rate anywhere over five orders of magnitude without affecting the fold change, threshold, or sensitivity of the system response.*

5.2 Parameters that cannot be engineered in isolation

ktr, the rate of transcription of the induced PstS promoter controlling expression of reporter is an example of a parameter that changes the system response, but cannot be engineered at this time without modifying additional parameters. Figure 5.3 shows the system response generated using

the literature derived value of *ktr* in red, the effect of a 100-fold increase in *ktr* in blue and the effect of a 100-fold decrease of *ktr* in purple. An increase in *ktr* increases the difference between the OFF and ON states significantly (blue line), but this increase only occurs if all other parameters are held constant. At this time we cannot engineer an increase in transcription rate into a promoter without potentially changing other promoter characteristics (for example, the basal rate of transcription, or the cooperativity of PhoB binding). This will be discussed further in section 5.4.

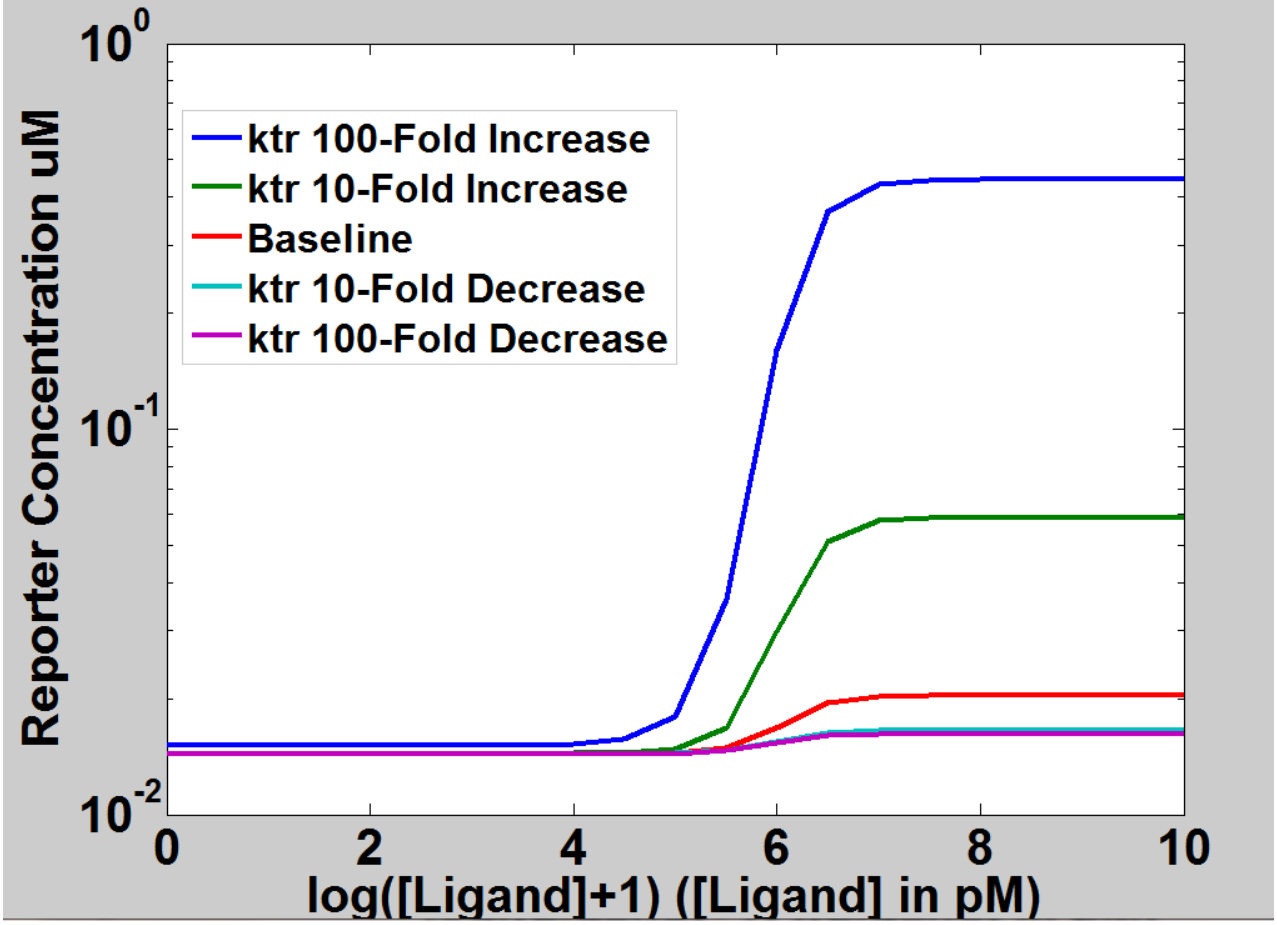


Figure 5.5. Effect of parameter ktr, the transcription rate of the induced PstS promoter, on system response. *X-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. As binding affinity increases (blue line), the maximum signal generated in the ON state increases, and the sensitivity of the system increases. As* ktr *decreases, maximum signal decreases as does sensitivity (purple line).*

kb1, the binding rate of RR to the phosphorylated, ligand-PBP bound HK, is another example of a parameter which cannot be engineered in isolation. Figure 5.6 shows the system response generated using the literature derived value of *kb1* in in red, the effect of a 100 fold increase in *kb1* in blue and the effect of a 100 fold decrease of *kb1* in purple.

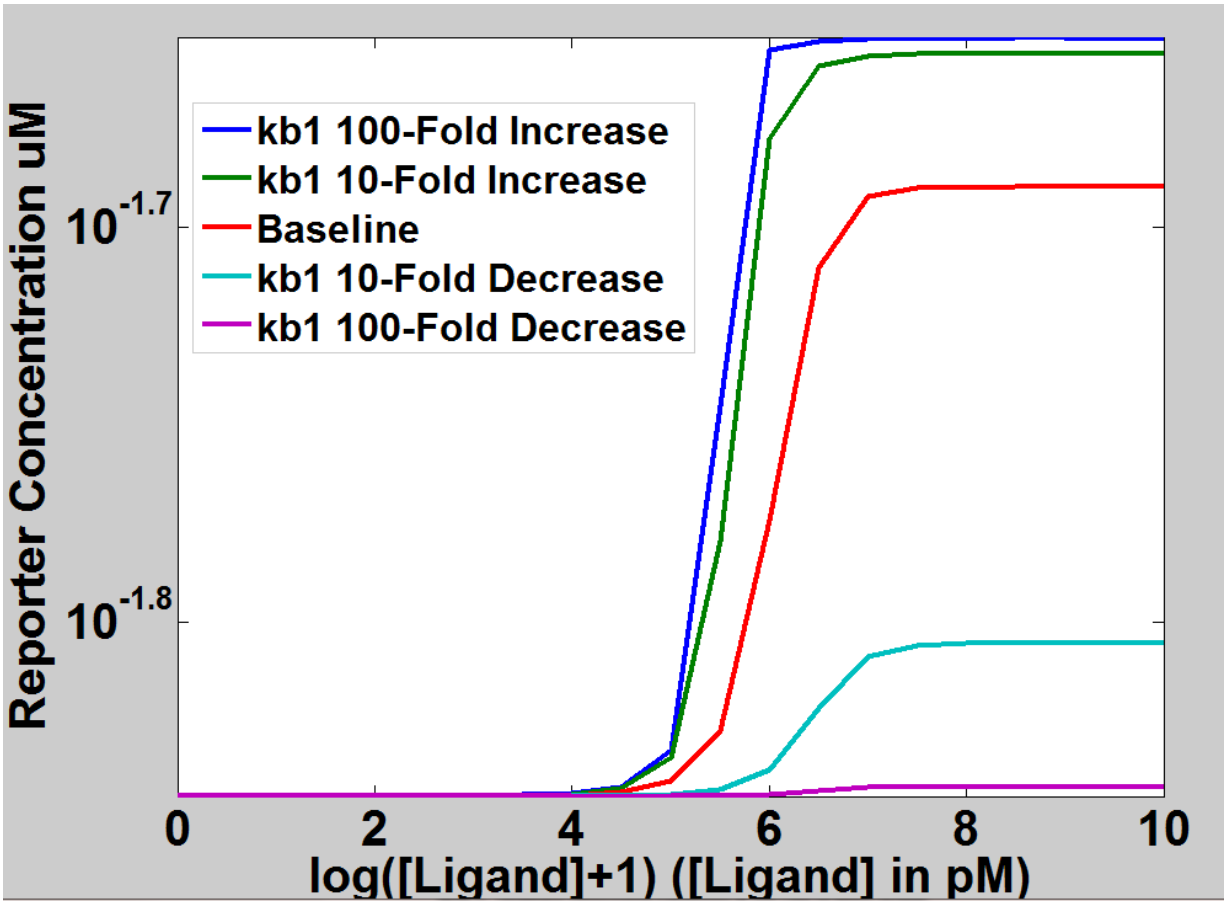


***Figure 5.6. Effect of the binding affinity of RR to HK,* kb1.** X-*axis represents the log of the ribose concentration in picomolar, and the y-axis shows the reporter in μM concentration as predicted by the model. As binding affinity increases (blue line), the maximum signal generated in the ON state increases, and the sensitivity of the system increases. As* kb1 *decreases, maximum signal decreases as does sensitivity (purple line).*

This result suggests that if the rate of binding of RR to HKp:L could be increased, RR binding to the HK would no longer be limiting, more RR would be phosphorylated, and a greater system

response generated in the 'ON state' (blue line). The binding of RR to HK is the result of specific molecular interactions that can be difficult to identify [102], but in the case of PhoB, the protein specificity interface has been identified [37]. Haldimann et al used directed evolution to establish new specificities for PhoB[37], and rational design has been used to decrease binding to the HK [66]. In addition, Skerker et al (2008) have rationally re-designed the HK EnvZ to phosphorylate a non-cognate RR[102]. However, in order to test the prediction modeled in Figure 5.6, the modification that increases PhoB binding must not change any other parameters, for example the rate of unbinding from the RR:HKp:L complex or phospho-transfer – all parameters other than *kb1* are held constant in this simulation. The intricacies of protein engineering are not fully understood[103], and engineering a better binding interaction between PhoB and the HK without impacting additional parameters is outside the scope of this work.

5.3 Parameters that impact the system and can be modified in isolation – Changing PBP concentration

The third class of parameters contains those parameters that have a large impact on system behavior and can be easily modified. Figure 5.7 shows the impact of increasing RBP concentration on the system. As the total amount of RBP increases, the maximum signal of the system increases (blue line).

Modifying the concentration of RBP would achieve one of the goals of this project – increasing the maximum response of the ON state. Furthermore, we could effect that change, without modifying any other experimental parameters, by changing the promoter controlling the expression of RBP.

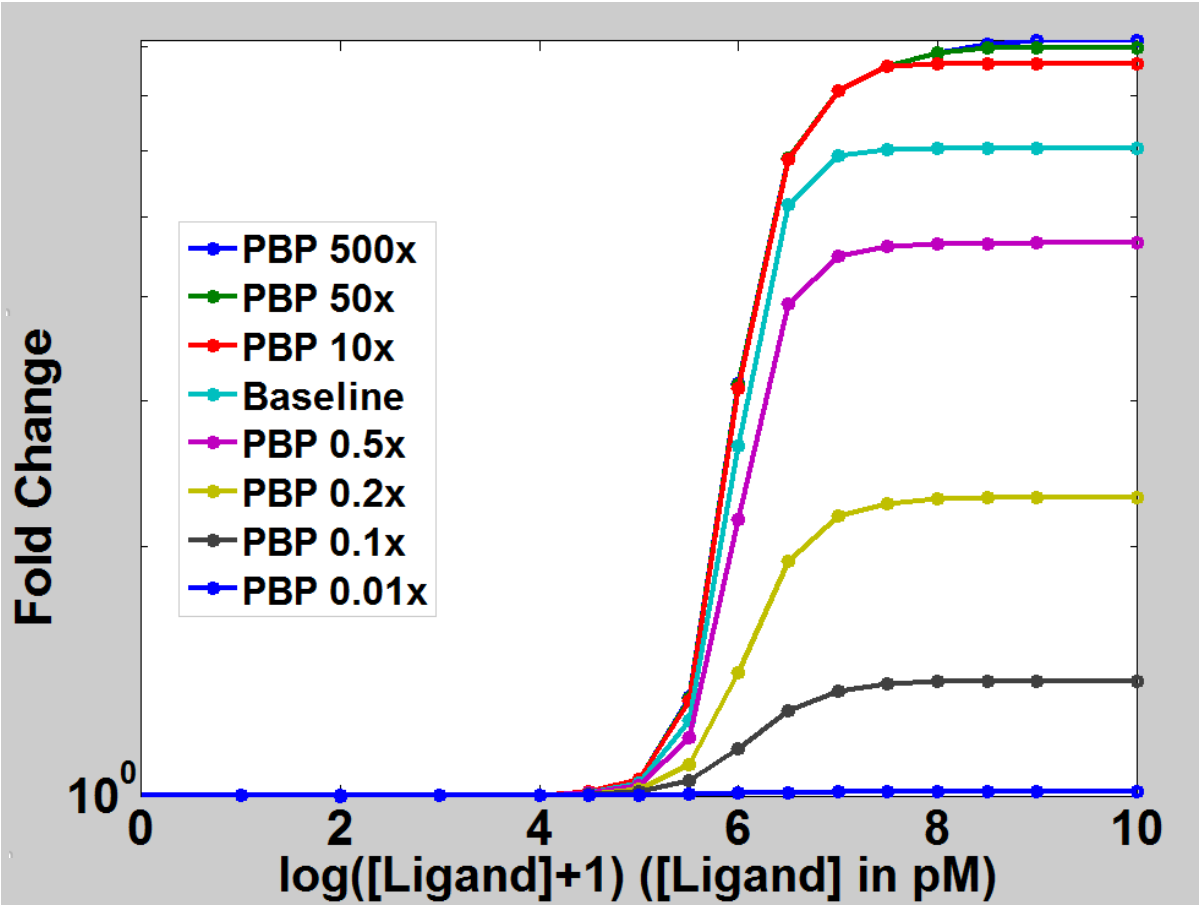


Figure 5.7. Model simulation of synthetic signaling system response as PBP concentration increases. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. As the concentration of PBP increases, the maximum signal increases. MODEL PREDICTIONS: Predicts system response in the presence and absence of ribose using initial conditions as described in Table 4.1 and 4.2, with the exception of the initial concentration of RR, which has been increased by 50%, and ktr, which has been increased 5-fold. Responses are normalized to the predicted response in the absence of ribose.*

The impact of this change in protein concentration depends on where our signaling system currently exists in this parameter space. If the concentration of PBP was already close to a maximum (*e.g.*, green), then additional increases in PBP concentration (blue line) would not greatly impact the system. Our initial assumption about the concentration of PBP in the

synthetic signaling system (red) suggests that is the case. However, our assumption may be incorrect.

Figure 5.8 shows an overlay of experimental results (blue, with error bars) on the model predictions. In this parameter regime, the PBP concentration which most closely matches my experimental results is a concentration five-fold lower than the initial assumption (mustard). If this is indeed the case, increasing PBP concentration (red) could have a substantial impact.

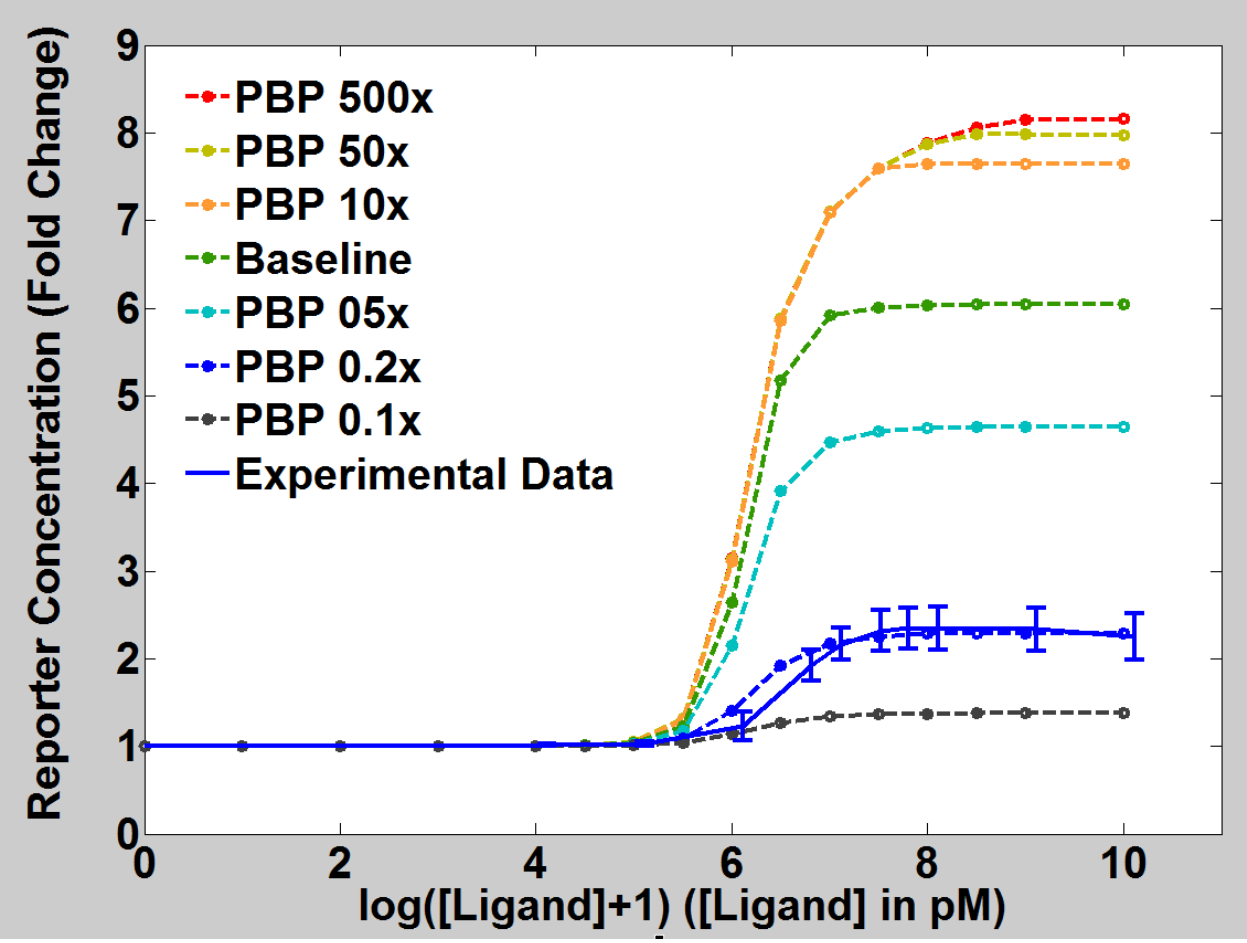


Figure 5.8. Model simulation of synthetic signaling system response as PBP concentration increases compared with experimental results. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the reporter concentration in fold change as predicted by the model. As the concentration of PBP increases from 0.1x of baseline value to 500x baseline value, the maximum signal increases. EXPERIMENTAL VALUES: Mean values and standard deviations of 13 synthetic signaling system responses generated over six months.*

Error bars show the standard deviation of fluorescence measurements from twelve different experiments. MODEL PREDICTIONS: Blue dashed line, predicted system response in the presence and absence of ribose using initial conditions as described in Table 4.1 and 4.2, with the exception of the initial concentration of RR, which has been increased by 50%. Responses are normalized to the predicted response in the absence of ribose.

The original synthetic signaling system uses P_{lacI} to express RBP. Based on the assumptions described above, our estimated baseline RBP concentration is 3.5 μM, or 150 molecules (or perhaps fewer). Wild type *E. coli*, however is known to express 37,000 molecules of RBP in the presence of ribose [97], a concentration approximately 200-fold greater than our existing system. The disparity between wild-type RBP levels and predicted levels of RBP in our synthetic signaling system suggests that experimentally increasing expression of RBP is a viable way to validate the model.

5.3.1 Inducible expression of RBP.

To validate the model prediction that an increase in RBP increases the maximum response of the signaling system in the presence of ligand, the level of PBP expression in the synthetic signaling system must be increased experimentally. One commonly used method to increase gene expression is through the use of inducible promoters like P_{BAD}[104].

P_{BAD} is repressed by the transcription factor araC[105]. araC is composed of a C-terminal DNA binding domain, which interacts with the P_{BAD} promoter and a N-terminal dimerization domain, which interacts with other araC molecules. In the absence of the inducer, arabinose, one araC molecule binds to each of two half P_{BAD} operator sites. The sites are separated by 210 base pairs, and the N-terminal domains of the bound araC interact and form a loop in the promoter that represses expression by P_{BAD}. In the presence of L-arabinose, araC binds the sugar and undergoes a conformational change that makes dimerization more favorable and encourages

binding to two adjacent half sites[105]. This releases the loop and allows the polymerase access to P_{BAD}, thereby inducing expression.

P_{BAD} has low levels of un-induced expression and high levels of induced expression, and is therefore a useful tool for increased protein expression[105]. In addition, it has already proven to be a useful tool for synthetic biology, enabling researchers to test a wide range of protein expression levels and thereby tune the function of the system (*e.g.*, Whitaker et al[106], Anderson et al[11]). However, the L-arabinose-inducible system does not induce a graded response, in which cells gradually increase production of protein with increasing concentrations of L-arabinose.

L-arabinose induces the expression of its own transport system in a positive feedback loop[105]. Cells that have an internal concentration of L-arabinose above a certain threshold express genes to actively import more L-arabinose, whereas cells with lower internal concentrations of L-arabinose do not. This yields an all-or-none response in which a population of cells grown in the presence of L-arabinose is split into two sub-populations: one fully induced, the other uninduced[107]. The relative fractions of these induced and uninduced sub-populations change with L-arabinose concentration, and when observing a population of cells over increasing L-arabinose concentrations, the combined expression levels of the sub-populations result in an apparently graded transition from OFF to ON. However, when observing individual cells, there are only cells in the ON state and cells in the OFF state, effectively generating only a single level of protein expression.

The cell line used to test the synthetic signaling system, BW23423, already contains an L-

arabinose inducible cassette controlling expression of the cytoplasmic portion of PhoR, *araBAD*$_{AH37}$*::* P_{araB}-NFLAG-*'phoR*$_{AH41,}$[37]. The cytoplasmic portion of PhoR autophosphorylates in the presence of ATP, and rapidly transfers that phosphate to PhoB[70]. Increased L-arabinose concentration should result in increased expression of cytoplasmic PhoR, and therefore increased PhoB phosphorylation and increased expression by PhoB inducible promoters. However, in prior experimental work, when the signaling system and reporter plasmids were co-expressed and exposed to L-arabinose, no such increase occurred (Tessa Albrecht, PhD, unpublished data). This suggested that either the cassette no longer functioned, or the cell line was unable to import L-arabinose.

I co-transformed BW23423 cells with a partial synthetic signaling system containing only RBP and PhoB (plasmid KLH546), and the reporter plasmid containing PhoB-inducible GFP. Figure 5.9 shows that cells cultured in high L-arabinose (blue) had increased GFP expression in comparison to cells grown at low arabinose concentrations (red), indicating that the L-arabinose inducible cytoplasmic PhoR cassette was indeed still functional. The cells also appear to exhibit an all-or-nothing response at intermediate concentrations of L-arabinose (yellow), suggesting that the L-arabinose transport positive feedback loop was likely intact.

This experiment suggested that both the cytoplasmic PhoR and HK72 are active in a culture transformed with the original signaling system and incubated in L-arabinose. Cells containing the original signaling system were grown in L-arabinose and glycerol backgrounds and tested at a range of ribose concentrations from 0mM to 13mM ribose.

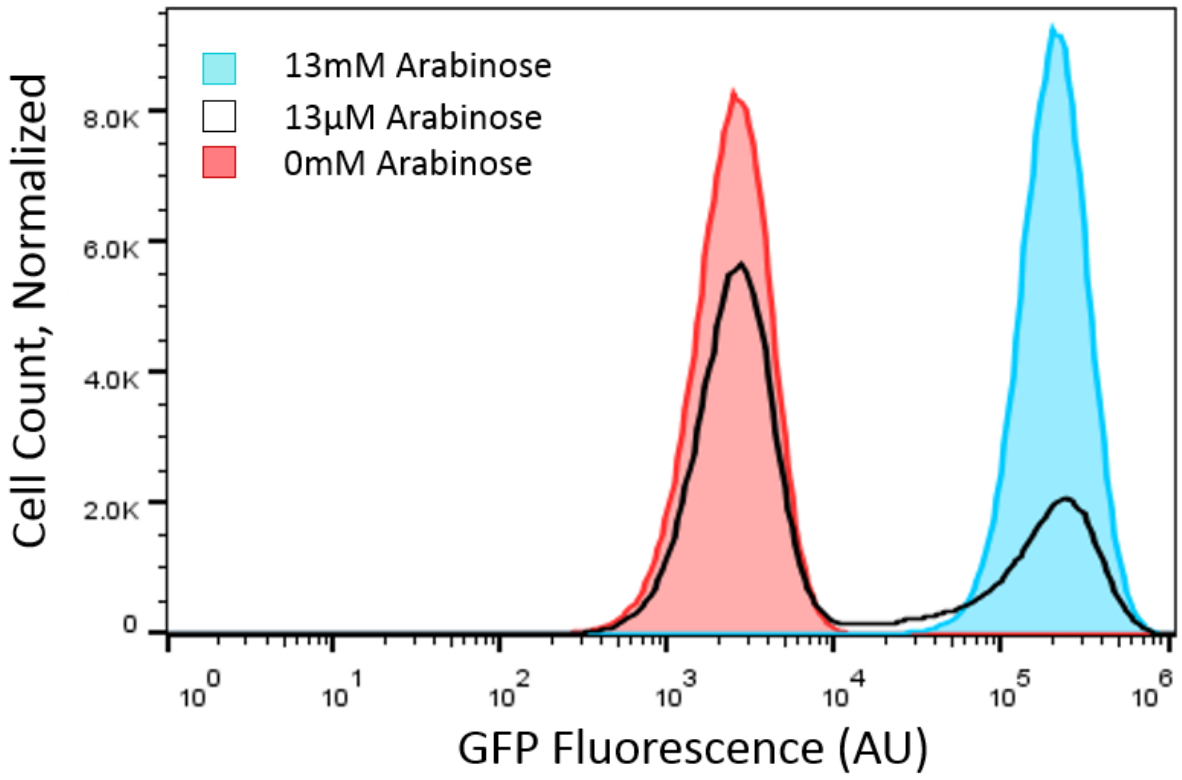


Figure 5.9 Histogram of GFP fluorescence in BW23423 cells containing only PhoB and reporter plasmids upon L-arabinose induction. *Y-axis is number of cells, normalized to mode. X-axis is GFP fluorescence in arbitrary units. Red: Cells incubated in the absence of L-arabinose. There is only one peak in this histogram. Black line: Cells inoculated in 13 µM L-arabinose. Note the small peak of ON state cells to the right of the cells in the OFF state. Blue: Cells incubated with 13 mM L-arabinose. Histograms are normalized to the mode and contain over 50,000 events. Cells were grown in M9CA+ for 6 hours.*

Figure 5.10 shows that in the absence of ligand, cells containing the synthetic signaling system grown in arabinose had roughly 2.5 times the background fluorescence of cells grown in glycerol or no additional sugar. However, the constitutive activity of cytoplasmic PhoR does not completely mask the ON and OFF states of the system generated by HK72.

One interesting possibility for the persistence of ON and OFF states in the presence of constitutive kinase activity involves the phosphatase activity of HK72. In the absence of ligand,

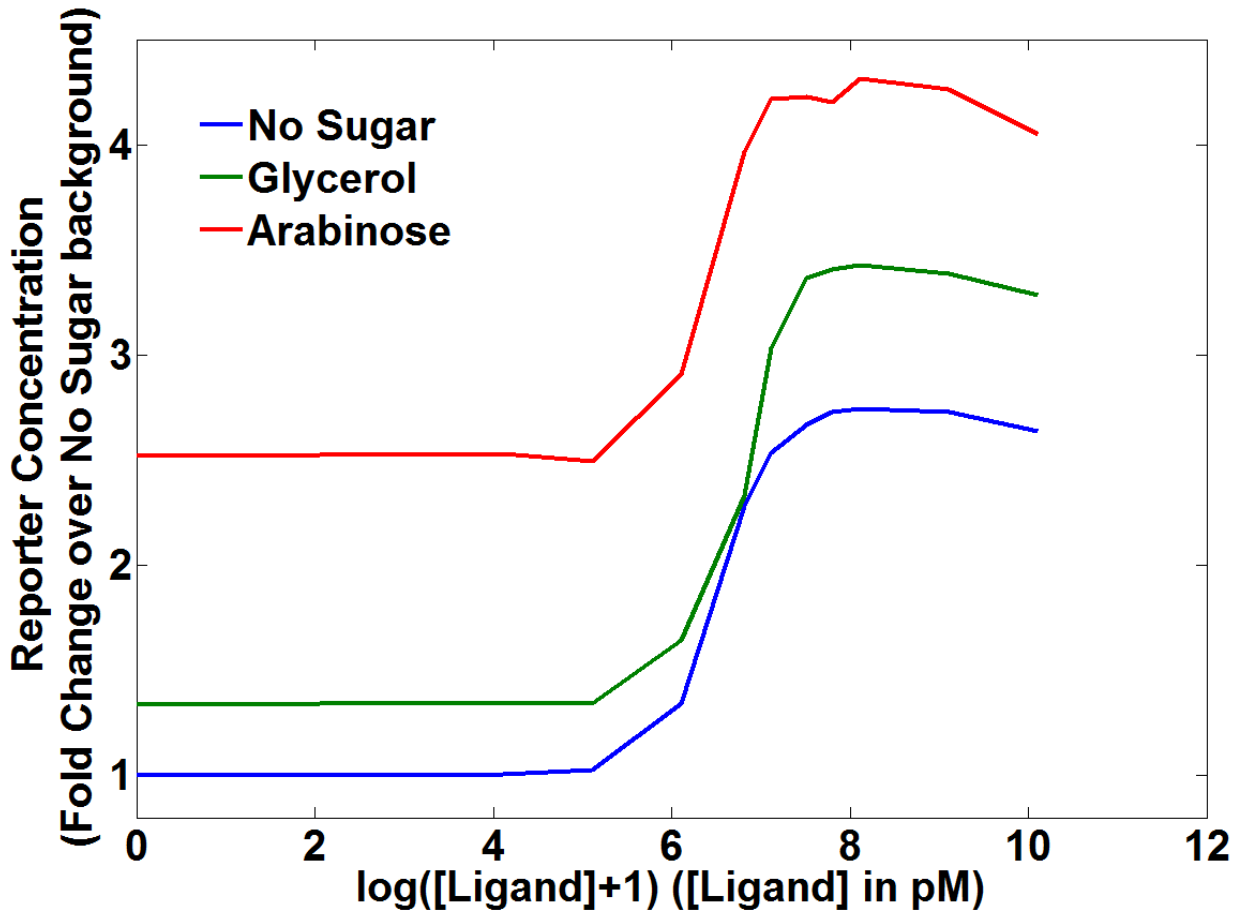


Figure 5.10. Average synthetic signaling system response to carbon source. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the fold change of fluorescent response over background. Raw fluorescence data is highly variable. The average value for each data point was normalized to the fluorescence of the average 'no sugar' culture in the absence of ribose. There are no error bars for this calculation. Cultures grown in M9CA were diluted to an OD600 below 0.01, inoculated with the standard series of ribose concentrations and GFP fluorescence was measured on the flow cytometer. The average geometric mean of the area and standard deviations of 50,000 events captured from at least three individual colonies transformed independently. 'No Sugar' has only the 0.4% glycerol added to the media before autoclaving. L-Arabinose and Glycerol cultures have had additional 13mM sugar added, as described in section 4.4.2.*

the equilibrium of HK autophosphorylation reaction favors the unphosphorylated state[55]. If the HK is unphosphorylated and therefore inactive, it is unlikely to catalyze PhoB phosphorylation. In that conformation, phosphatase activity may dominate. In the presence of ligand, wild-type HKs undergo a conformational change that results in autophosphorylation[55]. In this

conformation, kinase activity is likely to dominate. If HK72 shares this functionality, in the absence of ribose HK72 phosphatase activity may be able to depress the level of RRp generated by the constitutive kinase cytoplasmic PhoR. In the presence of ligand, the additional kinase activity of HK72 may boost the level of RRp.

While investigation into inducible promoters yielded an interesting insight about the phosphatase activity of the synthetic signaling system, it was not the optimal tool to regulate gene expression. Most sugar-dependent inducible systems, including the *lac* repressible system when induced with lactose, are involved in similar positive feedback loops[104], and would require additional modification of the BW23423 cell lines to remove the all-or-nothing phenotype. An IPTG inducible system is discussed as a potential avenue for future work at the end of this chapter.

5.3.2 Constitutively increasing RBP expression.

One alternative to inducible promoters is to use constitutive promoters of various strengths. A mutant of P_{lacI}, P_{LacIQ1} was shown by Glascock et al to be 50 times stronger than P_{lacI} as determined by repressor titration[87]. This increase in activity is the result of a 15 basepair deletion introducing the sigma70 consensus sequence at the -35 box of the promoter. This promoter has a known increase in activity over P_{lacI} and therefore would explicitly validate or reject the model prediction that an increase in RBP would result in increased maximum signal.

P_{LacIQ1} was functionally placed 5' to RBP via overlapping PCR, sequence verified, and the P_{LacIQ1}::RBP cassette was inserted into the signaling system plasmid. Based on the same assumptions used to establish the baseline concentration of RBP in the model, this would yield

approximately 30,000 molecules of RBP in the periplasm, similar to the endogenous expression level of 37,000 molecules.

5.3.4 Using RBS to increase protein expression

A constitutive promoter stronger than P_{LacIQ1} could not be identified in a literature search. In a library of synthetic promoters generated via degenerate PCR, the best performing promoter was only 27.5 times as strong as P_{lacI} [108]. Instead of additional promoters, I elected to use a translational mechanism available to increase protein production. Ribosome binding site (RBS) variants have been shown to increase or decrease translation in a predictable manner[85]. Using the Salis Lab RBS calculator, I analyzed the existing RBS and generated three new RBS for use with P_{LacIQ1} to express RBP with 10x, 100x, and 200x the predicted translational efficiency of the original P_{LacI}::RBP plasmid. The increased transcription rate of $P_{LacIQ1,}$ combined with the new 10x RBS is predicted to result in RBP expression 500-fold greater than the standard synthetic signaling system and 10-fold greater than RBP expression in wild-type chemotactic systems. 100x and 200x RBS are predicted to result in protein expression levels 100 and 200 times higher than wild-type RBP expression, respectively.

I attempted to build these three synthetic signaling system variants using the methods described in this Chapter, Strategy II. However, I was unable to clone signaling systems containing the 100x and 200x RBS. It is possible that the cells were unable to sustain those levels of expression. The synthetic signaling system variant containing 10xRBS was cloned, but only after extending the recovery period after transformation by an hour and conducting recovery and overnight plate growth at 30 degrees.

5.3.4 Results

Cells were co-transformed with each of the three synthetic signaling system plasmids, P_{lacI}::RBP (1x RBP), P_{lacIQ1}::RBP (referred to as 50x RBP), and P_{lacIQ1}::10xRBS::RBP (referred to as 500x RBP), and with the GFP reporter plasmid into LT002 cells and grown according to the methods described in Chapter 4. Figure 5.11 shows an increase in maximum signal of the synthetic signaling system in response to ribose as PBP concentration increases. The response of the original synthetic signaling system (1xPBP, blue) is two-fold above background in the presence of ribose. Increasing RBP expression with P_{LacIQ1} (50xRBP, green) resulted in an increase in maximum signal from 2-fold to approximately 7-fold above background fluorescence levels. Increasing the expression of RBP further, by introducing an RBS designed to improve translation, (500x RBP, red) resulted in a 12 fold-change over background in the presence of ribose. In addition, the apparent Kd of the synthetic signaling system shifted from 33 µM ribose for a system with low levels of ribose (blue) to 1.3 µM in a system with high levels of ribose (red).

5.3.5 Conclusions & Future Work

The model appears to fit the experimental data reasonably well. Experimental results for the P_{lacI}::RBP (1x RBP) synthetic signaling system and the P_{lacIQ1}::RBP (which putatively expresses 50x RBP), show very little discrepancy between the experimental results and the model. The close corroboration between model results and data suggest that the model is qualitatively valid, and trends derived from this model can be useful tools for directing experimental efforts.

The parameter used to validate the model in bacteria, the concentration of PBP, also functioned as a 'lever' to begin optimizing the system. By modifying this single parameter, I increased fold-change response in the presence of ligand from two-fold (1x RBP) to twelve-fold induction

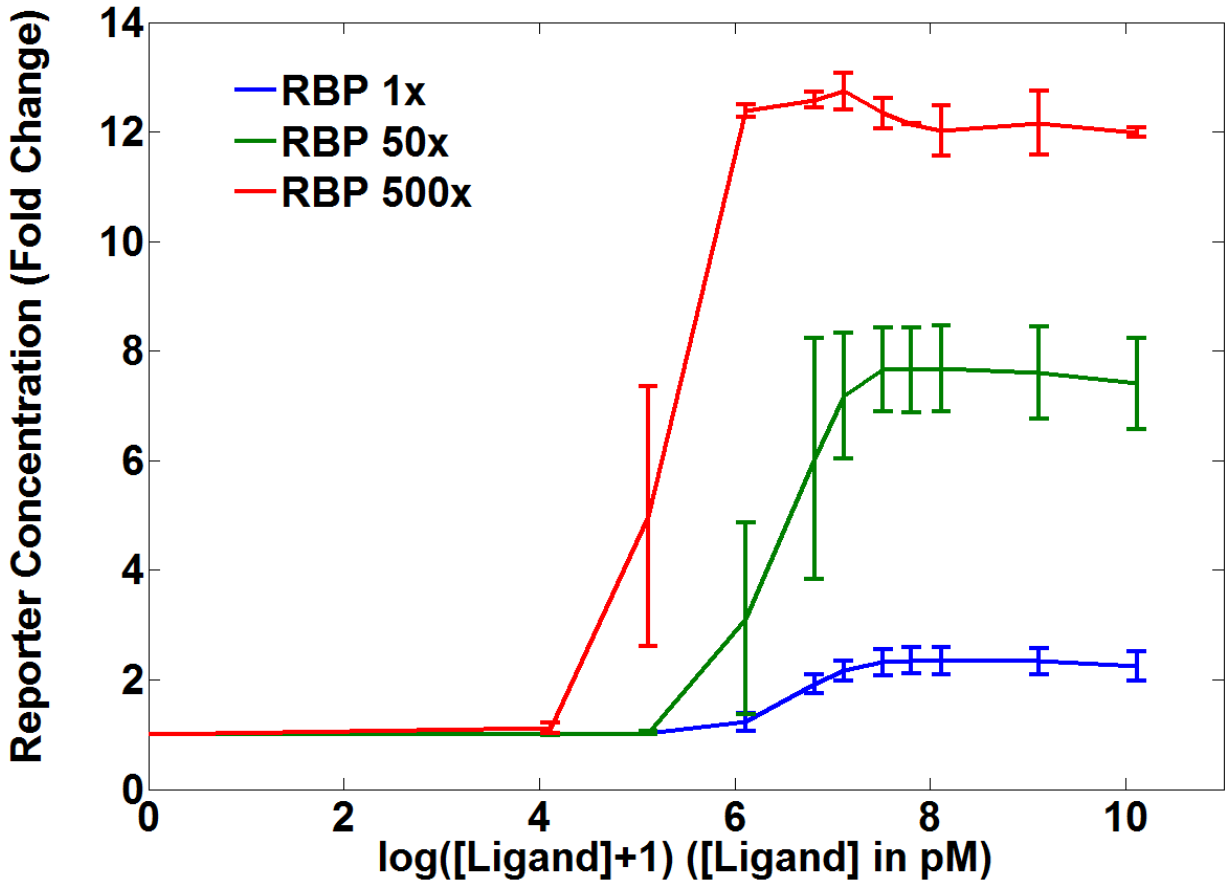


Figure 5.11a. Synthetic signaling system with three different concentrations of PBP. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the fold change in fluorescence (arbitrary units) TOP: LT002 cells were co-transformed with synthetic signaling system and reporter plasmids, incubated with a ribose concentration gradient as described in Chapter 3 and analyzed with the flow cytometer.*

Data is the average of the mean of 50,000 events from at least three independent experiments. As the concentration of PBP increases, the maximum signal increases and the threshold of the signaling system decreases as well (i.e. the signal responds at lower levels of ligand.)

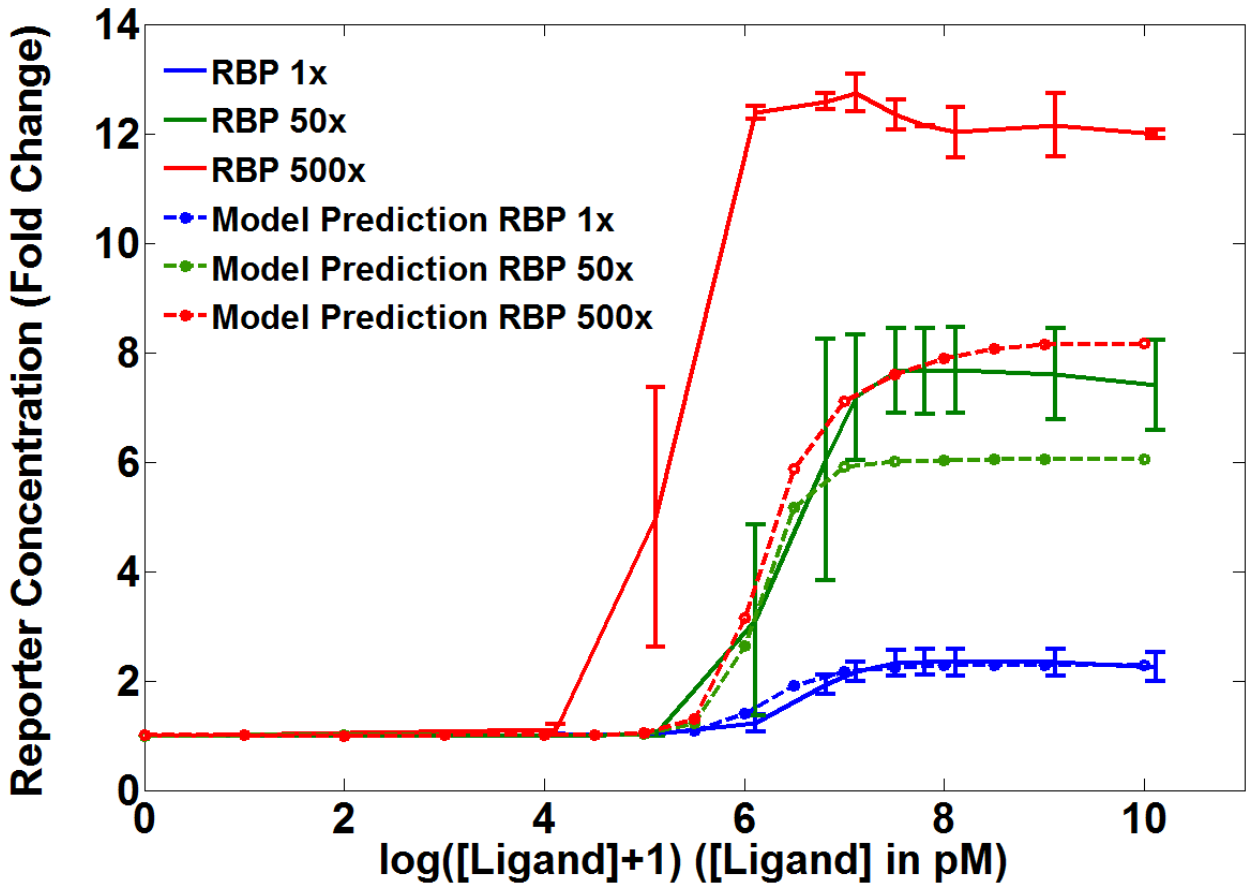


Figure 5.11b Detailed Model prediction of synthetic signaling system response upon increasing PBP, compared with experimental results. *The dashed blue line is the model predicted response at 0.2 fold the concentration of PBP I established. The dashed green line is the predicted response at 10 fold the initial concentration I established, 50 times the concentration of PBP used to model the blue trace. The red line is the predicted response at 500 times the concentration of PBP used to model the blue trace. Experimental results (solid lines, error bars) are overlaid on the model predictions. The model predicts increasing maximum signal as PBP increases, but reaches a ceiling that the experimental results do not show. The model does not predict the lower detection threshold.*

(500x RBP). An important next step is to experimentally determine the RBP concentration in each cell line and establish the actual quantity of RBP produced by these modifications. This will not only support the conclusion that the concentration of PBP impacts maximum signal and threshold of detection in the synthetic signaling system, it will provide a useful link between the experimental and model data. Data is currently displayed as normalized fold change response

(response in the on state normalized to response in the absence of ligand), but if it is possible to connect the model to experimental molecular quantities, it may be useful for downstream applications.

P_{lacIQ1}::10xRBS::RBP, on the other hand, which putatively expresses 500x RBP, has a higher maximum signal and a lower detection threshold than the model predictions. The model reaches a ceiling at about 10x PBP, after which additional increases in PBP do not significantly impact the maximum fold-change response of the synthetic signaling system, as seen in Figure 5.6. This suggests a limiting parameter in this regime of the model. One modifier is the quantity of response regulator in the system. An increase in response regulator is predicted to directly increase the maximum signal, as shown in Figure 5.12. This prediction will have to be experimentally validated, and there may be unintended consequences to this increase. One possibility is that an overabundance of response regulator initiates transcription in the absence of ligand. In this case, more RR will increase both background and maximum signal and so will not yield a true increase in fold change .

Figure 5.12 shows that by increasing the RR concentration, the discrepancy in maximum signal between the model and the experimental data can be addressed, but it does not address the lower threshold of the 500x RBP signaling system. In order to investigate this limiting parameter, it will be necessary to reaccomplish the global parameter scan at two PBP concentrations, one at the threshold and one above. When the limiting parameter is modified, it should result in a difference in response between the two PBP values.

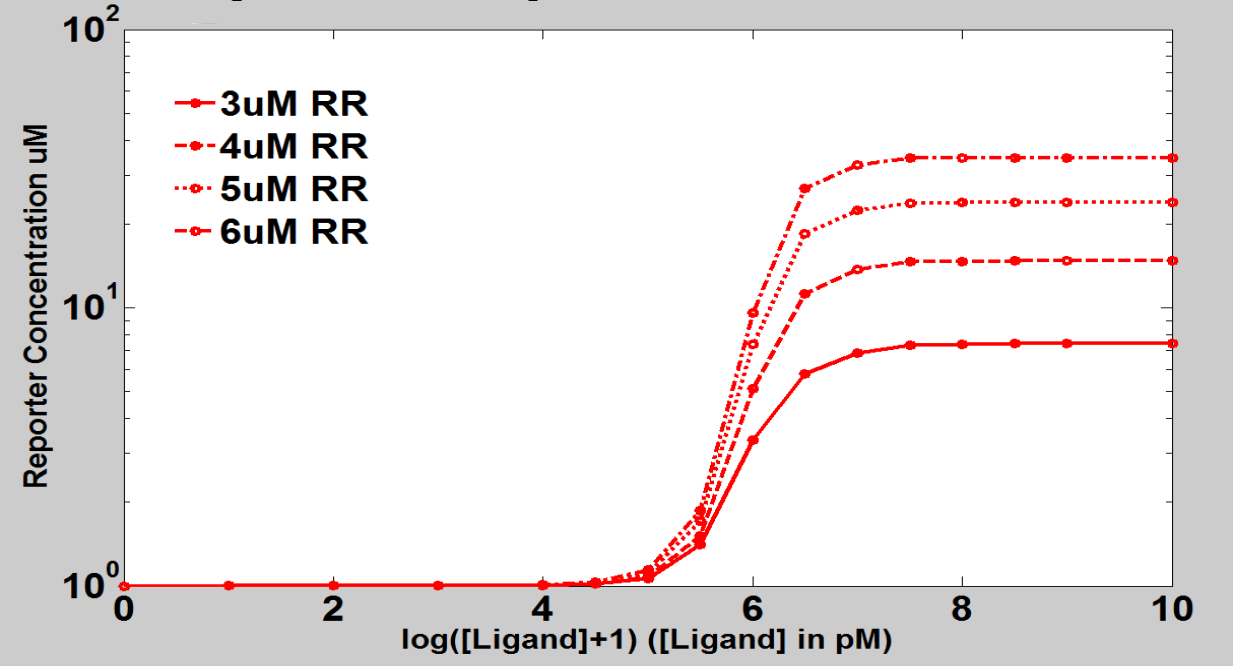


Figure 5.12. Model simulation of synthetic signaling system response as RR concentration increases. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. As the concentration of RR increases, the maximum signal increases. MODEL PREDICTIONS: predicts system response in the presence and absence of ribose using initial conditions as described in Table 4.1 and 4.2, with the exception of the initial concentration of RR, which has been increased by 50%, and ktr, which has been increased 5-fold. Responses are normalized to the predicted response in the absence of ribose.*

5.4 Revisiting parameter *ktr*

The model predicted that a change in fully induced rate of transcription (*ktr)* would have a substantial impact on the transcriptional response generated by the synthetic signaling system. I initially used this parameter to align the model with experimental results, changing the value of *ktr in silico*. However, if we were able to develop a promoter with a stronger rate of transcription when induced, but with a similar basal rate of transcription to PstS, we could use that promoter to again increase the maximum signal generated by the synthetic signaling system response.

PstS was selected to control the transcriptional response of the synthetic signaling system specifically because it has a low background, and strong induction (approximately 100 fold over background), and as such is one of the strongest endogenous PhoB responsive promoters[71]. I cannot, therefore, replace PstS with a stronger wild-type PhoB inducible promoter from the *E. coli* genome. However, I can use existing parts to engineer new promoters.

5.4.1 Building a synthetic Pho responsive promoter

The PhoB dimer regulates the expression of over 40 genes in the PHO regulon[66]. Once phosphorylated, PhoB dimerizes and binds to DNA sequences known as PHO-boxes[55]. Each promoter has a unique PhoBox, which interacts with a unique -10 element. In most of these promoters, one PhoBox replaces the -35 element, but the PstS, psiE, and ugpB promoters all contain two PhoBox. To determine if the PhoBox itself impacted transcription, Pete Bowerman, PhD synthesized new PstS promoters, replacing the PstS PhoBoxes with PhoBoxes from PhoA, PhoB, PhoE, PhoH, ugpB, as well as the consensus PhoB binding sequence. In cases where PstS PhoBoxes were replaced with the consensus PhoBox or PhoBoxes from promoters containing only a single PhoB binding site, both sites in PstS were replaced with the same PhoB binding site. I used these promoters to control the synthetic signaling system response, building reporter plasmids with the synthetic Psts.PhoBox promoters 5' of the GFP gene.

5.4.2 Results

All cells were co transformed with the 50x PBP, P_{LacIQ1}::RBP, synthetic signaling system, and one of the six synthetic PstS promoter::GFP plasmids described here. I induced the synthetic signaling system as described in section 4.3.1, and response to ribose induction was measured as described in section 4.3.2. KLH218 could not be co-transformed or tested. Figure 5.13 shows

the raw fluorescence values of the synthetic signaling system responses as translated through the various promoters.

Table 5.1 Reporter plasmids, controlled by synthetic PstS variants.

Plasmid Name	Forward Primer	Reverse Primer	Components	Description
KLH213	62	61	PstS.Consensus::GFP	Synthetic PstS promoter with two consensus PhoBoxes
KLH214	62	61	PstS.PhoH::GFP	Synthetic PstS promoter with two PhoH PhoBoxes
KLH215	62	61	PstS.PhoA::GFP	Synthetic Psts Promoter with two PhoA PhoBoxes
KLH216	62	61	PstS.ugpB::GFP	Synthetic PstS promoter with ugpB PhoBoxes
KLH217	62	61	PstS.PhoB::GFP	Synthetic PstS promoter with two PhoB PhoBoxes
KLH218	62	61	PstS.PhoE::GFP	Synthetic PstS promoter with two PhoE PhoBoxes

While these synthetic promoters express GFP at levels three to four times higher than the wild-type PstS promoter in the presence of ribose, they also have an apparent increase in the basal rate of transcription. In addition, none of the synthetic PstS promoters improved sensitivity of the synthetic signaling system. Figure 5.14 shows that of all the PstS promoters tested, the wild-type PstS promoter has the greatest fold change, although the Psts.PhoB promoter is a close second. Promoters are a function of the physical properties of the DNA sequence as well as the binding of transcription factors; by modifying the PhoB binding sites, I appear to have changed the basal transcription rate as well.

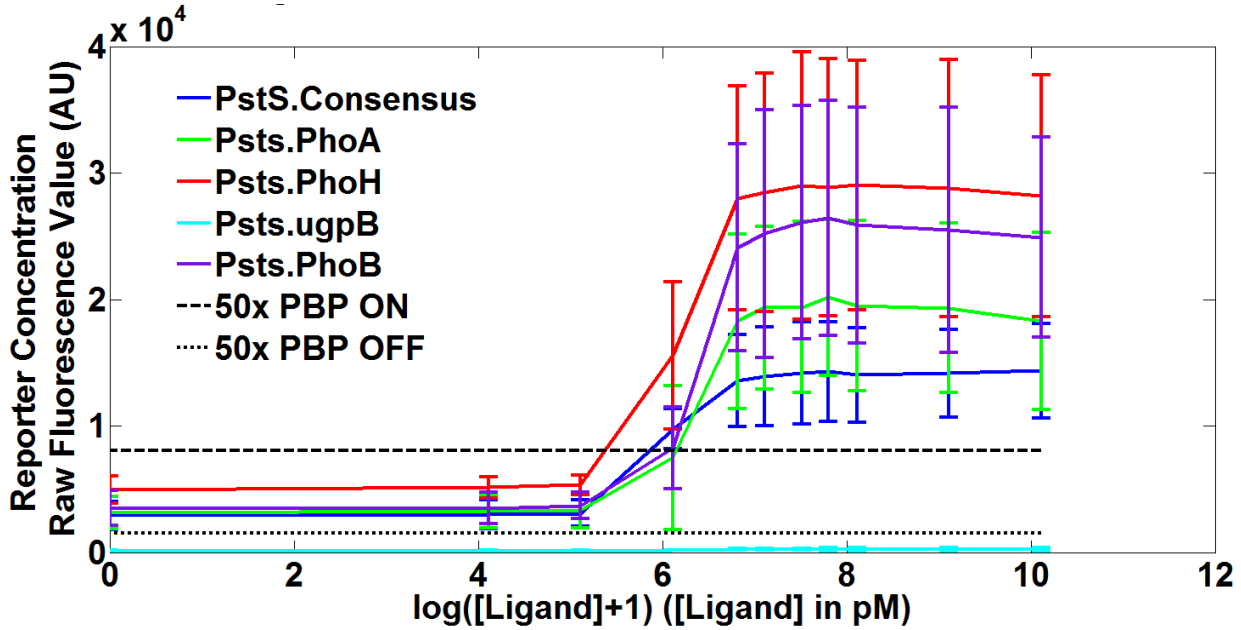


Figure 5.13: Raw data of synthetic PstS promoter activities. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows raw fluorescence units as captured by the flow cytometer. The dotted line (black, 50x PBP OFF) represents the average basal level of the original PstS promoter; the dashed line (black, 50x PBP ON) represents the average induced transcription level of the original PstS promoter. All promoters were co-transformed with the 50x signaling system into LT002 cells and inoculated as described in the methods. Cells were read on the flow cytometer, the geometric mean of the area was recorded. Each trace represents the average of at least 50,000 events generated per each of three unique transformation events. All transformation events were accomplished on different days.*

The model showed improved system behavior only for a promoter with a greater induced transcription rate, but an identical basal rate of transcription. No synthetic promoter tested has a similar basal rate to the PstS promoter, so I am unable to test our model prediction with these promoters. However, I can still draw useful information from these experiments. First, should I need to increase the output of the response in the bacterial system, PstS.PhoB is a good candidate to maintain fold change but increase total protein production. Second, there is no endogenous PhoB promoter in plants, and the Medford Lab has had to construct a plant synthetic promoter, PlantPho. Plant Pho uses four tandem consensus PhoBoxes upstream of a recognized minimal

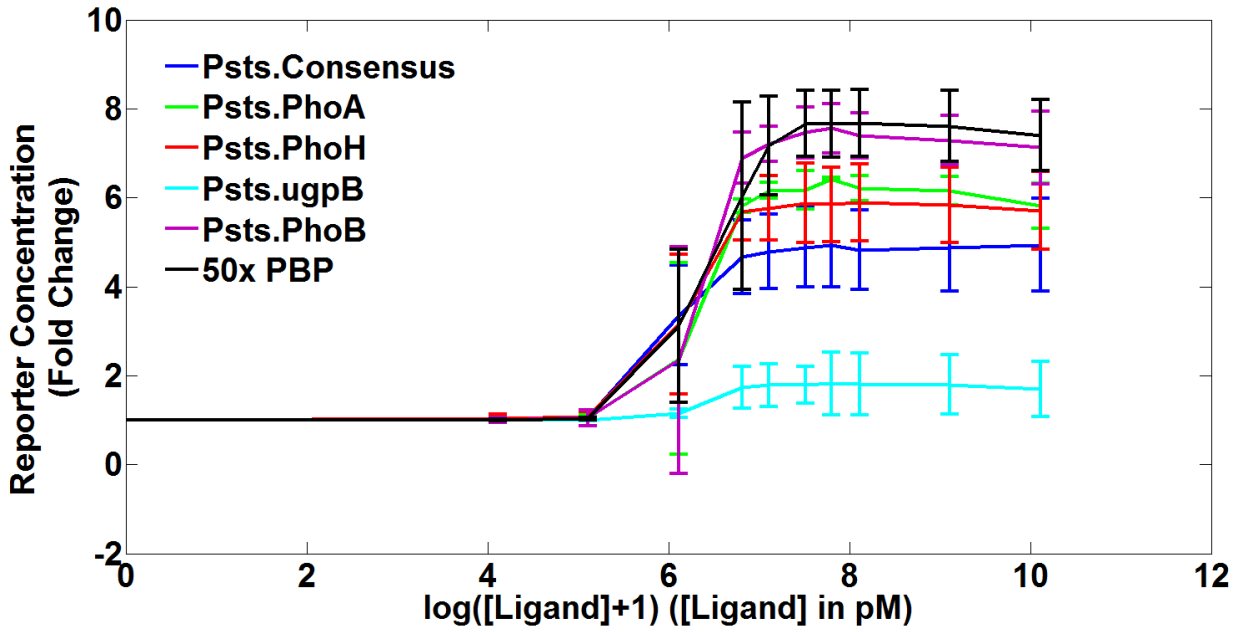


Figure 5.14: Fold change of response generated by synthetic signaling system using synthetic PstS promoters. *The x-axis represents the log of the ribose concentration in picomolar, and the y-axis shows raw fluorescence units as captured by the flow cytometer. All promoters were co-transformed with the 50x signaling system into LT002 cells and inoculated as described in the methods. Cells were read on the flow cytometer using a protocol with a lower voltage on the F1 channel, Ecoli_Gfp_BigResponders. The geometric mean of the area was recorded. Each trace represents the average of at least 50,000 events generated per each of three unique transformation events.*

plant promoter (the -46 region of the CAMV35S promoter). The PstS.Consensus promoter is one of the weakest tested here, both in terms of total protein output (raw GFP fluorescence) and fold change. Modifying the PlantPho promoter to incorporate PhoB or PhoH PhoBoxes may result in more robust plant signal.

5.5 Validating the model predictions in plants

The true test of this signaling system model will be the experimental demonstration of signal dependence on RBP in plants. Novel synthetic promoters have been developed by the Medford lab with higher constitutive expression than the 35S constitutive promoter from the cauliflower mosaic virus in transient protoplast assays (unpublished). Two of those promoters have been

incorporated into the synthetic signaling system and now control the expression of RBP in plants.

5.5.1 Experimental protocol in plants

While the potential exists to use flow cytometry to measure transcription activity in plant protoplasts (plant cells without their cell walls)[109], the synthetic signaling system relies on the external PBP to initiate the signal in the presence of ligand. If the cell wall is stripped away, the PBP will no longer be confined to the apoplast of the plant and will instead freely diffuse throughout the protoplast solution. This is likely to reduce RBP concentration to such an extent that we would be unable to gather any useful data on the synthetic signaling system.

Instead, we made use of another precise instrument, a charge-coupled device (CCD) camera capable of measuring the minute amounts of light produced by the enzyme luciferase. Luciferase is responsible for firefly and marine creature bioluminescence, generating light by oxidizing the substrate luciferin[110]. Luciferin provided exogenously to a plant that constitutively expresses luciferase results in distinct light production detectable by CCD camera[111]. Luciferase was therefore established as the PhoB inducible reporter for the plant synthetic signaling system.

The synthetic signaling circuit was introduced into a plant transformation vector and transformed into *Agrobacterium.* Standard floral dipping[112] was used to transfer the synthetic signaling system circuit to *Arabidopsis* embryos, inserting the synthetic signaling system randomly in the genome. These stable transformants are plated on MS+ media agar with 1% sucrose containing 100 µg/µl kanamycin and 100 µg/µl cefotaxime and germinated under a 16/8 hr light-dark cycle. KLH805, containing the synthetic signaling system with the strongest constitutive promoter expressing RBP showed a marked decrease in transformation efficiency (16 plants) when

compared to KLH804 (mid-level promoter, 40 plants) and KLH803 (CaMV 35S promoter, 49 plants). Forty KLH804 and KLH803 plants were selected for testing, and all 16 KLH805 plants were tested.

After 10 days, T0 plants were transferred to fresh MS+ media agar plates with 1% sucrose, 100 µg/µl kanamycin and 100 µg/µl cefotaxime, and 100 µg/µl luciferin. A small percentage of plants (two from KLH803 and one from KLH804) did not survive transplantation. Transplanted T0 plants grew for 8 days under a 16/8 hr light-dark cycle, until first and second leaf pairs had emerged on all seedlings. First and second leaves were then excised from the plant at the petiole with a No. 10 scalpel blade and placed, abaxial side down, onto MS+ agar plates with sucrose, 100 µg/µl luciferin, and 200 mM ribose (ligand) or sorbitol (control). Leaves were placed flat to ensure good contact with the media, and the petioles were inserted into the agar to ensure the vasculature of the leaf had access to the ligand.

Leaves were dark-adapted for 30 minutes to eliminate background fluorescence from photosynthetic activity, then imaged with the luciferase camera 4 hours after induction and every 20 hours thereafter. In between images, leaves were placed on a light rack with a 16hr/8hr light/dark cycle, and prior to each image leaves were dark adapted for 30 minutes. The images shown in Figure 5.15 were taken at 44 hours after induction and show accumulated light detected over 15 minutes of imaging.

Several variables must be taken into account when interpreting the results of this assay. The promoters used to increase expression of RBP have been demonstrated in transient protoplast assays (unpublished), but have not been stably integrated in plants before. Transient assays do not always reflect the performance of genetic elements in plants. For example, the 35S promoter

5.5.2 Results

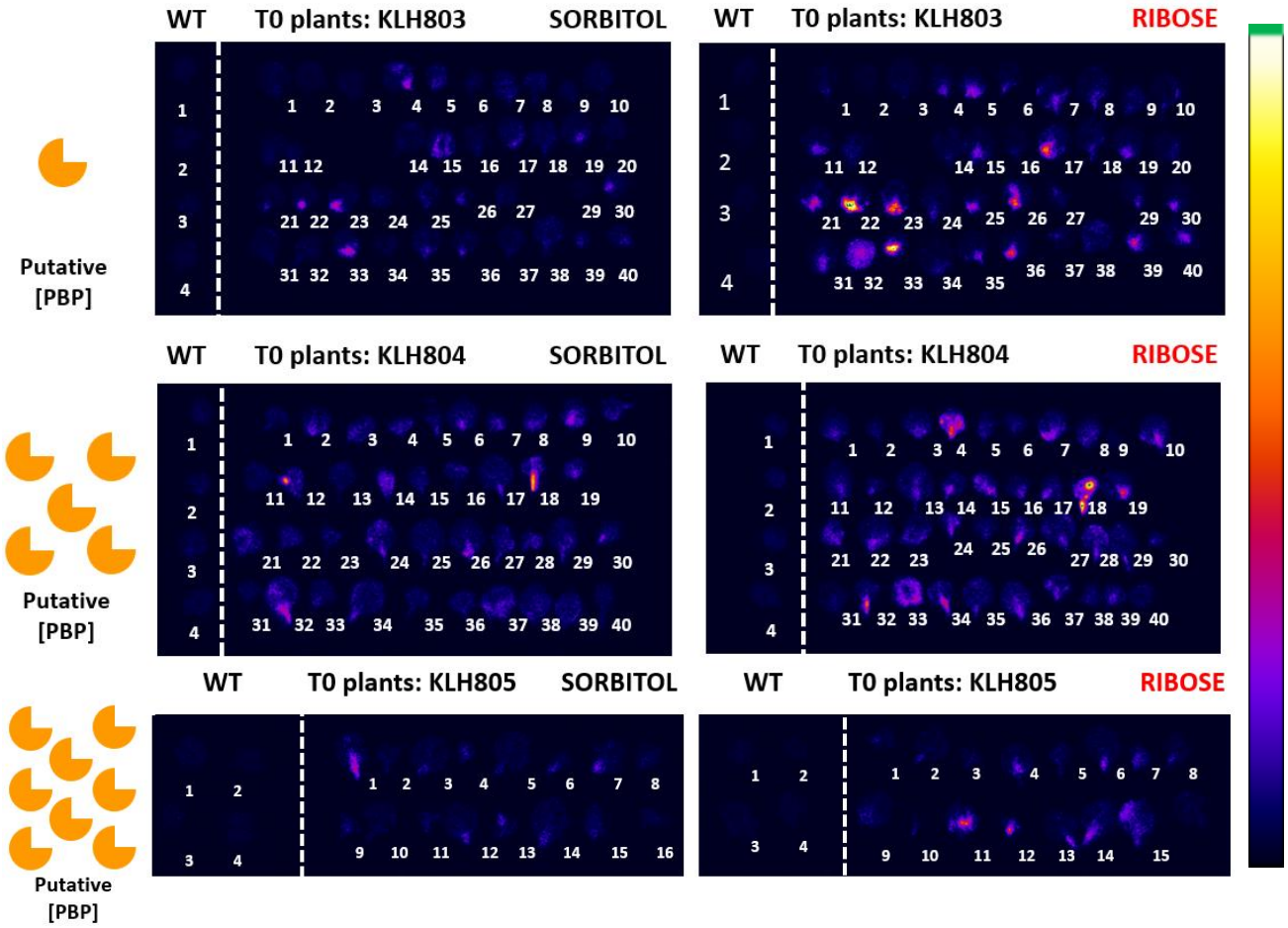


Figure 5.15. Luciferase activity imaging of T0 plant lines containing (from top to bottom) plasmids KLH803, KLH804, KLH805. *On the left is the putative concentration of PBP produced by the synthetic signaling system in each line. On the right is the false colored luciferase scale. Black and blue are no to low levels of light detection, green is light saturation. Wild-type leaves (left of the dotted line in all luciferase images) are uniformly on the lower end of the light detection scale, and expected for a plant line that does not express luciferase. Leaves from T0 plants show luciferase activity at various levels (purple, pink, and yellow). Matching numbers on sorbitol and ribose plates indicate leaves from the same T0 plant.*

from cauliflower mosaic virus is fairly weak in monocot protoplasts, but it expresses robustly in stable transformants[113]. In addition, *Agrobacterium* transformation inserts DNA into the *Arabidopsis* genome in a random fashion, resulting in T0 plants with a wide range of expression

levels[114]. Therefore, it is unlikely that we would see the a uniform increase in luciferase activity levels for plants expressing synthetic signaling systems with RBP. Furthermore, differences in the way the petiole was cut or percent of leaf contacting the plate may vary, resulting in different levels of luciferin uptake.

Mathematical analysis of the data comes with its own challenges. The luciferase camera assigns a value to each pixel based on the quantity of light detected in that portion of the visual field. That value can be analyzed to quantitatively determine the relative light production in various portions of the image. However, each leaf is a different size, and a simple summation of pixel values may result in discrepancies. For example, a large leaf expressing a low level of luciferase may have the same total light detected by the camera as a small leaf expressing a lot of luciferase. Normalization by area is also misleading as light is not produced uniformly throught each leaf, generally producing more light closer to the vascular tissue.

5.5.3 Analyzing the T1 generation of synthetic signaling system

Despite these complications, these plants can still be used to validate or invalidate the hypothesis that increased RBP increases synthetic signaling system activity in the presence of ribose in plants. Based on visual comparisons of luciferase activity of leaves from the same plant on ribose and sorbitol, 20 KLH803 plants, 26 KLH804 plants, and 7 KLH805 plants were transferred to soil and allowed to set seed. The T1 generation will yield plants heterozygous and homozygous for the synthetic signaling system, allowing us to test multiple plants with similar genetic background, thereby mitigating the sources of human error noted above and generating better data on synthetuic signaling system activity.

5.6 Future Work

5.6.1 Improving plant experiments

While reverse transcriptase PCR (RT-PCR) could be performed on these lines to correlate signaling strength with PBP expression level, that experiment would not take into account the decreased or increased expression of other synthetic signaling system elements. RBP and all other elements of the synthetic signaling system are incorporated into one T-DNA. If that T-DNA is inserted in an area of the *Arabidopsis* genome that is weakly expressed, all of the signaling system components will be weakly expressed. The stoichiometry of the components will remain the same, regardless of the level of expression.

An alternative method of validation would be to control RBP expression with an inducible promoter. The small molecule 4-hydroxytamoxifen (4-OHT), for example, is taken up by plants and can effect a transcriptional response[115]. In an inducible system, the expression of RBP in the presence of the inducer will be greater than the expression of RBP of the same plant in the absence of inducer, regardless of where the T-DNA was inserted in the genome. This effectively controls for the variability of T-DNA insertion.

5.6.2 Improving bacterial test platforms by using inducible promoters

Similarly, inducible promoters could be used to fine-tune the synthetic signaling system in bacteria. The synthetic signaling system that showed the largest response in the presence of ribose expressed RBP under the control of the P_{LacIQ1} promoter, a 50-fold increase over the original system transcription strength, and was translated using an RBS predicted to be 10-fold stronger than the original system translation rate.

Cells transformed with these plasmids were very slow to grow, and often colonies would generate no fluorescent signal at all. Previous work in the Medford Lab has shown that cells producing maltose binding protein (MBP) at very high levels were found to recombine, precluding experimental work (unpublished data). It is possible that in the synthetic signaling system the high metabolic load incurred by overproduction of RBP was unsustainable, or the high concentrations of the protein itself may have a toxic effect on the cells.

An inducible promoter would enable an exploration of the full range of PBP production with one promoter, identifying the optimal concentration of PBP and thereby directing the development of an appropriate constitutive promoter. pBAD cannot be used in this cell line, but the IPTG inducible system may be useful. IPTG cannot be metabolized by *E. coli,* and induces expression of the P_{lac} promoter without initiating the positive feedback loop common to sugar based repressible gene expression systems[104].

5.6.3 Improving bacterial test platforms by incorporating orthogonal components

The synthetic signaling system uses bacterial components derived from the chassis they are tested in. While I am using cell lines with deletions in proteins known to activate synthetic signaling system components[37], there may be endogenous interactions unaccounted for. Parts from other bacterial strains and species could be incorporated into the synthetic signaling system to make it more orthogonal to both our bacterial and plant chassis.

The *B. subtilis* PhoR/PhoP phosphate sensing system shares particularly high homology with the *E. coli* PhoR/PhoB TCS[116]. Twenty-seven percent of *B. subtilis* PhoR residues are identical to the *E.coli* PhoR, and another 20% are conserved. In particular, 66% of the cytoplasmic portion of *B. subtilis* PhoR is identical to *E. Coli* PhoR. *B. subtilis* PhoR's cognate response regulator is

PhoP, the *B. subtilis* homolog to PhoB. Similar to PhoB, PhoP is phosphorylated by its cognate HK in low inorganic phosphate conditions, dimerizes and regulates expression of the *pho* operon in *B. subtilis* [117].

PhoP is unable to activate the *E. coli pho* operon, and PhoB cannot activate the *B. subtilis pho* operon[117], suggesting that a synthetic signaling system based on the *B. subtilis* TCS will be relatively independent of interference from wild-type *E. coli* proteins. As an additional benefit, *B. subtilis* PhoP has been shown to increase the activity of the *B. subtilis* PstS promoter 5000 fold[118], whereas the Pho responsive promoter used in the synthetic signaling system, *E. coli* PstS, is only induced 100 fold by *E. coli* PhoB[71]. *B. subtilis* components in the synthetic signaling system therefore may increase orthogonality and potentially increase the difference between the ON and OFF states of the synthetic signaling system.

5.6.4 Understanding the applicability of the model to other systems.

One of the benefits of models is their ability to make predictions about similar systems with different parameters. The modularity of the synthetic signaling system is a great advantage in building detectors for relevant compounds, and there are numerous HK fusions and PBP re-designs that could be developed to build detectors with new capabilities. The model can be adapted to reflect the binding affinities and rates of the new system, but these changes will only result in an effective, predictive model when the new system has the same underlying architecture as the original synthetic signaling system.

One variant of the synthetic signaling system demonstrates the ability of the model to predict the behavior of multiple systems under certain circumstances. In this system, RBP was replaced with a glucose/galactose binding protein (GBP), as shown in Figure 5.16. Like RBP, GBP also effects

chemotaxis through Trg[61]. Therefore the only difference between a glucose/galactose synthetic signaling system and a ribose synthetic signaling system is the PBP.

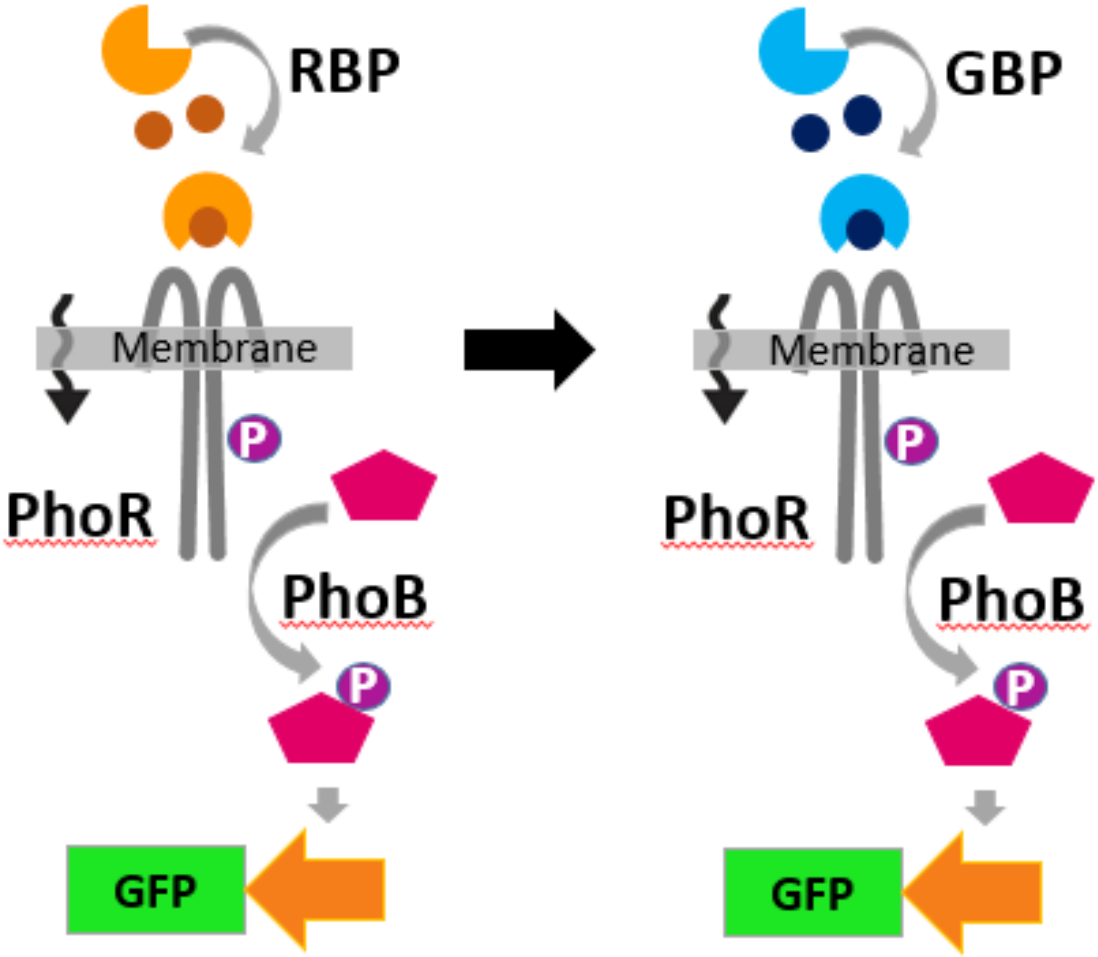


Figure 5.16. Schematic of ribose and glucose/galactose inducible synthetic signaling system. *All components are identical save for the PBP – Ribose binding protein (yellow) has been replaced with the glucose/galactose binding protein (light blue).*

Preliminary experiments have indicated that the maximum signal of the GBP synthetic signaling system in the presence of galactose is also PBP-concentration dependent. Figure 5.17 shows that the response generated in the presence of galactose is similar to that of the RBP inducible system in the presence of ribose.

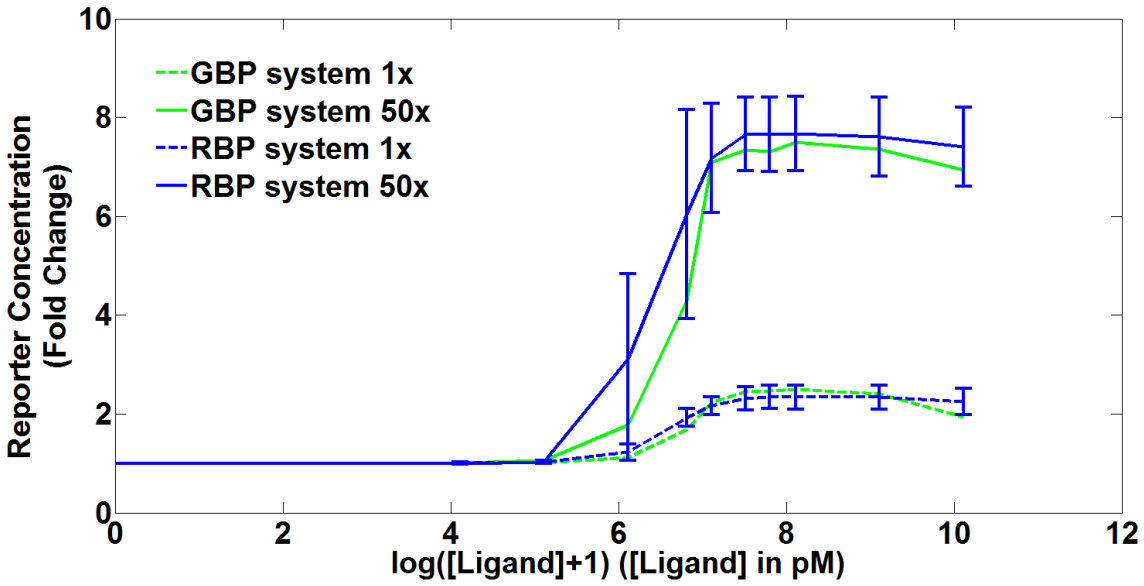


Figure 5.17. Response of RBP and GBP signaling systems in the presence of ribose and galactose. *1x systems have* P_{lacI} *controlling expression of the PBP, 50x systems use* P_{LacIQ1} to express the PBP. *The x-axis represents the log of the ligand (ribose for RBP systems and galactose for GBP systems) concentration in picomolar, and the y-axis shows fold change over background fluorescence in the absence of ligand. GBP data is a single replicate, RBP data is the average of three replicates from at least three independent experiments. Each replicate is the mean of 50,000 events*

As predicted by the model, the maximum response of the synthetic signaling system to ligand increases as PBP increases for both RBP and GBP synthetic signaling systems. Therefore, the model appears to effectively capture the behavior of the RBP system in the presence of ribose and the GBP system in the presence of galactose. This suggests that the RBP and GBP synthetic signaling systems are relatively unaffected by endogenous fluctuations in these cases, as only the synthetic signaling system is encompassed by the model. However, the GBP system has a very different response in the presence of another common sugar, glucose.

The ribose inducible system response was not dramatically affected by the carbon source, with the exception of L-arabinose (Figure 4.6). However, the impact of L-arabinose was likely the result of the induced cytoplasmic PhoR rather than an effect on the components of the RBP synthetic signaling system. When cells containing the GBP synthetic signaling system are incubated with glucose, the ligand-dependent response is no longer predicted by the model.

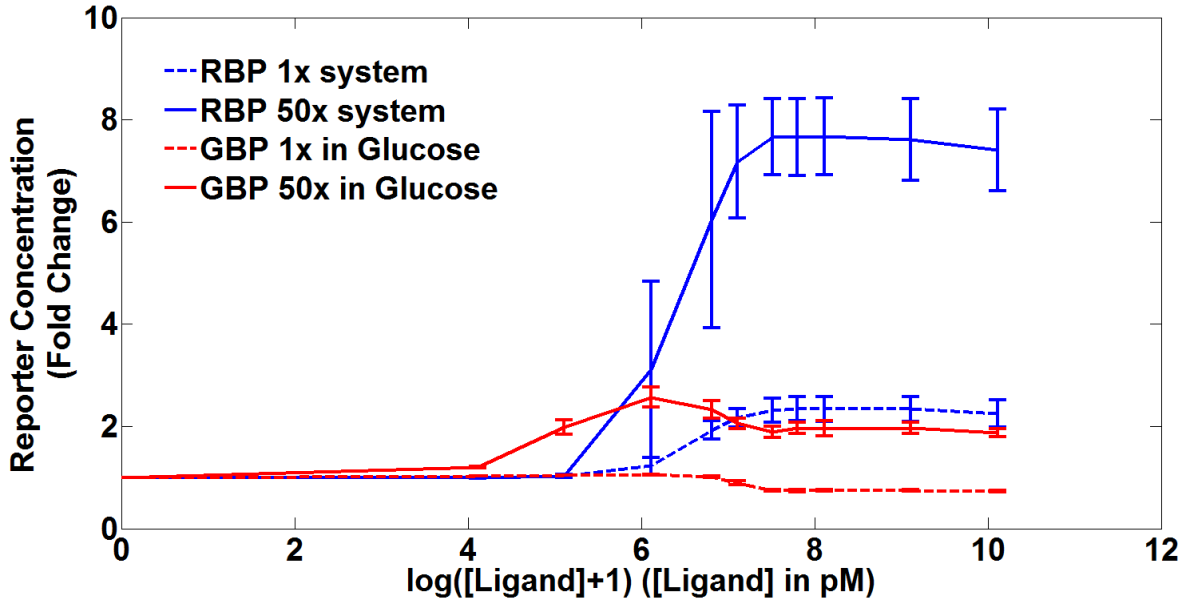


Figure 5.18. Response of RBP and GBP signaling systems in the presence of ribose and glucose. *1x systems have* P_{lacI} *controlling expression of the PBP, 50x systems use* P_{LacIQ1}. *The x-axis represents the log of the ligand (ribose for RBP systems and glucose for GBP systems) concentration in picomolar, and the y-axis shows fold change over background of arbitrary fluorescence units as captured by the flow cytometer.*

In the RBP system, the low PBP response (RBP 1x system, blue dashed line) increases in maximum signal and sensitivity as PBP concentration increases (RBP 1x system, blue solid line), with very little change otherwise in the transition between ON and OFF states. In contrast, the GBP system (GBP 1x, red dashed line) exhibits *decreased* GFP expression at high levels of glucose. In addition, this response is not enhanced by an increase in PBP. Rather, a new behavior

is observed in the 50x GBP system (red solid line), where the response peaks at intermediate levels of glucose and reaches a plateau at higher concentrations of glucose.

Glucose is a central metabolite of *E. coli*, and as such, is involved in reactions crucial to cell survival and growth. Because the GBP synthetic systems have been shown to function in accordance with the model in the presence of galactose, it is likely that changes to endogenous processes are responsible for this deviation from the expected behavior. One possibility is that *E. coli* responds to different concentrations of glucose in different ways, impacting the synthetic signaling system differently at different points in the ligand gradient. Death & Ferenci suggested that a starvation response is triggered when *E. coli* is grown in micromolar glucose concentrations[119]. They tracked extracellular glucose concentration and glucose transport operon expression of *E. coli* cultures grown in a carbon limited environment in a chemostat, and found that the glucose transport system is rapidly upregulated in micromolar glucose. The same system is repressed when cultures are exposed to millimolar concentrations of glucose. Interestingly, at high concentrations of PBP, the GBP synthetic signaling system reaches a peak at micromolar concentrations of glucose and reaches a plateau at millimolar concentrations of glucose. The physiological response of *E. coli* to various concentrations of glucose may be impacting the output of the synthetic signaling system.

Another variant of the synthetic signaling system uses a different PBP, maltose binding protein (MBP, Figure 5.19, purple). Maltose binding protein interacts with Tar, not Trg, and so a new HK had to be developed as well, fusing the extracellular sensing domain of Tar (Figure 5.19, dark blue) to the cytoplasmic portion of PhoR. The model is unable to predict the behavior of this system, as there is very little difference between the ON and OFF states (unpublished data).

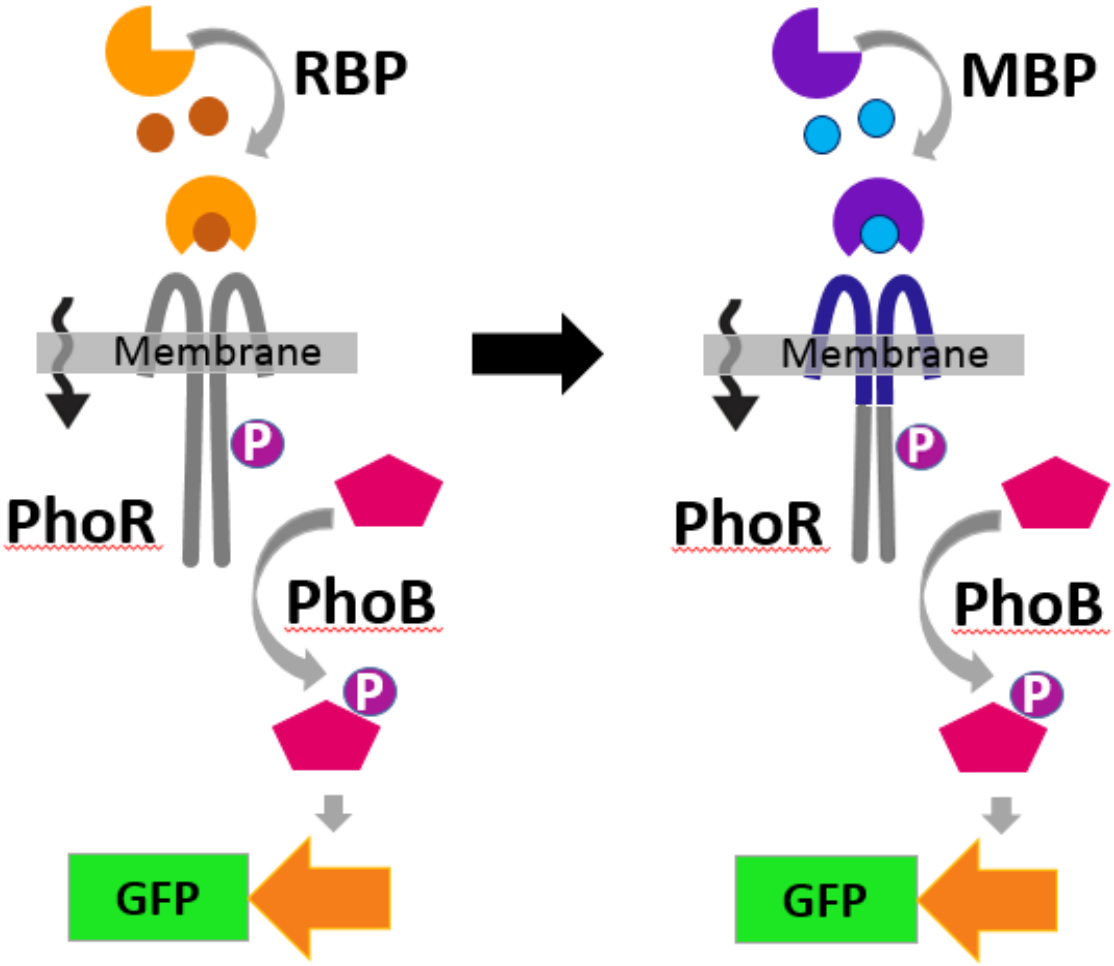


Figure 5.19. Schematic of RBP and MBP synthetic signaling circuits. *RBP has been replaced with MBP (purple), which binds maltose (light blue). The fusion HK has been adapted to bind the maltose-MBP complex, by replacing the extracellular chemotactic sensor domain from Trg with one from Tar. All other components, including the majority of the cytoplasmic PhoR portion of the HK, are identical to those used in the ribose inducible system.*

If the underlying mechanisms of this new HK are the same as the original signaling system, the model can suggest rationales for this regime: the HK autophosphorylation rate could have increased as a result of the new fusion and the system is always ON, or perhaps the interaction between the HK and the RR has changed. If however, this signaling system does not share the same mechanisms (*i.e.*, the extracellular chemotactic sensor domain binds ligand directly rather

than through a PBP, or PBP bound to ligand preferentially interacts with a transporter instead of the signaling system), the model will be insufficient. To effectively apply the existing model to a new synthetic signaling system, work must be done to identify whether a system lies within the model framework but has undesirable parameters or it involves another mechanism altogether. Alternatively, a new model can be developed to incorporate new architectures or answer new questions.

CHAPTER 6: BUILDING MODELS FOR SPECIFIC APPLICATIONS

The original RBP synthetic signaling system showed a two-fold induction of GFP upon exposure to ribose (Figure 4.3). While two-fold induction is a statistically significant result, a broader separation between the ON and OFF states is desirable. If the background is very close to the response, it is possible that stochastic mechanisms will transition the system from the OFF state to the ON state in the absence of ligand[120]. A false positive of this sort would limit the effectiveness of the synthetic signaling system as a detector.

In order to build a robust detection system, it is necessary to increase the difference between the background of the synthetic signaling system and the response. This can be achieved by decreasing background, increasing response, or some combination of the two. A method to increase maximum response, increasing the concentration of PBP, has already been explored in this thesis. In this chapter I consider the alternative, reducing expression of GFP in the absence of ligand by first identifying the source of basal transcription, and then developing a model to test approaches *in silico.*

6.1 Identifying the source of background transcription

While the PstS promoter has been shown to be tightly regulated[71,72] it is possible that the PstS promoter is 'leaky' enough to cause the background fluorescence the synthetic signaling system produces in the absence of ligand. If this is the case, then cells transformed with only the GFP plasmid (*i.e.*, no synthetic ribose signaling system) should exhibit the same level of fluorescence as those transformed with the synthetic signaling system in the absence of ribose. Figure 6.1 instead shows that cells containing only the GFP reporter plasmid (green) have a very similar

profile to cells grown without any plasmid (black). This suggests that the PstS promoter is not responsible for the majority of basal expression in the synthetic signaling system.

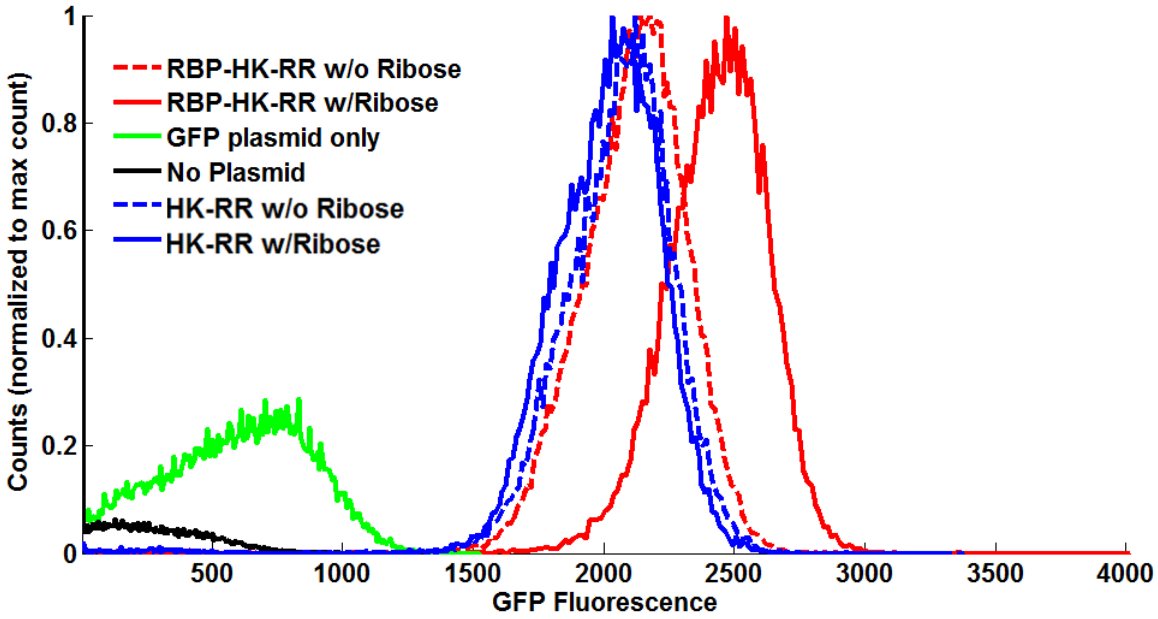


Figure 6.1. Histograms of cells containing the synthetic signaling circuit, a partial synthetic signaling circuit, the reporter plasmid only, or no plasmid, with and without ribose. *X-axis shows the amount of GFP fluorescence in arbitrary units, y-axis is number of cells. Cells without any plasmid (black) show very little GFP fluorescence (far left of graph). Cells with the full signaling system and GFP reporter plasmid are shown in red. The basal level of transcription (cells grown without ribose) is shown in dashed lines, and the induced transcription level is shown as a solid line. Cells containing only GFP plasmid (green), show slightly more fluorescence than the cells without plasmid. Cells with the GFP reporter plasmid and a truncated signaling system (blue, HK and RR only), show nearly as much basal transcription as the full signaling system. Histograms are normalized to the mode and contain at least 50,000 events*

An alternative explanation for the basal level of GFP expression is that the PstS promoter is being activated by some basal amount of phosphorylated response regulator. PhoB can be phosphorylated in the absence of ribose by either the synthetic signaling system HK, or by endogenous *E. coli* components capable of cross talk with PhoB. To distinguish between these two sources of phosphorylation, I constructed plasmids without one of the three signaling

components and compared the response of these knockouts with the response of the complete synthetic signaling system.

In systems lacking either the response regulator or the histidine kinase, basal fluorescence is approximately twice that of a cell containing only the reporter plasmid. This can be considered the basal level of endogenous interaction, *i.e.*, the noise due to the histidine kinase interacting with an endogenous transcription factor that triggers the system, or cross talk between PhoB and other histidine kinases (data not shown). However, Figure 6.1 shows that systems without a RBP (*i.e.*, incapable of binding ribose and initiating signal transduction through the synthetic signaling system), but containing both a functional HK and RR show levels of GFP fluorescence (blue, figure 6.1) very similar to the OFF state of the complete synthetic signaling system (bed, Figure 6.1). Therefore, the source of the high background level is primarily due to ligand-independent interactions between the synthetic signaling system HK and RR. However, non-cognate HKs or other phosphorylating factors could still play a small role in phosphorylating PhoB in bacteria.

6.2 Building the R^3 Model to explore a specific experimental strategy

Antunes et al described a plant line containing a variant of the synthetic signaling system designed to detect TNT in their 2011 paper[13]. The best responding plant line was found to have multiple copies of the reporter cassette, and therefore, multiple binding sites for the response regulator. The question was posed: "Would deliberately increasing the number of reporter cassettes result in a greater fold-change in the response of the synthetic signaling system?" A greater number of binding sites may saturate out the PhoB concentration at low levels of phosphorylation, effectively decreasing background. In addition, more fully bound promoters expressing reporter could generate more GFP in the presence of ribose. However, these

assumptions rely on a large number of variables, including the amount of response regulator available to bind to the promoters, the leakiness of the promoter, and the affinity of the response regulator for the promoter. I modified our existing model to explore these permutations. This modified model is referred to as the R^3 model.

6.3 R^3 Model Assumptions

A model for a particular system can indicate gaps in knowledge or suggest areas of investigation, as seen in the discussion of PBP concentration and its impact on the synthetic signaling system response. The detailed model of the synthetic signaling systems makes assumptions about the starting concentrations of the three primary components of the system, and determines all other species by the Law of Mass Action. Sometimes a direct link between cause and effect is clear – for example an increase in GFP expression in this model is most likely the result of an increase in phosphorylated PhoB concentration. The increase in phosphorylated PhoB, however, can be generated many different ways: changing the binding rate, changing the phosphotransfer rate, increasing the occupancy of the HK by PBP, and increasing HK autophosphorylation, among other parameters. Changing one parameter results in global changes to the model – phosphorylated PhoB may increase, but the HKp:RR complex may also increase, reducing the pool of available RR and limiting the formation of the HK:RR complex. These 'unintended consequences' will also affect the output of the model. Therefore, in order to answer a question about a sub-species of the synthetic signaling system, it is sometimes more appropriate to reduce the scope of the model and modify the sub-species directly.

I began by assuming that the synthetic signaling system would run to some steady state when exposed to a given concentration of ribose. At steady state, the total concentration of response regulator would be split into two populations, phosphorylated PhoB and unphosphorylated PhoB.

Instead of establishing the phosphorylated PhoB level through the Law of Mass Action equations in the detailed model, I instead initialize the model with concentrations of RRp and RR derived from the total amount of RR in the detailed model, and allow downstream reactions to occur using the same Law of Mass Action equations used in the detailed model. However, while the detailed model had been validated qualitatively, the explicit molecular steady state concentration of subspecies, and the ratio of one subspecies to another may not be accurate.

Ideally, I would observe changes in the pool of phospho-PhoB at varying concentrations of ribose directly, and thereby establish an experimental basis for the ratio of RRp to RR. However, previous work by the Medford lab was unable to replicate a method of phosphorylation specific gel retardation. Therefore, I could not accurately identify the percentage of response regulator phosphorylated in the OFF state as compared to the ON state.

In spite of this, I could make an assumption about the pool of RRp. The RBP synthetic signaling system has been shown to increase expression of GFP in the presence of ribose (Figure 4.3), suggesting that the concentration of RRp increases in the ON state. Hence, I defined the ON state as the ratio of RRp:RR is greater than 1, and an RRp:RR ratio less than 1 was defined as the OFF state.

Because I could not experimentally define a specific regime of RRp concentrations and ratios, the model needed to explore the impact of adding additional PhoB inducible promoters to the system in a variety of RRp contexts. To do so, every ON state concentration of RRp had to be compared to every OFF state level of RRp which met the assumptions above. To simplify the parameter space I considered fractions of a total conserved pool of RR. For example, an OFF state would be defined as having 10% of the total RR concentration phosphorylated. The response of the signaling system in the OFF state was then simulated and compared with the

response of the synthetic signaling system in each of the corresponding ON states, defined as 20% RRp, 30% RRp, and so on up to 90% phosphorylation. In order to use this model to explore the impact of additional PhoB inducible promoters, it was run iteratively and modified to increase the number of reporter cassettes in each iteration. The fold change response for each ON and OFF state pair is plotted as a function of number of reporter cassettes in Figure 6.2.

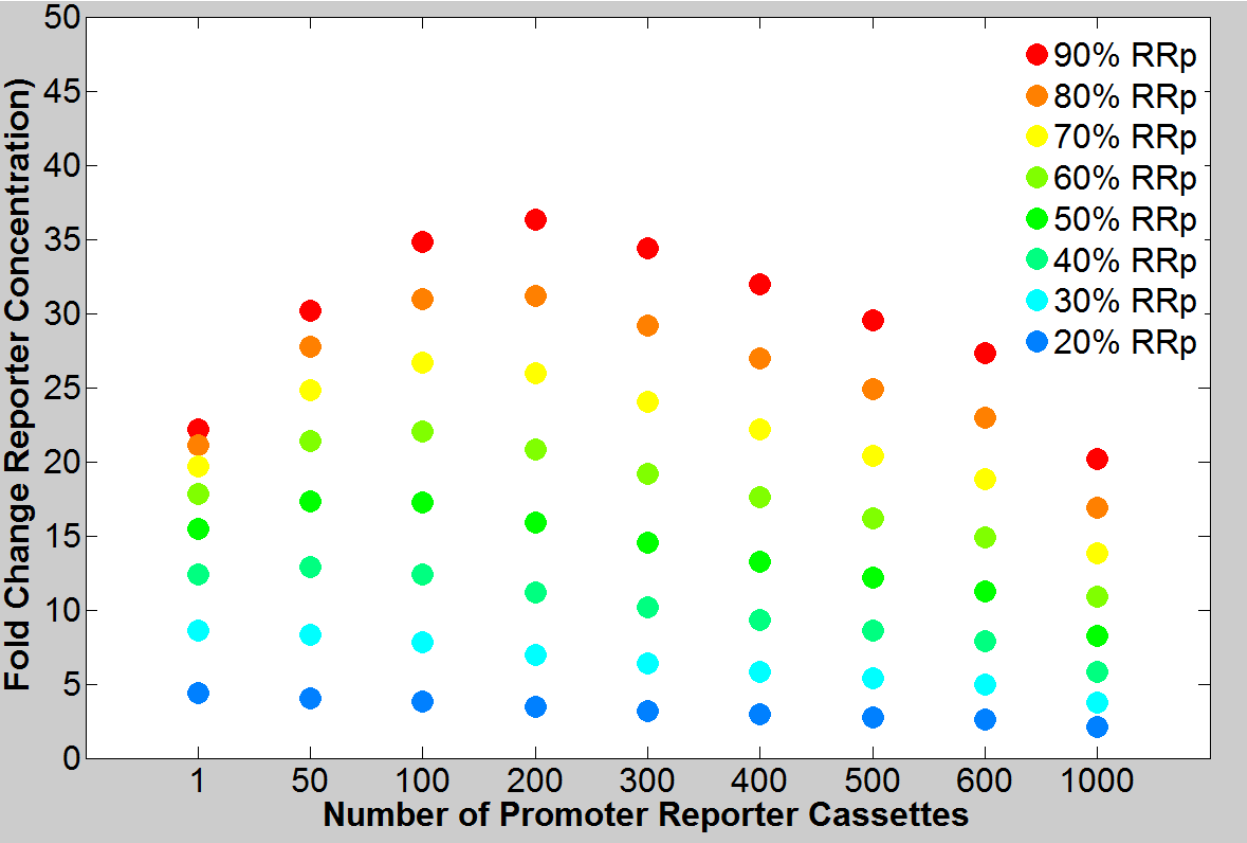


Figure 6.2. R^3 model results showing the fold change response of systems as a function of increasing PstS::GFP cassettes. *X axis represents the number of PstS::GFP reporter cassettes used in the simulation (i.e. the number of promoters controlling expression of GFP). Note the order of magnitude increases at the end of the scale. Y-axis represents total fold change for the system. For simplicity, the OFF state is held constant at 10% total RR is phosphorylated, but the trend described here pertains to all other combinations of ON and OFF states. Each ON state is denoted by a color. When the percentage of phosphorylated response regulator increases from 10% in the absence of ribose to 20% in the presence of ribose (dark blue), the number of promoter/reporter cassettes available to bind RR-P does not impact the fold change (i.e. the fold change of the system with one promoter reporter cassette is equal to the fold change obtained with six promoter reporter cassettes). When the percentage of phosphorylated response*

regulator increases from 10% in the absence of ribose to 90% in the presence of ribose (red), the number of promoter/reporter cassettes available to bind RRp does impact the fold change (i.e. the fold change generated with six promoter reporter cassettes is greater than the fold change obtained with one promoter reporter cassettes). The components have been scaled to clearly demonstrate the broad trend shown by the model: the concentration of RRp (40µM) is higher than predicted (3µM), as are the number of promoter reporter cassettes (100 in this model, plasmid copy number is approximately 20).

If the model was simulated with so many reporter cassettes that the phosphorylated response regulator could never saturate the promoters, the model returned a fold change of one in all scenarios, indicating that the off and the on states were the same. In addition, if the model had so much phosphorylated response regulator (or so few reporter cassettes), that the promoter was always saturated, the model also returned a fold change of one, indicating that the off and on states were again, the same.

Figure 6.2 shows a distinct trend for intermediate values of *PstS::GFP* reporter cassettes and RRp concentration: For very large changes in the pool of phosphorylated PhoB (*e.g.*, 10% of the response regulator phosphorylated in the OFF state, 90% phosphorylated in the ON state), increasing the number of reporter cassettes resulted in increased fold change between the off and the on states (red dots). In systems with small changes in the pool of phosphorylated PhoB (*e.g.*, 10% of the response regulator phosphorylated in the OFF state, 20% phosphorylated in the ON state), additional promoters did not noticeably increase the fold change between the OFF and the ON states (dark blue dots). There is a peak of fold-change response at some number of promoters for a particular RRp concentration (200 promoters for 90% phosphorylation in the ON state), but the number of promoters changes with the RRp concentration (50 promoters for 50% phosphorylation in the ON state).

Synthetic signaling system response to a particular number of PstS::GFP cassettes is also dependent on the total quantity of RR in the system. Consider the OFF state where 10% of 1µM

RR is phosphorylated and the OFF state where 10% of 10 μM RR is phosphorylated. The second case has an order of magnitude more molecules of RRp available to bind accessible promoters. The first case may show a maximum in response with two PstS::GFP cassettes, but those two cassettes will be saturated in the second case, and result in no difference between OFF and ON states. Therefore I am unable to identify an optimal number of promoter cassettes to capture the maximum difference between OFF and ON states of the RBP synthetic signaling system without knowing both the total concentration of RR in the synthetic signaling system and the proportion of RRp to RR in each state.

In order to generate Figure 6.2, I had to scale the components involved in the model. The effect shown above was almost imperceptible at the levels of RR defined in the full synthetic signaling model (3μM total RR), and more reasonable promoter quantities, as shown in Figure 6.3. At higher concentrations of promoter, the fold change trends towards one. However, the model predicts a fairly robust system at lower concentrations of promoter – a second or third reporter cassette does not significantly change the system response.

This is in direct contrast to data generated using the original model which was based on literature values from other dynamic models. In that model, maximum fold-change response in a system with large changes in the pool of phosphorylated response regulator could be modified by the introduction of additional reporter cassettes. Figure 6.4 shows the systems with large changes in response regulator (red) increase fold-change response as the number of promoters increases. Similar to Figure 6.2, when there are so many promoters the response regulators cannot effectively saturate a promoter and express GFP, there is no difference between the ON and OFF state of the system and the fold-change response trends towards one.

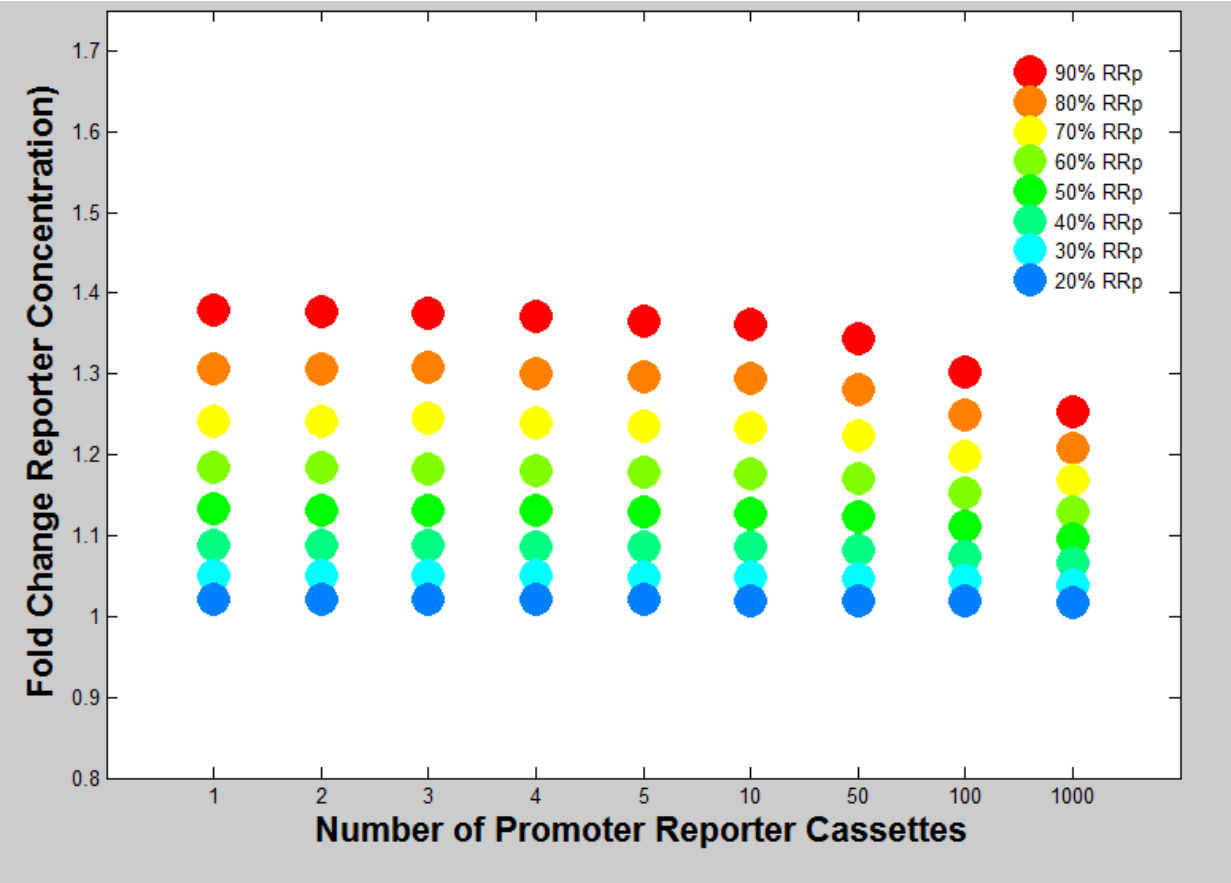


Figure 6.3. R^3 model results showing the fold change response of systems as a function of increasing PstS::GFP cassettes at concentrations relevant to detailed model. *X axis represents the number of PstS::GFP reporter cassettes used in the simulation (i.e. the number of promoters controlling expression of GFP). Note the order of magnitude increases at the end of the scale. Y-axis represents total fold change for the system.*

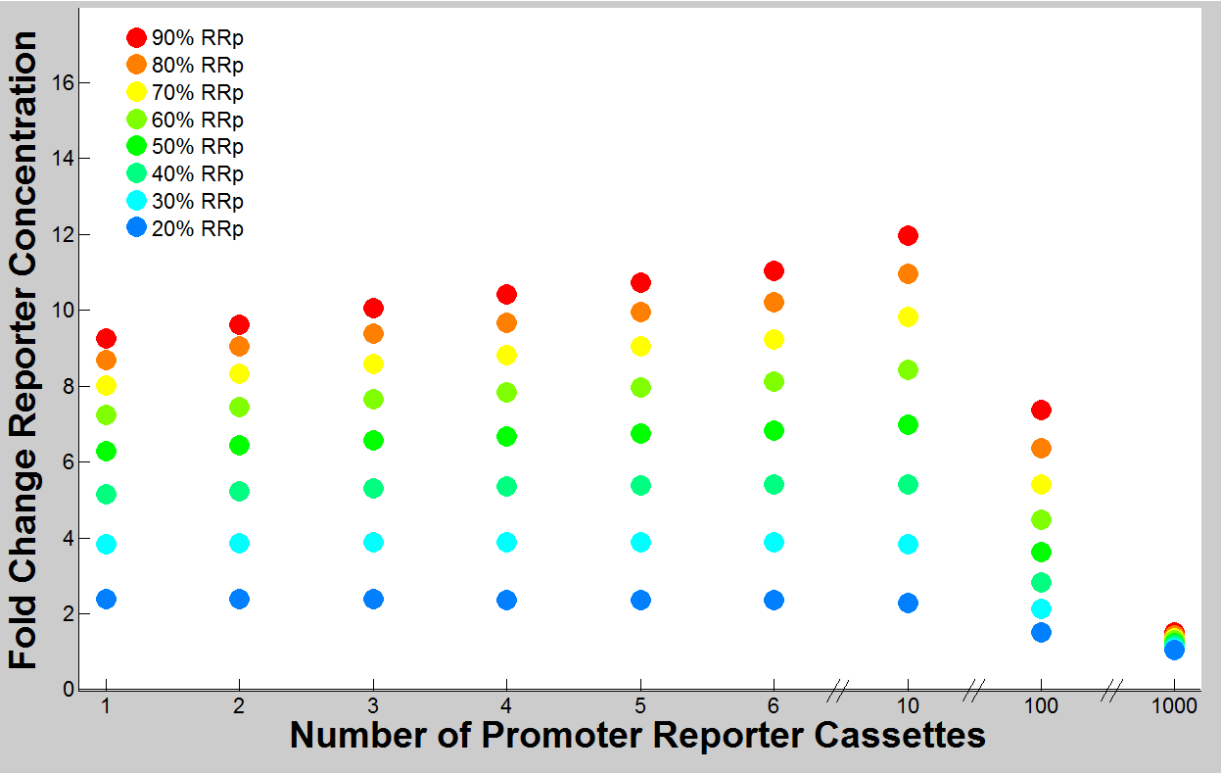


Figure 6.4. R^3 model based on original synthetic signaling model with 0.8μM total RR. *Results show the fold change response of systems as a function of increasing PstS::GFP cassettes at concentrations relevant to detailed model. X axis represents the number of PstS::GFP reporter cassettes used in the simulation (i.e. the number of promoters controlling expression of GFP). Note the order of magnitude increases at the end of the scale. Y-axis represents total fold change for the system.*

Despite the conflicts in scale, I decided to investigate these model predictions for the following reasons. The R^3 model is truncated, and as such does not contain additional phosphorylation or dephosphorylation pathways which may serve to maintain the pool of phosphorylated RR in the full signaling system. Therefore, it is possible that the concentration of RR used in the detailed model is not an accurate representation of the system in this model. In addition, although the scales are different, both old and new models showed a similar trend, suggesting that maximum fold-change response of systems with a large change in phosphorylated RR could be increased with additional reporter cassettes to some optimum level. The previous model had also indicated

that an increase in PBP would result in an increase in maximum signal, a finding which has now been validated experimentally. While the model prediction was not as clear-cut as it was for PBP concentration, an experiment will discern whether the new model is accurate in predicting robustness at the current estimated concentration of RR, or if the larger trend of increasing maximum fold-change for systems with sizeable changes in phosphorylated RR concentrations is more realistic.

6.4 Experimental Results

To investigate this question, I needed a way to generate different concentrations of RRp in the ON and OFF states of the synthetic signaling system, without changing any other parameters. As described, increasing the total quantity of RR would not be an effective comparison. However, the P_{LacIQ1}::RBP (50x RBP) and P_{LacI}::RBP (1x RBP) synthetic signaling systems could be of use.

In section 3, I showed that by increasing the amount of RBP of the synthetic signaling system, the maximum response in the presence of ligand also increased. Ostensibly, the increased number of PBP available in the P_{LacIQ1}::RBP signaling system increases the concentration of PBP-ligand complex at equilibrium, which in turn results in a larger fraction of PBP bound to the HK. Based on our understanding of the synthetic signaling system, a greater number of extracellular binding events initiates more signal transduction events and phosphorylates a larger fraction of the total HK, thereby generating more RRp. Thus, I can use the RBP concentration as a proxy for the level of phosphorylated response regulator in the ON state. However, this is an assumption, and cannot be used to quantify the amount of phosphorylated response regulator in the ON state.

To that end, I incorporated a second PstS::GFP cassette into the reporter plasmid as described in Appendix A.2, Strategy VII. Cells co-transformed with the dual promoter reporter cassettes and a synthetic signaling system plasmid were tested over a ribose concentration gradient.

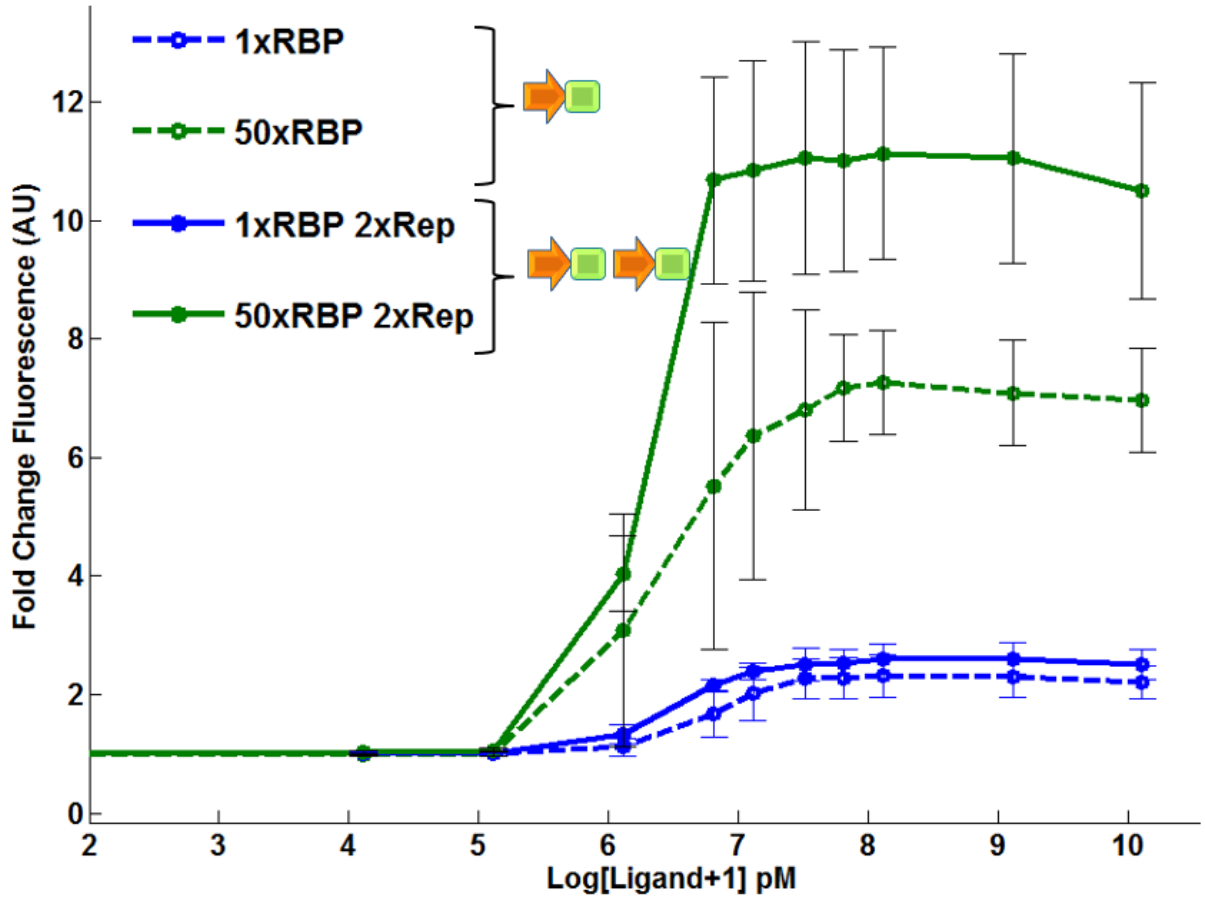


Figure 6.5. Expermiental determination of system response to multiple reporter cassettes.*X-axis represents log of ribose concentration in picomolar, y-axis shows fold change over background fluorescence in the absence of ribose, in arbitrary fluorescent units. The 1x signaling system was transformed with one GFP cassette (blue, dashed line) and two PstS::GFP cassettes (blue, solid line). The 50x system was transformed with one GFP cassette (green, dashed line) and two PstS::GFP cassettes (green, solid line).*

6.5 Conclusions

Figure 6.5 shows corroboration with the broader trend suggested by the models. The system expected to have a lower level of RRp in the ON state (characterized by less fluorescence, *i.e.*,

lower maximum fold-change), the 1x, P_{LacI}::RBP system, did not show an increased response in the presence of an additional promoter reporter cassette (one PstS::GFP cassette: dashed blue line, two PstS::GFP cassettes: solid blue line). Similarly, the system presumed to have a higher level of RRp in the ON state (characterized by greater fluorescence, *i.e.*, higher maximum fold change), the 50x P_{LacIq1}::RBP system, showed a substantial increase in response when cotransformed with the dual PstS::GFP cassette plasmid (solid green line).

The increase in maximum fold-change over background of cells transformed with the 50x P_{LacIQ1}::RBP system and dual PstS::GFP cassette is comparable to the increase obtained by expressing the PBP at unsustainable levels with the P_{LacIq1}::RBS::RBP (Figure 5.11, red line). However, in contrast to the cells transformed with the P_{LacIq1}::RBS::RBP plasmid, cells co-transformed with the P_{LacIQ1}::RBP plasmid and dual PstS::GFP cassettes had growth rates comparable to those of co-transformed with the P_{LacIQ1}::RBP plasmid and a single PstS::GFP cassette. This suggests that increasing the number of reporter cassettes is a viable method of increasing the difference between ON and OFF states in the synthetic signaling system, when the concentration of RRp is sufficiently large in the ON state.

6.6 Future Work

6.6.1 Understanding the impact of RR on the synthetic signaling system activity

The discrepancy between predictions generated from the original model of the synthetic signaling system and the updated model (as discussed in section 6.3) is representative of a significant difference between the two. In the original model, the concentration of response regulator had very little impact on the synthetic signaling system after a certain concentration. Figure 6.6 shows that the baseline concentration of RR (red) generates a very similar response to

that obtained with ten or one hundred times greater concentrations of RR (green and blue respectively).

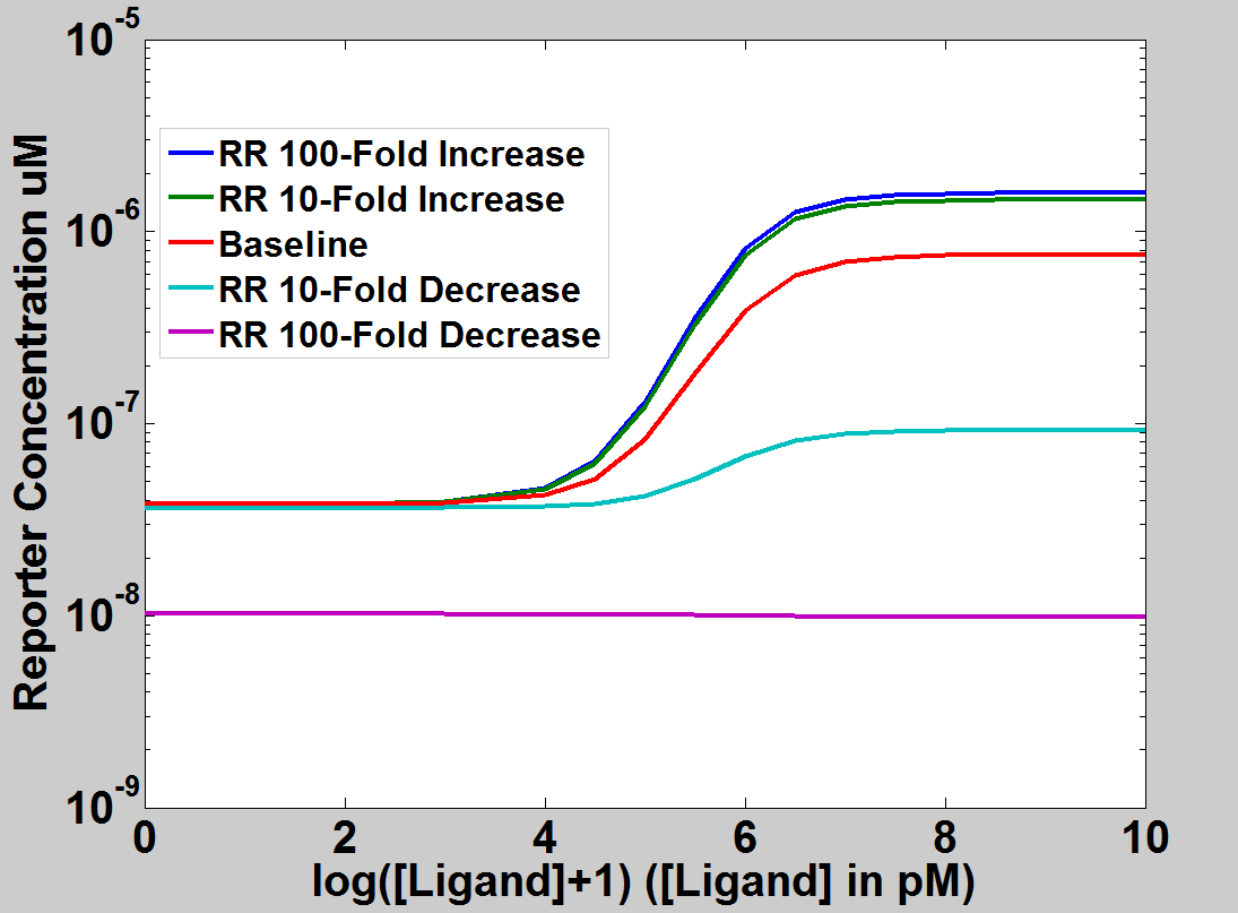


Figure 6.6. OLD MODEL Effect of RR on the synthetic signaling response. *The x-axis is log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. The baseline value (red) generates a response similar to that generated by concentrations ten or one hundred times greater.*

The new model has a different relationship to the concentration of RR. Figure 6.7 shows that the model predicts a specific peak of activity if the RR concentration is increased by ten-fold, but continuing to increase expression to one-hundred-fold will only serve to dampen the response. This prediction can be validated, because the concentration of RR can be modified with the same

tools used to increase PBP concentration. It is a significant departure from the old model, and may serve as another 'lever' to optimize the system.

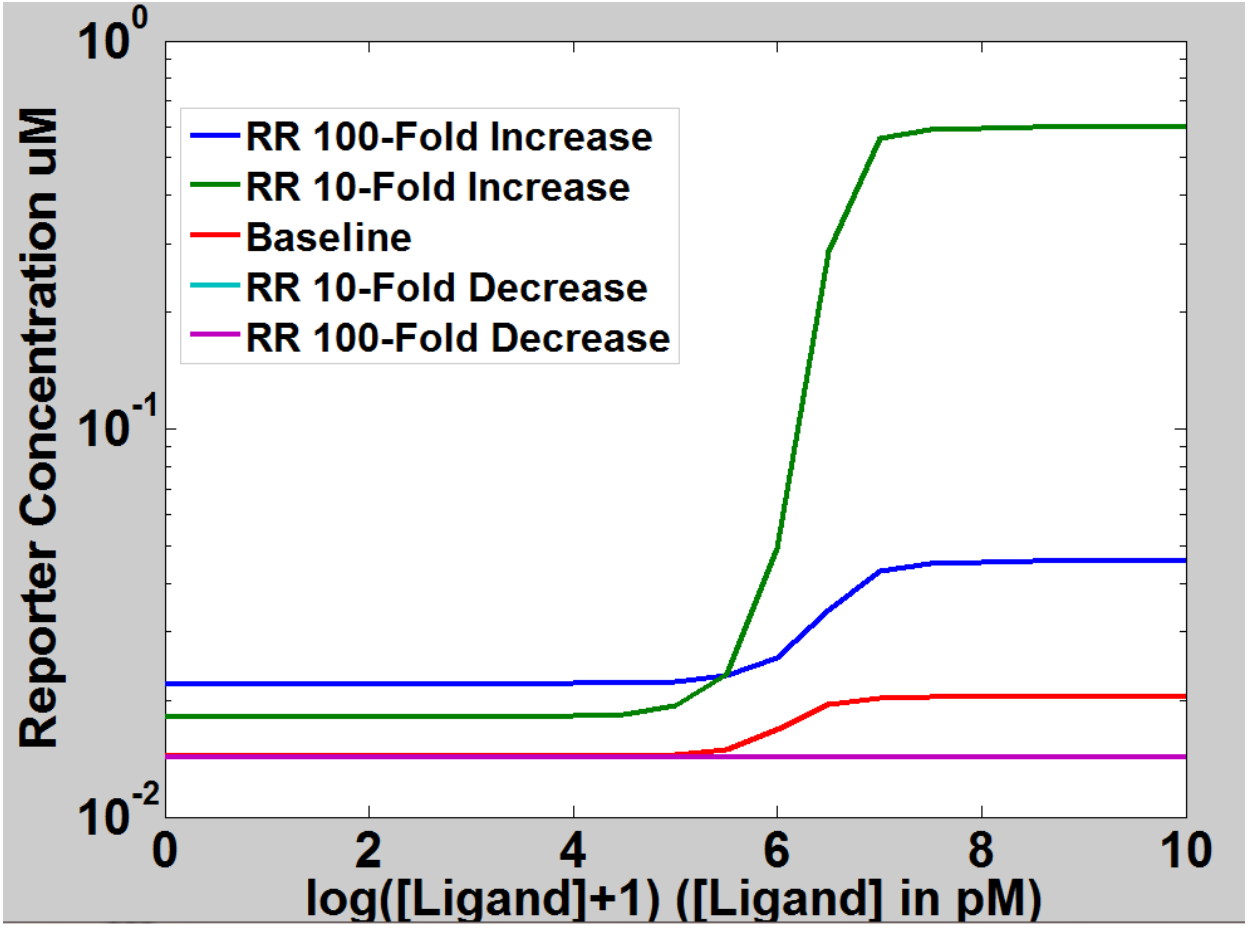


Figure 6.7. NEW MODEL Effect of RR on the synthetic signaling response. *The x-axis is log of the ribose concentration in picomolar, and the y-axis shows the reporter in µM concentration as predicted by the model. There is a predicted peak of activity at RR concentrations ten-fold (green) above the baseline level (red)*

6.6.2 Understanding the synthetic signaling system output

While the trend suggested by the models was qualitatively correct, it suggested a much smaller increase in fold change with the addition of multiple reporter cassettes than the experimental results generated. Because so much is unknown about the concentration of RRp in the ON state

and the OFF state, it is difficult to pinpoint the nature of the discrepancy between the model and the experiment. In order to better understand the output of the synthetic signaling system, an assay for measuring the relative concentrations of RRp in the ON and OFF states of the system must be developed. This assay could be biochemical, like the phosphorylation-gel retardation assay or mass spectrometry, or it could be measured *in vivo* through the use of an inducible system.

The cytoplasmic PhoR in the genome of BW23423 cells is capable of phosphorylating PhoB in the absence of HK72 (Figure 5.9). By placing PhoB under the control of an IPTG inducible system, it would be possible to quantify an 'amount' of phospho-PhoB responsible for a particular level of reporter expression. Both PhoB and a fluorescent protein can be placed under the control of the repressible P_{lac} promoter. Another fluorescent protein is placed under the control of a PhoB-inducible promoter. In the presence of ligand, PhoB and the fluorescent protein will be expressed, with the fluorescent protein controlled by P_{lac} acting as a quantifiable proxy for the amount of PhoB in the system. Assuming that the constitutive kinase cytoplasmic PhoR phosphorylates all (or a consistent fraction) of the PhoB produced, it may be possible to correlate the amount of phosphorylated PhoB with the output generated. The output could then be mapped onto synthetic signaling system responses in the OFF and ON states, thereby suggesting an approximate level of phosphorylated PhoB in each state.

Quantitatively understanding the output of a system is essential for building more complex circuits[121], for example, connecting the signaling system to a gene product other than a reporter. The synthetic signaling system has the potential to become a useful tool for agriculture and bio-manufacturing, in addition to detection, but in order to impact endogenous networks, the output of the synthetic signaling system has to be a gene product that interacts with those networks.

The RBP system, for example, has been connected to a circuit that degrades chlorophyll in the presence of ligand and turns white, generating a response visible to the naked eye (unpublished data, Medford Lab). However, this wasn't possible until a new, stronger promoter controlling the expression of the reporter was introduced. Understanding the output of the synthetic signaling system and tuning the transcriptional response appropriately will facilitate the engineering of other downstream functions.

6.6.3 Building a better plant detector

A robust detection circuit benefits from 'switch-like' behavior, where there is a sharp transition from the OFF to the ON state as a result of a small change in ligand concentration. The existing synthetic signaling circuit, without the optimization proposed here, switches from OFF to ON over an approximately 20-fold concentration. After optimizing the system by increasing PBP concentration, the system switches from OFF to ON over approximately a 10-fold change in concentration (Figure 5.8).

One avenue to improve the switch-like behavior of the synthetic signaling system is to build in bistability, in which a system has two stable states and switches between them in response to stimulus. In bistable systems, the output of the system depends on the state of the system upon receipt of a signal (*i.e.*, a system in the ON state will remain ON through the bistable region while a system in the OFF state will remain OFF) [57]. This effectively acts as a memory of the prior state of the system, an attribute also known as hysteresis.

Bistable systems have been found in both native and synthetic systems. Perhaps the most famous synthetic bistable system is the 'toggle-switch', developed by Gardner et al in 2000[10]. Figure 6.8 outlines a schematic of this system, in which two repressors mutually repress their promoters,

and two inducible promoters can be switched on to initiate the production of the repressor for the desired state. When repressor 1 is high (yellow rectangle), it will bind the repressible promoter (yellow arrow) controlling the expression of repressor 2 (green rectangle). Therefore, repressor 2 will not be expressed, and the promoter of repressor 1(green arrow) will be uninhibited, maintaining the high concentration of repressor 1 (yellow rectangle). Upon induction of repressor 2 from a different, inducible promoter, it begins binding to the repressible promoter controlling expression of repressor 1. As a result, repressor 1 concentrations decrease, and the repressible promoter of repressor 2 becomes more active, generating more repressor 2. This increased concentration of repressor 2 further shuts down the promoter controlling the expression of repressor 1. Assuming continued expression of repressor 2, this cycle continues until repressor 1 can no longer effectively inhibit transcription of repressor 2. This results in a change in state from high repressor 1 to high repressor 2

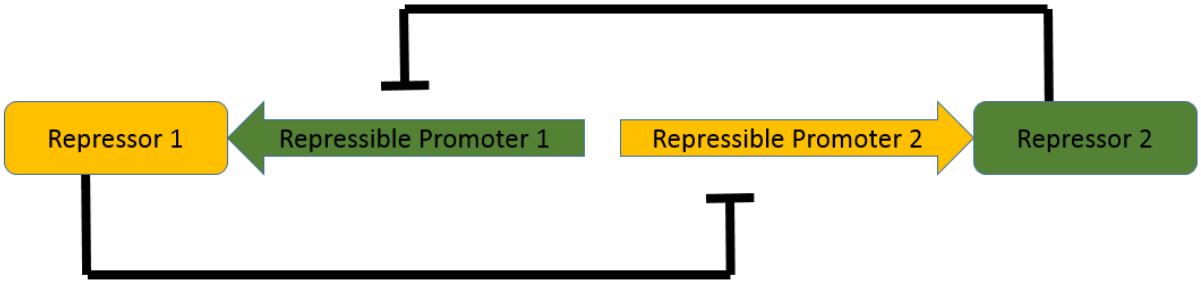


Figure 6.8. Schematic of the toggle switch. *Two mutually repressible promoters express repressors. When the concentration of Repressor 2 (green block) is high, the activity of repressible promoter 1 (green arrow) is low. When Repressor 1(yellow block) is present, the activity of repressible promoter 2 (yellow arrow) is repressed. It is possible to control the state of this system by endogenous application of Repressor 1 or Repressor 2.*

Both switch-like behavior and memory are traits which can be useful in a detector plant. A bi-stable system has a built-in threshold where small fluctuations in signal are unlikely to trigger the switch to the alternate state. In a synthetic signaling system this could help prevent false

positives and make the system a more reliable detector. Memory could enhance the detection of transient bursts of ligand. If the burst is strong enough to trigger the ON state, that state will persist even as the signal fades, ensuring detection.

Bifunctional TCS like HK72 demonstrate both kinase and phosphatase activity on the same molecule. These TCS can also be bistable systems under specific circumstances. First, the HK and RR must participate in unproductive binding – that is, an unphosphorylated HK can bind an unphosphorylated RR for an extended period of time[57]. When they dissociate, there is no net change in the component concentrations (nothing has been phosphorylated or dephosphorylated). However, while participating in the dead-end complex, the HK is unable to participate in productive signaling. Second, the system must have an alternative dephosphorylation pathway for the response regulator, either by an alternative phosphatase which constitutively dephosphorylates the RR or RR autophosphatase activity.

These 'dead-end' complexes have been suggested to occur in the wild-type EnvZ/OmpR system[57]. PhoB has also been shown to participate in dead-end complexes with non-cognate receptors in vitro[122]. Therefore, it may be reasonable to assume that the PhoB/PhoR system may also maintain dead end complexes. Should this be the case, it may be possible to add a phosphatase to the synthetic signaling system and achieve bistability without additional genetic components or time.

However, a literature search could not identify a PhoB phosphatase other than PhoR. Nonetheless, there are several candidates to introduce a new phosphatase to the system. One option is to use a non-functional histidine kinase, one of the 'broken' variants generated while developing HK72. These HKs do not respond to ligand, and are thought to be locked in to

phosphatase activity. Alternatively, the cytoplasmic DHP domain could be used. The PhoR DHP domain has been shown to function as a PhoB phosphatase *in vitro*[80].

Each method has particular considerations. The 'locked off' HK must not dimerize with the functional HK (*i.e.*, it must not contain a PhoR dimerization domain), but it must still interact with PhoB. In addition, the 'locked off' HK will be produced in greater quantities than the functional HK, possibly diluting the PBP-Ligand complex available to activate the functional HK. The cytoplasmic PhoR DHP domain has also been shown to form inclusion bodies when expressed at high levels[80], and is unlikely to be functional should this occur. It is additionally possible that PhoR and PhoB do not form unproductive complexes, in which case no additional phosphatase will result in bistability.

CHAPTER 7: CONCLUSIONS AND FUTURE WORK

7.1 Experimental Results

As synthetic biology matures as a field and addresses questions of greater significance and complexity (e.g. Anderson et al[11], Antunes et al[13], and Xie et al [12]), engineered gene circuits will continue to expand in size and complexity[1,26]. Experimentally modifying networks such as these parameter by parameter, gene by gene, in order to gain a complete understanding of the system is inefficient and costly. Predictive modeling allows researchers to rigorously analyze the output of a complex genetic network as a function of molecular-level modification *in silico,* thereby gaining deeper insights into the behavior of the system and assisting experimental design. This thesis demonstrates the use of modeling as an effective tool for efficiently directing experimental work and optimizing a complex synthetic gene circuit in bacteria.

The Medford lab synthetic signaling system described in Chapter 2 is capable of transferring a signal from the exterior of the cell to the interior of the cell, thereby inducing a transcriptional response in a ligand-dependent manner. In order to refine this system, (*i.e.*, to increase maximum fold-change in signal in the presence of ligand, to lower the detection threshold, and make the transition from OFF state to ON state more switch-like) I had to establish a comprehensive understanding of the system, build a model and explore the parameter space *in silico*, then validate that model. Because the components used in this circuit have been shown to function in both plants and bacteria, I was able to use a bacterial test platform to optimize this synthetic signaling system.

In Chapter 3 I describe the construction and exploration of a dynamic model of the synthetic signaling system. The synthetic signaling system is constructed with components derived from bacterial wild-type TCS and chemotactic systems that are associated with a large body of literature. I was therefore able to set many synthetic signaling system parameters to experimentally determined values for the wild-type precursor of the synthetic signaling system component. Literature values for other, closely related systems, were used to estimate parameters which were otherwise undetermined, limiting the assumptions needed to complete the model.

I establish experimental methods and characterize the synthetic signaling system in Chapter 4. The system is sensitive to abiotic conditions: temperature increases total expression of reporter in both the OFF and the ON states, but does not affect the transition from OFF to ON or the maximum fold-change response of the system in the presence of ligand. The system responds slightly differently to different carbon sources, with cells grown in glycerol showing the best response to ligand and cells grown in L-arabinose yielding the worst response. In addition, the RBP synthetic signaling system must be tested in minimal media, as rich media reduced activity. Time after induction is also an important measurement parameter, with short induction times (approximately 3 hours) showing the synthetic signaling system responding at very low levels of ligand, and maximum fold-change induction peaking at intermediate times (approximately 6 to 9 hours).

Chapter 5 describes experimental validation of the model. I first performed a global analysis of the synthetic signaling system *in silico* to predict the parameters most likely to improve signaling. One of those parameters, PBP concentration, both increased maximum fold-change induction in the synthetic signaling system and could be engineered without impacting additional parameters. I therefore constructed three plasmids containing synthetic signaling systems with

varying constitutive expression of PBP and tested them across a ribose concentration gradient. I was able to validate this model in bacteria, demonstrating that by increasing the concentration of a single component of the synthetic signaling system, the PBP, I could increase fold-change response in the presence of ligand six-fold (from two-fold to twelve-fold induction). This change also resulted in a 100-fold increase in sensitivity.

Upon validation, a synthetic signaling system putatively expressing increased PBP was developed for plants using novel constitutive promoters developed by the Medford lab. *Arabidopsis* plants were transformed with this construct and T0 plants generated. While the T0 generation did not show a definitive increase in activation, more accurate data can be gathered from the T1 generation. In addition, an alternative method for testing the model prediction in plants is proposed. Should the model be validated in plants as well, this work will not only enhance the performance of the synthetic signaling system *in planta*, it will advance the current state of plant synthetic biology by providing a new tool to the community: prokaryotic testing and optimization of synthetic gene circuits for use in plants.

A model specific to a particular experimental approach is developed and validated experimentally in Chapter 6. Introducing additional reporter cassettes controlled by PhoB inducible promoters had the potential to increase the maximum fold change response, however the concentration of phosphorylated RR in the system in the ON state as compared to the OFF state was unknown. In order to determine the feasibility of this approach I designed a model which explored all possible combinations of ON and OFF state phosphorylated RR concentrations and simulated fold-change response *in silico.* The model suggested that systems with large changes in RR phosphorylation from the OFF to the ON state would benefit from additional reporter cassettes, a prediction I then experimentally validated.

In this work I have demonstrated the effectiveness of using predictive models in combination with experimental results to optimize an existing synthetic signaling circuit in a bacterial system. The prototype synthetic signaling system developed by the Medford Lab has now been refined through research directed by model predictions. The improved detection and increased response in the presence of ligand moves this technology closer to real-world use.

7.2 Future Work

The synthetic signaling system links a specific extracellular ligand binding event to an internal transcriptional response. In the majority of the work in this thesis, ribose-binding has been linked to GFP expression, but both the input and output of this system are modular, expanding the potential applications for such a system. In addition, the process of using predictive modeling and experimental validation in bacterial systems to optimize a synthetic signaling system for plants has broad implications for the field of plant synthetic biology.

7.3 Plant based detectors

The optimized RBP synthetic signaling system presented here is a step towards a functional plant detector technology. Chemical and biological monitoring is currently impractical in many situations, but a plant based detector system is cheap to generate and can be remotely monitored, efficiently expanding detection capabilities[13]. Paired with receptors re-designed to bind ligands of interest to human health (*e.g.*, pollutants, explosives, or chemical agents), this is a potentially powerful detection technology. A detector plant could therefore be used to monitor pipelines across the country, rapidly and cheaply identifying leaks. Or, a detector plant could be deployed in a minefield in Cambodia, turning white when roots encounter ordinance underground. The possible applications are limited only by our ability to engineer the binding pocket of the PBP while retaining signaling through the system. In addition, multiple variants of the synthetic

signaling system, each detecting a unique substance and generating a unique response, could be deployed into the same plant, thereby increasing the number of compounds detected by this technology.

7.4 Downstream applications

In addition to detection of compounds by expression of GFP, the output of the detector plant could be linked to any transcriptional response, using a PstS-like inducible promoter. This allows the system to perform additional useful work. Consider a detector plant used to detect a common contaminant (e.g. arsenic, toluene, lead, DDT, mercury, etc.). The synthetic signaling system then be used to trigger a phytoremediation process upon detection of these compounds, either by breaking them down or transporting them to the aerial tissue where they can be harvested and removed from the soil.

Additionally, plants containing multiple synthetic signaling systems, each with a unique input and output, could be used to perform complex logic in plants. Plants already perform signal integration in a variety of contexts: floral organs are determined as a result of the interactions of at least three different gene families[18], auxin and cytokinin hormone levels play a substantial role in determining plant cell development[29,47], and an array of photoreceptors detects the intensity, quality, and direction of light[123], enabling the plant to modify growth appropriately. Logic gates integrate digital inputs (*i.e.*, the presence or absence of a signal) into a digital output (*i.e.*, an ON or OFF state). Synthetic genetic logic gates, including the NOR gate, from which all other logic gates can be composed, have been successfully constructed in prokaryotic systems[121] as well as mammalian systems[12], however logic gates have not yet been published in plants. As in prokaryotic and mammalian systems, synthetic genetic logic gates in plants can be used to establish multivariable control of outputs from transcription of individual genes (*e.g.*, Wang et

al[35,121]) to cell fate[12]. This type of control could be of interest to agricultural and bio-manufacturing industries.

7.5 Plant synthetic biology

The impact of this work is not limited to the Medford lab synthetic signaling system and its various applications. While the ease of bacterial transformation and testing has encouraged prokaryotic synthetic biology, the difficulties associated with plant work have inhibited synthetic biology in plants. However, the potential of plant synthetic biology is such that advances in this field can have enormous impact worldwide, from food to fuel to the manufacturing of high value molecules. If it can be established that model predictions of an orthogonal synthetic gene network are applicable in both plant and bacterial chassis, the use of bacterial platforms to evaluate and optimize synthetic gene networks destined for use in plants will encourage the budding field of plant synthetic biology.

Appendix

A.1 Construction of synthetic signaling system plasmids

Strategy I. This strategy was used to investigate the impact of additional genetic material included in the original synthetic signaling system, KJM114. KJM114 included genetic material unnecessary for synthetic signaling expression in bacteria: wild-RBP was expressed with additional residues at the C-terminus (a FLAG tag, DYKDDDDK), HK72 included the leader sequence of the Arabidopsis Flagellin Sensitive 2 receptor (FLS) MKLLSKTFLILTLTFFFFGIALAK. These residues comprise a eukaryotic plasma membrane targeting sequence, and are locate 5' of the N terminus of HK57. In addition, PhoB was not wild-type. A point mutation had altered amino acid 227 from threonine to asparagine. Each of these issues were addressed using PCR.

DNA segments free of additional tags were developed by PCR of the appropriate open reading frame. Primers 60 and 36 were used to amplify the DNA segment P_{lacI}::RBP with the addition of an EcoRI site immediately 3' of the RBP stop codon. The PCR product was digested with EcoRI and Eco91I (an isoschizomer of BsteII) and ligated into KJM114 to form KLH545, the synthetic signaling system without a Flag Tag on the RBP. Primers 63 and 64 were used to amplify the DNA segment HK72 with the addition of an XhoI site immediately 5' of the start codon. XhoI does not appear in KJM114, so an additional PCR step using primers 54 and 53 to amplify the DNA segment P_{lacI}::PhoB::IntergenicRegion was required. This PCR product included an XhoI site immediately 3' of the intergenic region and a HindIII 3' of the XhoI site. The resulting PCR product was digested using HindIII and NheI and ligated into KJM114 to form KLH546. The

HK72 amplicon was digested with XhoI and HindIII and ligated into KLH546 to form KLH551, a plasmid identical to KJM114 with the addition of an XhoI site and without the extraneous FLS.

Similarly, primers 53 and 89 were used to correct the point mutation at PhoB amino acid 227 and amplify the DNA segment P_{lacI}::PhoB::IntergenicRegion::XhoI. This fragment was digested with XhoI and NheI and ligated into KLH551 to form KLH581.

All primers were designed with a Tm of 60 ˚C, simplifying the PCR reactions. 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of the forward and 1.25 µL of a 10 µM solution of the reverse primers (as described above), 1.25 µL of a 10 µM solution of dNTPs were mixed with 60 ng KJM114 per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds), annealing (54 ˚C, for 20 seconds), and extension (72˚C for 1.5 min).

Strategy II. This strategy was used to construct signaling systems with different concentrations of PBP. Using plasmid KJM114 (a low copy plasmid) as a template, one PCR step was carried out for each of the planned constructs to create a DNA segment with a unique promoter controlling the expression of RBP. 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of the reverse primer and 1.25 µL of a 10 µM solution of forward primer (listed in Table 4.1), 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 60 ng/ µL solution of KJM114 per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds) , annealing (51˚C, for 20 seconds), and extension (72˚C for

1.5 min). The PCR product was digested with BsteII and EcoRI and ligated into KJM114 creating the plasmids described in Table XX

Strategy III. This strategy was used to introduce a new periplasmic binding protein to the signaling system. The extracellular sensor domain of HK72 can interact with both ribose binding protein and glucose binding protein. Using genomic DNA from *E. coli* strain BW23423 as a template, two PCR steps were carried out to create a DNA segment with P_{lacI} controlling the expression of GBP. The first PCR was used to isolate the GBP sequence: 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 7 and 1.25 µL of a 10 µM solution of Primer 8, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 60 ng/µL solution of KJM114 per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds) , annealing (50˚C, for 20 seconds), and extension (72˚C for 1.5 min). A second PCR step was used to attach the P_{lacI} promoter 5' of GBP: 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 26 and 1.25 µL of a 10 µM solution of Primer 5, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 60 ng/µL solution of KJM114 per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds), annealing (50˚C, for 20 seconds), and extension (72˚C for 1.5 min). The final PCR product was digested with Eco91I and EcoR1 and ligated into KJM114 to form KLH109, the synthetic signaling system with a GBP input (Figure 7.1).

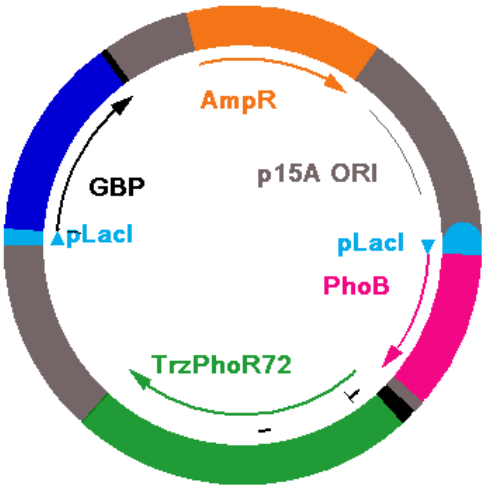


Figure 7.1. Vector map of KLH109, the 1x GBP signaling system plasmid. *GBP and the PhoB/PhoR operon are both controlled by the weakly constitutive* P_{lacI}*. The vector is the low copy plasmid, pACYC177.*

The LacIQ1 promoter was also used to increase GBP expression by approximately 50 fold. This promoter was attached as described in Strategy II through PCR amplification by Nikolai Braun.. The PCR product P_{lacI} Q1::GBP was digested with EcoRI and Eco91I and ligated into KJM114 to form NAB251, the synthetic signaling system with increased concentration of GBP.

Strategy IV. This strategy was used to generate null plasmids, signaling systems without one of the three signaling components. RBP was deleted by digesting KJM114 with PsiI and AfeI, dropping out the DNA sequence containing P_{lacI}::RBP, and then re-ligating the backbone to form KLH524.

The HK deletion was generated in the process of deleting the FLS from HK72. As described in Strategy I, a single PCR step was accomplished, using primers 54 and 53 to amplify the DNA segment P_{lacI}::PhoB::IntergenicRegion. This PCR product included an XhoI site immediately 3' of the intergenic region and a HindIII 3' of the XhoI site. The resulting PCR product was digested using HindIII and NheI and ligated into KJM114 to form KLH546, the synthetic signaling system without HK72.

The PhoB deletion was generated in two parts. First, as described in Strategy I the deletion of the FLS required an HK72 amplicon digested with XhoI and HindIII and ligated into KLH546 to form KLH551, a plasmid identical to KJM114 with the addition of an XhoI site and without the

extraneous FLS. Using KLH524 as a template, an XhoI site was added immediately 3' of the P_{lacI} promoter via primer 84, and the DNA sequence consisting of P_{lacI}::XhoI was amplified using the following recipe: 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 53 and 1.25 µL of a 10 µM solution of Primer 84, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 60 ng/µL solution of KLH524 per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds), annealing (55˚C, for 20 seconds), and extension (72˚C for 45 seconds). The resulting PCR product was digested with XhoI and NheI and ligated into KLH551 to form KLH575, the synthetic signaling system without PhoB.

Strategy V. This strategy was used to introduce new features to the signaling plasmid. Desired gene sequences are PCR amplified, and KJM114 is digested using the blunt enzyme PsiI or AfeI. KJM114 is dephosphorylated using Antarctic Phosphatase (New England Biolabs, Ipswich, MA), and the PCR product and dephosphorylated vector are ligated. The resulting plasmid is then screened for directionality if necessary.

To introduce additional PhoB binding sites which did not induce the expression of any genes, I ordered a Geneblock (IDTDNA, Coralville, Iowa) containing 7 PhoBoxes spaced four nucleotides apart (endogenous PhoBox spacing) or 6 nucleotides apart (to allow for restriction sites). The minigene was ligated into pJet to develop KLH610. KLH610 was digested using AfeI and PsiI and ligated into KJM114 (digested with PsiI and dephosphorylated) to yield KLH614, the synthetic signaling system with additional binding sites. 3 separate 7 PhoBox inserts were incorporated during the blunt end cloning process, resulting in 21 additional PhoB binding sites.

To introduce another PhoB dephosphorylation pathway to the synthetic signaling system, a non-functional HK fusion was introduced to the synthetic signaling system. This fusion, EGA had been shown to be unresponsive to ligand, in prior work. First, a PCR step was accomplished to generate a Trz DNA fragment from plasmid JLF7 (constructed previously) with a 5' XhoI site and a 3'HindIII site. 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 94 and 1.25 µL of a 10 µM solution of Primer 95, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 100 ng/µL solution of JLF7 plasmid per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). This fragment and KLH575 were digested with XhoI and HindIII and ligated together, forming KLH597 JLF7 expressed directly by the P_{lacI} promoter. A minigene was ordered, with residues 275-277 of JLF7 modified to the amino acids EGA. This minigene and KLH597 were cut with MfeI and NdeI, and the minigene was ligated into the KLH597 backbone to form KLH587, HK EGA expressed directly by the P_{lacI} Promoter. Finally, a PCR step was accomplished using primers 53 and 94 to amplify the DNA segment P_{lacI}::EGA. KJM114 was digested with PsiI and dephosphorylated, and ligated to the PCR product to yield KLH592, the synthetic signaling system with an additional, nonfunctional histidine kinase capable of interacting with PhoB.

Strategy VI. This strategy was used to generate plant synthetic signaling systems with various levels of PBP using strong constitutive promoters developed by Alberto Donayre-Torres, PhD, of the Medford Lab (pNML1 & pNML2). First, a NotI site was introduced 5' of the plant codon optimized RBP using primer 70. The DNA sequence consisting of NotI::RBP was amplified using the following recipe: 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 70 and 1.25 µL of a 10 µM solution of Primer 71, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 80 ng/µL solution of PAB12 (containing the

original plant synthetic signaling circuit) per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds), annealing (62˚C, for 20 seconds), and extension (72˚C for 45 seconds). In order to place RBP under the control of the new promoters, plasmid KDA35 (containing the pNML1 promoter) was digested with NotI and EcoRV, and the PCR product was digested with NotI. The PCR product was ligated into KDA35 to form plasmid KLH560, a plasmid containing plant codon optimized RBP under the control of pNML1, a promoter stronger than the CaMV35S promoter in transient protoplast assays.

KLH560 and AJD170 were then digested with SacI and NotI in order to place RBP under the control of the pNML2 promoter. The promoter from AJD170 was ligated into KLH560 to create KLH577, a plasmid containing codon optimized RBP under the control of pNML2, a stronger promoter than pNML1 in transient protoplast assays. AJD143 was digested using PvuII and Eco53KI to extract the second half of the promoter system, and KLH577 and KLH560 were linearized with Eco53K1, a blunt isoschizomer of SacI. KLH560 and KLH577 were dephosphorylated with Antarctic Phosphatase and ligated to the insert, forming plasmids containing RBP under the control of the pNML1 system (KLH561) and the pNML2 system (KLH578).

In order to build a complete synthetic signaling system, KJM51 (complete synthetic signaling system) and KLH561(plant codon optimized RBP under the control of the pNML1 system) and KLH578 (plant codon optimized RBP under the control of the pNML1 system), were digested with SacI and ApaI. The PBP fragment from KLH561 or 578 was ligated to the vector KJM51 to form KLH801 and KLH802 respectively.

In order to insert these transgenes into the plant, the synthetic signaling system needed to be placed on a pCambia vector. This was accomplished by using SacI and ApaI to digest KJM51 (the complete synthetic signaling system), KLH561 (the complete synthetic signaling system with plant codon optimized RBP under the control of the pNML1 system), and KLH578 (the complete synthetic signaling system with plant codon optimized RBP under the control of the pNML2 system), as well as a pCambia vector modified (by Pete Bowerman, PhD) to contain ApaI. The three synthetic signaling systems were ligated into the pCambia vector, forming KLH803, KLH804, and KLH805 respectively. These clones required additional recovery time and overnight incubation at 30˚C.

A.2 Construction of reporter plasmids

Strategy VII. This strategy was used to investigate alternative PhoB inducible promoters to control expression of the reporter gene. Synthetic PstS promoters with modified PhoBoxes were synthesized (IDTDNA, Coralville, Iowa) for digestion and ligation into a luciferase reporter plasmid. The promoters were synthesized two to a plasmid: SNK1, which contained Psts.PhoA and Psts.consensus promoters; SNK2 with Psts.PhoB and PstS.PhoE promoters; and SNK3, with PstS.PhoH and PstS.ughP promoters. Each plasmid was individually digested with XbaI and NcoI to remove one of the promoters, and re-ligated, forming KLH201, PstS.Consensus, KLH202, PstS.PhoH, KLH203, Psts.PhoA, KLH204, PstS.ughP, KLH205, PstS.PhoB, and KLH206, Psts.PhoH. One PCR step was used to add the GFP sequence 3’ of each promoter via primer 61. Promoter::GFP sequences were amplified as follows: 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 61 and 1.25 µL of a 10 µM solution of Primer 62, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 45 ng/µL solution of TAA002 per the Herculase II Fusion DNA Polymerase protocol (Agilent,

Santa Clara, CA). The resulting PCR product was digested with XhoI and NheI and ligated into KLH551 to form KLH575. The reaction mixture was heated to 95˚C for 2 minutes prior to 30 cycles of denaturation (95 ˚C for 30 seconds), annealing (57˚C, for 20 seconds), and extension (72˚C for 45 seconds). The final PCR product was digested with AatII and NcoI and ligated into TAA001 to form KLH213, KLH214, KLH215, KLH216, KLH217, KLH218 (Table XX).

Strategy VIII. This strategy was used to increase the number of PhoB-inducible reporter cassettes in the reporter plasmid. pBR322 has few available cloning sites, and one PCR step was accomplished to introduced the necessary sites. Using Primers 57 and 58, a DNA segment containing PstS::GFP was amplified, with the addition of a PstI site 5' of the PstS promoter and an AseI site 3' of GFP. 34 µL diH2O, 10 µL *HerculaseII* 5x buffer solution, 1.25 µL of a 10 µM solution of Primer 57 and 1.25 µL of a 10 µM solution of Primer58, 1.25 µL of a 10 µM solution of dNTPs were mixed with 1 µL of a 60 ng/µL solution of TAA002 per the Herculase II Fusion DNA Polymerase protocol (Agilent, Santa Clara, CA). The resulting PCR product and TAA002 was digested using AatII and EcoRI and ligated into TAA002 to develop a dual cassette reporter plasmid, KLH557.

A.3 Media

LB Media

LB media was prepared as follows: To 2000mL diH_20 add 40 g Tryptone (Fluka analytical via Sigma Aldrich, St. Louis MO), 20g yeast extract (Teknova, Hollister CA), 40g NaCL (Sigma Aldrich, St. Louis MO). pH balance to 7.0 with 3M NaOH, and add diH_20 to 4000 mL.

2XYT Media

2XYT media was prepared as follows: To 3800 mL di H_20 add 50 g Tryptone (Fluka Analytical via Sigma Aldrich, St. Louis MO), 20 g yeast extract (Teknova, Hollister, CA), 20 g NaCl (Sigma Aldrich, St. Louis, MO). pH balance to 7.0 with 3M NaOh, and add diH_20 to 4000 mL.

M9CA Media

M9CA media was prepared as follows: To 2000 mL diH_20 add 16 mL 100% glycerol (*i.e.*,5.4 mM glycerol), 50g M9CA Medium Powder (Amresco, Solon OH), add diH_20 to 4000 mL

M9CA + Media

To 500 mL autoclaved M9CA media, add the following: 1 mL 1M $MgS0^4$, 50 µL $CaCl_2$, 500 µL 50mg/ µL carbinicillin, 500 µL 20mg/ µL tetracycline, 11.0 mL 5% glycerol (sterile filtered).

MS Media

MS media was prepared as follows: To 3200 mL diH_20, add 40 g Sucrose (Sigma-Aldrich, St. Louis, MO), 17.5 g MS basal medium (Sigma-Aldrich, St. Louis, MO), 2 g MES (Sigma-Aldrich, St. Louis, MO). pH to 5.7 with KOH, and add diH_20 to 4000 mL. Add 3g Phyto Agar per 500 ml (Plant Media, Dublin, OH) as necessary

New Infiltration Media

New Infiltration Media was prepared as follows: To 1800 mL diH_20, add 100g Sucrose (Sigma-Aldrich, St. Louis, MO) and ___G $MgCl_2$. Add diH_20 to 2000 mL and autoclave. To 500 mL autoclaved media add 25 µL 0.005% Silwet-77, & 2.5 µL BAP.

A.4 Plant Transformation Protocol

1. Grow a starter culture of Agrobacterium (2mL) containing the desired binary plasmid on LB media supplemented with the appropriate antibiotics (20mg/L Gentamicin for the Agrobact.), at 30^{o}C / 2 days.
2. Inoculate this starter culture into 250 mL of LB media (containing the same antibiotics) contained in 1L flasks and grow them at 30^{o}C for about 1 day (according to Andrew Bent the O.D. of the bacteria does not influence the efficiency of transformation.)
3. Transfer the media to centrifuge bottles, and spin the cells at 6000 rpm for 12 minutes at 4^{o}C.
4. Discard the supernatant and resuspend the bacterial cells in 500 mL of 'New' Infiltration Media containing 25 □L 0.005% Silwet L-77 (Lehle Seeds) & 25 uL BAP.
5. Transfer the bacteria-containing infiltration media to a 4 L plastic beaker.
6. Invert the pots containing the Arabidopsis plants, and dip them in the solution for about 1 minute, making sure the flower buds are immersed in the solution (it is not necessary to dip the leaves, but it is ok if they are).
7. After dipping the plants, transfer them to a flat, with the pots lying on their sides.
8. Cover the flat with Saran Wrap and take it back to the growth chamber at 25^{o}C.
9. After ON, remove the Saran Wrap and place the plants upright.
10. Water the plants regularly until the siliques are formed.
11. Harvest Seeds.
12. Select for transformants according to resistance conferred by the plasmid.

A.5 Primers

Primer Number	Primer Sequence (5’ to 3’)
31	CTGAATTCCTAGTGGTGATGGTGATGATGAAGGCCTGATCCGCCCTGCTTAACAACCAGTTTCAGA TCAAC
60	GAATTGAATTCCTACTGCTTAACAACCAGTTTCAGATCAACC
63	GCCGCCATATGCTGTAAATGACGACTTAAG
64	TGATTTCTTAACTCGAGTATCTTAATGAATACAACTCCCTC
53	CACTGACACCTTCATCAGTGCC
89	ATCGTTTTTCAACCCGCTTTTAACGCC
43	AGGATCTGCATGCATTTACGTTGACACCACCTTTCGCGGTATGGCATGATAGCGCCCACATCACAC ACTTTTTAGGTCCGAATCGGCATGAACATGAAAAAACTGGCTACCCTGG
37	GCGACACGGAAATGTTGAATACTCAT
33	TATTTGAGCTCCCTTCTGATGAAGCGTCAGC
54	ATGGACAAGCTTCTCGAGGTAAGATACTCCAGTTAAGAAATCATAAGCCC
84	CACCATCTCGAGATTCACCACCCTGAATTGACTCTCTTCC
94	GGCAAAAGCTTTACCCTTCTTTTGTCGTGCC
95	TTACACTCGAGTATCTTAATGAATACAACTCCCTCACAGCGATTAGG
70	ATGTATGCGGCCGCATGGAAAGACCTTTTGGATGCTTC
71	CCAGAGAAATGTTCTGGCACCTGC
57	TAGATTCTGCAGTCTCTCTGTCATAAAACTGTCATATTCCTTACATATAACTGTC
58	ATGGCAATTAATGCGGTGTCATCTATGTTACTAGATCGGG
61	TGTTGGCCATGGAACAGGTAGTTTTCCAGTAGTGCAAATAAATTTAAGGGTAAGTTTTCCGTATGT TGCATCACCTTCACCCTCTCCACTGACAGAAAATTTGTGCCCATTAACATCACCATCTAATTCAACAA GAATTGGGACAACTCCAGTGAAAAGTTCTTCTCCTTTACGCATAATGTCTCCTGGGAGGATTCATA
62	AAATTAGACGTCATGGTCGGCTAGCTCTCT

REFERENCES

1 Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2** (2006).

2 Jacob, F. & Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of Molecular Biology* **3**, 318-356, doi:http://dx.doi.org/10.1016/S0022-2836(61)80072-7 (1961).

3 Watson, J. D. & FHC, C. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature* **171**, 2, doi:10.1038/171737a0 (1953).

4 Brenner, S., Jacob, F. & Meselson, M. An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. *Nature* **190**, 6 (1961).

5 Meyer, B. J., Kleid, D. G. & Ptashne, M. Lambda repressor turns off transcription of its own gene. *Proceedings of the National Academy of Sciences* **72**, 4785-4789 (1975).

6 Ohshima, Y., Matsuura, M. & Horiuchi, T. Conformational change of the lac repressor induced with the inducer. *Biochemical and Biophysical Research Communications* **47**, 1444-1450, doi:http://dx.doi.org/10.1016/0006-291X(72)90234-3 (1972).

7 Goodwin, B. C. Oscillatory behavior in enzymatic control processes. *Advances in Enzyme Regulation* **3**, 425-437, doi:http://dx.doi.org/10.1016/0065-2571(65)90067-1 (1965).

8 Lodish, H. *et al. Molecular Cell Biology*. 6th Edition edn, 2 (W. H. Freeman, 2008).

9 Elowitz, M. B. & Leibler, S. A synthetic socilatory network of transcriptional regulators. *Nature* **403**, 4 (2000).

10 Gardner, T. S., Cantor, C. R. & Collins, J. J. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342, doi:http://www.nature.com/nature/journal/v403/n6767/suppinfo/403339a0_S1.html (2000).

11 Anderson, J. C., Clarke, E. J., Arkin, A. P. & Voigt, C. A. Environmentally Controlled Invasion of Cancer Cells by Engineered Bacteria. *Journal of Molecular Biology* **355**, 619-627, doi:http://dx.doi.org/10.1016/j.jmb.2005.10.076 (2006).

12 Xie, Z., Wroblewska, L., Prochazka, L., Weiss, R. & Benenson, Y. Multi-Input RNAi-Based Logic Circuit for Identification of Specific Cancer Cells. *Science* **333**, 1307-1311, doi:10.1126/science.1205527 (2011).

13 Antunes, M. S. *et al.* Programmable Ligand Detection System in Plants through a Synthetic Signal Transduction Pathway. *PLoS ONE* **6**, doi:10.1371/journal.pone.0016292 (2011).

14 Hu, P., Janga, S., M, B., Diaz-Mejia, J. & Butland, G. Global Functional Atlas of Escherichia coli Encompassing Previously Uncharacterized Proteins. *PLoS Biology* **7**, doi:10.1371/journal.pbio.1000096 (2009).

15 Endler, L. *et al.* Designing and encoding models for synthetic biology. *Journal of The Royal Society Interface*, doi:10.1098/rsif.2009.0035.focus (2009).

16 Eschenlauer, A. C. & Reznikoff, W. S. Escherichia coli catabolite gene activator protein mutants defective in positive control of lac operon transcription. *Journal of Bacteriology* **173**, 5024-5029 (1991).

17 Yanofsky, M. F. *et al.* The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors. *Nature* **346**, 35-39 (1990).

18 Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. Genetic interactions among floral homeotic genes of Arabidopsis. *Development* **112**, 1-20 (1991).
19 Bowman, J. L., Smyth, D. R. & Meyerowitz, E. M. The ABC model of flower development: then and now. *Development* **139**, 4095-4098, doi:10.1242/dev.083972 (2012).
20 Yang, X., Lau, K.-Y., Sevim, V. & Tang, C. Design Principles of the Yeast G1/S switch. *PLoS Biology* **11** (2013).
21 Rausenberger, J. *et al.* Photoconversion and Nuclear Trafficking Cycles Determine Phytochrome A's Response Profile to Far-Red Light. *Cell* **146**, 813-825, doi:http://dx.doi.org/10.1016/j.cell.2011.07.023 (2011).
22 Alon, U. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. 1 edn, (Taylor & Francis, 2006).
23 Antunes, M. S. *et al.* Engineering key components in a synthetic eukaryotic signal transduction pathway. *Mol Syst Biol* **5**, doi:http://www.nature.com/msb/journal/v5/n1/suppinfo/msb200928_S1.html (2009).
24 Endy, D. Foundations for engineering biology. *Nature* **438**, 449-453 (2005).
25 Voigt, C. A. Genetic parts to program bacteria. *Current Opinion in Biotechnology* **17**, 548-557, doi:http://dx.doi.org/10.1016/j.copbio.2006.09.001 (2006).
26 Purnick, P. E. M. & Weiss, R. The second wave of synthetic biology: from modules to systems. *Nature Reviews Molecular Cell Biology* **10**, 13 (2009).
27 Quatrano, R. S., Assmann, S. M., Sanders, D. & Eckardt, N. A. *The Plant Cell Online* **14**, S1, doi:10.1105/tpc.141350 (2002).
28 Middleton, A. M., King, J. R., Bennett, M. J. & Owen, M. R. Mathematical Modelling of the Aux/IAA Negative Feedback Loop. *Bull. Math. Biol.* **72**, 1383-1407, doi:10.1007/s11538-009-9497-4 (2010).
29 Muraro, D. *et al.* The influence of cytokinin–auxin cross-regulation on cell-fate determination in Arabidopsis thaliana root development. *Journal of Theoretical Biology* **283**, 152-167, doi:http://dx.doi.org/10.1016/j.jtbi.2011.05.011 (2011).
30 Kramer, B. P. *et al.* An engineered epigenetic transgene switch in mammalian cells. *Nature Biotechnology* **22**, 4 (2004).
31 Haynes, K. A. & Silver, P. A. Eukaryotic systems broaden the scope of synthetic biology. *The Journal of Cell Biology* **187**, 589-596, doi:10.1083/jcb.200908138 (2009).
32 Ajo-Franklin, C. M. *et al.* Rational design of memory in eukaryotic cells. *Genes & Development* **21**, 2271-2276, doi:10.1101/gad.1586107 (2007).
33 Edamatsu, H., Kaziro, Y. & Itoh, H. Inducible high-level expression vector for mammalian cells, pEF-LAC carrying human elongation factor 1α promoter and lac operator. *Gene* **187**, 289-294, doi:http://dx.doi.org/10.1016/S0378-1119(96)00768-8 (1997).
34 Berens, C. & Hillen, W. Gene regulation by tetracyclines. *European Journal of Biochemistry* **270**, 3109-3121, doi:10.1046/j.1432-1033.2003.03694.x (2003).
35 Wang, B., Kitney, R. I., Joly, N. & Buck, M. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. *Nat Commun* **2**, 508, doi:http://www.nature.com/ncomms/journal/v2/n10/suppinfo/ncomms1516_S1.html (2011).
36 O'Shaughnessy, E. C., Palani, S., Collins, J. J. & Sarkar, C. A. Tunable Signal Processing in Synthetic MAP Kinase Cascades. *Cell* **144**, 119-131, doi:http://dx.doi.org/10.1016/j.cell.2010.12.014 (2011).
37 Haldimann, A. *et al.* Altered recognition mutants of the response regulator PhoB: A new genetic strategy for studying protein–protein interactions. *Proceedings of the National Academy of Sciences* **93**, 14361-14366 (1996).

38 Rao, C. V. Expanding the synthetic biology toolbox: engineering orthogonal regulators of gene expression. *Current Opinion in Biotechnology* **23**, 689-694, doi:http://dx.doi.org/10.1016/j.copbio.2011.12.015 (2012).
39 Ruder, W. C., Lu, T. & Collins, J. J. Synthetic Biology Moving into the Clinic. *Science* **333**, 1248-1252, doi:10.1126/science.1206843 (2011).
40 Khalil, A. S. & Collins, J. J. Synthetic Biology: applications come of age. *Nature Reviews Genetics* **11**, 13 (2010).
41 Cheng, A. A. & Lu, T. K. Synthetic Biology: An Emerging Engineering Discipline. *Annual Review of Biomedical Engineering* **14**, 155-178, doi:doi:10.1146/annurev-bioeng-071811-150118 (2012).
42 Weber, W. *et al.* Streptomyces-derived quorum-sensing systems engineered for adjustable transgene expression in mammalian cells and mice. *Nucleic Acids Research* **31**, e71, doi:10.1093/nar/gng071 (2003).
43 Nevozhay, D., Zal, T. & Balazsi, G. Transferring a synthetic gene circuit from yeast to mammalian cells. *Nature Communications* **4** (2013).
44 Koornneef, M. & Meinke, D. The development of Arabidopsis as a model plant. *The Plant Journal* **61**, 909-921, doi:10.1111/j.1365-313X.2009.04086.x (2010).
45 Zurbriggen, M. D., Moor, A. & Weber, W. Plant and bacterial systems biology as platform for plant synthetic bio(techno)logy. *Journal of Biotechnology* **160**, 80-90, doi:http://dx.doi.org/10.1016/j.jbiotec.2012.01.014 (2012).
46 Rodriguez, M. C., Petersen, M. & Mundy, J. Mitogen-activated protein kinase signaling in plants. *Annu Rev Plant Biol* **61**, 621-649 (2010).
47 Benjamins, R. & Scheres, B. Auxin: The Looping Star in Plant Development. *Annual Review of Plant Biology* **59**, 443-465, doi:doi:10.1146/annurev.arplant.58.032806.103805 (2008).
48 Fankhauser, C. & Staiger, D. Photoreceptors in Arabidopsis thaliana: light perception, signal transduction and entrainment of the endogenous clock. *Planta* **216**, 1-16, doi:10.1007/s00425-002-0831-4 (2002).
49 Aoyama, T. & Chua, N.-H. A glucocorticoid-mediated transcriptional induction system in transgenic plants. *The Plant Journal* **11**, 605-612, doi:10.1046/j.1365-313X.1997.11030605.x (1997).
50 Padidam, M., Gore, M., Lily Lu, D. & Smirnova, O. Chemical-Inducible, Ecdysone Receptor-Based Gene Expression System for Plants. *Transgenic Res* **12**, 101-109, doi:10.1023/A:1022113817892 (2003).
51 Love, J., Allen, G. C., Gatz, C. & Thompson, W. F. Differential Top10 promoter regulation by six tetracycline analogues in plant cells. *Journal of Experimental Botany* **53**, 1871-1877, doi:10.1093/jxb/erf050 (2002).
52 Looger, L. L., Dwyer, M. A., Smith, J. J. & Hellinga, H. W. Computational design of receptor and sensor proteins with novel functions. *Nature* **423**, 185-190, doi:http://www.nature.com/nature/journal/v423/n6936/suppinfo/nature01556_S1.html (2003).
53 Dwyer, M. A. & Hellinga, H. W. Periplasmic binding proteins: a versatile superfamily for protein engineering. *Current Opinion in Structural Biology* **14**, 495-504, doi:http://dx.doi.org/10.1016/j.sbi.2004.07.004 (2004).
54 Morey, K. J. *et al.* in *Methods in Enzymology* Vol. Volume 497 (ed Voigt Chris) 581-602 (Academic Press, 2011).
55 Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-Component Signal Transduction. *Annual Review of Biochemistry* **69**, 183-215, doi:doi:10.1146/annurev.biochem.69.1.183 (2000).

56 Batchelor, E. & Goulian, M. Robustness and the cycle of phosphorylation and dephosphorylation in a two-component regulatory system. *Proceedings of the National Academy of Sciences* **100**, 691-696, doi:10.1073/pnas.0234782100 (2003).
57 Igoshin, O. A., Alves, R. & Savageau, M. A. Hysteretic and graded responses in bacterial two-component signal transduction. *Molecular Microbiology* **68**, 1196-1215, doi:10.1111/j.1365-2958.2008.06221.x (2008).
58 Russo, F. D. & Silhavy, T. J. The essential tension: opposed reactions in bacterial two-component regulatory systems. *Trends in Microbiology* **1**, 306-310, doi:http://dx.doi.org/10.1016/0966-842X(93)90007-E (1993).
59 Schaller, G. E., Kieber, J. J. & Shiu, S.-H. Two-Component Signaling Elements and Histidyl-Aspartyl Phosphorelays†. *The Arabidopsis Book*, e0112, doi:10.1199/tab.0112 (2008).
60 Falke, J. J., Bass, R. B., Butler, S. L., Chervitz, S. A. & Danielson, M. A. THE TWO-COMPONENT SIGNALING PATHWAY OF BACTERIAL CHEMOTAXIS: A Molecular View of Signal Transduction by Receptors, Kinases, and Adaptation Enzymes. *Annual Review of Cell and Developmental Biology* **13**, 457-512, doi:doi:10.1146/annurev.cellbio.13.1.457 (1997).
61 Yaghmai, R. & Hazelbauer, G. L. Strategies for differential sensory responses mediated through the same transmembrane receptor. *The EMBO journal* **12**, 1897-1905 (1993).
62 Marvin, J. S. & Hellinga, H. W. Conversion of a maltose receptor into a zinc biosensor by computational design. *Proceedings of the National Academy of Sciences* **98**, 4955-4960, doi:10.1073/pnas.091083898 (2001).
63 Rodrigo, G. *et al.* Vanillin cell sensor. *Synthetic Biology, IET* **1**, 74-78, doi:10.1049/iet-stb:20060003 (2007).
64 Utsumi, R. *et al.* Activation of bacterial porin gene expression by a chimeric signal transducer in response to aspartate. *Science* **245**, 4 (1989).
65 Wanner, B. L. Gene regulation by phosphate in enteric bacteria. *Journal of Cellular Biochemistry* **51**, 47-54, doi:10.1002/jcb.240510110 (1993).
66 Mack, T. R., Gao, R. & Stock, A. M. Probing the Roles of the Two Different Dimers Mediated by the Receiver Domain of the Response Regulator PhoB. *Journal of Molecular Biology* **389**, 349-364, doi:http://dx.doi.org/10.1016/j.jmb.2009.04.014 (2009).
67 Albrecht, T. *Characterization of a Synthetic Signal Transduction System* Ph.D thesis, Colorado State University, (2012).
68 Yamada, H. *et al.* The Arabidopsis AHK4 Histidine Kinase is a Cytokinin-Binding Receptor that Transduces Cytokinin Signals Across the Membrane. *Plant and Cell Physiology* **42**, 1017-1023, doi:10.1093/pcp/pce127 (2001).
69 Van Dien, S. J. & Keasling, J. D. A Dynamic Model of theEscherichia coliPhosphate-Starvation Response. *Journal of Theoretical Biology* **190**, 37-49, doi:http://dx.doi.org/10.1006/jtbi.1997.0524 (1998).
70 Gao, R. & Stock, A. M. Probing kinase and phosphatase activities of two-component systems in vivo with concentration-dependent phosphorylation profiling. *Proceedings of the National Academy of Sciences* **110**, 672-677, doi:10.1073/pnas.1214587110 (2013).
71 Makino, K. *et al.* Regulation of the phosphate regulon of Escherichia coli: Activation of pstS transcription by PhoB protein in vitro. *Journal of Molecular Biology* **203**, 85-95, doi:http://dx.doi.org/10.1016/0022-2836(88)90093-9 (1988).
72 Kimura, S., Makino, K., Shinagawa, H., Amemura, M. & Nakata, A. Regulation of the phosphate regulon of Escherichia coli: Characterization of the promoter of the pstS gene. *Molec. Gen. Genet.* **215**, 374-380, doi:10.1007/BF00427032 (1989).

73 Pelaz, S., Tapia-López, R., Alvarez-Buylla, E. R. & Yanofsky, M. F. Conversion of leaves into petals in Arabidopsis. *Current Biology* **11**, 182-184, doi:http://dx.doi.org/10.1016/S0960-9822(01)00024-0 (2001).
74 van Mourik, S. *et al.* Continuous-time modeling of cell fate determination in Arabidopsis flowers. *BMC Systems Biology* **4**, 101 (2010).
75 Alvarez-Buylla, E. R., Azpeitia, E., Barrio, R., Benítez, M. & Padilla-Longoria, P. From ABC genes to regulatory networks, epigenetic landscapes and flower morphogenesis: Making biological sense of theoretical approaches. *Seminars in Cell & Developmental Biology* **21**, 108-117, doi:http://dx.doi.org/10.1016/j.semcdb.2009.11.010 (2010).
76 Yadav, R. K. *et al.* Plant stem cell maintenance involves direct transcriptional repression of differentiation program. *Mol Syst Biol* **9**, doi:http://www.nature.com/msb/journal/v9/n1/suppinfo/msb20138_S1.html (2013).
77 Clairambault, J. in *Encylopedia of Systems Biology* Vol. 2 (eds Werner Dubitzky, Olaf Wolkenhauer, Kwang-Hyun Cho, & Hiroki Yokota) 1109 (Springer Science+Business Media LLC, New York, 2013).
78 Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic Gene Expression in a Single Cell. *Science* **297**, 1183-1186, doi:10.1126/science.1070919 (2002).
79 McNaught, A. D. & Wilkinson, A. in *the "Gold Book"* (Blackwell Scientfic Publications, Oxford, 1997).
80 Carmany, D. O., Hollingsworth, K. & McCleary, W. R. Genetic and Biochemical Studies of Phosphatase Activity of PhoR. *Journal of Bacteriology* **185**, 1112-1115, doi:10.1128/jb.185.3.1112-1115.2003 (2003).
81 Yoshida, T., Cai, S. j. & Inouye, M. Interaction of EnvZ, a sensory histidine kinase, with phosphorylated OmpR, the cognate response regulator. *Molecular Microbiology* **46**, 1283-1294, doi:10.1046/j.1365-2958.2002.03240.x (2002).
82 Dyson, F. A meeting with Enrico Fermi. *Nature* **427**

1(2004).
83 Muller-Hill, B. *The Lac Operon*. 207 (de Gruyter, 1996).
84 Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2 edn, 1.12-1.13 (Cold Spring Harbor Laboratory Press, 1989).
85 Salis, H. M., Mirsky, E. A. & Voigt, C. A. Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotech* **27**, 946-950, doi:http://www.nature.com/nbt/journal/v27/n10/suppinfo/nbt.1568_S1.html (2009).
86 Cayley, S., Lewis, B. A., Guttman, H. J. & Record Jr, M. T. Characterization of the cytoplasm of Escherichia coli K-12 as a function of external osmolarity: Implications for protein-DNA interactions in vivo. *Journal of Molecular Biology* **222**, 281-300, doi:http://dx.doi.org/10.1016/0022-2836(91)90212-O (1991).
87 Glascock, C. B. & J. Weickert, M. Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in Escherichia coli. *Gene* **223**, 221-231, doi:http://dx.doi.org/10.1016/S0378-1119(98)00240-6 (1998).
88 Lopilato, J. E., Garwin, J. L., Emr, S. D., Silhavy, T. J. & Beckwith, J. R. D-ribose metabolism in Escherichia coli K-12: genetics, regulation, and transport. *Journal of Bacteriology* **158**, 665-673 (1984).
89 Michalodimitrakis, K. M., Sourjik, V. & Serrano, L. Plasticity in amino acid sensing of the chimeric receptor Taz. *Molecular Microbiology* **58**, 257-266, doi:10.1111/j.1365-2958.2005.04821.x (2005).

90 Parkinson, J. S. Signaling Mechanisms of HAMP Domains in Chemoreceptors and Sensor Kinases. *Annual Review of Microbiology* **64**, 101-122, doi:doi:10.1146/annurev.micro.112408.134215 (2010).
91 Manson, M. D. Transmembrane Signaling Is Anything but Rigid. *Journal of Bacteriology* **193**, 5059-5061, doi:10.1128/jb.05874-11 (2011).
92 Ferris, Hedda U. *et al.* The Mechanisms of HAMP-Mediated Signaling in Transmembrane Receptors. *Structure* **19**, 378-385, doi:http://dx.doi.org/10.1016/j.str.2011.01.006 (2011).
93 Zhou, L., Gregori, G., Blackman, J. M., Robinson, J. P. & Wanner, B. L. Stochastic activation of the response regulator PhoB by noncognate histidine kinases. *Journal of Integrative Bioinformatics* **2** (2005).
94 McCleary, W. R. The activation of PhoB by acetylphosphate. *Molecular Microbiology* **20**, 1155-1163, doi:10.1111/j.1365-2958.1996.tb02636.x (1996).
95 Blanco Alexandre, G., Canals, A. & Coll, M. Vol. 393 1165 (2012).
96 Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proceedings of the National Academy of Sciences* **97**, 6640-6645, doi:10.1073/pnas.120163297 (2000).
97 Koman, A., Harayama, S. & Hazelbauer, G. L. Relation of Chemotactic Response to the Amount of Receptor: Evidence for Different Efficiencies of Signal Transduction. *Journal of Bacteriology* **138**, 739-747 (1979).
98 Liang, J. C., Chang, A. L., Kennedy, A. B. & Smolke, C. D. A high-throughput, quantitative cell-based screen for efficient tailoring of RNA device activity. *Nucleic Acids Research* **40**, e154, doi:10.1093/nar/gks636 (2012).
99 Hooshangi, S., Thiberge, S. & Weiss, R. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3581-3586, doi:10.1073/pnas.0408507102 (2005).
100 Robinson, J. P. *Handbook of flow cytometry methods*. (Wiley-Liss, Inc., 1993).
101 Moser, F. Sample Dilution Affects Fluorescence Measurements in the LSR Fortessa Flow Cytometer. *SBC Technical Report* (2013).
102 Skerker, J. M. *et al.* Rewiring the Specificity of Two-Component Signal Transduction Systems. *Cell* **133**, 1043-1054, doi:http://dx.doi.org/10.1016/j.cell.2008.04.040 (2008).
103 Schreier, B., Stumpp, C., Wiesner, S. & Höcker, B. Computational design of ligand binding is not a solved problem. *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.0907950106 (2009).
104 Keasling, J. D. Gene-expression tools for the metabolic engineering of bacteria. *Trends in Biotechnology* **17**, 452-460, doi:http://dx.doi.org/10.1016/S0167-7799(99)01376-1 (1999).
105 Schleif, R. Regulation of the l-arabinose operon of Escherichia coli. *Trends in Genetics* **16**, 559-565, doi:http://dx.doi.org/10.1016/S0168-9525(00)02153-3 (2000).
106 Whitaker, W. R., Davis, S. A., Arkin, A. P. & Dueber, J. E. Engineering robust control of two-component system phosphotransfer using modular scaffolds. *Proceedings of the National Academy of Sciences* **109**, 18090-18095, doi:10.1073/pnas.1209230109 (2012).
107 Siegele, D. A. & Hu, J. C. Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proceedings of the National Academy of Sciences* **94**, 8168-8172 (1997).
108 De Mey, M., Maertens, J., Lequeux, G., Soetaert, W. & Vandamme, E. Construction and model-based analysis of a promoter library for E. coli: an indispensable tool for metabolic engineering. *BMC Biotechnology* **7**, doi:doi:10.1186/1472-6750-7-34 (2007).
109 Dolezel, J., Greilhuber, J. & Suda, J. 479 (WILEY-VCH Verlag GmbH & Co. , Weinheim, 2007).

110 Koncz, C., Langridge, W. H. R., Olsson, O., Schell, J. & Szalay, A. A. Bacterial and firefly luciferase genes in transgenic plants: Advantages and disadvantages of a reporter gene. *Developmental Genetics* **11**, 224-232, doi:10.1002/dvg.1020110308 (1990).
111 OW, D. W. *et al.* Transient and Stable Expression of the Firefly Luciferase Gene in Plant Cells and Transgenic Plants. *Science* **234**, 856-859, doi:10.1126/science.234.4778.856 (1986).
112 Clough, S. J. & Bent, A. F. Floral dip: a simplified method forAgrobacterium-mediated transformation ofArabidopsis thaliana. *The Plant Journal* **16**, 735-743, doi:10.1046/j.1365-313x.1998.00343.x (1998).
113 Wang, K., Hererra-strella, A. & van Montagu, M. *Transformation of Plants and Soil Microorganisms*. (Cambridge University Press, 2004).
114 Kim, S.-I. & Gelvin, S. B. Genome-wide analysis of Agrobacterium T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. *The Plant Journal* **51**, 779-791, doi:10.1111/j.1365-313X.2007.03183.x (2007).
115 Antunes, M. S. *et al.* A synthetic de-greening gene circuit provides a reporting system that is remotely detectable and has a re-set capacity. *Plant Biotechnology Journal* **4**, 605-622, doi:10.1111/j.1467-7652.2006.00205.x (2006).
116 Seki, T., Yoshikawa, H., Takahashi, H. & Saito, H. Nucleotide sequence of the Bacillus subtilis phoR gene. *Journal of Bacteriology* **170**, 5935-5938 (1988).
117 Liu, W. & Hulett, F. M. Bacillus subtilis PhoP binds to the phoB tandem promoter exclusively within the phosphate starvation-inducible promoter. *Journal of Bacteriology* **179**, 6302-6310 (1997).
118 Qi, Y., Kobayashi, Y. & Hulett, F. M. The pst operon of Bacillus subtilis has a phosphate-regulated promoter and is involved in phosphate transport but not in regulation of the pho regulon. *Journal of Bacteriology* **179**, 2534-2539 (1997).
119 Death, A. & Ferenci, T. Between feast and famine: endogenous inducer synthesis in the adaptation of Escherichia coli to growth with limiting carbohydrates. *Journal of Bacteriology* **176**, 5101-5107 (1994).
120 Raj, A. & van Oudenaarden, A. Nature, Nurture, or Chance: Stochastic Gene Expression and Its Consequences. *Cell* **135**, 216-226, doi:http://dx.doi.org/10.1016/j.cell.2008.09.050 (2008).
121 Tamsir, A., Tabor, J. & Voigt, C. A. Robust multicellular computing using genetically encoded nor gates and chemical 'wires'. *Nature* **469**, 4, doi:http://dx.doi.org/10.1038/nature09565 (2011).
122 Fisher, S. L., Kim, S.-K., Wanner, B. L. & Walsh, C. T. Kinetic Comparison of the Specificity of the Vancomycin Resistance Kinase VanS for Two Response Regulators, VanR and PhoB†. *Biochemistry* **35**, 4732-4740, doi:10.1021/bi9525435 (1996).
123 Devlin, P. F., Christie, J. M. & Terry, M. J. Many hands make light work. *Journal of Experimental Botany* **58**, 3071-3077, doi:10.1093/jxb/erm251 (2007).